Notes de cours de l'ISIMA, deuxième année
http://www.isima.fr/~leborgne

Intégrale de Lebesgue, introduction

Gilles Leborgne

28 octobre 2024

Table des matières

0 Rappels

Soit E un ensemble et $\mathcal{P}(E)$ l'ensemble des sous-ensembles de E (donc $A \in \mathcal{P}(E) \Leftrightarrow A \subset E$).

0.1 Fonctions indicatrices, en escalier, étagées

Définition 0.1 Soit $A \subset E$. On appelle fonction indicatrice de A (ou fonction caractéristique de A) la fonction $1_A : E \to \mathbb{R}$ définie par :

$$1_A(x) \begin{cases} = 1 & \text{si } x \in A \quad \text{(présence)}, \\ = 0 & \text{si } x \notin A \quad \text{(absence)}. \end{cases} \tag{0.1}$$

(NB, vocabulaire en optimisation : la fonction caractéristique de A est $\chi_A : E \to \overline{\mathbb{R}}$ définie par $\chi_A(x) = 0$ si $x \in A$, et $\chi_A(x) = \infty$ si $x \notin A$.)

Définition 0.2 Une fonction en escalier est une fonction $f : E \to \mathbb{R}$ t.q.

$$\exists n \in \mathbb{N}^*,\ \exists (a_i, b_i, c_i)_{i=1,...,n} \in \mathbb{R}^{3n} \text{ t.q. } \forall i \in [1,n]_{\mathbb{N}},\ a_i < b_i,\ f = \sum_{i=1}^{n} c_i 1_{(a_i,b_i)}, \tag{0.2}$$

où (a_i, b_i) désigne un intervalle ouvert ou fermé ou semi-ouvert.

Définition 0.3 Une fonction $f : E \to \mathbb{R}$ est étagée (is a simple function) ssi f est une combinaison linéaire de fonctions indicatrices d'ensembles :

$$\exists n \in \mathbb{N}^*, \quad \exists (c_i)_{i=1,...,n} \in \mathbb{R}^n, \quad \exists (A_i)_{i=1,...,n} \in \mathcal{P}(E)^n, \quad f = \sum_{i=1}^{n} c_i 1_{A_i}. \tag{0.3}$$

(Ici les ensembles A_i ne sont pas nécessairement disjoints.)

Définition équivalente : Une fonction $f : E \to \mathbb{R}$ est étagée ssi son image $\mathrm{Im} f$ est un ensemble fini : $\exists n \in \mathbb{N}^*$, $\exists (c_i)_{i=1,...,n} \in \mathbb{R}^n$, $\mathrm{Im} f = \{c_1, ..., c_n\}$; donc notant $A_i = f^{-1}(c_i) = \{x \in E : f(x) = c_i\}$, on a (0.3).

Exemple 0.4 La fonction $1_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ est étagée (non en escalier) : $n = 1$, $c_1 = 1$ et $A_1 = \mathbb{Q}$ dans (0.3). ∴

Remarque 0.5 Pour l'intégrale de Lebesgue les fonctions en escalier ne sont pas suffisantes : on aura besoin des images réciproques $1_A^{-1}(J)$ où J est un intervalle de $\mathbb{R}$ (alors que l'intégrale de Riemann est basée sur les images directes des $1_A(I)$ où I est un intervalle de $\mathbb{R}$). En particulier on aura besoin de : ∴

Proposition 0.6 *Si $f : E \to \overline{\mathbb{R}_+} = [0, \infty]$ ("fonction à valeurs positives"), alors il existe une suite croissante $(f_n)_{n \in \mathbb{N}^*}$ de fonctions étagées qui converge simplement vers f :*

$$\forall x \in E, \quad \lim_{n \to \infty} f_n(x) = f(x),$$

i.e., $\forall x \in E$, $\forall \varepsilon > 0$, $\exists N_{x,\varepsilon} \in \mathbb{N}$, $\forall n \geq_{x,\varepsilon}$, $|f_n(x) - f(x)| < \varepsilon$. Idem pour f de signe quelconque.

Si de plus f est bornée alors la convergence est uniforme : $||f - f_n||_\infty \underset{n \to \infty}{\longrightarrow} 0$, où $||g||_\infty \overset{\text{déf}}{=} \sup_{x \in E} |g(x)|$.

Preuve. $f \geq 0$: pour $n \in \mathbb{N}^*$ et $i = 1, ..., n2^n$ on pose :

$$B_{n,i} = f^{-1}([\frac{i-1}{2^n}, \frac{i}{2^n}[) = \{x : \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n}\} \quad \text{et} \quad F_n = f^{-1}([n, \infty[) = \{x : f(x) \geq n\}. \tag{0.4}$$

Donc, pour tout n, $(\bigcup_{i=1}^{n2^n} B_{n,i}) \bigcup F_n$ est une partition de E. Puis on définit les fonctions étagées f_n par :

$$f_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} 1_{B_{n,i}} + n 1_{F_n} \qquad (= \sum_{i=2}^{n2^n} \frac{i-1}{2^n} 1_{B_{n,i}} + n 1_{F_n}). \tag{0.5}$$

Faire un dessin. Il est immédiat que $f_n \leq f$, que $F_{n+1} \subset F_n$, que

$$\text{pour} \quad j = 2i-1, 2i, \quad \frac{i-1}{2^n} \leq \frac{j-1}{2^{n+1}}, \quad \frac{j}{2^{n+1}} \leq \frac{i}{2n}, \quad \text{donc} \quad B_{n+1,j} \subset B_{n,i}, \tag{0.6}$$

et donc $f_n(x) \leq f_{n+1}(x)$ pour tout $x \in E$: la suite (f_n) est croissante majorée par f.

Soit $x \in E$ t.q. $f(x) < \infty$, puis n t.q. $f(x) < n$, puis le $i \in [1, n2^n]_{\mathbb{N}}$ t.q. $x \in B_{n,i}$ i.e.

$$f_n(x) = \frac{i-1}{2^n} \le f(x) < \frac{i}{2^n}. \tag{0.7}$$

Alors $|f(x) - f_n(x)| = f(x) - f_n(x) < \frac{1}{2^n}$, donc (f_n) converge simplement vers f en x. Et si $f(x) = \infty$ on a $f_n(x) = n \longrightarrow_{n\to\infty} \infty$ en x, donc (f_n) converge simplement vers f. Donc (f_n) converge simplement vers f dans E.

Cas f de signe quelconque : on a $f = f_+ - f_-$ où f_+ et f_- sont les fonctions positives définies par $f_+ = \sup(f, 0)$ et $f_- = -\min(f, 0)$; on construit les suites (f_{n+}) et (f_{n-}) associés, et on pose $f_n = f_{n+} - f_{n-}$. La suite (f_n) convient.

Cas f bornée ≥ 0. Notons $C = ||f||_\infty = \sup_{x\in E} |f(x)|$. Alors pour tout $n > C$ on a $F_n = \emptyset$, et donc $f_n = \sum_{i=1}^{n2^n} \frac{i-1}{n} 1_{B_{n,i}}$ pour tout $n > C$. Donc pour tout $x \in E$ on a $|f(x) - f_n(x)| < \frac{1}{2^n}$ dès que $n > C$, par définition des $B_{n,i}$. D'où $||f - f_n||_\infty \longrightarrow_{n\to\infty} 0$. Cas f signe quelconque : démarche précédente. ▪

Exercice 0.7 Soit $f = 1_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ la fonction indicatrice de $\mathbb{Q}$ l'ensemble des rationnels. Que valent F_n et les B_n dans la preuve précédente, et que vaut donc f_n ?

Réponse. (La fonction $1_{\mathbb{Q}}$ est déjà une fonction étagée au sens de la définition 0.3.) $F_1 = \mathbb{Q}$, $f_1 = 1_{\mathbb{Q}} = f$. Et $(f_n)_{\mathbb{N}^*}$ croissante $\le f$, donc $f_n = f = 1_{\mathbb{Q}}$ pour tout n.

Vérif : Pour $n \ge 2$, $F_n = \emptyset$ et $B_{n,i} = \emptyset$ sauf $B_{n,1} = \{x : 0 \le 1_{\mathbb{Q}}(x) < \frac{1}{2^n}\} = \mathbb{R} - \mathbb{Q}$ et $B_{n,2^n+1} = \{x : 1 \le 1_{\mathbb{Q}}(x) < \frac{2^n+1}{2^n}\} = \mathbb{Q}$, donc $f_n = 1_{B_{n,2^n+1}} = 1_{\mathbb{Q}} = f$ pour $n \ge 2$. ▪

Exercice 0.8 Soit $A, B \in \mathcal{P}(E)$. Montrer : $1_{A\cup B} = 1_A 1_B$ ssi $A = B$.

Réponse. On a $1_A 1_B = 1_{A\cap B}$ car $1_A(x) 1_B(x) = 1$ ssi $x \in A$ et $x \in B$.

Donc $1_{A\cup B} = 1_A 1_B$ ssi $A \cup B = A \cap B$.

Et $A \cup B = A \cap B$ ssi $A \subset A \cap B$ et $B \subset A \cap B$, ssi $A \subset B$ et $B \subset A$, ssi $A = B$. ▪

0.2 Espace topologique

Soit un ensemble E.

Définition 0.9 Une topologie sur E est un ensemble de sous-ensembles de E, i.e. $\mathcal{O} \subset \mathcal{P}(E)$, tel que :

1. $\emptyset$ et E sont éléments de $\mathcal{O}$,
2. $\mathcal{O}$ est stable par intersection ***finie***, i.e., pour tout $n \in \mathbb{N}^*$ et pour toute famille (finie) $(U_i)_{i=1,\dots,n} \in \mathcal{O}^n$ on a $\bigcap_{i=1}^n U_i \in \mathcal{O}$,
3. $\mathcal{O}$ est stable par union quelconque, i.e., pour toute famille $(U_i)_{i\in I}$ où I est un ensemble quelconque on a $\bigcup_{i\in I} U_i \in \mathcal{O}$.

Les éléments U de $\mathcal{O}$ sont appelés les ouverts de E (ou les ensembles ouverts dans E).

Le couple $(E, \mathcal{O})$ est appelé espace topologique.

Exemple 0.10 Topologie $\mathcal{O}_{\mathbb{R}}$ usuelle de $\mathbb{R}$: celle engendrée par les intervalles ouverts $]a, b[$ pour $a < b$. Elle est appelée topologie borélienne, ou topologie usuelle. Et par exemple $U_i =]-\frac{1}{i}, 1[$ donne $\bigcap_{i=1}^\infty U_i = [0, 1[$ qui n'est pas ouvert : d'où le ***finie*** dans le 2. de la définition 0.9. ▪

Exemple 0.11 Topologie usuelle de $\mathbb{R}^2$: celle engendrée par les rectangles ouverts $]a, b[\times]c, d[$ pour $a < b$ et $c < d$ (les pavés ouverts). Elle est également appelée topologie borélienne, ou topologie usuelle. Idem dans $\mathbb{R}^n$ à l'aide des pavés ouverts. ▪

Exemple 0.12 Pour tout E, $\mathcal{O} = \{\emptyset, E\}$ est une topologie appelée la topologie grossière. Exclue dans la suite. ▪

Exemple 0.13 Pour tout E, $\mathcal{O} = \mathcal{P}(E)$ est une topologie appelée la topologie discrète. Utilisée pour la mesure de Dirac, la mesure de comptage, les probabilités discrètes... ▪

Définition 0.14 Si $(E, \mathcal{O}_E)$ et $(Z, \mathcal{O}_Z)$ sont deux espaces topologiques, une application $f : E \to Z$ est dite continue ssi l'image réciproque de tout ouvert de Z est un ouvert de E, i.e. ssi $\forall V \in \mathcal{O}_Z$ on a $f^{-1}(V) \in \mathcal{O}_E$.

Définition 0.15 Soit I un ensemble, soit $(A_i)_{i\in I}$ une famille de sous-ensembles de E, et soit $F = \bigcup_I A_i \subset E$. La topologie engendrée par la famille $(A_i)_{i\in I}$ est la plus petite des topologies sur l'ensemble F pour laquelle les A_i sont ouverts (i.e. l'intersection de toutes les topologies pour lesquelles les A_i sont ouverts).

Exercice 0.16 Soit $(A_i)_{i\in I}$ une famille de sous-ensembles de E contenant $\emptyset$, et soit $\mathcal{O}_F$ la topologie engendrée par $(A_i)_{i\in I}$ dans $F=\bigcup_I A_i$. Montrer que $U\in\mathcal{O}_F$ ssi U est une union quelconque d'intersections finies de A_i.

Réponse. Soit $G=\{B\subset E : \exists n\in\mathbb{N},\ \exists(i_k)_{k=1,...,n}\in I^n,\ B=\bigcap_{k=1}^n A_{i_k}\}$ l'ensemble constitué de toutes les intersections finies de A_i, et soit $H=\{Z\subset E : \exists J \text{ ensemble },\ \forall j\in J,\ \exists B_j\in G,\ Z=\bigcup_{j\in J}B_j\}$ l'ensemble constitué des unions quelconques d'éléments de G.

Il s'agit de montrer que $\mathcal{O}_F=H$.

$\supset$: soit $Z\in H$; donc $Z=\bigcup_{j\in J}B_j$ où $B_j=\bigcap_{k=1}^{n_j}A_{i_k}$ pour tout j. Donc $B_j\in\mathcal{O}_F$ pour tout j (intersection finie), donc $Z\in\mathcal{O}_F$ (union quelconque). Donc $H\subset\mathcal{O}_F$.

$\subset$: montrons que H est une topologie : si c'est le cas, comme elle contient F, elle contient la plus petite des topologies engendrées, donc $H\supset\mathcal{O}_F$.

On a $\emptyset\in H$ par hypothèse. Comme $F=\bigcup_I A_i$, on a $F\in H$. Il est immédiat que H est stable par union quelconque puisqu'une union quelconque d'union quelconques est encore une union quelconque. Il reste à montrer que H est stable par intersection finie. Soient $Z=\bigcup_{\alpha\in J_1}B_\alpha$ et $Y=\bigcup_{\beta\in J_2}C_\beta$ deux éléments quelconques de H, les B_α et les C_β étant des intersections finies de A_i. Comme l'union et l'intersection sont distributives l'une par rapport à l'autre, on a $Z\cap Y=(\bigcup_{\alpha\in J_1}B_\alpha)\bigcap(\bigcup_{\beta\in J_2}C_\beta)=\bigcup_{(\alpha,\beta)\in J_1\times J_2}(B_\alpha\bigcap C_\beta)$ qui est bien une union quelconque d'intersections finies. Pour un nombre fini d'intersection, on fait une récurrence immédiate. ▪

Exemple 0.17 $\mathcal{O}_\mathbb{R}$ étant la topologie (usuelle) engendrée par les intervalles ouverts de $\mathbb{R}$, on déduit de l'exercice précédent que tout ouvert est une union quelconque d'intervalles ouverts. En effet, une intersection finie d'intervalles est un intervalle : soit $X=]a_1,b_1[\cap]a_2,b_2[$. Si $a_1\le a_2$, on a soit $b_1\le a_2$ et alors $X=\emptyset$, soit $a_2\le b_1\le b_2$ et alors $X=]a_2,b_1[$, soit $b_2\le b_1$ et alors $X=]a_1,b_1[$. Et idem si $a_2\le a_1$. Et récurrence immédiate. ▪

Exercice 0.18 $\mathcal{O}_\mathbb{R}$ étant la topologie (usuelle) engendrée par les intervalles ouverts de $\mathbb{R}$, on a mieux : montrer : si $U\in\mathcal{O}_\mathbb{R}$ (un ouvert), alors U est une union au plus dénombrable (unique) d'intervalles ouverts disjoints. (Ainsi $]0,1[\cup]2,3[$ est un ouvert réunion de 2 intervalles ouverts ; et si U est ouvert alors U est “au pire” une réunion dénombrable d'intervalles ouverts.)

Réponse. Soit $x\in U$. Comme U est ouvert et les intervalles ouverts formant une base de voisinage, il existe un intervalle ouvert inclus dans U contenant x. Notons $]a_x,b_x[$ le plus grand intervalle ouvert dans U contenant x, où $a_x,b_x\in\overline{\mathbb{R}}$. D'où $U=\bigcup_{x\in U}]a_x,b_x[$ (en particulier si U est connexe alors U est constitué d'un seul intervalle). Et dans chaque $]a_x,b_x[$, $\exists q_x\in\mathbb{Q}$ t.q. $q_x\in]a_x,b_x[$, avec immédiatement $I_{q_x}=I_x$. Donc $U=\bigcup_{q_x\in U}I_{q_x}$. Comme $\mathbb{Q}$ est dénombrable, l'union est au plus dénombrable. Et par construction les ouverts sont disjoints. Et l'unicité d'une telle union est immédiate. ▪

Exercice 0.19 Malheureusement le résultat précédent ne tient pas dans $\mathbb{R}^2$ (et plus généralement dans $\mathbb{R}^n$ pour $n\ge 2$) au sens : les pavés ouverts $]a_1,b_1[\times]a_2,b_2[$ de $\mathbb{R}^2$ engendrent la topologie usuelle de $\mathbb{R}^2$, mais un ouvert de $\mathbb{R}^2$ n'est pas une union dénombrable de pavés ouverts (non dégénérés) disjoints. Le montrer pour le cercle unité. (Puis voir l'exercice suivant.)

Réponse. Soit $D=\{\vec{x} : ||\vec{x}||^2<1\}$. Supposons qu'il existe une union $\bigcup_{\mathbb{N}^*}P_i$ constituée de pavés ouverts disjoints non dégénérés telle que $\bigcup_{\mathbb{N}^*}P_i=D$. Notons $P_1=]x_1,x_2[\times]y_1,y_2[$, avec $(x_2-x_1)(y_2-y_1)\ne 0$ car P_1 non dégénéré. Et on a $(x_1,\frac{y_1+y_2}{2})\notin P_1$ (car P_1 est ouvert et $(x_1,\frac{y_1+y_2}{2})$ est sur son bord, dessin), avec $(x_1,\frac{y_1+y_2}{2})\in D$ (strictement convexe), et $D=\bigcup_{\mathbb{N}^*}P_i$, donc $\exists k\in\mathbb{N}^*$, $k>1$, $(x_1,\frac{y_1+y_2}{2})\in P_k$. Avec $P_1\cap P_k\ne\emptyset$ (union disjointe) donc $(x_1,\frac{y_1+y_2}{2})$ sur le bord de P_k, donc non dans P_k (ouvert). Absurde. ▪

Exercice 0.20 Définition : Soit $(F_i=[a_i,b_i])_{i\in I}$ une famille d'intervalles fermés t.q. $a_i<b_i$ pour tout i. Cette famille est dite presque disjointe ssi les intervalles ouverts $]a_i,b_i[$ sont deux à deux disjoints.

Montrer que si $U\in\mathcal{O}_\mathbb{R}$ (i.e. U est un ouvert), alors il existe une famille dénombrable $(F_i=[a_i,b_i])_{i\in\mathbb{N}^*}$ presque disjointe d'intervalles fermés t.q. $U=\bigcup_{i\in\mathbb{N}^*}F_i$. Et cette propriété est conservée dans $\mathbb{R}^n$ où les F_i sont des pavés.

Réponse. On pave $\mathbb{R}^n$ par les pavés cubiques $[a_1,a_1+q]\times...\times[a_n,a_n+q]$ où les a_i sont des multiples de 2^{-m} et $q=2^{-m}$ (pavés de volume $2^{-m})^n$). (Voir par exemple Rudin [17] p. 48.) ▪

0.3 Intégrale de Riemann

Rappel, l'intégrale de Riemann étant supposée connue. Soit $f:[a,b]\to\mathbb{R}$ continue, $-\infty<a<b<\infty$. Soit $n\in\mathbb{N}^*$, $h=\frac{b-a}{n}$, et $x_i=a+ih$ pour $i=0,1,...,n$. Donc $[a,b]=\bigcup_{i=1}^n[x_{i-1},x_i]$ avec $x_i-x_{i-1}=h$ pour tout i. Dessin. Soit $\xi_i\in[x_{i-1},x_i]$ un point dans le i-ème intervalle $[x_{i-1},x_i]$. On approxime f par la fonction en escalier définie par : $x\in]x_{i-1},x_i[\Rightarrow f_n(x)=f(\xi_i)$ (la valeur de f_n au point x_i n'a pas d'importance). Donc $f_n=\sum_{i=1}^n f(\xi_i)1_{]x_{i-1},x_i[}$, définie “presque partout” dans $[a,b]$ (voir la suite).

L'aire $A(f)=^{\text{noté}}\int_a^b f(x)\,dx$ (notation de Riemann) “sous la courbe f” est approximée par $A_n(f)=\sum_{i=1}^n(x_i-x_{i-1})f(\xi_i)$, somme des aires des rectangles de largeur (x_i-x_{i-1}) et de hauteur $f(\xi_i)$. Et $A_n\longrightarrow_{n\to\infty}A$.

1 Ensembles, tribus et mesures

1.1 Tribus, espaces et ensembles mesurables, tribu borélienne

Soit E un ensemble. Pour $A \subset E$, on note $A^C = E - A$ le complémentaire de A dans E.

Définition 1.1 Étant donné un ensemble E, une tribu (ou σ-algèbre) est un ensemble $\mathcal{A} \subset \mathcal{P}(E)$ de parties de E tel que :

1. $\emptyset$ et E sont éléments de $\mathcal{A}$,
2. $\mathcal{A}$ est stable par complémentation, i.e., si $A \in \mathcal{A}$, alors $A^C \in \mathcal{A}$.
3. $\mathcal{A}$ est stable par union et intersection dénombrable, i.e. si $(A_i)_{i\in\mathbb{N}^*} \in \mathcal{A}^{\mathbb{N}^*}$, alors $\bigcap_{i\in\mathbb{N}^*} A_i \in \mathcal{A}$ et $\bigcup_{i\in\mathbb{N}^*} A_i \in \mathcal{A}$,

Remarque 1.2 Dans la définition précédente, il y a des redondances : une définition “minimale” équivalente est (i) $E \in \mathcal{A}$, (ii) stabilité par complémentation, (iii) stabilité par union dénombrable. En effet (ii) et (i) impliquent $E^C = \emptyset \in \mathcal{A}$ et (ii) et (iii) impliquent $(\bigcup_{i\in\mathbb{N}^*} A_i)^C = \bigcap_{i\in\mathbb{N}^*} A_i^C \in \mathcal{A}$. ▟

Définition 1.3 Un ensemble E muni d'une tribu $\mathcal{A}$ constitue un espace mesurable noté $(E, \mathcal{A})$ (espace sur lequel on va pouvoir construire des mesures).

Et les ensembles $A \in \mathcal{A}$ sont appelés les ensembles mesurables de E.

Exemple 1.4 $\mathcal{A} = \{\emptyset, E\}$ est la tribu grossière d E. ▟

Exemple 1.5 $\mathcal{A} = \mathcal{P}(E)$ est la tribu discrète de E (elle contient tous les singletons $\{x\}$ pour $x \in E$). ▟

Exemple 1.6 Pour $A \subset E$, $A \neq \emptyset$, $A \neq E$, $\mathcal{A} = \{\emptyset, A, A^C, E\}$ est une tribu (importante en probabilités). ▟

Exercice 1.7 Montrer que si $A, B \in \mathcal{A}$, alors $B - A \in \mathcal{A}$.

Réponse. $B - A = B \bigcap A^C$. ▟

Proposition 1.8 *Une intersection quelconque de tribus de E est une tribu de E.*

Preuve. Un élément de l'intersection appartient à toutes les tribus et donc vérifie 1., 2., 3. ▟

Définition 1.9 Soit $F = (A_i)_{i\in I}$ une famille de sous-ensembles de E, où I est un ensemble quelconque (dénombrable ou non). L'intersection de toutes les tribus contenant $(A_i)_{i\in I}$ est appelée tribu engendrée par la famille $(A_i)_{i\in I}$, souvent notée $\sigma(F)$ (σ-algèbre engendrée par la famille F). Construite à lan remarque 1.13.

Définition 1.10 Si $(E, \mathcal{O})$ est un espace topologique, on appelle tribu borélienne la tribu $\sigma(\mathcal{O})$ engendrée par les ouverts de E (c'est la plus petite tribu contenant la topologie $\mathcal{O}$, ou encore, c'est l'intersection de toutes les tribus qui contiennent la topologie $\mathcal{O}$). Ses éléments sont appelés les boréliens.

Exemple 1.11 Cas particulier du langage : la tribu borélienne de $\mathbb{R}$, notée $\mathcal{A}_\mathbb{R}$, est la tribu engendrée par les intervalles ouverts. Donc dans $\mathbb{R}$, l'usage du mot tribu borélienne indique que $\mathbb{R}$ a été préalablement muni de sa topologie usuelle, celle engendrée par les intervalles ouverts.

On vérifie immédiatement que les intervalles ouverts, fermés, semi-fermés, bornés ou non, sont dans la tribu borélienne (en particulier, les singletons $A = \{x\} = [x, x]$ sont dans la tribu borélienne). Et on voit que $\mathcal{A}_\mathbb{R}$ est également, par exemple, la tribu engendrée par les intervalles $]-\infty, b[$ pour $b \in \mathbb{R}$.

(La tribu borélienne de $\mathbb{R}$ est strictement incluse dans $\mathcal{P}(\mathbb{R})$, voir remarque 1.37. On montre également que la tribu borélienne de $\mathbb{R}$ a “seulement” la puissance du continu, dit de cardinal $2^\mathbb{N}$, contrairement à $P(\mathbb{R})$.) ▟

Proposition 1.12 *(Tribu par restriction.) Soit $(E, \mathcal{A}_E)$ un ensemble mesurable, soit $F \subset E$, et soit $\mathcal{A}_F = \{A \cap F : A \in \mathcal{A}_E\}$. Alors $\mathcal{A}_F$ est une tribu sur F, et $(F, \mathcal{A}_F)$ est un ensemble mesurable.*

Preuve. 1- On a $A = \emptyset \in \mathcal{A}_E$, donc $\emptyset \cap F = \emptyset \in \mathcal{A}_F$. 2- Soit $B \in \mathcal{A}_F$, donc $B = A \cap F$ avec $A \in \mathcal{A}_E$. On a $A^C \in \mathcal{A}_E$ (tribu), donc $F - B = F - (A \cap F) = F \cap A^C \in \mathcal{A}_F$: le complémentaire de B dans F appartient à $\mathcal{A}_F$. 3- Soit $(B_i)_{\mathbb{N}^*}$ une famille dans $\mathcal{A}_F$. Donc $B_i = A_i \cap F$ pour tout i où $A_i \in \mathcal{A}_E$, donc $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n (A_i \cap F) = (\bigcup_{i=1}^n A_i) \cap F \in \mathcal{A}_F$ (distributivité). ▟

Remarque 1.13 La construction d'une tribu engendrée $\sigma(Z)$ par une famille $Z = (A_i)_{i\in I}$ de sous-ensembles de E n'est pas aussi simple que celle de la construction d'une topologie engendrée par cette même famille, cf. exercice 0.16. Cette construction peut être faite par "récurrence transfinie" (ou "induction transfinie" ou "construction transfinie").

Exemple pour la tribu borélienne de $\mathbb{R}$ on a $Z = \mathcal{O}$ l'ensemble des ouverts. Les étapes sont :

0- si Z est une famille d'ensembles, on note Z_σ l'ensemble des unions dénombrables et Z_δ l'ensemble des intersections dénombrables.

1- Soit $Z_0 = \{A \subset \mathbb{R} : A \in \mathcal{O}$ ou $A^C \in \mathcal{O}\}$ l'ensemble des ouverts et des fermés ("$= \mathcal{O}\bigcup(\mathcal{O})^C$").

2- Soit $Z_1 = (Z_0)_\sigma \bigcup (Z_0)_\delta$ et par récurrence $Z_{n+1} = (Z_n)_\sigma \bigcup (Z_n)_\delta$.

3- On montre que $\sigma(Z) = \bigcup_{n=1}^\infty Z_n$, i.e. que tout élément de la tribu appartient à un Z_n (quitte à prendre n suffisamment grand) : le sens $\supset$ est immédiat. Le sens $\subset$ est moins immédiat : il s'agit de montrer que $Z_\infty = \bigcup_{n=1}^\infty Z_n$ est bien une tribu : il est immédiat que toute autre tribu contenant $\mathcal{O}$ doit contenir Z_∞. Tout d'abord il est non vide car contient $Z_0 \supset \mathcal{O}$.

La stabilité par passage au complémentaire : si $A \in Z_\infty$ alors il existe $n \in \mathbb{N}$ tel que $A \in Z_n$. Montrons que $A^C \in Z_n$ (donc $\in Z_\infty$) : c'est vrai pour $n=0$ par construction de Z_0. Supposons que ce soit vrai pour $n{-}1$ (i.e. si $B \in Z_{n-1}$ alors $B^C \in Z_{n-1}$). Si $A \in Z_n$ alors A est de la forme $A = (\bigcup_\mathbb{N} B_i)\bigcup(\bigcap_\mathbb{N} C_j)$ où les $B_i, C_j \in Z_{n-1}$, donc $A^C = (\bigcup_\mathbb{N} B_i)^C \bigcap (\bigcap_\mathbb{N} C_j)^C = (\bigcap_\mathbb{N} B_i^C)\bigcap(\bigcup_\mathbb{N} C_j^C) \subset (\bigcap_\mathbb{N} B_i^C)\bigcup(\bigcup_\mathbb{N} C_j^C) \in (Z_{n-1})_\delta \bigcup (Z_{n-1})_\sigma = Z_n$: on a bien $A^C \in Z_n$.

La stabilité par union dénombrable : soit $(A_i)_{i\in\mathbb{N}}$ une famille dénombrable dans Z_∞. Notons $\alpha_i \in \mathbb{N}$ tel que $A_i \in Z_{\alpha_i}$. On a alors $\bigcup_{i=0}^\infty A_i \in \bigcup_{i=0}^\infty Z_{\alpha_i} \subset Z_\infty$.

Rappel : le translaté de t d'un ensemble A est l'ensemble $t{+}A = \{x = t{+}a,\ a \in A\}$ (i.e. $x \in t{+}A$ ssi $\exists a \in A$ t.q. $x = t{+}a$). En particulier le translaté d'un intervalle est un intervalle. On en déduit que le translaté d'un borélien dans $\mathbb{R}$ est un borélien. ■

1.2 Mesure d'ensembles, espace mesuré

Soit $(E, \mathcal{A})$ un ensemble mesurable.

Définition 1.14 Une mesure positive (ou plus simplement une mesure) sur la tribu $\mathcal{A}$ est une application positive non nulle $\mu : \left\{ \begin{array}{l} \mathcal{A} \to \overline{\mathbb{R}_+} = [0,\infty] \\ A \to \mu(A) \end{array} \right\}$ t.q. :

1. $\mu(\emptyset) = 0$,
2. μ est σ-additive (propriété d'additivité dénombrable), i.e., pour toute famille dénombrable $(A_i)_{i\in\mathbb{N}^*}$ d'ensembles mesurables 2 à 2 disjoints on a $\mu(\bigcup_{i\in\mathbb{N}^*} A_i) = \sum_{i=1}^\infty \mu(A_i)$:
$$\text{si } (A_i)_{i\in\mathbb{N}^*} \in \mathcal{A}^{\mathbb{N}^*} \text{ et si, } \forall i,j \in \mathbb{N}^*,\ A_i \bigcap A_j = \emptyset, \text{ alors } \mu(\bigcup_{i\in\mathbb{N}^*} A_i) = \sum_{i=1}^\infty \mu(A_i). \tag{1.1}$$
3. E est réunion dénombrable d'éléments de $\mathcal{A}$ de mesure finie, i.e., $E = \bigcup_{i=1}^\infty A_i$ avec $\mu(A_i) < \infty$ pour tout $i \in \mathbb{N}^*$.

Et $\mu(A)$ est appelée la mesure de l'ensemble mesurable $A \in \mathcal{A}$.

Définition 1.15 Un espace mesurable $(E, \mathcal{A})$ muni d'une mesure μ est noté $(E, \mathcal{A}, \mu)$, et est appelé espace mesuré.

Cas particulier : si $\mu(E) = 1$, alors μ est appelée mesure de probabilité, et $(E, \mathcal{A}, \mu)$ est appelé espace probabilisé.

Exercice 1.16 Montrer que 2. implique 1. (donc 1. est redondant dans la définition 1.14).

Réponse. On prend $A_3 = A_4 = ... = \emptyset$ dans 2, et donc quand $A_1 \cap A_2 = \emptyset$ on a $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$.

Puis soit $A_1 \in \mathcal{A}$ t.q. $\mu(A_1) < \infty$. Un tel A_1 existe car une mesure vérifie 3.. Et $A_2 = \emptyset \in \mathcal{A}$. Et A_1 et $\emptyset$ sont disjoints ($A_1 \cap \emptyset = \emptyset$). D'où $\mu(A_1) = \mu(A_1 \cup \emptyset) = \mu(A_1) + \mu(\emptyset)$, d'où $\mu(\emptyset) = 0$. ■

Exercice 1.17 Montrer que si $A, B \in \mathcal{A}$ et $A \subset B$, alors $\mu(A) \le \mu(B)$.

Réponse. $B = A \bigcup (B - A)$ avec $A \bigcap (B - A) = \emptyset$. ■

Exemple 1.18 $E = \mathbb{R}$ muni de sa tribu borélienne $\mathcal{A}_\mathbb{R}$: la mesure usuelle μ_ℓ, dite mesure de Lebesgue, est définie par : $\forall a, b \in \mathbb{R}$, $a \le b$,
$$\mu_\ell([a,b]) = b - a \qquad \text{(la longueur de } [a,b]). \tag{1.2}$$
En particulier $\mu_\ell([a,a]) = 0 = \mu_\ell(\{a\})$ pour tout a : on dit que μ_ℓ est une mesure "diffuse" (ou sans atome ou de densité). Vérifier que μ_ℓ est une mesure sur $\mathcal{A}_\mathbb{R}$ n'est pas simple (le point 3. de la définition 1.14 est immédiat : $\mathbb{R} = \bigcup_{n\in\mathbb{Z}} [n-1, n]$).

Point le 2. de la définition 1.14, voir par exemple Muthukumar [10], Rudin [17], Villani [18]... (Le § 8.6 donne une idée de la démonstration à l'aide de la mesure extérieure).

Ainsi l'espace $(\mathbb{R}, \mathcal{A}, \mu_\ell)$ est un espace mesuré. ▪

Exemple 1.19 Sur la tribu $\mathcal{A} = \{\emptyset, E\}$, la fonction μ définie par $\mu(\emptyset) = 0$ et $\mu(E) = 1$ est une mesure.

En revanche, toujours sur $\mathcal{A} = \{\emptyset, E\}$, la fonction définie par $\mu(\emptyset) = 0$ et $\mu(E) = \infty$ n'est pas une mesure (le point 3. n'est pas satisfait). ▪

Exemple 1.20 (Mesure de Dirac $\mu = \delta_x$) Soit E un ensemble muni de sa tribu discrète $\mathcal{A} = \mathcal{P}(E)$, et soit un point $x \in E$. Pour $A \in \mathcal{A}$ on pose :

$$\delta_x(A) \begin{cases} = 1 & \text{si } x \in A, \\ = 0 & \text{si } x \notin A, \end{cases} \qquad \text{i.e.} \qquad \delta_x(A) = 1_A(x), \tag{1.3}$$

où 1_A est la fonction indicatrice de A, cf. (0.1). On vérifie que $\mu = \delta_x$ est bien une mesure : le point 3. est donné par $\delta_x(E) = \delta_x(\{x\}) = 1$. Le point 2. est donné par : si $(A_i)_{\mathbb{N}^*}$ est une famille d'ensembles mesurables 2 à 2 disjoints, alors un au plus des A_i contient x.

Remarque : si $c \in E$ alors δ_c est une mesure de probabilité car $\delta_c(E) = 1$, i.e. $1_E(c) = 1$.

Comme $\delta_x(\{x\}) = 1$, la mesure δ_x n'est pas une mesure "diffuse" (la mesure d'un singleton n'est pas nécessairement nulle), contrairement à la mesure de Lebesgue. ▪

Exemple 1.21 (Mesure de comptage) On prend $E = (x_i)_{i\in\mathbb{N}^*}$ un ensemble dénombrable et $\mathcal{A} = \mathcal{P}(E)$ (tribu discrète). On pose :

$$\mu = \sum_{i=1}^{\infty} \delta_{x_i}, \qquad \text{i.e.} \qquad \forall A \in \mathcal{A}, \quad \mu(A) = \sum_{i=1}^{\infty} 1_A(x_i). \tag{1.4}$$

$\mu(A)$ compte le nombre de points x_i dans A. On vérifie que c'est une mesure. ▪

Exemple 1.22 (Mesure atomique) On prend E un ensemble quelconque et $\mathcal{A} = \mathcal{P}(E)$. On se donne une suite $(x_i)_{i\in\mathbb{N}^*}$ de points de E, et une suite $(\lambda_i)_{i\in\mathbb{N}^*}$ de réels, et on pose

$$\mu = \sum_{i=1}^{\infty} \lambda_i \delta_{x_i}, \qquad \text{i.e.} \qquad \forall A \in \mathcal{A}, \quad \mu(A) = \sum_{i=1}^{\infty} \lambda_i 1_A(x_i). \tag{1.5}$$

On vérifie que c'est une mesure. En particulier δ_x est une mesure atomique. ▪

Exemple 1.23 (Mesure de Stieljes) On prend $E = \mathbb{R}$, $\mathcal{A}$ la tribu borélienne et $F : \mathbb{R} \to \mathbb{R}$ une fonction réelle croissante. On pose, pour tout $a, b \in \mathbb{R}$ si $a < b$:

$$\mu_F([a, b]) \stackrel{\text{déf}}{=} F(b+) - F(a-), \tag{1.6}$$

$F(b+) = \lim_{h>0,h\to 0} F(b + h)$ et $F(a-) = \lim_{h>0,h\to 0} F(a - h)$. Faire un dessin. Et F étant croissante, on a $F(b+) = \inf_{x>b} F(x)$ et $F(a-) = \sup_{x<a} F(x)$ qui existent toujours.

F est appelée primitive de μ_F (définie à une constante près).

Cas particulier F est continue : $\mu_F([a, b]) =^{\text{déf}} F(b) - F(a)$.

Et en particulier lorsque $F(x) = x$, alors μ_F est la mesure de Lebesgue. ▪

Exercice 1.24 Vérifiez que pour une mesure de Stieljes donnée et $b > a$ on a :

$$\begin{cases} \mu_F([a, b[) = F(b-) - F(a-), \\ \mu_F(]a, b]) = F(b+) - F(a+), \\ \mu_F(]a, b[) = F(b-) - F(a+), \\ \mu_F(\{a\}) = F(a+) - F(a-) \stackrel{\text{noté}}{=} [F](a) \quad (\text{saut de } F \text{ en } a). \end{cases} \tag{1.7}$$

Réponse. Par définition $\mu_F(\{a\}) = \mu_F([a, a]) = F(a+) - F(a-)$, i.e. $(1.7)_4$.

Et $[a, b] = \{a\} \bigcup]a, b]$, union disjointe, donne $\mu_F([a, b]) = \mu_F(\{a\}) + \mu_F(]a, b])$, soit $F(b+) - F(a-) = F(a+) - F(a-) + \mu_F(]a, b])$. D'où $(1.7)_2$. Idem pour les autres cas.

Autre démonstration. Pour tout $h > 0$ tel que $b-h > a$:

$$\mu_F([a, b-h]) = F((b-h)+) - F(a-) \xrightarrow[h\to 0]{} F(b-) - F(a-),$$

car $F((b-h)+) \le F(b-)$ pour tout $h>0$, et $F((b-h)+) \ge F(b - \varepsilon)$ pour tout ε dès que $h < \varepsilon$. De même pour les autres cas. ▪

Exemple 1.25 Mesure de probabilité : c'est le cas de la mesure de Stieljes où F (croissante) est continue à droite et est telle que $F(-\infty) = 0$ et $F(\infty) = 1$. En particulier on a :

$$\mu_F(]-\infty, t]) = F(t). \tag{1.8}$$

Et F est appelée fonction de répartition de la probabilité μ.

Exemple : $F = 1_{[c,\infty[}$ est la fonction de répartition de la probabilité δ_c sur $\mathbb{R}$, et F est une primitive de δ_c, voir cours de distributions. ∴

1.3 Propriété essentielle : une mesure est monotone

Définition 1.26 Soit $(A_n)_{1\leq n\leq\infty} \in \mathcal{P}(E)^{\mathbb{N}^*}$ une suite de sous-ensembles de E.

Elle est : croissante ssi $\forall n \in \mathbb{N}^*$, $A_{n+1}\supset A_n$; décroissante ssi $\forall n \in \mathbb{N}^*$, $A_{n+1}\subset A_n$; monotone ssi elle est croissante ou décroissante.

Si $(A_n)_{\mathbb{N}^*}$ est croissante, on note $\displaystyle\bigcup_{n\in\mathbb{N}^*} A_n \overset{\text{noté}}{=} \lim_{n\to\infty} A_n$.

Si $(A_n)_{\mathbb{N}^*}$ est décroissante, on note $\displaystyle\bigcap_{n\in\mathbb{N}^*} A_n \overset{\text{noté}}{=} \lim_{n\to\infty} A_n$.

Proposition 1.27 et définition. *Soit $(E, \mathcal{A}, \mu)$ est un espace mesuré et $(A_n)_{n\in\mathbb{N}^*}$ une suite dans $\mathcal{A}$.*

1- Si $(A_n)_{\mathbb{N}^*}$ est croissante alors $\mu(\displaystyle\bigcup_{n\in\mathbb{N}^*} A_n) = \lim_{n\to\infty} \mu(A_n)$, i.e.

$$(A_n)_{n\in\mathbb{N}^*} \text{ croissante} \quad \Rightarrow \quad \mu(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mu(A_n). \tag{1.9}$$

On dit qu'une mesure μ est croissante. Donc $\lim_{n\to\infty} \mu(A_n) = \mu(\lim_{n\to\infty} A_n)$ pour les suites croissantes, et on dit : "on passe à la limite sous μ pour les suites croissantes".

2- Si $(A_n)_{n\in\mathbb{N}^*}$ est décroissante **et** *si de plus $\mu(A_1) < \infty$ alors on a également (1.9).*

Preuve. 1- $\bigcup_{n\in\mathbb{N}^*} A_n \in \mathcal{A}$ car $\mathcal{A}$ est une tribu. On pose $B_1 = A_1$ et $B_{n+1} = A_{n+1} - A_n$ pour $n\geq 1$ (les B_n sont les "couches successives de l'oignon A"). On a $B_i \in \mathcal{A}$ car $\mathcal{A}$ est une tribu, $A_n = \bigcup_{k=1}^n B_k$ (immédiat), et les B_k sont 2 à 2 disjoints (immédiat), donc $\bigcup_{n\in\mathbb{N}^*} A_n = \bigcup_{k\in\mathbb{N}^*} B_k$ et $\mu(\bigcup_{n\in\mathbb{N}^*} A_n) = \sum_{k\in\mathbb{N}^*} \mu(B_k)$ car μ (mesure) est σ-additive, avec $\sum_{k\in\mathbb{N}^*} \mu(B_k) = \lim_{n\to\infty}(\sum_{k=1}^n \mu(B_k))$ (série de termes positifs) $= \lim_{n\to\infty}(\mu(\sum_{k=1}^n B_k))$ (somme finie) $= \lim_{n\to\infty} \mu(A_n)$, cqfd.

2- Démonstration similaire. N.B. : l'hypothèse $\mu(A_1) < \infty$ est nécessaire : sinon eg. prendre $A_n = \{n, n+1, ...\}$, on a $(A_n)_{\mathbb{N}^*}$ décroissante, $A_n \longrightarrow_{n\to\infty} \emptyset$, donc $\mu(\lim_{n\to\infty} A_n) = 0$, alors que $\mu(A_n) = \infty$ pour tout n. ∴

Remarque 1.28 On considère $(\mathbb{R}, \mathcal{A}_\mathbb{R})$ l'espace borélien réel usuel, et μ_ℓ la mesure de Lebesgue sur $\mathbb{R}$. Alors si $A \in \mathcal{A}_\mathbb{R}$, on a $\mu_\ell(A) = \inf\{\mu_\ell(U) : U \in \mathcal{O},\ U \supset A\}$.

Pour le voir, on définit $\mu : \mathcal{A}_\mathbb{R} \to \mathbb{R}_+$ par $\mu(A) = \inf\{\mu_\ell(U) : U \in \mathcal{O},\ U \supset A\}$. On montre que μ est bien une mesure, et qu'elle est égale à μ_ℓ sur les intervalles ouverts. ∴

1.4 Ensembles μ-négligeables, ensemble de Cantor

1.4.1 Ensembles μ-négligeables

Soit $(E, \mathcal{A}, \mu)$ un ensemble mesuré et $A \in \mathcal{A}$

Définition 1.29 A est μ-négligeable ssi $\mu(A) = 0$.

Exemple 1.30 Pour la mesure de Dirac δ_x, si $x \notin A$, alors $\delta_x(A) = 0$ et A est δ_x-négligeable. Et si $x \in A$, alors $\delta_x(A) = 1$ et A n'est pas négligeable. ∴

Exemple 1.31 Pour la mesure de Lebesgue μ_ℓ, tout ensemble dénombrable est négligeable (eg. $\mu_\ell(\mathbb{Q}) = 0$). Et un ouvert non vide n'est jamais négligeable (il contient un intervalle ouvert non vide). ∴

1.4.2 Ensemble de Cantor

Pour la mesure de Lebesgue μ_ℓ, il existe des ensembles négligeables qui sont non dénombrables. Exemple : l'ensemble de Cantor, construit comme limite de la suite de fermés emboités dont les premiers termes sont :

$$\begin{aligned} A_0 &= [0,1] \\ A_1 &= [0,\frac{1}{3}]\bigcup[\frac{2}{3},1] \\ A_2 &= [0,\frac{1}{9}]\bigcup[\frac{2}{9},\frac{1}{3}]\bigcup[\frac{2}{3},\frac{7}{9}]\bigcup[\frac{8}{9},1] \end{aligned} \tag{1.10}$$

(On enlève $\frac{1}{3}$ au centre de chaque intervalle restant. Faire un dessin.)

La suite (A_n) est décroissante. L'ensemble $A_\infty = \lim_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} A_n$ est mesurable borélien comme intersection dénombrable de fermés. Et il est immédiat que $\mu_\ell(A_n) = \frac{2}{3}\mu_\ell(A_{n-1}) = (\frac{2}{3})^n$ et donc qu'à la limite $\mu_\ell(A_\infty) = 0$, i.e. A_∞ est μ_ℓ-négligeable.

Et A_∞ est non dénombrable : A_∞ contient les bords de A_n pour tout n, et sachant que A_0 a 2 bords et A_{n+1} a deux fois plus de bords que A_n, on déduit que A_n a 2^{n+1} bords. D'où A_∞ contient au moins $2^{\mathbb{N}}$ éléments (puissance du continu voir annexe A) et est donc non dénombrable.

Remarque 1.32 A_∞ est compact comme intersection dénombrable $\bigcap_{n\in\mathbb{N}} A_n$ de compacts (chaque A_n est par construction fermé et borné). Et A_∞ n'a pas d'intérieur (il ne contient aucun ouvert, ou encore son complémentaire est partout dense) : s'il contenait un ouvert, il contiendrait un intervalle ouvert $]x-\varepsilon, x+\varepsilon[$ pour $x \in \mathbb{R}$ et un $\varepsilon > 0$. Or par construction A_n contient des intervalles disjoints de largeur $\frac{1}{3^n}$. Donc quel que soit $\varepsilon > 0$ en choisissant n t.q. $\frac{1}{3^n} < \varepsilon$, A_n ne contient pas de tel intervalle, et $A_\infty \subset A_n$. ▪

Remarque 1.33 L'ensemble de Cantor permet de définir une fonction f continue croissante sur $[0,1]$ telle que $f(0) = 0$ et $f(1) = 1$ dérivable presque partout, i.e. dérivable en tout point de $[0,1] - A$ avec A négligeable, et telle que $f'(x) = 0$ presque partout, i.e. pour tout $x \in [0,1] - A$, où A est négligeable.

C'est la fonction définie qui sur $[\frac{1}{3}, \frac{2}{3}]$ vaut $\frac{1}{2}$ (intervalle manquant de A_1), qui sur $[\frac{1}{9}, \frac{2}{9}]$ et $[\frac{7}{9}, \frac{8}{9}]$ vaut respectivement $\frac{1}{4}$ et $\frac{3}{4}$ (cf intervalles manquant de A_2)... Et f est bien constante au voisinage de chaque point de $[0,1] - A$. Faire un dessin.

Pour la mesure de Lebesgue, f est donc dérivable "partout où on la voit", car un ensemble A négligeable (= de mesure nulle) est un ensemble "invisible" pour l'instrument de mesure "la mesure de Lebesgue". Et partout où on la voit on a $f'(x) = 0$ (on dira plus loin que $f' = 0$ preque partout), mais f n'est pas constante (bien que f soit continue). En particulier $f(x) \neq \int_0^x f'(t)\,dt$: donc f n'est pas la primitive de sa dérivée (là où elle existe), bien que $\int_0^x f'(t)\,dt$ soit bien définie au sens de l'intégrale de Lebesgue (voir plus loin).

Cette fonction de Cantor est un exemple de fonction qui est continue mais qui n'est pas absolument continue, i.e. telle que f n'est pas l'intégrale de sa dérivée : $f(x) - f(0) \neq \int_0^x f'(t)\,dt$. Attention, ici la dérivée n'est définie que presque partout, et la notation de l'intégrale est celle de Lebesgue (l'intégrale de Riemann est insuffisante). ▪

Exercice 1.34 On a vu que l'ensemble de Cantor est un ensemble non dénombrable d'intérieur vide et de mesure nulle. Construire un ensemble B non dénombrable d'intérieur vide et de mesure non nulle.

Réponse. On reprend le procédé de Cantor en partant de $B_0 = [0,1]$, mais au lieu supprimer un intervalle de largeur $\frac{1}{3}$ au milieu, on supprime un intervalle de largeur $1-c_1$ où $c_1 \in]0,1[$. On obtient l'ensemble $B_1 = [0, \frac{c_1}{2}] \cup [1-\frac{c_1}{2}, 1]$ avec $\mu_\ell(B_1) = c_1$ et B_1 contient deux intervalles de longueur $\frac{c_1}{2}$, donc $\mu_\ell(B_1) = c_1$. Puis on continue en supprimant un intervalle de largeur $1-c_2$,où $c_2 \in]0,1[$, pour obtenir $B_2 = [0, \frac{c_2}{2}\frac{c_1}{2}] \cup [\frac{c_1}{2} - \frac{c_2}{2}\frac{c_1}{2}, \frac{c_1}{2}] \cup ...$ avec donc $\mu_\ell(B_2) = c_1 c_2$, et B_2 ne contient que des intervalles de longueur $\frac{c_1}{2}\frac{c_2}{2}...$

On a construit un ensemble B_n tel que $\mu_\ell(B_n) = c_1 c_2 ... c_n$ et B_n ne contient que des intervalles de longueur $\frac{c_1 \dots c_n}{2^n}$. À la limite on obtient l'ensemble B_∞ tel que $\mu_\ell(B_\infty) = c_1 c_2 ... c_n ....$ Soit $c \in]0,1[$. Choisissons les $c_i \in]0,1[$ tels que $\mu_\ell(B_\infty) = c$ (en particulier tous les B_n vérifient $\mu_\ell(B_n) \geq c$). Donc $\prod_{i=1}^\infty c_i = c$, donc $\sum_{i=1}^\infty \ln(c_i) = \ln(c)$, donc $\sum_{i=1}^\infty \frac{\ln(c_i)}{\ln(c)} = 1 = \sum_{i=1}^\infty \frac{1}{2^i}$. On choisit les c_i t.q. $\frac{\ln(c_i)}{\ln(c)} = \frac{1}{2^i}$, soit $c_i = c^{\frac{1}{2^i}}$. Comme B_∞ a le même nombre de points que l'ensemble de Cantor, cet ensemble est non dénombrable. Et comme pour l'ensemble de Cantor, B_∞ n'a pas d'intérieur puisque B_n ne contient que des intervalles de largeur $\frac{c_1 \dots c_n}{2^n} < \frac{1}{2^n}$ (pour Cantor on a $c_n = \frac{2}{3}$ pour tout n). Donc $B = B_\infty$ convient. ▪

1.5 Tribu étendue $\mathcal{A}_\mu$

Il est pratique mathématiquement de considérer une tribu "étendue" :

Définition 1.35 Soit $B \in \mathcal{P}(E)$. S'il existe $A \in \mathcal{A}$ μ-négligeable ($\mu(A) = 0$) t.q. $B \subset A$, alors B est également dit μ-négligeable (bien que non mesurable a priori), et on note $\mu(B) = 0$.

Un ensemble $B \in \mathcal{P}(E)$ est dit μ-mesurable s'il est égal μ-presque-partout à un ensemble A mesurable ($A \in \mathcal{A}$), i.e. si $A \bigcap B^C$ et $B \bigcap A^C$ sont μ-négligeables (faire un dessin), i.e. si :

$$\mu(A \bigcap (E - B)) = 0 = \mu(B \bigcap (E - A)). \tag{1.11}$$

Et on dit alors que $B = A$ μ-presque-partout, et on définit la mesure de B par $\mu(B) = \mu(A)$.

Et si B est μ-négligeable, on dit que μ est portée par $B^C = E - B$.

Proposition 1.36 *Les ensembles μ-mesurables engendrent une nouvelle tribu $\mathcal{A}_\mu$ (qui contient $\mathcal{A}$).*

Preuve. On le vérifie facilement. Exercice. ▪

Remarque 1.37 En général, $\mathcal{A}_\mu \subsetneq \mathcal{P}(E)$, i.e. il existe en général des sous-ensembles de E qui sont non μ-mesurables.

Exemple (Vitali, 1905) dans $\mathbb{R}$ muni de sa tribu borélienne et de la mesure de Lebesgue μ_ℓ :

(i) On considère la relation d'équivalence : $x\mathcal{R}y$ ssi $y-x \in \mathbb{Q}$ (réflexive, symétrique, transitive : immédiat). Pour $x \in \mathbb{R}$, on note $\dot{x} = \{x+q : q \in \mathbb{Q}\}$ la classe d'équivalence de x.

(ii) L'axiome du choix permet de former un sous-ensemble A de $\mathbb{R}$ en choisissant un et un seul point a_x dans chaque classe d'équivalence $\dot{x}$. (Et A est l'ensemble de Vitali= l'ensemble des représentants choisis.)

(ii)' Supposons que A est μ_ℓ-mesurable.

(iii) Montrons alors que $\bigcup_{q\in\mathbb{Q}}(q+A) = \{q+A : q \in \mathbb{Q}\} =^{\text{noté}} P$ est une partition dénombrable de $\mathbb{R}$.

(iii)$_1$ L'union vaut $\mathbb{R}$: soit $x \in \mathbb{R}$, soit a_x le représentant de $\dot{x}$ dans A, donc $q = x-a_x \in \mathbb{Q}$, $x = q+a_x \in q+A$, donc $x \in P$, donc $P \supset \mathbb{R}$, donc $P = \mathbb{R}$.

(iii)$_2$ L'union est une partition : sinon $\exists q_1,q_2 \in \mathbb{Q}$, $q_1 \neq q_2$, et $\exists x \in (q_1+A)\cap(q_2+A)$; donc $x = q_1+a_1 = q_2+a_2$ où $a_1,a_2 \in A$, donc $a_2-a_1 = q_2-q_1 \in \mathbb{Q}$, donc $a_1\mathcal{R}a_2$, donc $\dot{a}_1 = \dot{a}_2$, donc $a_1 = a_2$ (le représentant choisi par construction de A), donc $q_2 = q_1$. Absurde puisque $q_1 \neq q_2$. Donc l'union est une partition.

(iv) Supposons que A est mesurable, i.e. $A \in \mathcal{A}$ (tribu borélienne de $\mathbb{R}$). Donc tout translaté $B_q = q+A \in \mathcal{A}$ est un borélien, voir remarque 1.13. Donc si $A \in \mathcal{A}_\mu$ (tribu borélienne étendue de $\mathbb{R}$) alors $B_q = q+A \in \mathcal{A}_\mu$. donc l'union dénombrable $P = \bigcup_{q\in\mathbb{Q}}(q+A) \in \mathcal{A}_\mu$.

On a $[0,1[= [0,1[\cap\mathbb{R} = [0,1[\cap P = [0,1[\bigcap(\bigcup_{q\in\mathbb{Q}}(q+A)) = \bigcup_{q\in\mathbb{Q}}([0,1[\bigcap(q+A))$, union disjointe de $[0,1[$ (car $[0,1[\bigcap(q+A) \subset (q+A)$ et(iii)). Donc

$$
\begin{aligned}
1 = \mu_\ell([0,1[) &= \sum_{q\in\mathbb{Q}} \mu_\ell\Big([0,1[\bigcap(q+A)\Big) = \sum_{q\in\mathbb{Q}} \mu_\ell\Big([-q,1-q[\bigcap A\Big) \\
&= \sum_{q\in[0,1[\bigcap\mathbb{Q}} \{\sum_{k\in\mathbb{Z}} \mu_\ell\Big([-q+k,1-q+k[\bigcap A\Big)\} = \sum_{q\in[0,1[\cap\mathbb{Q}} \mu_\ell\{\bigcup_{k\in\mathbb{Z}}\Big([-q+k,1-q+k[\bigcap A\Big)\},
\end{aligned}
$$

car l'union "$\bigcup_{k\in\mathbb{Z}}$" est disjointe (puisque si $k_1 \neq k_2$ alors $[-q+k_1,1-q+k_1[\bigcap[-q+k_2,1-q+k_2[= \emptyset$). Et donc (distributivité de $\cap$) :

$$
1 = \sum_{q\in[0,1[\cap\mathbb{Q}} \mu_\ell(\mathbb{R}\bigcap A) = \sum_{q\in[0,1[\cap\mathbb{Q}} \mu_\ell(A). \tag{1.12}
$$

(v) (1.12) est absurde : on ne peut n'y avoir $\mu_\ell(A) = 0$, ni avoir $\mu_\ell(A) \neq 0$: donc (ii)' est absurde, donc A n'est pas μ_ℓ-mesurable. ∴

Remarque 1.38 Dans les applications "ingénieur" des mathématiques appliquées, les sous-ensembles de $\mathbb{R}$ seront tous mesurables pour la mesure de Lebesgue : on ne sait pas exhiber explicitement un ensemble qui n'est pas mesurable (même si on sait qu'il en existe grâce à l'axiome du choix, cf. remarque précédente). ∴

1.6 Notion de μ-presque partout (μ-p.p.)

Rappel : Soit E un ensemble. Soit Z une affirmation (une assertion) sur E. Soit :

$$
V = \{x \in E : Z(x) \text{ vrai}\} \quad \text{et} \quad F = \{x \in E : Z(x) \text{ faux}\} \quad (\text{donc } F = V^C). \tag{1.13}
$$

On dit que Z est vraie sur E, ssi Z est "vraie partout", i.e. ssi elle est vraie en tous les points de E, i.e. ssi $V = E$, i.e. :

$$
\text{"}Z \text{ vrai"} \iff F = \emptyset \tag{1.14}
$$

Définition 1.39 Z est dite vraie μ-p.p. (μ-presque-partout) ssi $F \in \mathcal{M}_{\mu 0}$, i.e. ssi F est μ-mesurable et μ-négligeable :

$$
\text{"}Z \text{ vrai } \mu\text{-p.p."} \iff \mu(F) = 0, \tag{1.15}
$$

à comparer avec (1.14). On dit également "Z μ-p.p. vrai".

Interprétation de (1.15) : la mesure μ "ne voit pas" l'ensemble des points où c'est faux : $\mu(F) = 0$. Donc "Z est vraie pour μ".

Exercice 1.40 Soit μ_ℓ la mesure de Lebesgue sur $\mathbb{R}$. Montrer que $\mu_\ell(\mathbb{N}) = 0$ et $\mu_\ell(\mathbb{Q}) = 0$. En déduire que la propriété "tous les réels sont irrationnels" est une propriété vraie μ_ℓ-p.p..

Réponse. On a $\mathbb{N} \subset \mathbb{R}$ où $\mathbb{N} = \bigcup_{n\in\mathbb{N}}\{n\}$ est une union dénombrable d'ensembles (les singletons) deux à deux disjoints, donc $\mu_\ell(\mathbb{N}) = \sum_{\mathbb{N}} \mu_\ell(\{n\}) = \sum_{\mathbb{N}} 0 = 0$ car μ_ℓ est une mesure.

On a $\mathbb{Q} \subset \mathbb{R}$ et $\mathbb{Q}$ dénombrable : il existe une bijection $f : \mathbb{N} \to \mathbb{Q}$. Donc $\mu_\ell(\mathbb{Q}) = \mu_\ell(f(\mathbb{N})) = \mu_\ell(\bigcup_{n\in\mathbb{N}}\{f(n)\}) = \sum_{n\in\mathbb{N}} \mu_\ell(\{f(n)\}) = 0$, car union dénombrable de singletons distincts 2 à 2. ∴

L'ensemble F n'est pas toujours facile à caractériser. On utilise souvent un ensemble A plus grand que F, t.q. $\mu(A)=0$ et A facile à caractériser. On note :

$$\mathcal{M}_{\mu 0}=\{A\in\mathcal{A}:\mu(A)=0\},\tag{1.16}$$

l'ensemble des ensembles μ-négligeables. Donc si $A\in\mathcal{M}_{\mu 0}$ alors A est un ensemble qu'on "ne voit pas" quand on le regarde avec la mesure μ.

Proposition 1.41

$$\text{“}Z \text{ vrai } \mu\text{-p.p.”} \iff \exists A\in\mathcal{M}_{\mu 0} \text{ t.q. : } \forall x\in E-A \text{ on a “}Z(x)\text{ vrai”}.\tag{1.17}$$

(Un tel A contient F.)

Preuve. $F^C=\{x\in E:\ Z(x)\text{ vrai}\})$ est mesurable car complémentaire d'un ensemble mesurable.

$\Rightarrow$: supposons $\big(Z$ vrai μ-p.p.$\big)$, i.e. $\mu(F)=0$, cf. (1.15). On a, $\forall x\in F^C,\ Z(x)$ vrai. Donc il existe bien un $A\in\mathcal{A}_\mu$, à savoir $A=F$, t.q. $\mu(A)=0$ et, $\forall x\in A^C,\ Z(x)$ vrai.

$\Leftarrow$: supposons le membre de droite de (1.17) : il existe $A\in\mathcal{A}_\mu$ t.q. $\mu(A)=0$ et, $\forall x\in A^C,\ Z(x)$ vrai. Donc $A^C\subset V=F^C$, donc $F\subset A$, donc $\mu(F)\le\mu(A)=0$, donc $\mu(F)=0$ (une mesure est positive). ∎

Rappel : Soit Z une affirmation (une assertion) sur E.

$$\text{“}Z\text{ est faux”} \overset{\text{déf}}{=} \text{négation de “}Z\text{ est vrai”} \iff \exists x\in E,\ Z(x)\text{ faux} \iff F\neq\emptyset,\tag{1.18}$$

cf. (1.14).

Définition 1.42 Notation : On dit que Z est faux au sens μ-p.p. ssi $\mu(F)\neq 0$, et on note :

$$\text{“}Z\text{ est faux }\mu\text{-p.p.”} \overset{\text{déf}}{=} \text{négation de “}Z\text{ est vrai }\mu\text{-p.p.”} \iff \mu(F)\neq 0,\tag{1.19}$$

cf. (1.15). On dit également "Z μ-p.p. faux" (donc Z est faux sur un ensemble qu'on voit).

Proposition 1.43 *Z est faux μ-p.p. ssi il existe un ensemble non négligeable sur lequel A est faux :*

$$\text{“}Z\text{ est faux }\mu\text{-p.p.”} \iff \exists A\in\mathcal{A} \text{ t.q. } \mu(A)\neq 0 \text{ et, } \forall x\in A,\ \text{“}Z(x)\text{ faux”}.\tag{1.20}$$

Interprétation : "Z est faux μ-p.p." ssi la mesure μ "voit" un ensemble A sur lequel A est faux (ce n'est pas nécessairement F tout entier).

Preuve. $\Rightarrow$: supposons "Z est faux μ-p.p.". Donc $\mu(F)\neq 0$, cf. (1.15). Donc il existe A, à savoir $A=F$, t.q. $\mu(A)\neq 0$ et $\forall x\in A=F$ on a "$Z(x)$ faux".

$\Leftarrow$: supposons le membre de droite de (1.20). On a A t.q. $\mu(A)>0$ et $\forall x\in A$, "$Z(x)$ faux", donc $A\subset F$, donc $\mu(A)\le\mu(F)$, donc $\mu(F)>0$, et on utilise (1.15). ∎

Remarque 1.44 On a, par négation de (1.17) :

$$\text{“}Z\text{ est faux }\mu\text{-p.p.”} \iff \forall A\in\mathcal{M}_{\mu 0},\ \exists x\in E-A,\ \text{“}Z(x)\text{ faux”}.\tag{1.21}$$

Ce résultat sera peu utilisé. ∎

1.7 Fonctions égales μ-p.p.

Rappel : avec (1.14) on a : $\big(f=g\big)$ ssi $\big(\forall x\in E,\ f(x)=g(x)\big)$, i.e. :

$$\begin{aligned}\big(f=g\big) &\iff \{x\in E: f(x)\neq g(x)\}=\emptyset\\ &\iff \forall x\in E,\ f(x)=g(x).\end{aligned}\tag{1.22}$$

Soit f et $g:E\to\mathbb{R}$ deux fonctions μ-mesurables. On a, cf. (1.15) :

$$\begin{aligned}\big(f=g\ \mu\text{-p.p.}\big) &\iff \mu(\{x\in E: f(x)\neq g(x)\})=0\\ &\iff \exists A\in\mathcal{M}_{\mu 0},\ \forall x\in A^C,\ f(x)=g(x),\end{aligned}\tag{1.23}$$

et on dit alors que $f=g$ μ-p.p., ou encore $f=g$ à un ensemble μ-négligeable près. À comparer avec (1.22).

Exemple 1.45 $1_{\mathbb{Q}} = 0$ μ_ℓ-p.p. (la fonction $1_{\mathbb{Q}}$ est négligeable pour la mesure de Lebesgue).

Soit $f = 1_{[2,3]}$ et $g = 1_{]2,3[}$. On a $f = g$ partout sauf pour $x \in \{2,3\}$. On a $\mu_\ell(\{2,3\}) = 0$, donc $f = g$ μ_ℓ-p.p. (pour la mesure de Lebesgue).

En revanche $f \neq g$ δ_2-p.p. (pour la mesure de Dirac δ_2). En effet $f(2) \neq g(2)$ avec $\delta_2(\{2\}) = 1 \neq 0$; le singleton $\{2\}$ n'est pas négligeable pour la mesure δ_2. ▟

Exemple 1.46 Et (1.20) donne :

$$\begin{aligned}\Big(f \neq g \ \mu\text{-p.p.}\Big) \quad &\Longleftrightarrow \quad \mu(\{x \in E : f(x) \neq g(x)\}) \neq 0 \\ &\Longleftrightarrow \quad \exists A \notin \mathcal{M}_{\mu 0} \text{ t.q. } \forall x \in A,\ f(x) \neq g(x),\end{aligned} \tag{1.24}$$

i.e. "la mesure μ voit un ensemble A (on a $\mu(A) \neq 0$) sur lequel $f \neq g$". À comparer avec la négation de (1.22) :

$$\Big(f \neq g\Big) \iff \exists x \in E,\ f(x) \neq g(x). \tag{1.25}$$

Exemple pour $f = 1_{[2,3]}$ et $g = 1_{]2,3[}$ avec $\mu = \delta_2 + \delta_3$: ici $\mathcal{M}_{\mu 0} = \{A \subset \mathbb{R} : 2,3 \notin A\}$. Et il existe $A \notin \mathcal{M}_{\mu 0}$, par exemple $A = \{2\}$, t.q. $f(x) \neq g(x)$ pour tout $x \in A$ (i.e. pour $x = 2$). ▟

Exercice 1.47 Vérifier que la négation de (1.23) est :

$$\Big(f \neq g \ \mu\text{-p.p.}\Big) \iff \forall A \in \mathcal{M}_{\mu 0},\ \exists x \in A^C,\ f(x) \neq g(x). \tag{1.26}$$

Appliquer à l'exemple $f = 1_{[2,3]}$ et $g = 1_{]2,3[}$ avec $\mu = \delta_2$.

Réponse. Le contraire de $\big(\exists z \in Z$ t.q. "Propriété"$\big)$ est : $\big(\forall z \in Z$ on a "Propriété fausse"$\big)$.

Le contraire de la propriété : $\big(\forall x \in A^C$ on a $f(x) = g(x)\big)$ est : $\big(\exists x \in A^C,\ f(x) \neq g(x)\big)$. D'où (1.26).

Exemple pour $f = 1_{[2,3]}$ et $g = 1_{]2,3[}$ avec $\mu = \delta_2$: la propriété $\big(\forall A \in \mathcal{M}_{\mu 0}\big)$ s'écrit dans ce cas : $\big(\forall A \subset \mathbb{R}$ t.q. $2 \notin A\big)$. Donc pour $A \in \mathcal{M}_{\mu 0}$ on a $2 \in A^C$, et $f(2) \neq g(2)$. ▟

Exemple 1.48 Convergence μ-p.p.. Rappel : une suite $(f_n)_{\mathbb{N}^*}$ de fonctions converge simplement vers une fonction f ssi elle converge "en tout point" :

$$\Big(f_n \underset{n\to\infty}{\longrightarrow} f \text{ simplement}\Big) \quad \Longleftrightarrow \quad \forall x \in E, \quad f_n(x) \underset{n\to\infty}{\longrightarrow} f(x). \tag{1.27}$$

Une suite (f_n) de fonctions converge μ-p.p. (converge "presque partout") vers une fonction f ssi :

$$\begin{aligned}\Big(f_n \to f \ \mu\text{-p.p.}\Big) \quad &\Longleftrightarrow \quad \mu\{x \in E \ : \ f_n(x) \underset{n\to\infty}{\not\longrightarrow} f(x)\} = 0 \\ &\Longleftrightarrow \quad \exists A \in \mathcal{M}_{\mu 0}, \text{ t.q. } \forall x \in A^C,\ f_n(x) \underset{n\to\infty}{\longrightarrow} f(x),\end{aligned} \tag{1.28}$$

i.e. ssi l'ensemble $\{x \in E : f_n(x) \not\to f(x)\}$ est μ-négligeable.

En particulier, si la suite converge simplement alors la suite converge μ-p.p.. La réciproque est fausse : exemple avec la mesure de Lebesgue : $f_n = 1_{]0,\frac{1}{n}]}$ et $f(0) = (-1)^n$: on a $f_n \longrightarrow_{n\to\infty} 0$ μ_ℓ-p.p., alors que $(f_n(0))_{\mathbb{N}^*}$ ne converge pas. ▟

1.8 Fonctions nulles μ-p.p. (fonctions μ-négligeables)

On introduit la relation d'équivalence des fonctions égales μ-presque partout : si f et g sont mesurables, on pose :

$$\begin{aligned} f\mathcal{R}g \quad &\Longleftrightarrow \quad f = g \ \mu\text{-p.p.}, \\ &\Longleftrightarrow \quad \exists A \in \mathcal{M}_{\mu 0}, \quad \forall x \in E - A, \quad f(x) = g(x), \\ &\Longleftrightarrow \quad f = g + h \quad : \quad \exists A \in \mathcal{M}_{\mu 0}, \quad \forall x \in E - A, \quad h(x) = 0. \end{aligned} \tag{1.29}$$

i.e., les deux fonctions f et g sont égales à un ensemble de mesure nulle près. Et $\mathcal{R}$ est bien une relation d'équivalence dans $\mathcal{F}_M$ (réflexive, symétrique, transitive : immédiat). Et la classe d'équivalence d'une fonction f mesurable est l'ensemble :

$$\text{classe}(f) = \{g \in \mathcal{F}_M : g\mathcal{R}f\} \overset{\text{noté}}{=} \dot{f}.$$

En particulier, la classe des fonctions nulles presque partout, dite "ensemble des fonctions négligeables" (pour la mesure μ), est :

$$\mathcal{N} = \text{classe}(0) = \dot{0} = \{f \in \mathcal{F}_M : f = 0 \ \mu\text{-p.p.}\}. \tag{1.30}$$

Exercice 1.49 Montrer que $f \in \mathcal{N}$ ssi $\exists A \in \mathcal{M}_{\mu 0}$ t.q. $1_{E-A}f = 0$.

Réponse. C'est $(1.29)_2$ avec $g = 0$. ▟

Donc :

$$f\mathcal{R}g \iff f-g\in\dot{0}. \tag{1.31}$$

Remarque 1.50 Interprétation : f et g sont μ-p.p. égales ssi l'instrument de mesure μ ne peut pas les différencier. ▟

Exemple 1.51 Pour la mesure de Lebesgue la fonction $1_{\mathbb{Q}}$ est nulle μ_ℓ-p.p. : on a $1_{\mathbb{Q}}\in\mathcal{N}=\dot{0}$, i.e. $\dot{1}_{\mathbb{Q}}=\dot{0}$, i.e. $1_{\mathbb{Q}}\mathcal{R}0$. Ainsi pour la mesure de Lebesgue on a $\mu_\ell(1_{\mathbb{Q}})=\mu_\ell(0)$, et l'instrument de mesure μ_ℓ ne fait pas la différence entre $1_{\mathbb{Q}}$ et la fonction nulle. De même si on remplace $\mathbb{Q}$ par un ensemble fini ou dénombrable, ou par l'ensemble de Cantor par exemple. ▟

Exercice 1.52 Montrer que f "non nulle" μ-p.p. ssi :

$$\exists B\notin\mathcal{M}_{\mu 0},\quad \forall x\in B,\quad f(x)\neq 0, \tag{1.32}$$

i.e. f est non nulle sur tout un ensemble vu par la mesure μ.

Réponse. f nulle p.p. : $\exists A\in\mathcal{M}_{\mu 0}$, $\forall x\in A^C$, $f(x)=0$. Négation :

$$\forall A\in\mathcal{M}_{\mu 0},\quad \exists x\in A^C,\quad f(x)\neq 0. \tag{1.33}$$

Supposons (1.32). Soit B t.q. (1.32). Montrons (1.33). Soit $A\in\mathcal{M}_{\mu 0}$, donc 1- soit $A\cap B=\emptyset$ auquel cas $B\subset A^C$ et c'est trivial, 2- soit $A\cap B\neq\emptyset$ auquel cas $A^C\cap B\neq\emptyset$, car $A^C\cap B=\emptyset$ implique $B\subset A$, donc $\mu(B)=0$, absurde : il existe donc $x\in A^C\cap B$ t.q. $f(x)\neq 0$.

Supposons (1.33). Montrons (1.32). Soit $N=\{x: f(x)\neq 0\}$ (l'ensemble des non-zéros de f). Montrons que $\mu(N)\neq 0$, et on prendra alors $B=N$. Sinon $\mu(N)=0$, donc f est nulle p.p.. C'est la négation de (1.33) : absurde. Donc $\mu(N)\neq 0$. ▟

2 Fonction $\mathcal{A}$-mesurable

2.1 Définition

On se donne un espace mesurable $(E,\mathcal{A}_E)$. Pour le moment on ne dispose pas de mesure.

On cherche à savoir quand une fonction $f:E\to\mathbb{R}$ est "mesurable" (ou "intégrable relativement à une mesure μ") : il va falloir que f ait un sens sur les ensembles mesurables $A\subset\mathcal{A}_E$. On pourra alors, pour $A\in\mathcal{A}_E$ et après introduction d'une mesure μ, définir $\int_A f\,d\mu$, en commençant par définir $\int_A 1_A\,d\mu =^{\text{déf}} \mu(A)$, puis en approchant f par des fonctions étagées.

Définition 2.1 Soit $\mathbb{R}$ muni de sa tribu borélienne $\mathcal{A}_{\mathbb{R}}$. Soit $f:E\to\mathbb{R}$.

1- f est dite $\mathcal{A}_E$-mesurable ssi $f^{-1}(]-\infty,a])\in\mathcal{A}_E$ pour tout $a\in\mathbb{R}$.

1'- f est dite $\mathcal{A}_E$-mesurable ssi $f^{-1}(]-\infty,a[)\in\mathcal{A}_E$ pour tout $a\in\mathbb{R}$.

2- Ou de manière équivalente ssi pour tout intervalle I de $\mathbb{R}$ on a $f^{-1}(I)\in\mathcal{A}_E$.

3- Ou de manière équivalente ssi pour tout U ouvert de $\mathbb{R}$ on a $f^{-1}(U)\in\mathcal{A}_E$ (l'ensemble $f^{-1}(U)$ appartient à $\mathcal{A}_E$).

4- Ou de manière équivalente ssi pour tout borélien I de $\mathbb{R}$ on a $f^{-1}(I)\in\mathcal{A}_E$.

Lorsque le choix de la tribu $\mathcal{A}_E$ est implicite, on dit simplement que f est mesurable au lieu de $\mathcal{A}_E$-mesurable.

Et si $f:E\to\overline{\mathbb{R}}$ alors, notant $D_f=\{x\in E: |f(x)|<\infty\}=f^{-1}(\mathbb{R})$, la fonction f est $\mathcal{A}_E$-mesurable ssi $D_f\in\mathcal{A}_E$ (est mesurable) et $f_{|D_f}$ (la restriction de f à D_f) est $\mathcal{A}_E$-mesurable, où on a muni D_f de la tribu $\mathcal{A}_E$ restreinte à D_f (voir proposition 1.12).

Exercice 2.2 Montrer que les différentes définitions 1-, 1'-, 2-, 3-, 4- de la définition 2.1 sont effectivement toutes équivalentes.

Réponse. 1- $\Leftrightarrow$ 1' : pour tout $a\in\mathbb{R}$ on a $]-\infty,a[$ et $]-\infty,a]$ dans la tribu borélienne de $\mathbb{R}$ (ce sont des intervalles) et $\{a\}=]-\infty,a]\cap[a,\infty|$ également (intersection de deux boréliens. Supposons 1-. Alors comme $]-\infty,a[=]-\infty,a]-\{a\}$ on a $f^{-1}(]-\infty,a[)=f^{-1}(]-\infty,a])-f^{-1}(\{a\})$ dans la tribu borélienne, cf. exercice 1.7.

Supposons 1'- : alors $]-\infty,a]=]-\infty,a[\cap\{a\}$ union disjointe donc $f^{-1}(]-\infty,a])=f^{-1}(]-\infty,a[)\cup f^{-1}(\{a\})$ dans la tribu borélienne par stabilité d'une tribu.

4- $\Rightarrow$ 3- $\Rightarrow$ 2- $\Rightarrow$ 1- : immédiat.

1- $\Rightarrow$ 2- : si $I=[a,b]$ intervalle, alors $I=]-\infty,b]-]-\infty,a[$, donc $f^{-1}(I)=f^{-1}(]-\infty,b])-f^{-1}]-\infty,a[\in\mathcal{A}_E$, cf. exercice 1.7.

2- $\Rightarrow$ 3- : les boules ouvertes, ici les intervalles, forment une base de l'ensemble des ouverts, donc un ouvert U est de la forme $U=\bigcup_{i\in J}I_i$ où les I_i sont des intervalles et J est au plus dénombrable. Donc $f^{-1}(U)\subset\bigcup_{i\in J}f^{-1}(I_i)\in\mathcal{A}_E$ par stabilité d'une tribu.

3- $\Rightarrow$ 4- : idem. ▟

On notera :

$$\mathcal{F}_M(E,\mathcal{A}_E) \stackrel{\text{noté}}{=} \mathcal{F}_M \tag{2.1}$$

l'ensemble des fonctions mesurables, la dernière notation s'il n'y a pas d'ambiguïté.

Définition 2.3 De manière plus générale, si $(E,\mathcal{A}_E)$ est une espace mesurable, et si $(F,\mathcal{O}_F)$ est un espace topologique, alors $f : E \to F$ est $\mathcal{A}_E$-mesurable ssi $f^{-1}(U) \in \mathcal{A}_E$ pour tout $U \in \mathcal{O}_F$.

Remarque 2.4 Le nom $\mathcal{A}_E$-mesurable peut prêter à confusion : il n'y a pas besoin de mesure pour définir une fonction mesurable : mais f étant mesurable, on va pouvoir la mesurer, i.e. définir une mesure sur l'espace mesurable $(E,\mathcal{A}_E)$ puis définir une mesure de f. La démarche est la même que lors de la définition des espaces mesurables et des espaces mesurés. ▪

Exemple 2.5 Si $A \in \mathcal{A}_E$, alors 1_A la fonction indicatrice de A est $\mathcal{A}_E$-mesurable.

En effet, on a $1_A^{-1}(\{0\}) = A$, $1_A^{-1}(\{1\}) = E-A$ et $1_A^{-1}(\{y\}) = \emptyset$ si $y \notin \{0,1\}$, d'où : on a donc $1_A^{-1}(]-\infty,a])$ égal soit à $\emptyset$, à A, à $E-A$ ou à E, et comme par hypothèse $A \in \mathcal{A}$ on a bien $1_A^{-1}(]-\infty,a]) \in \mathcal{A}$.

Les fonctions indicatrices 1_A, pour les $A \in \mathcal{A}_E$, sont les fonctions de base qui serviront à construire les fonctions étagées puis les fonctions intégrables. ▪

Exercice 2.6 Soit E un ensemble de plus de trois éléments, soit $A \subset E$ t.q. $A \neq \emptyset$ et $A \neq E$, et soit $\mathcal{T}(A) = \{\emptyset, A, A^C, E\}$ la tribu engendrée par A. Montrer que les seules fonctions $f : E \to \mathbb{R}$ qui sont $\mathcal{T}(A)$-mesurables sont les fonctions $f = \alpha 1_A + \beta 1_{A^C}$ pour $\alpha,\beta \in \mathbb{R}$.

Réponse. $\Rightarrow$. $f=0$ vérifie $f^{-1}(]-\infty,a[) = \emptyset$ si $a<0$ et $=E$ si $a \geq 0$, donc $f=0$ est $\mathcal{T}(A)$-mesurable.

Soit $f = \alpha 1_A + \beta 1_{A^C}$ avec $\alpha < \beta$, alors $f^{-1}(]-\infty,a[) = \emptyset$ si $a<\alpha$, $=A$ si $a \in [\alpha,\beta[$, et $=E$ si $b \geq \beta$. Donc f est $\mathcal{T}(A)$-mesurable. Calculs similaires si $\alpha=\beta$ ou $\alpha>\beta$.

$\Leftarrow$. Soit f $\mathcal{T}(A)$-mesurable. Supposons qu'il existe $\alpha,\beta \in \overline{\mathbb{R}}$ avec $\alpha<\beta$ t.q. $f(A) =]\alpha,\beta[$. Notons $f^{-1}(]\alpha,\beta[) = A \cup B$ avec $B \subset A^C$. On doit avoir $A \cup B \in \mathcal{T}(A)$. On n'a pas $A \cup B = \emptyset$ car $A \neq \emptyset$. On n'a pas $A \cup B = A^C$ car $A \neq \emptyset$. Supposons $A \cup B = E$ i.e. $B = A^C$. Par définition de α,β on a $f^{-1}(]\frac{\alpha+\beta}{2},\beta[) \subsetneq A$, donc f n'est pas $\mathcal{T}(A)$-mesurable. Contraire à l'hypothèse. Il reste le cas $A \cup B = A$ i.e. $B = \emptyset$ auquel cas $f^{-1}(]\alpha,\beta[) = A$, et $f^{-1}(]\frac{\alpha+\beta}{2},\beta[) \subsetneq A$, donc f non $\mathcal{T}(A)$-mesurable. Donc l'hypothèse $\alpha<\beta$ t.q. $f(A) =]\alpha,\beta[$ n'est pas possible. Même démarche pour $\alpha<\beta$ et $f(A) = [\alpha,\beta]$ ou intervalle semi-ouvert : impossible.

Donc il reste le cas $\alpha=\beta$ et $f(A) = \{\alpha\}$, donc $f = \alpha 1_A + g 1_B$ avec g fonction et $B \subset A^C$. Comme A et A^C jouent le même rôle (poser $f(A^C) =]\alpha,\beta[$) on en déduit que $f = h 1_A + \beta 1_{A^C}$ avec h fonction et $\beta \in \mathbb{R}$. Donc $f = \alpha 1_A + \beta 1_B$. ▪

Remarque 2.7 Dans le langage des probabilités, étant donné un espace mesurable $(E,\mathcal{A}_E)$: une fonction mesurable est appelée variable aléatoire.

Une variable aléatoire est donc une fonction $X : E \to \mathbb{R}$ telle que : pour tout I intervalle de $\mathbb{R}$, $X^{-1}(I)$ est mesurable (i.e. $X^{-1}(I) \in \mathcal{A}_E$).

Si de plus on se donne une mesure μ définie sur $(E,\mathcal{A}_E)$ telle que $\mu(E)=1$ (une mesure de probabilité) qui fait de $(E,\mathcal{A},\mu)$ un espace probabilisé, on peut alors calculer le pourcentage des points $x \in E$ tels que $X(x)$ appartienne à un intervalle donné $I \subset \mathbb{R}$.

L'ensemble de tels x est donnée par $X^{-1}(I) = \{x \in E : X(x) \in I\} =^{\text{noté}} [X \in I]$, et la mesure de cet ensemble est $\mu(X^{-1}(I)) =^{\text{noté}} \mu([X \in I])$ appelée probabilité de “trouver X dans I”, i.e. $\mu([X \in I])$ est le “pourcentage” (relatif à la mesure μ) de points $x \in E$ dont l'image $X(x)$ est dans I. ▪

2.2 Fonctions continues : mesurables

Proposition 2.8 *“Toute fonction continue est mesurable”. Cela signifie précisément : si $(E,\mathcal{O}_E)$ et $(F,\mathcal{O}_F)$ sont deux espaces topologiques et si $f : E \to F$ est une fonction continue, alors f est mesurable lorsque E est muni de sa tribu borélienne $\sigma(\mathcal{O}_E)$.*

Preuve. Pour tout ouvert U de F, f étant continue, l'ensemble $f^{-1}(U)$ est un ouvert de E, donc est dans la tribu de E. Donc toute fonction continue est mesurable. ▪

Exercice 2.9 Montrer que la tribu borélienne $\mathcal{A}_\mathbb{R}$ de $\mathbb{R}$ est stable par translation, i.e. si $r \in \mathbb{R}$ et si $A \in \mathcal{A}_\mathbb{R}$ alors $r + A \in \mathcal{A}_\mathbb{R}$.

Réponse. Soit $r \in \mathbb{R}$. Soit $f : x \in \mathbb{R} \to y = x - r \in \mathbb{R}$ la translation de r : c'est une fonction continue (l'image réciproque d'un ouvert est un ouvert : l'ouvert translaté), donc mesurable (proposition précédente). Donc si A est un borélien alors $f^{-1}(A)$ est un borélien. Donc $f^{-1}(A) = \{x \in \mathbb{R} : \exists y \in A \text{ t.q. } x = y + r\} = r + A$ est un borélien. ▪

Proposition 2.10 *Soit $(E,\mathcal{A}_E)$ un espace mesurable et soient $(F,\mathcal{O}_F)$ et $(G,\mathcal{O}_G)$ deux espaces topologiques. Si $f : E \to F$ est mesurable et si $g : F \to G$ est continue, alors $g \circ f : E \to G$ est mesurable.*

Preuve. On a $(g \circ f)^{-1}(C) = \{x \in E : g(f(x)) \in C\} = \{x \in E : f(x) \in g^{-1}(C)\} = f^{-1}(g^{-1}(C))$ pour tout sous-ensemble $C \subset G$. Soit W un ouvert de G. Alors $V = g^{-1}(W)$ est ouvert dans F car g est continue. donc $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W)) = f^{-1}(V)$ est mesurable car f est mesurable. ∎

Corollaire 2.11 *Soit $(E, \mathcal{A}_E)$ un espace mesurable et on considère $(\mathbb{R}, \mathcal{O}_\mathbb{R})$ l'espace topologique réel usuel. Soient $u, v : E \to \mathbb{R}$ deux fonctions mesurables réelles, et soit $\Phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ une fonction continue. Alors la fonction $h : E \times E \to \mathbb{R}$ définie par :*

$$h(x) = \Phi(u(x), v(x)) \tag{2.2}$$

est mesurable. En particulier, la somme $u + v$ et le produit uv de deux fonctions mesurables u et v est une fonction mesurable.

Preuve. Soit $f : E \to \mathbb{R}^2$ définie par $f(x) = (u(x), v(x))$. Montrons que f est mesurable. Soit $I =]a, b[$, $J =]c, d[$ et le rectangle $R = I \times J$. Alors $f^{-1}(R) = u^{-1}(I) \cap v^{-1}(J)$ est mesurable comme intersection d'ensembles mesurables. Et tout ouvert $V \subset \mathbb{R}^2$ est réunion dénombrable $V = \bigcup_{i=1}^\infty R_i$ de rectangles de type $R = I \times J$. Comme $f^{-1}(V) = f^{-1}(\bigcup_{i=1}^\infty R_i) = \bigcup_{i=1}^\infty f^{-1}(R_i)$, on a bien $f^{-1}(V)$ mesurable. Et Φ continue donne $h = \Phi \circ f$ mesurable, cf. proposition 2.10.

Puis la fonction somme $\Phi(x, y) = S(x, y) = x+y$ et la fonction produit $\Phi(x, y) = P(x, y) = xy$ sont continues sur $\mathbb{R}^2$: si $x_0, y_0 \in \mathbb{R}$ sont fixés, si $h, k \in \mathbb{R}$, alors

$|S(x_0+h, y_0+k) - S(x_0, y_0)| = |h-k| \longrightarrow_{h,k\to 0} 0$ (immédiat), et

$|P(x_0+h, y_0+k) - P(x_0, y_0)| \le \max(|h|, |k|, |hk|)(|x_0|+|y_0|) \longrightarrow_{h,k\to 0} 0$ (immédiat). ∎

2.3 Stabilité de l'ensemble des fonctions mesurables

Corollaire 2.12 *L'ensemble des fonctions réelles $\mathcal{A}_E$-mesurables est un espace vectoriel, sous-espace vectoriel de l'ensemble des fonctions $E \to \mathbb{R}$. De plus, cet ensemble est stable par :*

** sup et inf :* $\sup(f, g)(x) = \sup(f(x), g(x))$.

** passage aux limites simples de suites de fonctions :* $\sup(f_n)$, $\inf(f_n)$, $\liminf(f_n)$ *et* $\limsup(f_n)$ *(voir plus loin (5.14) et (5.22) pour le rappel de la définition de* $\liminf$ *et* $\limsup$*).*

Preuve. Si f est mesurable et $\lambda \in \mathbb{R}$, comme $g_\lambda : \mu \in \mathbb{R} \to g_\lambda(\mu) = \lambda\mu$ est continue, $g_\lambda \circ f = \lambda f$ est mesurable.

Si $f, g : E \to \mathbb{R}$ sont mesurables, comme $\Phi = + : (x, y) \in E \times E \to x + y \in \mathbb{R}$ est continue, $h : x \to h(x) = \Phi(f(x), g(x)) = f(x) + g(x)$ est mesurable.

Si f et g sont mesurable, soit $h = \inf(f, g)$, i.e. $h(x) = \min(f(x), g(x)) \in \overline{\mathbb{R}}$ pour tout $x \in E$. Alors $h^{-1}(]-\infty, a[) = \{x : f(x), g(x) < a\} = f^{-1}(]-\infty, a[) \bigcap g^{-1}(]-\infty, a[)$ intersection de deux ensembles mesurables, donc est mesurable. Idem pour $\sup(f, g) = -\inf(-f, -g)$.

Soit $(f_n)_\mathbb{N}$ une suite de fonctions réelles mesurables sur E et soit $g = \inf(f_n)$. On a $g(x) = \inf_\mathbb{N}(f_n(x)) \in \overline{\mathbb{R}}$, donc $g^{-1}(]-\infty, a[) = \{x : g(x) < a\} = \{x : \forall n \in \mathbb{N},\ f_n(x) < a\} = \{x : \forall n \in \mathbb{N},\ x \in f_n^{-1}(]-\infty, a[)\} = \bigcap_\mathbb{N} f_n^{-1}(]-\infty, a[)$ intersection dénombrable d'ensembles mesurables. Donc $g = \inf(f_n)$ est mesurable. De même $\sup(f_n)$ est mesurable.

Et $\liminf(f_n) \stackrel{\text{déf}}{=} \sup_{k\in\mathbb{N}}(g_k)$, où on a posé $g_k = \inf_{n\ge k}(f_n)$ pour $k \in \mathbb{N}$. Comme les g_k sont mesurables (cf. ci-dessus), on a $\sup_{k\in\mathbb{N}}(g_k)$ est mesurable (cf. ci-dessus), donc $\liminf(f_n)$ est mesurable. De même $\limsup(f_n)$ est mesurable. ∎

3 Intégration de fonctions positives

3.1 Mesure de fonctions étagées

On se donne un ensemble mesuré $(E, \mathcal{A}_E, \mu)$. On commence par remarquer que :

Proposition 3.1 *Si $n \in \mathbb{N}^*$, si $c_1, ..., c_n \in \mathbb{R}$ et $A_1, ..., A_n \in \mathcal{A}_E$ (mesurables), alors la fonction étagée $f = \sum_{i=1}^n c_i 1_{A_i}$ est mesurable.*

Preuve. Une somme finie de fonctions mesurables est mesurable, cf. corollaire 2.12 (l'ensemble des fonctions mesurables est un espace vectoriel). ∎

La mesure n'a été jusqu'ici définie que pour les ensembles. Pour mesurer les fonctions on va se ramener à la mesure des ensembles à l'aide de :

Définition 3.2 Pour $A \in \mathcal{A}_E$ (ensemble mesurable) et $1_A : E \to \mathbb{R}$ sa fonction indicatrice, on définit la mesure $\tilde{\mu}(1_A)$ de la fonction mesurable 1_A par :

$$\tilde{\mu}(1_A) \stackrel{\text{déf}}{=} \mu(A) \stackrel{\text{noté}}{=} \int_E 1_A \, d\mu \stackrel{\text{noté}}{=} \int_A d\mu \stackrel{\text{noté}}{=} \int_{x \in A} d\mu(x). \tag{3.1}$$

Notation. Pour alléger les notations, on notera abusivement $\tilde{\mu} = \mu$:

$$\tilde{\mu}(1_A) \stackrel{\text{noté}}{=} \mu(1_A) \qquad (\stackrel{\text{déf}}{=} \mu(A)). \tag{3.2}$$

bien que μ soit définie sur les ensembles alors que $\tilde{\mu}$ est définie sur les fonctions (mais il n'y a pas d'ambiguïté dans l'utilisation). Cas particulier d'une mesure diffuse (mesure de densité) : on note :

$$\tilde{\mu}(1_A) \stackrel{\text{noté}}{=} \int_E 1_A \, d\mu \stackrel{\text{noté}}{=} \int_A d\mu \stackrel{\text{noté}}{=} \int_{x \in E} 1_A(x) \, d\mu(x) \qquad (= \mu(A)). \tag{3.3}$$

Définition 3.3 Soit $n \in \mathbb{N}^*$, $c_1, ..., c_n \in \mathbb{R}$, $A_1, ..., A_n \in \mathcal{A}_E$ (tous mesurables) et la fonction étagée $f = \sum_{i=1}^n c_i 1_{A_i}$. La mesure de f est définie dans $\overline{\mathbb{R}}$ par :

$$\tilde{\mu}(f) = \tilde{\mu}(\sum_{i=1}^n c_i 1_{A_i}) \stackrel{\text{déf}}{=} \sum_{i=1}^n c_i \mu(A_i) \stackrel{\text{noté}}{=} \mu(f). \tag{3.4}$$

Cas particulier d'une mesure diffuse (mesure de densité) : on note :

$$\mu(f) \stackrel{\text{noté}}{=} \int_E f \, d\mu \stackrel{\text{noté}}{=} \int_{x \in E} f(x) \, d\mu(x), \quad \text{donc} \quad = \sum_{i=1}^n c_i \int_E 1_{A_i} d\mu = \sum_{i=1}^n c_i \int_{A_i} d\mu. \tag{3.5}$$

En particulier pour la mesure de Lebesgue μ_ℓ on note également $d\mu(x) = dx$ quand le nom de la variable est x.

Exercice 3.4 Montrer : si $A \in \mathcal{A}$ est négligeable et si f est une fonction étagée alors $\mu(1_A f) = 0$.

Réponse. La fonction $1_A f$ est mesurable comme produit de fonctions mesurables et est étagée puisque $= \sum_{i=1}^n c_i 1_{A_i \bigcap A}$. Et $\mu(1_A f) = \sum_{i=1}^n c_i \mu(A_i \bigcap A) = \sum_{i=1}^n c_i 0 = 0$ car $0 \le \mu(A_i \bigcap A) \le \mu(A) = 0$. ▪

Exercice 3.5 Soit $E = [0,1]$ muni de la tribu borélienne $\mathcal{A}_E = E \cap \mathcal{A}_\mathbb{R}$, cf proposition 1.12, et de la mesure borélienne μ_ℓ. Soit $B \in \mathcal{A}_E$(mesurable) t.q. $B^o = \emptyset$ (intérieur vide), B non dénombrable, et $0 < \mu_\ell(B) < 1$, cf. exercice 1.34.

1- Rappeler ce qu'est une fonction intégrable au sens de Riemann.

2- Rappeler que si $A^o = \emptyset$ (intérieur vide), alors son complémentaire A^C est dense dans E.

3- Montrer que la fonction 1_B n'est pas intégrable au sens de Riemann (n'est égale presque partout à aucune fonction Riemann-intégrable).

Réponse. 1- Rappel : f intégrable au sens de Riemann sur $[0,1]$ veut dire : f est bornée sur $[0,1]$ et f peut être encadrée par deux fonctions f_m et f_M en escalier, $f_m \le f \le f_M$, avec $\mu_\ell(f_M - f_m)$ aussi petit que souhaité (voir remarque 0.5). Donc pour tout $\varepsilon > 0$, $\exists n \in \mathbb{N}$, $\exists (x_i)_{i=1,...,n-1} \in \mathbb{R}^{n-1}$ avec $x_0 = 0$ et $x_n = 1$ et $x_{i-1} < x_i$ pour tout i (donne une partition de $[0,1]$), t.q. si on pose $m_i = \min_{x \in]x_{i-1}, x_i[} f(x)$ et $M_i = \max_{x \in]x_{i-1}, x_i[} f(x)$ puis $f_m = \sum_i m_i 1_{[x_{i-1}, x_i[}$ et $f_M = \sum_i M_i 1_{[x_{i-1}, x_i[}$ alors $\mu_\ell(f_M - f_m) < \varepsilon$ où $\mu_\ell(f_M - f_m) = \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) \stackrel{\text{noté}}{=} \int_0^1 (f_M(x) - f_m(x)) \, dx$.

Et on note alors $\mu_\ell(f) = \int f \, dx$ la valeur limite de l'encadrement obtenu par "raffinement" de la partition : on part d'un découpage $P_n = \bigcup_{i=1}^n [x_{n,i-1}, x_{n,i}]$ donné et le découpage raffiné $P_{n+1} = \bigcup_{i=1}^{n+1} [x_{n+1,i-1}, x_{n+1,i}]$ est t.q. $\forall i = 0, ..., n$, $\exists j = 0, ..., n+1$ t.q. $x_{n+1,j} = x_{n,i}$, et on suppose que $\sup_{i=1,...,n}(x_{n,i} - x_{n,i-1}) \to 0$ quand $r \to 0$ (la taille des mailles tend vers 0).

2- Soit $A \subset [0,1]$ un ensemble d'intérieur vide. Montrons A^C dense dans E, i.e. $\forall x \in [0,1]$, $\forall \varepsilon > 0$, $\exists y \in A^C$ t.q. $y \in B(x, \varepsilon)$. Sinon $\exists x \in [0,1]$, $\exists \varepsilon > 0$ t.q. $B(x,\varepsilon) \cap A^C = \emptyset$, donc $B(x, \varepsilon) \subset A$, donc A n'est pas d'intérieur vide. Absurde.

3- $f = 1_B$ μ_ℓ-p.p. sur $[0,1]$ équivaut à : $f(x) = 1_B(x)$ pour tout $x \in A^C = [0,1] - A$ où $\mu_\ell(A) = 0$. Donc en particulier $f(x) = 0$ pour tout $x \in B^C \cap A^C$. Et $B^C \cap A^C \neq \emptyset$: sinon $B^C \cap A^C = \emptyset$ et donc $\mu_\ell(B^C \cup A^C) = \mu_\ell(B^C) + \mu_\ell(A^C)$, avec $\mu_\ell(A^C) = 1 - \mu_\ell(A) = 1$ (car A est négligeable) et avec $\mu_\ell(B^C) = 1 - \mu_\ell(B) > 0$ (par hypothèse). Donc $\mu_\ell(B^C \cup A^C) > 1$: absurde puisque $\mu(E) = 1$. Donc il existe un $x \in B^C \cap A^C$. Donc $\min_{x \in [0,1]} f(x) \le 0$.

Ce qu'on vient de faire sur $[0,1]$ peut être fait sur chaque intervalle $]x_{i-1}, x_i[$: donc pour tout découpage $[0,1] = \bigcup_{i=1}^n [x_{i-1}, x_i]$ où $x_{i-1} < x_i$ pour tout $i = 1, ..., n$, on a $\min_{]x_{i-1}, x_i[} f(x) = m_i \le 0$ pour tout i. Donc $\mu_\ell(f_m) \le 0$ pour toute fonction en escalier minimisant f au sens de Riemann. De même $\mu_\ell(f_M) \ge 1$. Ce pour toute partition de $[0,1]$: donc f n'est Riemann intégrable : il n'existe pas de fonction intégrable au sens de Riemann presque partout égale à la fonction 1_B. ▪

3.2 Intégrale de fonctions $\mathcal{A}$-mesurables positives

On s'intéresse d'abord aux fonctions positives (il n'y aura pas d'ambiguïté de type "$\infty - \infty = ?$") à valeurs dans $\overline{\mathbb{R}_+} = \mathbb{R}_+ \bigcup \infty$.

Théorème 3.6 et définition *Une fonction $\mathcal{A}$-mesurable positive $f : E \to \overline{\mathbb{R}_+}$ est limite simple d'une suite croissante $(f_n)_{\mathbb{N}^*}$ de fonctions étagées positives mesurables :*

$$\forall x \in E,\ f_n(x) \underset{n\to\infty}{\longrightarrow} f(x), \quad \text{où} \quad \forall n \in \mathbb{N}^*,\ f_n \geq 0, \quad \text{et} \quad (f_n)_{\mathbb{N}^*} \nearrow . \tag{3.6}$$

On définit sa mesure par (dans $\overline{\mathbb{R}_+}$) :

$$\mu(f) \stackrel{\text{déf}}{=} \sup_{n\in\mathbb{N}^*} \mu(f_n). \tag{3.7}$$

Et ce nombre est indépendant de la suite croissante (f_n) choisie. D'où la définition alternative :

$$\mu(f) \stackrel{\text{déf}}{=} \sup_{g \text{ mesurable étagée, } 0\leq g\leq f} \mu(g). \tag{3.8}$$

Et l'application μ ainsi prolongée (aux fonctions $\mathcal{A}$-mesurables positives) est :

— *semi-linéaire, i.e., $\mu(f+g) = \mu(f) + \mu(g)$, et pour tout $\lambda > 0$, $\mu(\lambda f) = \lambda\mu(f)$,*

— *monotone, i.e., si $f \leq g$ alors $\mu(f) \leq \mu(g)$,*

— *continue croissante, i.e., pour toute suite croissante $(f_n)_{\mathbb{N}^*}$ de fonctions positives, on a :*

$$\mu(\lim_{n\to\infty}(f_n)) = \lim_{n\to\infty}(\mu(f_n)) \quad \in \overline{\mathbb{R}_+}, \tag{3.9}$$

soit $\mu(\sup(f_n)) = \sup(\mu(f_n))$ (on passe à la limite sous le signe $\mu =^{\text{noté}} \int$ pour les suites croissantes de fonctions positives).

Preuve. On prend la suite $(f_n)_{\mathbb{N}^*}$ définie en (0.5). Comme f est mesurable, les F_n et B_{ni} sont mesurables, et donc les fonctions étagées f_n sont mesurables. Comme $(f_n)_{\mathbb{N}^*}$ est croissante, la suite $(\mu(f_n))_{\mathbb{N}}$ est croissante dans $\overline{\mathbb{R}}$, donc le $\sup_{n\in\mathbb{N}} \mu(f_n)$ existe dans $\overline{\mathbb{R}_+}$, et donc (3.7) a un sens.

Soit $(g_n)_{\mathbb{N}^*}$ une autre suite croissante de fonctions positives mesurables étagées qui converge simplement vers f. Chaque g_n est de type $g_n = \sum_{i=1}^{m_n} d_i 1_{D_i}$ (somme finie) où les $d_i \geq 0$, où les $D_i = f^{-1}(\{d_i\})$ sont mesurables et disjoints 2 à 2, cf. (0.3), et où $g_n \leq f$.

Pour n fixé, soit un entier $N_n =^{\text{noté}} N$ t.q. $N > 1 + \max_{i=1,\dots,m_n}(d_i)$.

$\bigcup_{i=1}^{N2^N} [\frac{i-1}{N2^N}, \frac{i}{N2^N}[$ est une partition de $[0,N[$. Donc, $\forall i \in [1,m_n]_{\mathbb{N}}$, $\exists k_i \in [1, N2^N-1]_{\mathbb{N}^*}$ t.q. $d_i \in [\frac{k_i-1}{2^N}, \frac{k_i}{2^N}[$. Donc $D_i = f^{-1}(\{d_i\}) \subset B_{N,k_i}$, cf. (0.4). Donc $g_n \leq f_N$. Donc $\mu(g_n) \leq \mu(f_N)$. Donc $\sup_{n\in\mathbb{N}} \mu(g_n) \leq \sup_{N\in\mathbb{N}} \mu(f_N)$.

De même en inversant les rôles des f_n et g_n. D'où le sup est indépendant de la suite choisie.

Le reste s'en déduit. ∎

Notations : La suite $(f_n)_{\mathbb{N}^*}$ étant croissante, on note :

$$\sup_{n\in\mathbb{N}^*} \mu(f_n) \stackrel{\text{noté}}{=} \lim_{n\to\infty} \mu(f_n) \quad (\text{quand } (f_n)_{\mathbb{N}^*} \geq 0 \text{ et } \nearrow). \tag{3.10}$$

Donc (3.7) se lit également (passage à la limite sous le signe μ pour $(f_n) \geq 0$ et $\nearrow$)

$$\mu(\lim_{n\to\infty} f_n) = \lim_{n\in\mathbb{N}^*} \mu(f_n) \quad (\text{quand } (f_n)_{\mathbb{N}^*} \geq 0 \text{ et } \nearrow). \tag{3.11}$$

Donc, avec (3.5) (passage à la limite sous le signe $\int$ pour $(f_n) \geq 0$ et $\nearrow$),

$$\int_E (\lim_{n\to\infty} f_n)\, d\mu = \lim_{n\to\infty} (\int_E f_n\, d\mu) \quad (\text{quand } (f_n)_{\mathbb{N}^*} \geq 0 \text{ et } \nearrow). \tag{3.12}$$

Et pour f mesurable et $A \in \mathcal{A}$ (donc pour 1_A mesurable), le produit $1_A f$ est mesurable, voir proposition 2.12, et on note alors :

$$\mu(1_A f) = \int_A f\, d\mu \quad (= \int_E 1_A f\, d\mu). \tag{3.13}$$

Exercice 3.7 Montrer que si $A \in \mathcal{A}$ est négligeable, alors pour toute fonction f mesurable positive de mesure finie (i.e. $\mu(f) < \infty$) on a $\mu(1_A f) = 0$.

Réponse. La fonction $1_A f$ est mesurable comme produit de fonctions mesurables, cf. corollaire 2.11. Puis on applique l'exercice 3.4. Soit (f_n) une suite croissante de fonctions étagées telles que $f_n \to f$. On a alors $1_A f_n \to 1_A f$ puisque si $x \in E - A$ on a bien $0 \to 0$ et si $x \in A$ on a bien $f_n(x) \to f(x)$. Et si $(f_n)_{\mathbb{N}^*}$ est positive croissante, alors il est immédiat que $(1_A f_n)_{\mathbb{N}^*}$ est positive croissante. Et le théorème 3.6 donne $\mu(1_A f) = \lim \mu(1_A f_n) = \lim 0 = 0$, cf. exercice 3.4. ∎

Et le théorème 3.6 donne :

Proposition 3.8 *(Relation de Chasles.) Si f est une fonction positive mesurable sur E, et si $A \in \mathcal{A}$, alors :*

$$\mu(f) = \mu(1_A f) + \mu(1_{E-A} f). \tag{3.14}$$

En particulier si $A \in \mathcal{M}_{\mu 0}$ alors $\mu(f) = \mu(1_{E-A} f)$.

Preuve. En effet, $f = 1_A f + 1_{E-A} f$ et on approche séparément $1_A f$ sur A et $1_{E-A} f$ sur $E - A$ par une suite de fonctions étagées croissante. ▪

Exercice 3.9 Montrer : pour $f : E \to \mathbb{R}_+$ mesurable (positive), si $\mu(f) = 0$ alors $f = 0$ μ-p.p..

(Ou encore : si $f > 0$ μ-p.p. alors $\mu(f) > 0$.)

Réponse. Soit $A_n = \{x \in E : f(x) > \frac{1}{n}\} = f^{-1}(]\frac{1}{n}, \infty[)$: A_n est une suite croissante. Par définition de A_n on a $\frac{1}{n} 1_{A_n} \le f 1_{A_n} \le f$ (vrai pour $x \in A_n$ et pour $x \in E - A_n$). D'où $\frac{1}{n}\mu(A_n) \le \mu(f) = 0$. D'où $\mu(A_n) = 0$. D'où, la suite (A_n) étant croissante, $0 = \lim_{n\to\infty} \mu(A_n) = \mu(\lim_{n\to\infty} A_n)$, cf. proposition 1.27. Et comme $\lim_{n\to\infty} A_n = \bigcup_{\mathbb{N}} A_n = \{x \in E : f(x) > 0\}$, on a obtenu $\mu(\{x \in E : f(x) > 0\}) = 0$, donc $f \le 0$ μ-p.p., et comme $f \ge 0$, on a $f = 0$ μ-p.p... ▪

Exercice 3.10 Soit $f : E \to \mathbb{R}$. Montrer : si pour tout $A \in \mathcal{A}$ on a $\int_A f\, d\mu = 0$ alors $f = 0$ μ-p.p.

Réponse. Soit $f_+ = \sup(0, f)$ la partie positive de f et $A = f_+^{-1}(]0, \infty[)$. Alors $1_A f = f_+$ et $0 = \int_A f\, d\mu = \int_E 1_A f\, d\mu = \int_E f_+\, d\mu$, d'où $f_+ = 0$ μ-p.p., cf. exercice 3.9. Idem pour $f_- = -(\inf(0, f))$ (fonction positive). Puis $f = f_+ - f_-$. ▪

3.3 Fonction intégrale

Définition 3.11 Une fonction mesurable positive $f : E \to \mathbb{R}_+$ est intégrable ssi $\mu(f) \in \mathbb{R}$, i.e. ssi $\mu(f) < \infty$.

Exercice 3.12 Montrer : toute fonction intégrable positive $f : E \to \overline{\mathbb{R}_+}$ est finie presque partout, i.e. : si $f : E \to \overline{\mathbb{R}_+}$ est mesurable (positive) et μ-intégrable ($\mu(f) < \infty$) alors :

$$\mu(\{f = \infty\}) = 0, \tag{3.15}$$

où $\mu(\{f = \infty\}) =^{\text{déf}} \mu(\{x : f(x) = \infty\})$. Autrement dit, $\forall x \in E - A$, on a $f(x) < \infty$, où $A = f^{-1}(\{\infty\}) \in \mathcal{M}_{\mu 0}$. (Exemple avec $f : x \in [-1, 1] \to f(x) = (|x|)^{-\frac{1}{2}} \in \overline{\mathbb{R}_+}$.)

Réponse. Soit $n > 0$ et $A_n = f^{-1}([n, \infty]) = \{x : f(x) \ge n\}$. Les A_n forment une suite décroissante avec $\lim_{n\to\infty} A_n = \{x : f(x) = \infty\} = \{f = \infty\}$.

Et on a $0 \le 1_{A_n} \le \frac{1}{n} f$ (trivialement vrai si $x \in A_n$ et si $x \in E - A_n$). Donc $0 \le \mu(A_n) \le \frac{1}{n}\mu(f)$, avec $\mu(f) < \infty$, en particulier $\mu(A_1) < \infty$, et donc $\lim_{n\to\infty} \mu(A_n) = 0$, donc $\mu(\lim_{n\to\infty} A_n) = 0$, cf. (1.9), i.e. $\mu(\{f = \infty\}) = 0$.

(Et si $f \not\ge 0$, on pose $f = f_+ - f_-$ avec $f_+ = \sup(0, f) \ge 0$ et $f_- = \sup(0, -f) \ge 0$.) ▪

Exercice 3.13 1- Soit $n \in \mathbb{N}^*$, soit $x_0, ..., x_n \in \mathbb{R}$ t.q. $x_{i-1} < x_i$ pour tout $i = 1, ...n$. Soit $\lambda_1, ..., \lambda_n \in \mathbb{R}_+$. Soit $f = \sum_{i=0}^n \lambda_i 1_{[x_{i-1}, x_i[}$ (fonction en escalier positive). Montrer : pour $x \in \mathbb{R}$ on a (avec δ_x est la mesure de Dirac en x) :

$$\delta_x(f) = f(x) = \delta_x(1_B f), \tag{3.16}$$

pour tout intervalle ouvert B contenant x ("aussi petit soit-il"). Et que cela reste vraie pour des fonctions en escalier de signe quelconque.

2- Montrer : si $f : [a, b] \to \mathbb{R}_+^*$ est continue en $x \in]a, b[$, alors on a (3.16). (Pour définir la mesure de Dirac $\delta_x(f)$ d'une fonction f on a besoin qu'elle soit définie partout, et pour les fonctions définies presque partout on définit $\delta_x(f)$ uniquement pour les fonctions continues en x.)

Réponse. 1- Soit $A \subset [a, b]$, A mesurable. On a $\delta_x(A) = 1_A(x)$, cf. (1.3). Donc $\delta_x(1_A) = 1_A(x)$, cf. (3.2). Donc $\delta_x(f) = \delta_x(\sum_{i=0}^n \lambda_i 1_{[x_{i-1}, x_i[}) = \sum_{i=0}^n \lambda_i \delta_x(1_{[x_{i-1}, x_i[})$ (une mesure est semi-linéaire, cf. proposition 3.6), donc $\delta_x(f) = \sum_{i=0}^n \lambda_i 1_{[x_{i-1}, x_i[}(x) = f(x)$.

Et pour B intervalle ouvert contenant x on a $1_B f$ en escalier, d'où (3.16).

Puis pour f de signe quelconque, on a $f + c1_E > 0$ pour c t.q. $c = \max(1, -\sup_i \lambda_i) > 0$. Avec $f(x) + c1_E(x) = (f + c1_E)(x) = \delta_x(f + c1_E) = \delta_x(f) + c\delta_x(1_E) = \delta_x(f) + c$. D'où $f(x) = \delta_x(f)$. Idem avec $1_B f$.

2- Soit f continue en x : $\forall \varepsilon > 0$, $\exists \eta_\varepsilon > 0$, t.q. $y \in [a, b]$ et $|y - x| < \eta_\varepsilon \Rightarrow |f(y) - f(x)| < \varepsilon$.

Soit $\varepsilon > 0$ et un η_ε associé, soit $B_\varepsilon = [x - \eta_\varepsilon, x + \eta_\varepsilon] \cap]a, b[$, soit $m_\varepsilon = \inf_{y \in B_\varepsilon} f(y)$ et $M_\varepsilon = \sup_{y \in B_\varepsilon} f(y)$: on a donc $m_\varepsilon \le f(y) \le M_\varepsilon$ pour tout $y \in B_\varepsilon$, soit $m_\varepsilon 1_{B_\varepsilon} \le 1_{B_\varepsilon} f \le M_\varepsilon 1_{B_\varepsilon}$. Donc $\delta_x(m_\varepsilon 1_{B_\varepsilon}) \le \delta_x(1_{B_\varepsilon} f) \le \delta_x(M_\varepsilon 1_{B_\varepsilon})$.

Quitte à considérer $f + c$ où $c = f(x) + 1$, on peut considérer f positive en x et prendre η_ε suffisamment petit pour que f soit positive dans B_ε.

Avec (3.16) on a donc $m_\varepsilon \le \delta_x(f) \le M_\varepsilon$, ce pour tout ε. Comme $m_\varepsilon \longrightarrow_{\varepsilon \to 0} f(x)$ et $M_\varepsilon \longrightarrow_{\varepsilon \to 0} f(x)$, on en déduit $\delta_x(f) = f(x)$.

Puis f est de signe quelconque, on pose $f = f_+ - f_-$ où $f_+ = \sup(0, f)$ et $f_- = -(\inf(0, f))$, avec $f_+ \ge 0$ et $f_- \ge 0$ auxquelles on applique le résultat précédent. ▪

3.4 * Support d'une fonction

On se restreint aux ensembles E qui sont des espaces topologiques métriques : on suppose qu'il existe une fonction distance $d : E \times E \to \mathbb{R}_+$, i.e. une fonction d telle que :
$d(x, y) \geq 0$ pour tout $x, y \in E$ (positivité),
$d(x, y) = 0 \Rightarrow x = y$ (séparation),
$d(x, y) = d(y, x)$ pour tout $x, y \in E$ (symétrie),
$d(x, z) \leq d(x, y) + d(y, z)$ pour tout $x, y, z \in E$ (inégalité triangulaire).
On dispose alors de la topologie sur E engendrée par les boules ouvertes $B(x, \varepsilon) = \{y \in E : d(x, y) < \varepsilon\}$, pour tout $x \in E$ et tout $\varepsilon > 0$.

3.4.1 Support d'une fonction définie partout

Définition 3.14 Soit $f : E \to \mathbb{R}$ et soit $x \in E$. On dit que f est nulle dans un voisinage ouvert de x ssi :

$$\exists \varepsilon = \varepsilon_x > 0, \quad f_{|B(x,\varepsilon)} = 0, \tag{3.17}$$

i.e. :

$$\exists \varepsilon = \varepsilon_x > 0, \quad \forall y \in B(x, \varepsilon), \quad f(y) = 0. \tag{3.18}$$

Notation. Pour $f : E \to \mathbb{R}$, on note :

$$\begin{aligned} \Omega_0(f) &= \{x \in E \ : \ f = 0 \text{ dans un voisinage ouvert de } x\} \\ &= \{x \in E \ : \exists \varepsilon_x > 0, \ \forall y \in B(x, \varepsilon_x), \ f(y) = 0\}. \end{aligned} \tag{3.19}$$

Exemple 3.15 Pour $f = 1_{[0,1[} : \mathbb{R} \to \mathbb{R}$ on a $\Omega_0(f) =]-\infty, 0[\cup]1, \infty[$ (ouvert).
Pour $f : \mathbb{R} \to \mathbb{R}$ donnée par $f(x) = x$ on a $\Omega_0(f) = \emptyset$.
Pour $f = 1_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ on a $\Omega_0(f) = \emptyset$ (la fonction $1_{\mathbb{Q}}$ n'est nulle dans aucun ouvert). ∎

Proposition 3.16 *$\Omega_0(f)$ est un ouvert. C'est le plus grand ouvert contenu dans $Z_0 = \{x \in E : f(x) = 0\}$ (l'ensemble des zéros de f).*

Preuve. Soit $x \in \Omega_0(f)$, donc $\exists \varepsilon > 0$ t.q. $B(x, \varepsilon) \subset \Omega_0(f)$. Donc $\Omega_0(f)$ est ouvert.
Soit U un ouvert t.q. $U \subset Z_0$. Donc pour $y \in U$ il existe $\varepsilon > 0$ t.q. $B(y, \varepsilon) \subset U$, donc $B(y, \varepsilon) \subset Z_0$, donc $y \in \Omega_0(f)$, cf. (3.19). Donc $U \subset \Omega_0(f)$. ∎

Définition 3.17 Si $f : E \to \mathbb{R}$, le support de f est le fermé :

$$\text{supp}_0(f) \stackrel{\text{déf}}{=} \mathbb{R} - \Omega_0(f). \tag{3.20}$$

Exemple 3.18 $\text{supp}_0(1_{[0,1[}) = [0, 1]$. Et $\text{supp}_0(1_{\mathbb{Q}}) = \mathbb{R}$. ∎

Proposition 3.19 *Si $f : E \to \mathbb{R}$, alors le support de f est le fermé :*

$$\begin{aligned} \text{supp}_0(f) &= \overline{\{x \in E : f(x) \neq 0\}}, \\ &= \{x \in E : \forall \varepsilon > 0, \ \exists y \in B(x, \varepsilon), \ f(y) \neq 0\}, \end{aligned} \tag{3.21}$$

adhérence de l'ensemble des points où f est non nulle.

Preuve. Cf (3.19). ∎

Remarque 3.20 Soit $f : \mathbb{R} \to \mathbb{R}$ donnée par $f(x) = x$. On a $\Omega_0(f) = \emptyset$ et $\text{supp}_0(f) = \mathbb{R}$ (bien que $f(0) = 0$) : pour le support de f, ce qui est important n'est pas l'annulation de f en un point, mais le fait qu'elle varie en ce point (si f n'est pas identiquement nulle dans un voisinage ouvert de x, alors il est aussi intéressant d'étudier les variations de f en x qu'en les autres points). ∎

3.4.2 Support d'une fonction définie presque partout

Si f est définie uniquement presque partout, la définition du support est moins directe.

On munit E de sa tribu borélienne $\mathcal{A}_E = \sigma(\mathcal{O}_E)$ engendrée par les boules ouvertes $B(x,\varepsilon)$, et d'une mesure μ sur $(E, \mathcal{A}_E)$.

Définition 3.21 Soit $f : E \to \mathbb{R}$ mesurable et soit $x \in \mathbb{R}$. On dit que f est nulle μ-p.p. dans un voisinage (ouvert) de x ssi :

$$\exists \varepsilon = \varepsilon_x > 0,\ f_{|B(x,\varepsilon)} = 0\ \mu\text{-p.p.}, \tag{3.22}$$

i.e. :

$$\exists \varepsilon = \varepsilon_x > 0,\ \exists A_\varepsilon \in \mathcal{A} \text{ t.q. } \mu(A_\varepsilon) = 0, \quad \text{t.q. } \forall y \in B(x,\varepsilon) - A_\varepsilon,\ f(y) = 0, \tag{3.23}$$

à comparer avec (3.18).

Notation. Soit :

$$\begin{aligned} \Omega(f) &= \{x \in E \ :\ f = 0\ \mu\text{-p.p. dans un voisinage de } x\} \\ &= \{x \in E \ :\ \exists \varepsilon_x > 0,\ \forall\ \mu\text{-p.p. } y \in B(x,\varepsilon_x),\ f(y) = 0\}. \end{aligned} \tag{3.24}$$

À comparer avec (3.19).

Exemple 3.22 La fonction $1_{\mathbb{Q}}$ vérifie $\Omega(f) = \mathbb{R}$ relativement à la mesure de Lebesgue : c'est une fonction nulle presque partout. ▟

Proposition 3.23 *$\Omega(f)$ est un ouvert.*

C'est le plus grand ouvert contenu dans $Z = \{x : f(x) = 0\ \mu\text{-p.p.}\}$.

Preuve. C'est un ouvert car une union quelconque d'ouverts est un ouvert.

Soit U un ouvert t.q. $U \subset Z$. Donc pour $y \in U$ il existe $\varepsilon > 0$ t.q. $B(y,\varepsilon) \subset U$, donc $B(y,\varepsilon) \subset Z$, donc $y \in \Omega(f)$, cf. (3.24). Donc $U \subset \Omega(f)$. ▟

Définition 3.24 On appelle support de la fonction f mesurable le fermé :

$$\text{supp}(f) = E - \Omega(f). \tag{3.25}$$

(Fermé car complémentaire d'un ouvert.)

Donc $x \in \text{supp}(f)$ ssi (contraire de (3.23)) :

$$\forall \varepsilon > 0,\ \forall A_\varepsilon\ \mu\text{-négligeable},\ \exists y \in B(x,\varepsilon) - A_\varepsilon : f(y) \neq 0. \tag{3.26}$$

Proposition 3.25 *Si $f = g$ μ-p.p. (i.e. f et g mesurables sont dans la même classe d'équivalence), alors* $\text{supp}(f) = \text{supp}(g)$.

Preuve. Il s'agit de montrer que $\Omega(f) = \Omega(g)$ quand $f = g$ p.p..

Soit $x \in \Omega(f)$: $\exists \varepsilon > 0$, $\exists A_\varepsilon$ mesurable avec $\mu(A_\varepsilon) = 0$, $\forall y \in B(x,\varepsilon) - A_\varepsilon$, $f(y) = 0$. Comme $f = g$ p.p., $\exists B_\varepsilon$ mesurable avec $\mu(B_\varepsilon) = 0$ t.q. $\forall y \in B(x,\varepsilon) - B_\varepsilon$, $f(y) = g(y)$. D'où $\forall y \in B(x,\varepsilon) - (B_\varepsilon \bigcup A_\varepsilon)$, $g(y) = 0$. Comme $\mu(A_\varepsilon \bigcup B_\varepsilon) = 0$, on a $g = 0$ p.p. dans $B(x,\varepsilon)$. D'où $x \in \Omega(g)$. De même si on part de g. D'où $\Omega(f) = \Omega(g)$. ▟

Exemple 3.26 $\text{supp}(1_{[-1,1[}) = [-1,1]$. Et $\text{supp}(1_{\mathbb{Q}}) = \emptyset$ (fonction nulle p.p.). Comparer avec l'exemple 3.18. ▟

Exercice 3.27 Montrer que pour $f \in \mathcal{F}_M$, on a $\text{supp} f = \text{supp} f_+ \bigcup \text{supp} f_-$. (Voir (4.1) pour les définitions.)

Réponse. On a $f_+ = \text{supp}(f,0)$. Soit $x \in \text{supp}(f_+)$. Avec (3.26) : $\forall \varepsilon$, $\forall A_\varepsilon$, $\exists y$, $f_+(y) \neq 0$, on a donc $f(y) \neq 0$, donc $x \in \text{supp}(f)$. Idem pour $x \in \text{supp}(f_-)$. D'où $\text{supp}(f_+) \bigcup \text{supp}(f_-) \subset \text{supp}(f)$.

Réciproquement. Soit $x \in \text{supp}(f)$: $\forall \varepsilon$, $\forall A_\varepsilon$, $\exists y$, $f(y) \neq 0$, donc soit $f(y) > 0$, soit $f(y) < 0$, donc $x \in \text{supp}(f_+)$ ou $\in \text{supp}(f_-)$. ▟

3.5 * Supremum essentiel et $||.||_\infty$

Une fonction $F : E \to \mathbb{R}$ est bornée ssi $\exists c > 0$, $\forall x \in E$, $|f(x)| \le c$. Et pour les fonctions bornées on définit :

$$||f||_\infty = \sup_{x\in E} |f(x)|. \tag{3.27}$$

Problème : ne considérer que les fonctions définies partout est trop restrictif.

Définition 3.28 Soit $f : E \to$ mesurable. Pour $h \ge 0$ (“hauteur”) on note :

$$F_h = \{x \in E : |f(x)| > h\} \overset{\text{noté}}{=} \{|f| > h\}. \tag{3.28}$$

On dit que f est essentiellement bornée pour la mesure μ ssi :

$$\exists h > 0, \quad \mu(F_h) = 0, \tag{3.29}$$

autrement dit, ssi il existe h assez grand tel la mesure μ “ne voit plus” F_h.

On note $\mathcal{B}_\mu$ l'ensemble des fonctions essentiellement bornées pour la mesure μ.

$$\mathcal{B}_\mu = \{f : E \to \mathbb{R},\ f \text{ mesurable t.q. (3.29)}\}. \tag{3.30}$$

Définition 3.29 Pour $f \in \mathcal{B}_\mu$ (essentiellement bornée) on note :

$$||f||_\infty = \min\{h > 0 \text{ t.q. (3.29)}\} \overset{\text{noté}}{=} \text{supess}_{x\in E}|f(x)|. \tag{3.31}$$

Exemple 3.30 Mesure de Lebesgue. Soit $n \in \mathbb{N}^*$ et f_n la fonction donnée par $f_n(x) = 0$ pour tout $x \neq 0$ et $f_n(0) = n$. On a $\text{supess}(|f|) = ||f_n||_\infty = 0$ pour tout n. En effet, pour tout $h > 0$ on a $\{|f| > h\} = \{0\}$, et $\mu_\ell(\{0\}) = 0$ (le singleton est négligeable pour la mesure de Lebesgue).

Ici la fonction f_n est vue par l'appareil de mesure μ_ℓ comme étant la fonction nulle. ∎

Lemme 3.31 $||f||_\infty = 0$ *ssi* $f = 0$ *μ-p.p.*.

Preuve. On note $f_+ = \sup(f, 0)$ (partie positive de f) et $f_- = \sup(-f, 0)$ (partie positive de $-f$, i.e. partie négative de f au signe près), donc $f = f_+ - f_-$ et $|f| = f_+ + f_-$.

Si $f = 0$ μ-p.p. alors $f_+ = 0$ μ-p.p. et $f_- = 0$ μ-p.p., donc $|f| = 0$ μ-p.p., donc $\mu(\{|f| > 0\}) = 0$, donc $||f||_\infty = 0$, cf. (3.31).

Réciproque. Si $||f||_\infty = 0$ alors $\mu(\{|f| > 0\}) = 0$, donc $|f| = 0$ μ-p.p., cf. exercice 3.9, donc $f_+ = 0$ μ-p.p. et $f_- = 0$ μ-p.p., donc $f = 0$ μ-p.p.. ∎

Proposition 3.32 *$\mathcal{B}_\mu$ est un espace vectoriel, et $||.||_\infty$ est une semi-norme sur $\mathcal{B}_\mu$.*

Preuve. $\mathcal{B}_\mu$ est un sous-espace vectoriel de l'ensemble des fonctions mesurables (immédiat).

Si $f, g \in \mathcal{B}_\mu$ et $\lambda \in \mathbb{R}$, alors 1- $||f||_\infty \ge 0$ (immédiat), 2- $||\lambda f||_\infty = |\lambda|\,||f||_\infty$ (immédiat), 3- $||f + g||_\infty \le ||f||_\infty + ||g||_\infty$ (immédiat). ∎

4 Espaces $\mathcal{L}^1$ et L^1

4.1 Fonction μ-mesurable

On considère un espace mesuré $(E, \mathcal{A}, \mu)$ et la tribu étendue $\mathcal{A}_\mu$, cf § 1.5 et proposition 1.36 (et voir exercice 4.2).

Définition 4.1 Une fonction réelle $f : E \to \mathbb{R}$ est dite μ-mesurable si elle est $\mathcal{A}_\mu$-mesurable.

Donc f est mesurable ssi $f^{-1}(I) \subset \mathcal{A}_\mu$ pour tout I intervalle de $\mathbb{R}$, cf. définition 2.1.

Exercice 4.2 Montrer que toute fonction μ-mesurable est μ-p.p. égale à une fonction $\mathcal{A}$-mesurable.

N.B. : actuellement, toutes les fonctions physiques sont $\mathcal{A}$-mesurables. ∎

4.2 Fonctions μ-intégrables et espace $\mathcal{L}^1(E,\mathcal{A},\mu)$

4.2.1 Fonctions à valeurs réelles

Pour f fonction à valeurs réelles on considère les fonctions positives :

$$f_+ = \sup(f,0) \qquad \text{et} \qquad f_- = \sup(-f,0) = -\inf(f,0). \tag{4.1}$$

(Faire un dessin.) f_+ et f_- sont bien des fonctions positives, et on a :

$$f = f_+ - f_- \qquad \text{et} \qquad |f| = f_+ + f_-. \tag{4.2}$$

Définition 4.3 f est dite μ-intégrable ssi :
1. f_+ et f_- sont μ-mesurables (i.e. f est μ-mesurable), et :
2. $\mu(f_+) < \infty$ et $\mu(f_-) < \infty$.

Comme $f_+ \geq 0$ et $f_- \geq 0$, on a $\mu(|f|) = \mu(f_+)+\mu(f_-)$ (relation de Chasles $\mu([f|) = \mu(1_{A+}f)+\mu(1_{A-}(-f))$ où $A_+ = \{x : f \geq 0\}$ et $A_- = \{x : f < 0\}$), et donc f est μ-intégrable ssi f est μ-mesurable et :

$$\mu(|f|) < \infty \qquad \text{i.e. si} \qquad \int_E |f|\,d\mu < \infty. \tag{4.3}$$

Remarque 4.4 "μ-intégrable" est la généralisation de "absolument intégrable" définie pour l'intégrable de Riemann. En particulier, l'intégrale de Lebesgue ne traite pas du cas des fonctions "semi-convergentes" de l'intégrale de Riemann, cf. intégrales impropres. ∴

Exemple 4.5 La fonction donnée par $f(t) = e^{-t}1_{\mathbb{R}_+}$ définie sur $\mathbb{R}$ est intégrable sur $\mathbb{R}$ pour la mesure de Lebesgue.

Pas la fonction donnée par $f(t) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} 1_{]n,n+1[}$ puisque $\int_{\mathbb{R}} f(t)\,dt = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$. Cette fonction, non nulle sur $]1,\infty[$, est seulement semi-convergente au sens de Riemann.

De même pour $f(t) = \sum_{n=1}^{\infty}(-1)^{n+1}(n+1)1_{]\frac{1}{n+1},\frac{1}{n}[}$ (non nulle sur $]0,1[$). ∴

Définition 4.6 Pour f μ-intégrable on note :

$$\mu(f) \stackrel{\text{déf}}{=} \mu(f_+) - \mu(f_-) \quad (\in \mathbb{R}), \tag{4.4}$$

la mesure de f (valeur finie).

Dans la suite on sous-entendra systématiquement "μ-mesurable" (cf. 1. de la définition 4.3) : si f_+ n'est pas μ-mesurable, considérer $\mu(f_+)$ n'a aucun sens.

Notation. L'espace des fonctions μ-intégrables est noté $\mathcal{L}^1(E,\mathcal{A},\mu)$:

$$\begin{aligned}\mathcal{L}^1(E,\mathcal{A},\mu) &= \{f : E \to \mathbb{R} \text{ t.q. } \mu(|f|) < \infty\} \\ &\stackrel{\text{noté}}{=} \{f : E \to \mathbb{R} \text{ t.q. } \int_E |f|\,d\mu < \infty\}.\end{aligned} \tag{4.5}$$

(Donc sous-entendu "l'ensemble des fonctions f μ-mesurables" telles que.)

Proposition 4.7 *$\mathcal{L}^1(E,\mathcal{A},\mu)$ est un espace vectoriel, et si $f,g \in \mathcal{L}^1$ et $\lambda \in \mathbb{R}$, alors $\mu(f+\lambda g) = \mu(f)+\lambda\mu(g)$.*

Preuve. $f = f_+ - f_-$ et $g = g_+ - g_-$ donne $f + \lambda g = (f_+ + \lambda g_+) - \lambda(f_- + \lambda g_-)$ et on a la semi-linéarité, cf. thm. 3.6. ∴

Remarque 4.8 Pour la mesure de comptage $\mu = \sum_{n=1}^{\infty}\delta_n$, les fonctions intégrables sont exactement les séries absolument convergentes : on a

$$\mu(f) = \sum_{n=1}^{\infty} f(n),$$

et f est μ-intégrable si et seulement si :

$$\mu(|f|) = \sum_{n=1}^{\infty} |f(n)| < \infty.$$

∴

Proposition 4.9 *(Relation de Chasles.) Si f est une fonction μ-intégrable sur E, et si $A \in \mathcal{A}$, alors :*

$$\mu(f) = \mu(1_A f) + \mu(1_{E-A} f).$$

Et si $f = 0$ sur A, alors $\mu(1_A f) = 0$.

Preuve. Ce résultat est vrai pour f_+ et f_- d'après la proposition 3.8. D'où le résultat par définition de $\mu(f)$. ∎

Proposition 4.10

— *Si $f \in \mathcal{L}^1(E)$ alors*

$$|\mu(f)| \le \mu(|f|), \quad \textit{noté} \quad |\int_E f\,d\mu| \le \int_E |f|\,d\mu. \tag{4.6}$$

— *Inégalité de la moyenne : si $f \in \mathcal{L}^1(E)$ et si $g \in \mathcal{L}^\infty(E)$ (i.e. g est mesurable et bornée μ-p.p. sur E), alors $fg \in \mathcal{L}^1(E)$ et, avec $||g||_\infty = \mathrm{supess}_{x\in E}|g(x)| =^{noté} \sup_{x\in E}|g(x)|$,*

$$|\mu(fg)| \le ||g||_\infty\, \mu(|f|), \quad \textit{noté} \quad |\int_E fg\,d\mu| \le ||g||_\infty \int_E |f|\,d\mu. \tag{4.7}$$

Preuve. Exercice. ∎

Remarque 4.11 On retrouve l'intégrale de Riemann $\int_a^b f(x)\,dx$ définie pour les fonctions continues par morceaux sur un intervalle compact $[a,b]$ (une telle fonction est absolument intégrable) en définissant, si $a<b$, avec ici $E =]a,b[$:

$$\begin{cases} \displaystyle\int_a^b f(x)\,dx \overset{\text{déf}}{=} \int_{[a,b]} f(x)\,dx = \mu_l(1_{[a,b]}f), \\ \text{et} \quad \displaystyle\int_b^a f(x)\,dx \overset{\text{déf}}{=} -\int_{[a,b]} f(x)\,dx = -\mu_l(1_{[a,b]}f). \end{cases}$$

∎

Exercice 4.12 Approximation d'une fonction $\mathcal{L}^1(\mathbb{R})$ par une fonction étagée : soit $f \in \mathcal{L}^1(\mathbb{R})$. Montrer : $\forall \varepsilon > 0$, $\exists g$ fonction étagée t.q. :

$$||f-g||_{\mathcal{L}^1} < \varepsilon. \tag{4.8}$$

Réponse. C'est vrai pour f_+ et pour f_-, cf. (3.7), donc pour $f = f_+ - f_-$. ∎

4.2.2 Fonctions à valeurs complexes

Soit une fonction complexe $f : E \to \mathbb{C}$, de parties réelles et imaginaires f_1 et $f_2 : E \to \mathbb{R}$:

$$f = f_1 + i f_2.$$

Son module $|f|$ est la fonction $|f| : x \in E \to |f|(x) \in \mathbb{R}$ donnée par :

$$|f|(x) \overset{\text{déf}}{=} |f(x)| = \sqrt{f_1(x)^2 + f_2(x)^2}. \tag{4.9}$$

La fonction conjuguée $\overline{f}$ de f est définie par $\overline{f} : x \in E \to \overline{f}(x) \in \mathbb{R}$ donnée par :

$$\overline{f}(x) \overset{\text{déf}}{=} \overline{f(x)} = f_1(x) - i f_2(x), \tag{4.10}$$

et en particulier $|f|^2(x) =^{\text{déf}} |f(x)|^2 = f(x)\overline{f}(x)$ pour $x \in E$.

Définition 4.13 La fonction complexe $f = f_1 + i f_2$ est dite μ-intégrable ssi :

1. f_1 et f_2 sont μ-mesurables (f est μ-mesurable), et :
2. $\mu(|f_1|) < \infty$ et $\mu(|f_2|) < \infty$.

Et on pose alors :

$$\mu(f) \overset{\text{déf}}{=} \mu(f_1) + i\mu(f_2), \tag{4.11}$$

valeur complexe (non réelle : on sort du cadre des "mesures positives").

(À comparer avec $\mu(|f|) = \mu(\sqrt{f_1^2 + f_2^2})$.)
Donc f est dite μ-intégrable si f est μ-mesurable et :

$$\int |f|\,d\mu < \infty.$$

Et les résultats précédents, donnés pour les fonctions à valeurs réelles, sont conservés (quand ils ont un sens).

4.3 Critère d'intégrabilité par domination

Un corollaire du théorème 3.6 donne immédiatement :

Corollaire 4.14 *(Intégrabilité par domination.) Si $f : E \to \mathbb{R}$ (où $\mathbb{C}$) est μ-mesurable et s'il existe $g \in \mathcal{L}^1(E;\mathbb{R}_+)$ (fonction à valeurs réelles positives qui est μ-intégrable) telle que $|f| \leq g$ μ-p.p. alors f est μ-intégrable.*

En effet si f est mesurable alors $|f|$ est mesurable positive de mesure $\mu(|f|)$, et $|f|$ est bornée par g donc $\mu(|f|) \leq \mu(g)$ (avec la monotonie pour les fonctions positives, théorème 3.6) avec $\mu(g) < \infty$ car g intégrable et donc $|f|$ est intégrable, et donc f est intégrable.

Exemple 4.15 Montrer : $f : x \in \mathbb{R}_+ \to f(x) = \sin(x)e^{-x}$ est dx-intégrable sur $\mathbb{R}_+$. ∴

4.4 Reste d'une intégrale

Pour simplifier l'écriture, on se place dans le cas $E = \mathbb{R}$, i.e. d'une fonction $f : \mathbb{R} \to \mathbb{R}$, et dans le cas $\mu = \mu_\ell$ mesure de Lebesgue ou dans le cas $\mu = \sum_{\mathbb{Z}} \delta_k$ mesure de comptage sur $\mathbb{Z}$.

Proposition 4.16 *Si $f \in \mathcal{L}^1(\mathbb{R})$, soit $\int_{\mathbb{R}} |f(x)|\,d\mu(x) < \infty$, alors :*

$$\forall \varepsilon > 0,\ \exists X_\varepsilon \geq 0,\ \int_{\mathbb{R}-[-X_\varepsilon,X_\varepsilon]} |f(x)|\,d\mu(x) < \varepsilon. \tag{4.12}$$

Et $\int_{\mathbb{R}-[-X_\varepsilon,X_\varepsilon]} |f(x)|\,d\mu(x)$ est appelé le reste de l'intégrale, et $\int_{\mathbb{R}} |f(x)|\,d\mu(x)$ est approché à ε-près par $\int_{-X_\varepsilon}^{X_\varepsilon} |f(x)|\,d\mu(x)$.

Preuve. Comme $\int_{\mathbb{R}} |f(x)|\,d\mu(x) < \infty$ et $|f| \geq 0$, on a $\int_{\mathbb{R}_+} |f(x)|\,d\mu(x) < \infty$. Soit $F(x) = \int_{t=0}^{x} |f(t)|\,d\mu(t)$. Comme une mesure est positive et que $|f|$ est positive, la fonction F est positive croissante sur $\mathbb{R}_+$, et donc bornée par $F_\infty = \lim_{Y\to\infty} \int_0^Y |f(x)|\,d\mu(x) = \int_0^\infty |f(x)|\,d\mu(x) < \infty$. Donc $\forall \varepsilon > 0$, $\exists Y_\varepsilon \geq 0$ t.q. $F_\infty - F(Y_\varepsilon) < \varepsilon/2$, par définition de la limite.

Idem du côté $-\infty$ en considérant $G(x) = \int_{t=x}^{0} |f(t)|\,d\mu(t)$ pour $x \leq 0$. Donc $\exists Z_\varepsilon \leq 0$ t.q. $G_\infty - G(Z_\varepsilon) < \varepsilon/2$. Et $X_\varepsilon = \max(Y_\varepsilon, -Z_\varepsilon)$ convient. ∴

4.5 Fonctions négligeables et semi-norme sur $\mathcal{L}^1(E,\mathcal{A},\mu)$

On reprend la relation d'équivalence $\mathcal{R}$, cf. (1.29), en la restreignant à $\mathcal{L}^1(E,\mathcal{A},\mu) =^{\text{noté}} \mathcal{L}^1$: pour $f \in \mathcal{L}^1$,

$$\dot{f} = \{g \in \mathcal{L}^1 : g = f\ \mu\text{-p.p.}\}. \tag{4.13}$$

En particulier,

$$\mathcal{N} = \dot{0} = \{f \in \mathcal{L}^1 : f = 0\ \mu\text{-p.p.}\} = \{\text{fonctions } \mathcal{L}^1 \text{ nulles } \mu\text{-p.p.}\}. \tag{4.14}$$

Proposition 4.17 *$\mathcal{N}$ est un sous-espace vectoriel de $\mathcal{L}^1$.*

Toute classe d'équivalence est un espace affine dans $\mathcal{L}^1$, à savoir : si F est une classe d'équivalence, si $f \in F$, donc $F = \dot{f}$, alors :

$$\dot{f} = f + \mathcal{N}. \tag{4.15}$$

Preuve. $\mathcal{N}$ sous-espace vectoriel de $\mathcal{L}^1$: si $f, g \in \mathcal{N}$ soient $A_f, A_g \in \mathcal{A}$ tels que $\mu(A_f) = 0$, $\mu(A_g) = 0$ et $f(x) = 0$ pour tout $x \in E - A_f$, $g(x) = 0$ pour tout $x \in E - A_g$. Alors $B = A_f \bigcup A_g$ vérifie $\mu(B) = 0$ et $(f + \lambda g)(x) = 0$ pour tout $x \in E - B$ et $\lambda \in \mathbb{R}$. Donc $f + \lambda g \in \mathcal{N}$.

Puis soit $f \in \dot{f}$; si $g \in \dot{f}$ alors, par définition de la classe d'équivalence, $g - f \in \mathcal{N}$ espace vectoriel, donc $g \in f + \mathcal{N}$, donc $\dot{f} \subset f + \mathcal{N}$. Et $f + \mathcal{N} \subset \dot{f}$ (immédiat), donc $\dot{f} = f + \mathcal{N}$ (espace affine). ∴

Proposition 4.18 *Si $f \in \mathcal{N}$ alors $\mu(f) = 0$ (la réciproque est trivialement fausse : prendre f impaire avec la mesure de Lebesgue sur $[-1, 1]$).*

Si $f, g \in \mathcal{L}^1$ vérifient $f = g$ μ-p.p. alors $\mu(f) = \mu(g)$ et $\mu(|f|) = \mu(|g|)$.

Preuve. Soit $f \in \mathcal{N}$: il existe $A \in \mathcal{A}$ tel que $\mu(A) = 0$ et $f(x) = 0$ pour tout $x \in E - A$, i.e. $1_{E-A}f = 0$ (application nulle). D'où $\mu(1_{E-A}f) = \mu(0) = 0$. Et comme A est négligeable on a $\mu(1_A f) = 0$ (exercice 3.7). Et comme $\mu(f) = \mu(1_A f) + \mu(1_{E-A}f)$ (proposition 3.8), on a $\mu(f) = 0 + 0 = 0$.

Si $f\mathcal{R}g$ alors $f - g = 0$ μ-p.p., donc $f_+ = g_+$ et $f_- = g_-$ μ-p.p. (immédiat), donc $\mu(f) = \mu(g)$ et $\mu(|f|) = \mu(|g|)$. ∴

Définition 4.19 On définit la fonction $||.||_{\mathcal{L}^1} : \mathcal{L}^1 \to \mathbb{R}$ par :

$$||f||_{\mathcal{L}^1} \stackrel{\text{déf}}{=} \mu(|f|) = \int |f|\,d\mu. \tag{4.16}$$

Proposition 4.20 *Une fonction mesurable f est μ-négligeable ssi $||f||_{\mathcal{L}^1} = 0$:*

$$f \in \mathcal{N} \quad\iff\quad ||f||_{\mathcal{L}^1} = 0. \tag{4.17}$$

Preuve. $\Rightarrow$: donné par la proposition 4.18.
Réciproque $\Leftarrow$: voir exercice 3.9 en remplaçant f par $|f|$, et $|f| = 0$ p.p. donne $f = 0$ p.p.. ∎

Remarque 4.21 Si au lieu de $\int |f|\,d\mu = 0$ on suppose seulement $\int f\,d\mu = 0$, on n'a pas $f = 0$ μ-p.p. en général (prendre f impaire avec la mesure de Lebesgue) ; ∎

Proposition 4.22 *La fonction $||.||_{\mathcal{L}^1} : \mathcal{L}^1 \to \mathbb{R}$ définie sur $\mathcal{L}^1$ par :*

$$||f||_{\mathcal{L}^1} = \mu(|f|) = \int |f|\,d\mu \tag{4.18}$$

est une semi-norme sur $\mathcal{L}^1$.

Preuve. Pour f et g dans $\mathcal{L}^1$ et λ réel on a : $\mu(|f+g|) \le \mu(|f|)+\mu(|g|)$ car $|f+g| \le |f|+|g|$, et $\mu(|\lambda f|) = |\lambda|\mu(|f|)$ car $|\lambda f| = |\lambda|\,|f|$ et μ semi-linéaire. ∎

$||.||_{\mathcal{L}^1}$ n'est pas une norme si $\mathcal{N} \ne \{0\}$, puisque $||f||_{\mathcal{L}^1} = 0$ équivaut à $f \in \mathcal{N}$ (proposition précédente 4.20). Par exemple, avec la mesure de Lebesgue, $\mu_\ell(1_{\mathbb{Q}}) = 0$ bien que $1_{\mathbb{Q}} \ne 0$.

4.6 Espace $L^1(E,\mathcal{A},\mu)$

On considère l'espace quotient, relativement à la relation d'équivalence (1.29) :

$$L^1(E,\mathcal{A},\mu) = \mathcal{L}^1(E,\mathcal{A},\mu)/\mathcal{R} = \{\text{les classes d'équivalence } \dot f \text{ pour } f \in \mathcal{L}^1(E,\mathcal{A},\mu)\}, \tag{4.19}$$

espace qui est également noté :

$$L^1(E,\mathcal{A},\mu) = \mathcal{L}^1(E,\mathcal{A},\mu)/\mathcal{N} = \{f+\mathcal{N},\ f \in \mathcal{L}^1(E,\mathcal{A},\mu)\}. \tag{4.20}$$

Si aucune confusion n'est possible (i.e. E, $\mathcal{A}$ et μ sont précisés par le contexte), on note simplement $L^1 = L^1(E,\mathcal{A},\mu)$. Donc L^1 est l'ensemble des ensembles définis par :

$$L^1 = \{F \subset \mathcal{L}^1 \text{ t.q. } \exists f \in \mathcal{L}^1,\ F = \dot f\}. \tag{4.21}$$

On notera $\dot f$ les éléments de L^1 (sous entendu : pour décrire une classe d'équivalence on prend un de ses représentants f, et on note $\dot f = f + \mathcal{N}$ la classe de f).

Proposition 4.23 $(L^1,+,.)$ *est un espace vectoriel, où + et . sont les opérations induites par celles de $\mathcal{L}^1$.*

Preuve. Par définition de + et . induites par $\mathcal{L}^1$, on a $\dot f + \dot g = \dot{\overbrace{f+g}}$ et $\lambda \dot f = \dot{\overbrace{\lambda f}}$. Ainsi $(L^1,+,.)$ a une structure d'espace vectoriel : vérification facile. ∎

On munit L^1 de la semi-norme induite par celle de $\mathcal{L}^1$: pour $\dot f \in L^1$:

$$\text{si } f \in \dot f,\quad ||\dot f||_{L^1} \stackrel{\text{déf}}{=} ||f||_{\mathcal{L}^1} \quad (\stackrel{\text{déf}}{=} \mu(|f|)). \tag{4.22}$$

Proposition 4.24 *$||\dot f||_{L^1}$ est indépendante du choix de f dans $\dot f$.*
Et $||.||_{L^1}$ donnée par (4.22) est une norme sur L^1.

Preuve. Soit $F \in L^1$ et soit $f, g \in F$: il s'agit de montrer que $||g||_{\mathcal{L}^1} = ||f||_{\mathcal{L}^1}$. Comme $F = f + \mathcal{N}$, il existe $h \in \mathcal{N}$ t.q. $g = f + h$, i.e. t.q. $f = g$ p.p.. Soit $A = \{x : f(x) \ne g(x)\} \in \mathcal{A}$: comme $f = g$ p.p., on a $\mu(A) = 0$ et $1_{E-A}f = 1_{E-A}g$, donc $\mu(|1_{E-A}f|) = \mu(|1_{E-A}g|)$. Avec $|f| = |1_{E-A}f| + |1_A f|$ (immédiat), donc $\mu(|f|) = \mu(|1_{E-A}f|) + \mu|1_A f|)$ (semi-linéarité), donc $\mu(|f|) = \mu(|1_{E-A}f|)$ car $\mu(|1_A f|) = 0$ (cf. exercice 3.7 avec la relation immédiate $|1_A f| = 1_A|f|$). De même $\mu(|g|) = \mu(|1_{E-A}g|)$. Donc $\mu(|f|) = \mu(|g|)$. Ainsi $||.||_{L^1} : L^1 \to \mathbb{R}$ est bien une fonction.

$||.||_{L^1}$ est une semi-norme sur L^1 : vérification facile car $||.||_{\mathcal{L}^1}$ est une semi-norme sur $\mathcal{L}^1$.

Et si $||\dot f||_{L^1} = 0$ on a $||f||_{\mathcal{L}^1} = 0$ pour $f \in \dot f$, soit $f \in \mathcal{N} = \dot 0$, donc $\dot f = \dot 0$ l'élément nul de L^1. Donc $||.||_{L^1}$ est une norme. ∎

Théorème 4.25 *(Théorème de Fischer–Riesz) L'espace vectoriel L^1 muni de la norme $||.||_1$ est complet : c'est donc un espace de Banach.*

Preuve. Voir paragraphe 8.5. ∎

Remarque 4.26 On parle abusivement d'une fonction $f \in L^1$ alors qu'un élément de L^1 n'est pas une fonction mais un ensemble (une classe) de fonctions. Cet abus permet de parler de la norme de la 'fonction' $f \in L^1$, alors qu'on devrait parler de la norme d'une fonction de la classe, cette norme étant indépendante du représentant choisi dans la classe. Et on notera abusivement :

$$||f||_{\mathcal{L}^1} \overset{\text{noté}}{=} ||f||_{L^1} \overset{\text{noté}}{=} ||f||_1. \tag{4.23}$$

∎

4.7 Convergence dans $C^0(\mathbb{R})\bigcap L^1(\mathbb{R})$

Cas de la mesure de Lebesgue.

Proposition 4.27 *Si $(f_n)_{\mathbb{N}^*}$ est une suite de fonctions dans $C^0(\mathbb{R})\bigcap L^1(\mathbb{R})$ qui converge vers g dans $(C^0(\mathbb{R}),||.||_\infty)$ (convergence uniforme) et qui converge vers h dans $(L^1(\mathbb{R}),||.||_{L^1})$, alors $h = g$ p.p. (en particulier h est continue p.p.).*

Cet énoncé s'écrit sans abus de vocabulaire :

Si $(f_n)_{\mathbb{N}^*}$ est une suite de fonctions dans $C^0(\mathbb{R})\bigcap \mathcal{L}^1(\mathbb{R})$ qui converge dans $(C^0(\mathbb{R}),||.||_\infty)$, et on note $g = \lim f_n$ dans C^0, et telle que la suite des classes $(\dot{f}_n)_{\mathbb{N}^*}$ converge dans $(L^1(\mathbb{R}),||.||_{L^1})$, et on note $\dot{h} = \lim \dot{f}_n$ dans L^1, alors $\dot{h} = \dot{g}$, et on peut prendre g comme représentant continu dans $\dot{h}$.

Preuve. Hypothèses : $\left\{\begin{aligned} &||f_n-g||_\infty = \sup_{\mathbb{R}} |f_n(x)-g(x)| \underset{n\to\infty}{\longrightarrow} 0,\\ &||\dot{f}_n-\dot{h}||_{L^1} = ||f_n-h||_{\mathcal{L}^1} = \int_{\mathbb{R}} |f_n(x)-h(x)|\,dx \underset{n\to\infty}{\longrightarrow} 0,\end{aligned}\right\}$ où on a noté h un représentant dans $\dot{h}$ (soit $h \in \dot{h}$).

Donc pour tout $K = [-r,r]$ compact, $r > 0$, comme $g \in L^1(K;\mathbb{R})$ (car continue), on a :

$$\begin{aligned} ||\dot{g}-\dot{h}||_{L^1(K)} = ||g-h||_{\mathcal{L}^1(K)} = \int_K |g(x)-h(x)|\,dx &\le \int_K |g(x)-f_n(x)|\,dx + \int_K |f_n(x)-h(x)|\,dx\\ &\le ||g-f_n||_\infty |K| + ||h-f_n||_{\mathcal{L}^1} \underset{n\to 0}{\longrightarrow} 0. \end{aligned}$$

Donc $\dot{g} = \dot{h} \in L^1(K;\mathbb{R})$, soit $g = h$ p.p. dans $[-r,r]$. Vrai pour tout $r > 0$, donc vrai dans $\mathbb{R}$. ∎

4.8 Espace $\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$

Définition 4.28 L'espace des fonctions localement intégrables sur $\mathbb{R}$ est l'ensemble des fonctions mesurables qui sont intégrables sur tout compact mesurable de $\mathbb{R}$:

$$\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : \forall K \subset \mathbb{R},\ K \text{ compact } \Rightarrow\ f_{|K} \in \mathcal{L}^1(K)\}. \tag{4.24}$$

Exemple 4.29 Mesure de Lebesgue : $f \in \mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$ ssi pour tout $a < b$ on a $\int_a^b |f(x)|\,dx < \infty$.

En particulier tout fonction $f \in \mathcal{L}^1(\mathbb{R})$ est dans $\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$: on a $\mathcal{L}^1(\mathbb{R}) \subset \mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$.

La fonction $1_{\mathbb{R}}$ est dans $\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$ mais n'est pas dans $\mathcal{L}^1(\mathbb{R})$: donc $\mathcal{L}^1(\mathbb{R}) \subsetneq \mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$. ∎

On reprend la relation d'équivalence $\mathcal{R}$ cette fois-ci restreinte à $\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$.

Proposition 4.30 $\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$ *est un s.e.v. de $\mathcal{F}_M$ (ensemble des fonctions mesurables).*

Preuve. Exercice. ∎

Définition 4.31 $L^1_{\mathrm{loc}}(\mathbb{R}) = \mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})/\mathcal{N}$ est l'ensemble des classes de fonctions $\mathcal{L}^1_{\mathrm{loc}}(\mathbb{R})$.

5 Théorèmes de convergence

Ces théorèmes permettent d'inverser signe $\int$ et lim (passage à la limite sous le signe $\int$), sous certaines conditions, en particulier sous la condition (assez faible) de domination :

$$\exists g \in \mathcal{L}^1, \quad \forall n \in \mathbb{N}^*, \quad |f_n| \le g \quad \Longrightarrow \quad \lim_{n\to\infty} (\int_E f_n)\, d\mu = \int_E (\lim_{n\to\infty} f_n)\, d\mu.$$

Comme $\int_E f\, d\mu = \int_{E-A} f\, d\mu$ quand A est μ-négligeable, les théorèmes suivants seront valides sous l'hypothèse (f_n) converge μ-p.p. vers $f = \lim f_n$, voir (1.28).

Rappel : pour l'intégrale de Riemann, le passage à la limite sous le signe $\int$ est en général démontré sous l'hypothèse (très forte) de convergence uniforme $||f - f_n||_\infty \longrightarrow_{n\to\infty} 0$.

5.1 Théorème de Beppo–Lévi (ou de convergence monotone)

Ce théorème est à la base de toutes les démonstrations des théorèmes de Lebesgue qui suivent.

Ici les fonctions considérées sont des fonctions $E \to \overline{\mathbb{R}}$ mesurables, où $\mathbb{R}$ est muni de sa tribu borélienne et $E = (E, \mathcal{A}, \mu)$ (voir exercice 3.12).

Rappel : une suite (f_n) de fonctions de E dans $\mathbb{R}$ est croissante ssi pour tout $n \in \mathbb{N}$ on a $f_n \le f_{n+1}$ (le graphe de f_{n+1} est au dessus du graphe de f_n), i.e. : $\forall x \in E$, $\forall n \in \mathbb{N}$, $f_n(x) \le f_{n+1}(x)$.

Théorème 5.1 Beppo–Lévi (théorème de de convergence monotone).
Soit $(f_n)_\mathbb{N}$ une suite croissante de fonctions réelles mesurables et positives, et, pour $x \in E$:

$$f(x) \overset{\text{déf}}{=} \sup_\mathbb{N} f_n(x) \overset{\text{noté}}{=} \lim_{n\to\infty} f_n(x) \quad \in [0,\infty]. \tag{5.1}$$

La fonction $f = \lim_{n\to\infty} f_n : E \to \overline{\mathbb{R}_+}$ ainsi définie est mesurable et positive, et, dans $[0,\infty]$:

$$\mu(f) = \lim_{n\to\infty} (\mu(f_n)) \quad \text{i.e.} \quad \mu(\lim_{n\to\infty} f_n) = \lim_{n\to\infty} (\mu(f_n)), \tag{5.2}$$

i.e. on peut donc inverser les signes μ et lim *dans le cas d'une suite croissante de fonctions positives. Avec la notation intégrale on a :*

$$\lim_{n\to\infty} (\int_E f_n\, d\mu) = \int_E (\lim_{n\to\infty} f_n)\, d\mu, \tag{5.3}$$

i.e. on peut passer à la limite sous le signe $\int$ (cas d'une suite croissante de fonctions positives).

Egalement vrai si $(f_n)_{\mathbb{N}^*}$ une suite décroissante de fonctions réelles mesurables et négatives (considérer $(-f_n)_{\mathbb{N}^*}$ qui est une suite croissante de fonctions réelles mesurables et positives).

Preuve. (Corollaire de (1.9) et de (3.8).) Avec le corollaire 2.12 on a f mesurable.

Soit $\alpha_n = \mu(f_n)$. La suite $(\alpha_n)_\mathbb{N}$ est positive croissante dans $[0,\infty]$ car $(f_n)_\mathbb{N}$ est croissante et une mesure est monotone, cf. prop. 3.6. Donc $(\alpha_n)_\mathbb{N}$ converge dans $\overline{\mathbb{R}_+}$. Soit $\alpha = \lim_n \alpha_n \in \overline{\mathbb{R}_+}$.

Comme $f_n \le f$ pour tout n, on a $\mu(f) \ge \mu(f_n) = \alpha_n$ pour tout n, donc $\mu(f) \ge \alpha$.

Soit $g \ge 0$ une fonction étagée telle que $0 \le g \le f$. Montrons que $\mu(g) \le \alpha$, ce qui donnera $\mu(f) \le \alpha$, cf. (3.8). Posons pour $c \in]0,1[$ et $n \in \mathbb{N}^*$:

$$E_{c,n} = \{x \in E : f_n(x) \ge cg(x)\}. \tag{5.4}$$

Comme $f_{n+1} \ge f_n$ pour tout n, la suite $(E_{c,n})_{n\in\mathbb{N}}$ est croissante, et :

$$\bigcup_{n\in\mathbb{N}} E_{c,n} = \lim_{n\to\infty} E_{c,n} = E, \quad \text{donc} \quad \lim_{n\to\infty} \mu(E_{c,n}) = \mu(E). \tag{5.5}$$

En effet, pour $x \in E$, ayant $f_n(x) \longrightarrow_{n\to\infty} f(x)$ et $c < 1$, on a $f_n(x) \ge c\, f(x)$ dès que n est assez grand, donc $x \in E_{c,n}$ dès que n est assez grand, et on a (1.9). Donc, pour tout $x \in E$:

$$(1_{E_{c,n}} g(x))_\mathbb{N} \underset{n\to\infty}{\longrightarrow} g(x), \quad \text{donc} \quad \mu(1_{E_{c,n}} g) \underset{n\to\infty}{\longrightarrow} \mu(1_E g) = \mu(g), \tag{5.6}$$

En effet, avec (5.5), $(1_{E_{c,n}} g)_\mathbb{N}$ est une suite croissante de fonctions étagées positives, et si $x \in E$ alors $x \in E_{c,n}$ pour n assez grand, cf. (5.5), donc $g(x) = 1_{E_{c,n}} g(x)$ pour n assez grand ; et on a (3.7).

Comme une mesure est croissante, on a pour tout $n \in \mathbb{N}$:

$$\alpha \ge \alpha_n = \mu(f_n) = \mu(1_E f_n) \ge \mu(1_{E_{c,n}} f_n) \ge \mu(c\, 1_{E_{c,n}} g) = c\, \mu(1_{E_{c,n}} g). \tag{5.7}$$

Donc $\forall c \in]0,1[$, $\alpha \ge c\, \mu(g)$, cf. (5.6), donc $\alpha \ge \mu(g)$, donc $\alpha \ge \mu(f)$, cf. (3.8). ▪

Remarque 5.2 Contre-exemple avec la mesure de Lebesgue $\mu_\ell = dx$: soit $f_n = n1_{]0,\frac{1}{n}]}$. On a $f_n \geq 0$ pour tout n et $f_n \underset{n\to\infty}{\longrightarrow} f = 0$ p.p., avec immédiatement :

$$\int_{\mathbb{R}} f_n(x)\,dx = 1 \quad \text{et} \quad \int_{\mathbb{R}} f(x)\,dx = 0. \tag{5.8}$$

Donc, $\int_{\mathbb{R}}(\lim f_n)(x)\,dx \neq \lim(\int_{\mathbb{R}} f_n(x)\,dx)$. Mais ici la suite (f_n) n'est pas croissante : pour $x = \frac{1}{2}(\frac{1}{n} + \frac{1}{n+1})$, on a $f_n(x) = 1 > 0 = f_{n+1}(x)$. Faire un dessin. L'hypothèse suite croissante (de fonctions positives) est primordiale. ∎

Corollaire 5.3 *Sous les hypothèses du théorème précédent, la fonction positive $f = \sup_{\mathbb{N}}(f_n)$ est intégrable si et seulement si la suite $(\mu(f_n))_{\mathbb{N}}$ est bornée supérieurement. Et on a alors :*

$$\mu(f) = \sup_{\mathbb{N}}(\mu(f_n)) \quad \textit{i.e.} \quad \sup_{\mathbb{N}}\Big(\int_E f_n\,d\mu\Big) = \int_E\Big(\sup_{\mathbb{N}} f_n\Big)\,d\mu \quad (\textit{dans } \mathbb{R}).$$

Preuve. Immédiat. ∎

Corollaire 5.4 *Soit $(g_n)_{\mathbb{N}}$ une suite de fonctions réelles mesurables et positives. Alors :*

$$\mu(\sum_{n=1}^{\infty} g_n) = \sum_{n=1}^{\infty}(\mu(g_n)). \tag{5.9}$$

Preuve. On pose $f_n = \sum_{k=1}^n g_k$. On a $(f_n)_{\mathbb{N}^*}$ suite croissante de fonctions positives. (5.2) donne :

$$\mu(\lim_{n\to\infty}\sum_{k=1}^n g_k) = \lim_{n\to\infty}(\mu(\sum_{k=1}^n g_k)), \quad \text{donc} \quad = \lim_{n\to\infty}(\sum_{k=1}^n \mu(g_k)). \tag{5.10}$$

D'où (5.9). ∎

Exemple 5.5 Avec $\mu = \mu_\ell$ la mesure de Lebesgue, et si les g_n sont des fonctions positives :

$$\int_E(\sum_{n=1}^{\infty} g_n(x))\,d\mu_\ell(x) = \sum_{n=1}^{\infty}(\int_E g_n(x)\,d\mu_\ell(x)), \tag{5.11}$$

puisque $\displaystyle\int_E(\lim_{n\to\infty}\sum_{k=1}^n g_k(x))\,d\mu_\ell(x) = \sum_{n=1}^{\infty}(\int_E g_n(x)\,d\mu_\ell(x))$.

Avec $\mu = \sum_{i=1}^{\infty}\delta_i$ (mesure de comptage), et si $b_{ij} \geq 0$ pour tout $i,j \in \mathbb{N}^*$, alors :

$$\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty} b_{ij}) = \sum_{j=1}^{\infty}(\sum_{i=1}^{\infty} b_{ij}). \tag{5.12}$$

En effet, on pose $b_j : \left\{\begin{array}{ll}\mathbb{N}^* & \to \mathbb{R}_+ \\ i & \to b_j(i) = b_{ij}\end{array}\right\}$, et on a $\displaystyle\sum_{i=1}^{\infty}\delta_i(\sum_{j=1}^{\infty} b_j) = \sum_{j=1}^{\infty}(\sum_{i=1}^{\infty}\delta_i(b_j))$. ∎

5.2 Lemme de Fatou

5.2.1 Rappel : lim inf

Soit $(\lambda_n)_{\mathbb{N}^*}$ une suite de réels. Soit la suite $(\mu_n)_{\mathbb{N}^*}$ de réels définie par :

$$\mu_n = \inf_{k\geq n}\lambda_k \quad (= \inf\{\lambda_n, \lambda_{n+1}, \lambda_{n+2}, ...\}). \tag{5.13}$$

Proposition 5.6 *La suite $(\mu_n)_{\mathbb{N}^*}$ est croissante dans $\overline{\mathbb{R}}$, donc convergente dans $\overline{\mathbb{R}}$.*

Preuve. $\mu_n = \inf_{k\geq n}\lambda_k = \min(\lambda_n, \inf_{k\geq n+1}\lambda_k) \leq \inf_{k\geq n+1}\lambda_k = \mu_{n+1}$. ∎

Définition 5.7 La limite inférieure de la suite $(\lambda_n)_{\mathbb{N}^*}$ est le réel :

$$\begin{aligned}\liminf_{n\to\infty}(\lambda_n) &\overset{\text{déf}}{=} \lim_{n\to\infty}(\mu_n) = \sup_{n\in\mathbb{N}^*}(\mu_n) \\ &= \sup_{n\in\mathbb{N}^*}(\inf_{k\geq n}\lambda_k) = \lim_{n\to\infty}(\inf_{k\geq n}\lambda_k).\end{aligned} \tag{5.14}$$

Exercice 5.8 Soit $(\lambda_n)_{\mathbb{N}^*}$ une suite qui a deux valeurs d'adhérences (distinctes) a_1 et a_2. Donc : $\forall\varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{N}^*$ t.q. $\forall n \geq N_\varepsilon$ on a $|\lambda_n - a_1| < \varepsilon$ ou $|\lambda_n - a_2| < \varepsilon$ (dessin).

1- Rappeler la définition 11- d'une limite, 12- d'une sous-suite, et 13- d'une valeur d'adhérence

2- Montrer qu'il existe deux sous-suites $(\alpha_k)_{k\in\mathbb{N}^*}$ et $(\beta_k)_{k\in\mathbb{N}^*}$ t.q. $(\alpha_k)_{\mathbb{N}^*}$ converge vers a_1 et $(\beta_k)_{\mathbb{N}^*}$ converge vers a_2 et t.q. $(\lambda_n)_{\mathbb{N}^*} = (\alpha_k)_{\mathbb{N}^*} \cup (\beta_k)_{\mathbb{N}^*}$.

3- Montrer que $\liminf(\lambda_n)$ est la plus petite des valeurs d'adhérences.

Réponse. 11- une suite (u_n) est convergente ssi $\exists\ell \in \mathbb{R}$ t.q. $\lim_{n\to\infty} u_n = \ell$, i.e., t.q. $\forall\varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{N}^*$, $\forall n \geq N_\varepsilon$, $|u_n - \ell| < \varepsilon$, et alors ℓ est appelée la limite de la suite (u_n) (à partir "d'un certain rang N_ε", toute les valeurs u_n de la suite sont "égales à ℓ à ε près".

12- Une sous-suite de $(u_n)_{n\in\mathbb{N}^*}$ est une suite $(v_k)_{k\in\mathbb{N}^*}$ t.q. il existe une fonction $n : \left\{\begin{array}{l} \mathbb{N}^* \to \mathbb{N}^* \\ k \to n(k) \overset{\text{noté}}{=} n_k \end{array}\right\}$ strictement croissante vérifiant $v_k = u_{n_k}$.

13- Un réel ℓ est une valeur d'adhérence de la suite (u_n) ssi une sous-suite de (u_n) converge vers ℓ, i.e., ssi $\forall\varepsilon > 0$, $\forall N \in \mathbb{N}^*$, $\exists n_{\varepsilon,N} > N$ t.q. $|u_{n_{\varepsilon,N}} - \ell| < \varepsilon$, i.e., ssi $\forall\varepsilon > 0$, Cardinal$(\{u_n : |u_n - \ell| < \varepsilon\}) = \infty$.

2- Quitte à renuméroter a_1 et a_2, supposons $a_1 < a_2$.

Soit $\varepsilon > 0$, avec $\varepsilon < \frac{a_2-a_1}{2}$. On prend $N_\varepsilon \in \mathbb{N}^*$ t.q. $\forall n > N_\varepsilon$, soit $|\lambda_n - a_1| < \varepsilon$, soit $|\lambda_n - a_2| < \varepsilon$.

On note $E_0 = [1, N_\varepsilon]_{\mathbb{N}}$, $E_1 = \{n > N_\varepsilon : |\lambda_n - a_1| < \varepsilon\}$, $E_2 = \{n > N_\varepsilon : |\lambda_n - a_2| < \varepsilon\}$. Donc $\mathbb{N}^* = E_0 \cup E_1 \cup E_2$ partition. Les ensembles E_1 et E_2 sont infinis (sinon il n'y a qu'une valeur d'adhérence) On note $E_1 = (i_k)_{k\in\mathbb{N}^*}$ avec $i : k \in \mathbb{N}^* \to i(k) = i_k \in \mathbb{N}^*$ fonction croissante stricte, et $E_2 = (j_k)_{k\in\mathbb{N}^*}$ avec $j : k \in \mathbb{N}^* \to j(k) = j_k \in \mathbb{N}^*$ fonction croissante stricte (possible car $\mathbb{N}$ est ordonné).

On pose $(\tilde\alpha_k)_{k\in\mathbb{N}^*} = (\lambda_{i_k})_{k\in\mathbb{N}^*}$ et $(\beta_k)_{k\in\mathbb{N}^*} = (\lambda_{j_k})_{k\in\mathbb{N}^*}$. On vérifie immédiatement que $(\tilde\alpha_k)_{\mathbb{N}^*}$ converge vers a_1 et que $(\beta_k)_{\mathbb{N}^*}$ converge vers a_2 (écrire la définition de la convergence).

Et on note $(\alpha_k)_{k\in\mathbb{N}^*}$ la suite définie par $\alpha_k = \lambda_k$ pour tout $k \in E_0$, et $\alpha_{N_\varepsilon+k} = \tilde\alpha_k$ pour tout $k \in \mathbb{N}^*$ (la suite $(\tilde\alpha_k)_{k\in\mathbb{N}^*}$ à laquelle on a ajouté $\lambda_1, ..., \lambda_{N_\varepsilon}$)). On a $(\lambda_n)_{\mathbb{N}^*} = (\alpha_k)_{\mathbb{N}^*} \cup (\beta_k)_{\mathbb{N}^*}$ et les suites (α_k) et $\beta_k)$ conviennent.

3- On note, cf. (5.13), $\mu_n = \inf_{k\geq n}\lambda_k$. Donc, pour $n > N_\varepsilon$, comme $\beta_k \geq \alpha_k$ pour tout k, on a $\mu_n = \inf_{k\in E_2}\alpha_k$, avec $|\mu_n - a_1| < \varepsilon$. Vrai pour tout ε, donc $\lim_{n\to\infty}\mu_n = a_1$. ▪

Exemple 5.9 La suite (λ_n) définie par $\lambda_{2n+1} = 1$ et $\lambda_{2n} = \frac{1}{2n}$ (faire un dessin) vérifie $\liminf(\lambda_n) = 0$: à la limite, l'inf de la suite est 0. ▪

Définition 5.10 De même :

$$\limsup_{n\to\infty}(\lambda_n) \overset{\text{déf}}{=} \lim_{n\in\mathbb{N}^*}(\sup_{k\geq n}\lambda_k) = \inf_{n\in\mathbb{N}^*}(\sup_{k\geq n}\lambda_k) = \inf_{n\in\mathbb{N}^*}(\nu_n), \tag{5.15}$$

où $\nu_n = \sup_{k\geq n}\lambda_k$ (et donc $(\nu_n)_{\mathbb{N}^*}$ est une suite décroissante).

Exercice 5.11 Montrer que dans $\overline{\mathbb{R}}$:

$$\liminf_{n\to\infty}(\lambda_n) \leq \limsup_{n\to\infty}(\lambda_n). \tag{5.16}$$

Réponse. On veut : $\sup_{n\in\mathbb{N}}(\inf_{k\geq n}\lambda_k) \leq \inf_{m\in\mathbb{N}}(\sup_{\ell\geq m}\lambda_\ell)$.

Soit, $\forall n \in \mathbb{N}$, $\inf_{k\geq n}\lambda_k \leq \inf_{m\in\mathbb{N}}(\sup_{\ell\geq m}\lambda_\ell)$.

Soit, $\forall n, m \in \mathbb{N}$, $\inf_{k\geq n}\lambda_k \leq \sup_{\ell\geq m}\lambda_\ell$.

Si $n \geq m$ alors $\inf_{k\geq n}\lambda_k \leq \lambda_n \leq \sup_{\ell\geq n}\lambda_\ell \leq \sup_{\ell\geq m}\lambda_\ell$: OK.

Si $n \leq m$ alors $\inf_{k\geq n}\lambda_k \leq \inf_{k\geq m}\lambda_k \leq \lambda_m \leq \sup_{\ell\geq m}\lambda_\ell$: OK. ▪

Exercice 5.12 Montrer : $-\inf(\lambda_n) = \sup(-\lambda_n))$ puis $-\liminf(\lambda_n) = \limsup(-\lambda_n)$.

Réponse. Soit $a = \inf_{\mathbb{N}^*}(\lambda_n)$. On a $a \leq \lambda_n$ pour tout n et $\exists(\lambda_{n_k})$ sous-suite extraite t.q. $\lambda_{n_k} \longrightarrow_{k\to\infty} a$. Donc $-a \geq -\lambda_n$ pour tout n et $-\lambda_{n_k} \longrightarrow_{k\to\infty} -a$, donc $-a = \sup_{\mathbb{N}^*}(-\lambda_n)$.

Donc $-\liminf(\lambda_n) \overset{\text{déf}}{=} -\sup_{n\in\mathbb{N}^*}(\inf_{k\geq n}(\lambda_k)) = +\inf_{n\in\mathbb{N}^*}(-\inf_{k\geq n}(\lambda_k)) = +\inf_{n\in\mathbb{N}^*}(+\sup_{k\geq n}(-\lambda_k))$. ▪

Exercice 5.13 Montrer que (λ_n) est convergente ssi :

$$\liminf_{n\to\infty}\lambda_n = \lim_{n\to\infty}\lambda_n = \limsup_{n\to\infty}\lambda_n. \tag{5.17}$$

Réponse. Soit $\lambda_\infty = \lim_{n\to\infty}\lambda_n$. On a : $\forall\varepsilon > 0$, $\exists N_\varepsilon > 0$, $\forall n > N_\varepsilon$, $|\lambda_n - \lambda_\infty| < \varepsilon$.

Soit $\mu_n = \inf_{k\geq n}(\lambda_k)$. Donc $\forall n > N_\varepsilon$, $|\mu_n - \lambda_\infty| < \varepsilon$. Donc la suite croissante $(\mu_n)_{\mathbb{N}^*}$ converge vers λ. Donc $\liminf_{n\to\infty}\lambda_n = \lim_{n\to\infty}\lambda_n$. Idem pour $\limsup$. ▪

Exercice 5.14 Si $c \in \mathbb{R}$ (une constante), montrer que :

$$\liminf_{n\to\infty}(\lambda_n + c) = \liminf_{n\to\infty}(\lambda_n) + c. \tag{5.18}$$

Réponse. $\sup_{n\in\mathbb{N}}(\inf_{k\geq n}(\lambda_k + c)) = \sup_{n\in\mathbb{N}}((\inf_{k\geq n}\lambda_k) + c) = \sup_{n\in\mathbb{N}}(\inf_{k\geq n}\lambda_k) + c$. ∴

Exercice 5.15 Si (μ_n) est convergente, montrer que :

$$\liminf_{n\to\infty}(\lambda_n + \mu_n) = \liminf_{n\to\infty}(\lambda_n) + \lim_{n\to\infty}(\mu_n). \tag{5.19}$$

Réponse. Pour $\varepsilon > 0$ soit n_ε t.q., $\forall k \geq n_\varepsilon$, $\mu_\infty - \varepsilon \leq \mu_k \leq \mu_\infty + \varepsilon$.

Donc pour $k \geq n$ on a $\lambda_k + \mu_\infty - \varepsilon \leq \lambda_k + \mu_k \leq \lambda_k + \mu_\infty + \varepsilon$.

Donc $\inf_{k\geq n}(\lambda_k) + \mu_\infty - \varepsilon \leq \inf_{k\geq n}(\lambda_k + \mu_k) \leq \inf_{k\geq n}(\lambda_k) + \mu_\infty + \varepsilon$. ∴

Exercice 5.16 Montrer : pour des suites minorées on a :

$$\inf_{n\to\infty}(\lambda_n) + \inf_{n\to\infty}(\mu_n) = \inf_{n\to\infty}(\lambda_n + \mu_n). \tag{5.20}$$

Réponse. Soit $a = \inf(\lambda_n)$, donc $\forall n$, $a \leq \lambda_n$, et $\forall \varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{N}$, $\forall n > N_\varepsilon$, $-\varepsilon < a - \lambda_n < \varepsilon$. Soit $b = \inf(\mu_n)$, donc $\forall n$, $b \leq \mu_n$, et $\forall \varepsilon > 0$, $\exists M_\varepsilon \in \mathbb{N}$, $\forall n > M_\varepsilon$, $-\varepsilon < b - \mu_n < \varepsilon$.

Donc $\forall n$, $a + b \leq \lambda_n + \mu_n$, et $\forall \varepsilon$, $\forall n > \max(N_{\frac{\varepsilon}{2}}, M_{\frac{\varepsilon}{2}})$, $-\varepsilon < a + b - \lambda_n - \mu_n < \varepsilon$. ∴

Exercice 5.17 Montrer :

$$\liminf_{n\to\infty}(\lambda_n) + \liminf_{n\to\infty}(\mu_n) \leq \liminf_{n\to\infty}(\lambda_n + \mu_n). \tag{5.21}$$

Donner un exemple où il n'y a pas égalité.

Réponse. Il faut montrer que $\sup_{n\in\mathbb{N}^*}(\inf_{k\geq n}(\lambda_k)) + \sup_{m\in\mathbb{N}^*}(\inf_{k\geq m}(\mu_k)) \leq \sup_{\ell\in\mathbb{N}^*}(\inf_{k\geq \ell}(\lambda_k + \mu_k))$.

C'est vrai si $\forall n \in \mathbb{N}^*$ on a $\inf_{k\geq n}(\lambda_k) + \inf_{k\geq n}(\mu_k) \leq \sup_{\ell\in\mathbb{N}^*}(\inf_{k\geq\ell}(\lambda_k + \mu_k))$.

C'est vrai si $\forall n \in \mathbb{N}^*$, $\exists \ell \in \mathbb{N}^*$ t.q. $\inf_{k\geq n}(\lambda_k) + \inf_{k\geq n}(\mu_k) \leq \inf_{k\geq\ell}(\lambda_k + \mu_k))$: il suffit de prendre $\ell = n$ car $\inf_{k\geq n}(\lambda_k) + \inf_{k\geq n}(\mu_k) \leq \lambda_n + \mu_n \leq \inf_{k\geq n}(\lambda_k + \mu_k)$.

Exemple de non égalité : $\lambda_{2n} = 0$ et $\lambda_{2n+1} = 1$ et $\mu_{2n} = 1$ et $\mu_{2n+1} = 0$: ici $\lambda_n + \mu_n = 1$ et $\liminf_{n\to\infty}(\lambda_n) = \liminf_{n\to\infty}(\mu_n) = 0$.

Ou encore $\mu_n = -\lambda_n$ avec $\lambda_{2n} = 0$ et $\lambda_{2n+1} = 1$ pour tout n : ici $\lambda_n + \mu_n = 0$, donc $\liminf_{n\to\infty}(\lambda_n + \mu_n) = 0$, alors que $\liminf_{n\to\infty}(\lambda_n) = 0$ et $\liminf_{n\to\infty}(\mu_n) = -1$. ∴

5.2.2 Lemme de Fatou

Soit $(f_n)_{\mathbb{N}^*}$ une suite de fonctions. Soit x fixé. Appliquons la définition 5.14 à la suite $(\lambda_n)_{\mathbb{N}^*} = (f_n(x))_{\mathbb{N}^*}$:

$$\liminf_{n\to\infty} f_n(x) \stackrel{\text{déf}}{=} \sup_{n\in\mathbb{N}^*}(\inf_{k\geq n} f_k(x)) = \lim_{n\to\infty}(\inf_{k\geq n} f_k(x)). \tag{5.22}$$

Autrement dit, si on définit les fonctions g_n par :

$$\forall x \in E, \quad g_n(x) = \inf_{k\geq n} f_k(x), \quad \text{donc} \quad (g_n)_{\mathbb{N}^*} \text{ est croissante,} \tag{5.23}$$

et on a :

$$\forall x \in E, \quad \liminf_{n\to\infty}(f_n(x)) \stackrel{\text{déf}}{=} \sup_{n\in\mathbb{N}^*}(g_n(x)) = \lim_{n\to\infty}(g_n(x)). \tag{5.24}$$

Exemple 5.18 Sur l'intervalle $[0,2]$ muni de la mesure de Lebesgue, soit $(f_n)_{\mathbb{N}^*}$ la suite définie par $f_{2m} = 1_{[0,1]} = f_2$, et $f_{2m+1} = 1_{]1,2]} = f_1$ (faire un dessin). Ici $g_n(x) = 0$ pour tout x et tout n (immédiat), et donc $\sup_n g_n = 0 = \liminf(f_n)$. ∴

Lemme 5.19 *(de Pierre Fatou.) Si (f_n) est une suite de fonctions positives et mesurables, alors $f = \liminf_{n\to\infty}(f_n)$ est mesurable positive, et :*

$$\mu(\liminf_{n\to\infty} f_n) \leq \liminf_{n\to\infty}(\mu(f_n)). \tag{5.25}$$

On n'a pas égalité en général (par exemple on peut avoir $\liminf_{n\to\infty} f_n = 0$ alors que $\mu(f_n) = 1$ pour tout n). Avec la notation intégrale :

$$\int_E (\liminf_{n\to\infty} f_n)d\mu \leq \liminf_{n\to\infty}(\int_E f_n\,d\mu), \tag{5.26}$$

et il n'y a pas inversion des signes $\int$ et $\liminf$, uniquement une minoration dans un sens.

De même $\mu(\limsup_{n\to\infty} f_n) \geq \limsup_{n\to\infty}(\mu(f_n))$.

Preuve. On utilise (5.23) et (5.24). On a $g_n = \inf_{k\geq n} f_k$ est mesurable, cf. corollaire 2.12.

On pose $g = \sup_{n\in\mathbb{N}^*} g_n = \sup_{n\in\mathbb{N}^*}(\inf_{k\geq n} f_k)$. Il s'agit de montrer $\mu(g) \leq \sup_{n\in\mathbb{N}^*}(\inf_{k\geq n}\mu(f_k))$.

Pour tout $k \geq n$ on a : $g_n \leq f_k$, donc $\mu(g_n) \leq \mu(f_k)$. Donc $\mu(g_n) \leq \inf_{k\geq n}\mu(f_k)$. Donc $\sup_{n\in\mathbb{N}^*}\mu(g_n) \leq \sup_{n\in\mathbb{N}^*}(\inf_{k\geq n}\mu(f_k))$.

Et le théorème de convergence monotone donne $\sup_{\mathbb{N}^*}\mu(g_n) = \mu(\sup_{\mathbb{N}^*} g_n) = \mu(g)$, car les g_n sont positives et $(g_n)_{\mathbb{N}^*}$ est croissante. Donc $\mu(g) \leq \sup_{n\in\mathbb{N}^*}(\inf_{k\geq n}\mu(f_k))$.

Exemple d'inégalité stricte : soit les f_n de l'exemple 5.18. On a $\liminf_{n\to\infty}(f_n) = 0$. Donc $\int_0^2(\liminf_{n\to\infty} f_n)dx = 0$. Et on a $\int_0^2 f_n\,dx = 1$ pour tout n. Donc $\liminf_{n\to\infty}(\int_0^2 f_n\,dx) = 1$.

Puis, pour lim sup, on a $\limsup_{n\to\infty}(f_n) = -\liminf_{n\to\infty}(-f_n)$, et (5.25) donne le résultat. ▪

Remarque 5.20 Exemples où (f_n) vérifie $\int_{\mathbb{R}} f_n(x)\,dx = 1$ pour tout n et $\liminf_{n\to\infty} f_n = 0$ (donc inégalité stricte dans (5.25)) :

$f_n(x) = n1_{]0,\frac{1}{n}]}$ (concentration de la masse en 0).

$g_n(x) = \frac{1}{n}1_{]0,n]}$ (dilution de la masse).

$h_n(x) = f_1(x-n)$ ("bosse glissante")... ▪

5.3 Convergence en moyenne pour la mesure μ

5.3.1 Convergence en moyenne

Soit $(E, \mathcal{A}, \mu)$ un espace mesuré.

On considère l'espace $\mathcal{L}^1(E, \mathcal{A}, \mu) =^{\text{noté}} \mathcal{L}^1$ des fonctions $f : E \to \mathbb{R}$ ou $\mathbb{C}$ qui sont μ-intégrables, i.e. $\mu(|f|) = ||f||_{\mathcal{L}^1} < \infty$. Et on notera $||.||_{\mathcal{L}^1} =^{\text{noté}} ||.||_1$.

Comme $||.||_1$ est une semi-norme dans $\mathcal{L}^1$, on considèrera les fonctions de $\mathcal{L}^1$ presque partout.

Ou encore, cf. remarque 4.26, on considèrera l'espace $L^1 = \mathcal{L}^1/\mathcal{N}$ avec $||.||_{L^1} =^{\text{noté}} ||.||_1$, où maintenant $||.||_{L^1}$ est une norme dans L^1, norme qui fait de L^1 un espace topologique séparé (topologie engendrée par les boules ouvertes $B(f,\varepsilon) = \{g \in L^1 : ||g-f||_1 < \varepsilon\}$ pour tout $f \in L^1$ et tout $\varepsilon > 0$).

Définition 5.21 La topologie induite par la norme $||.||_1$ sur L^1 (topologie engendrée par les boules $B(f,\varepsilon)$) est appelée topologie de la convergence en moyenne.

Les convergences dans $(\mathcal{L}^1, ||.||_1)$ et dans $(L^1, ||.||_1)$ sont appelées convergences en moyenne (relativement à la mesure μ).

Donc : une suite $(f_n)_{\mathbb{N}^*} \in (\mathcal{L}^1)^{\mathbb{N}^*}$ (suite de fonctions μ-intégrables) converge en moyenne pour la mesure μ vers une fonction $f \in \mathcal{L}^1$ (fonction μ-intégrable) ssi :

$$\mu(|f_n - f|) \underset{n\to\infty}{\longrightarrow} 0, \quad \text{i.e.} \quad ||f_n - f||_1 \underset{n\to\infty}{\longrightarrow} 0, \quad \text{i.e.} \quad \int_E |f_n - f|\,d\mu \underset{n\to\infty}{\longrightarrow} 0. \tag{5.27}$$

(Ne pas oublier les valeurs absolues ! C'est la convergence au sens de la norme $||.||_1$.)

Remarque 5.22 On parle abusivement de fonction $f \in L^1$, au sens où, de la classe de fonctions f, on choisit (arbitrairement) une fonction dans $\mathcal{L}^1$. Ce tant que cet abus de notation ne pose pas de problèmes (autres que des problèmes techniques dans les démonstrations). ▪

Proposition 5.23 *Soit (f_n) une suite de fonctions de $\mathcal{L}^1$, et soit $f \in \mathcal{L}^1$, avec $f_n \underset{n\to\infty}{\to} f$ en moyenne, cf. (5.27). On a :*

$$\mu(|f - f_n|) \underset{n\to\infty}{\to} 0 \implies \mu(f_n) \underset{n\to\infty}{\to} \mu(f) \quad \text{et} \quad \mu(|f_n|) \underset{n\to\infty}{\to} \mu(|f|). \tag{5.28}$$

Soit avec la notation intégrale :

$$\int_E |f - f_n|\,d\mu \underset{n\to\infty}{\longrightarrow} 0 \implies \int_E f_n\,d\mu \underset{n\to\infty}{\longrightarrow} \int_E f\,d\mu \quad \text{et} \quad \int_E |f_n|\,d\mu \underset{n\to\infty}{\longrightarrow} \int_E |f|\,d\mu. \tag{5.29}$$

Preuve. On a $|\mu(f_n) - \mu(f))| = |\mu(f_n - f)| \leq \mu(|f_n - f|) \to 0$ d'où la première affirmation, et on a "$\big||a| - |b|\big| \leq |a-b|$" pour tous réels a et b, d'où $|\mu(|f_n|) - \mu(|f|)| = |\mu(|f_n| - |f|)| \leq \mu(\big||f_n| - |f|\big|) \leq \mu(|f_n - f|) \to 0$ d'où la deuxième affirmation. ▪

Remarque 5.24 Par contre, la réciproque est fausse : $\mu(|f_n|) \underset{n\to\infty}{\to} \mu(|f|) \not\Rightarrow \mu(|f_n-f|) \underset{n\to\infty}{\to} 0$. Prendre $\mu = dx$, $f_n = 1_{[0,1]}$ (suite constante) et $f = 1_{[1,2]}$ pour lesquelles on a $\mu(|f_n|) = \mu(|f|) = 1$ mais $\mu(|f_n - f|) = 2 \neq 0$. Faire un dessin. ∴

Sous l'hypothèse supplémentaire de convergence presque partout on a quand même :

Proposition 5.25

$$\textit{Si} \quad \left\{ \begin{array}{ll} (i) & \mu(|f_n|) \underset{n\to\infty}{\to} \mu(|f|), \text{ et} \\ (ii) & (f_n) \text{ converge } \mu\text{-p.p. vers } f, \end{array} \right\} \quad \textit{alors} \quad \mu(|f_n - f|) \underset{n\to\infty}{\to} 0. \tag{5.30}$$

Preuve. Appliquons le lemme de Fatou 5.19 : on considère la suite de fonctions positives et intégrables $g_n = |f|+|f_n|-|f-f_n|$ qui converge μ-p.p. vers la fonction $|f|+|f|-0 = 2|f|$. Comme ici $\liminf(g_n) = \lim(g_n) = 2|f|$, le lemme de Fatou, cf. (5.25), indique alors que :

$$\mu(2|f|) \leq \liminf(\mu(|f| + |f_n| - |f - f_n|)) = \lim(\mu(|f| + |f_n|)) - \liminf(\mu(|f - f_n|)).$$

D'où $0 \leq -\liminf(\mu(|f - f_n|)) \leq 0$ et donc $\liminf(\mu(|f - f_n|)) = 0$.

De même avec $h_n = |f - f_n| - (|f| + |f_n|) \to -2|f|$: on a $-\mu(2|f|) \geq \limsup(\mu(|f - f_n| - (|f| + |f_n|)))$, d'où $\limsup(\mu(|f - f_n|)) = 0$. D'où $\lim(\mu(|f - f_n|)) = 0$, cf. exercice 5.13. ∴

Remarque 5.26 En revanche, on n'a rien si on suppose seulement $\mu(f_n) \to_{n\to\infty} \mu(f)$ (au lieu du (i)) même si f_n converge presque partout vers f : prendre la mesure de Lebesgue sur $]0,1[$, la suite de fonctions définie par $f_n(x) = n(1_{]0,\frac{1}{n}[} - 1_{]\frac{1}{n},\frac{2}{n}[})$ qui vérifie $\int_0^1 f_n(x)\,dx = 0$ et qui p.p. converge vers 0, alors que $\mu(|f_n|) = 2$. Faire un dessin.

Ou encore, sur $\mathbb{R}$, prendre la fonction qui vaut $\frac{1}{n}(1_{[0,n]} - 1_{[-n,0]})$. Faire un dessin. ∴

5.3.2 Convergence en moyenne vs convergence μ-presque partout

Remarque 5.27 Attention : la convergence p.p. n'implique pas la convergence en moyenne :

$$f_n \underset{n\to\infty}{\longrightarrow} f \text{ au sens p.p} \quad \not\Longrightarrow \quad f_n \underset{n\to\infty}{\longrightarrow} f \text{ en moyenne (au sens de } L^1).$$

Prendre $f_n = n1_{]0,\frac{1}{n}[}$:

1- on a f_n tend simplement vers la fonction nulle $f = 0$, avec $||f||_1 = 0$, alors que

2- $||f_n||_1 = 1$ pour tout n, et donc que $||f_n - f||_1 = ||f_n|| = 1 \not\to 0$. D'ailleurs la suite (f_n) ne converge pas dans $\mathcal{L}^1$, sa “limite” généralisée étant la masse de Dirac (qui n'est pas une fonction). ∴

Remarque 5.28 Attention : la convergence en moyenne n'implique pas la convergence p.p. :

$$f_n \underset{n\to\infty}{\longrightarrow} f \text{ au sens de } L^1 \quad \not\Longrightarrow \quad f_n \underset{n\to\infty}{\longrightarrow} f \text{ au sens p.p.}$$

Voir Bouyssel [3] p. 132 pour le contre-exemple : $f_n = 1_{A_n}$ où $A_n = [\frac{j_n}{2^{k_n}}, \frac{j_n+1}{2^{k_n}}[$ et où, à n donné, j_n et k_n sont les (seuls) entiers tels que $0 \leq j_n \leq 2^{k_n}$ et $j_n + 2^{k_n} = n$. ∴

Exercice 5.29 Complément de la remarque précédente. On a quand même le résultat suivant. Soit $(f_n)_{n\in\mathbb{N}^*}$ une suite qui converge vers f au sens L^1 (de la moyenne). Montrer qu'il existe une sous suite extraite $(f_{n_k})_{k\in\mathbb{N}^*}$ qui converge p.p. vers f.

Réponse. La convergence en moyenne implique : il existe une sous-suite $(f_{n_k})_{k\in\mathbb{N}^*}$ t.q. $\mu(|f-f_{n_k}|) < 2^{-k}$. D'où $\sum_k \mu(|f-f_{n_k}|) < \infty$. Donc $\mu(\sum_k |f-f_{n_k}|) < \infty$, cf. exercice 5.5, donc $F = \sum_k |f-f_{n_k}| \in L^1$, avec $F(x) = \sum_k |f(x)-f_{n_k}(x)|$ pour presque tout x. Donc $F(x) < \infty$ pour presque tout x, cf. exercice 3.12. Donc, pour presque tout x, le terme général de la série $F(x) = \sum_k |f-f_{n_k}|(x)$ tend vers 0 : $|f-f_{n_k}|(x) \underset{k\to\infty}{\longrightarrow} 0$. ∴

5.4 Théorème de Lebesgue (de convergence dominée)

C'est un théorème fondamental de la théorie de l'intégration.

5.4.1 Le théorème

Cadre : soit $(E,\mathcal{A},\mu)$ un espace mesuré, et $\mathcal{L}^1(E,\mathcal{A},\mu) =^{\text{noté}} \mathcal{L}^1$ l'ensemble des fonctions $f : E \to \mathbb{R}$ qui sont μ-mesurables de mesure $||f||_1 =^{\text{déf}} \mu(|f|) < \infty$.

Motivation : soit $(f_n)_{\mathbb{N}^*}$ une suite de fonctions réelles (ou complexes). Soient les intégrales :

$$F_n = \int_E f_n \, d\mu \quad \in \mathbb{R}.$$

Question : la suite $(F_n)_{\mathbb{N}^*}$ converge-t-elle dans $\mathbb{R}$? Et si oui, a-t-on :

$$\lim_{n\to\infty} F_n = \int_E (\lim_{n\to\infty} f_n) \, d\mu \quad ?$$

En d'autres termes, peut-on passer à la limite sous le signe $\int$:

$$\lim_{n\to\infty} \Big(\int_E f_n \, d\mu \Big) = \int_E (\lim_{n\to\infty} f_n) \, d\mu \quad ? \tag{5.31}$$

C'est faux en général : voir (5.8). C'est vrai sous les hypothèses de convergence p.p. et domination :

Théorème 5.30 *Théorème de Lebesgue (de convergence dominée). On suppose :*

1. *la suite $(f_n)_{\mathbb{N}^*}$ converge p.p. vers une fonction f (i.e., à x fixé, la suite $(f_n(x))_{\mathbb{N}^*}$ converge vers $f(x)$ (on peut remplacer $\mathbb{N}^*$ par tout ensemble dénombrable), ce pour μ-presque tout x), et*
2. *domination $\mathcal{L}^1$:*
$$\exists h \in \mathcal{L}^1, \quad \forall n \in \mathbb{N}^*, \quad |f_n| \le h \quad \mu\text{-p.p.}, \tag{5.32}$$
(domination indépendante de n par une fonction h intégrable) i.e. il existe $h \in \mathcal{L}^1(E)$ t.q. pour tout $n \in \mathbb{N}^*$ et pour μ-presque tout $x \in E$ on a $|f_n(x)| \le h(x)$.

Alors la fonction f est intégrable et de plus :

$$\lim_{n\to\infty} (\mu(f_n)) = \mu(f), \tag{5.33}$$

soit $\lim\limits_{n\to\infty} (\mu(f_n)) = \mu(\lim\limits_{n\to\infty} (f_n))$, et avec la notation intégrale :

$$\lim_{n\to\infty} \Big(\int_E f_n \, d\mu \Big) = \int_E \Big(\lim_{n\to\infty} f_n \Big) \, d\mu. \tag{5.34}$$

Preuve. On a $|f_n| \le h$ p.p. pour tout n avec h intégrable, donc f_n est intégrable (critère d'intégration par domination).

Et $f_n \to f$ p.p., donc $|f| \le h$ p.p. avec h intégrable, donc f est intégrable (critère d'intégration par domination).

Puis pour presque tout $x \in E$ on a $\lim_{n\to\infty} f_n(x) = f(x) = \liminf_{n\to\infty} f_n(x)$ (les lim inf et lim sont égales pour les suites convergentes). On applique le lemme de Fatou 5.19 (pour lequel on a besoin de fonctions positives) : on a $|f_n| \le h$, donc $-f_n \le h$ et $f_n \le h$, ce pour tout n et p.p., soit $h + f_n \ge 0$ et $h - f_n \ge 0$. Donc (5.25) donne :

$$\begin{aligned} &\mu(f+h) \le \liminf_{n\to\infty} \mu(f_n + h), \quad \text{d'où} \quad \mu(f) \le \liminf_{n\to\infty} \mu(f_n), \\ \text{et} \quad &\mu(h-f) \le \liminf_{n\to\infty} \mu(h - f_n), \quad \text{d'où} \quad \mu(f) \ge \liminf_{n\to\infty} \mu(f_n). \end{aligned}$$

D'où $\mu(f) = \liminf_{n\to\infty} \mu(f_n) = \lim_{n\to\infty} \mu(f_n)$. ∎

Remarque 5.31 Si la fonction dominante h existe, elle est nécessairement supérieure à $g = \sup_{\mathbb{N}} |f_n|$. En particulier si $g = \sup_{\mathbb{N}} |f_n|$ n'est pas intégrable (cas $\mu(g) = \infty$), on ne peut pas appliquer le théorème de convergence dominée (de Lebesgue). Exemple : voir remarque 5.2 où $g \ge \frac{1}{2x}$ (exercice).

On peut également noter que la suite (f_n) de la remarque 5.2 n'est pas une suite de Cauchy dans $\mathcal{L}^1$, le vérifier avec $||f_n - f_{2n}||_1$, et donc ne saurait converger au sens de (dans) $\mathcal{L}^1$. Cette suite est uniquement simplement convergente (converge ponctuelle).

(La suite (f_n) de cet exemple est convergente au sens des distributions vers la mesure de Dirac en 0 (qui n'est pas une fonction), voir cours de distribution.) ∎

5.4.2 Exemples

Exercice 5.32 Montrer 1- directement, puis 2- en appliquant le théorème de convergence dominée que :

$$\int_0^1 nxe^{-nx}\,dx = \frac{1}{n}(1-e^{-n}) - e^{-n} \underset{n\to\infty}{\longrightarrow} 0. \tag{5.35}$$

Puis démontrer que $\int_0^1 nx^k e^{-nx}\,dx \longrightarrow_{n\to\infty} 0$ pour tout $k \geq 1$.

Réponse. 1- Par intégration par parties : $\int_0^1 xne^{-nx}\,dx = -\int_0^1 -e^{-nx}\,dx + [-xe^{-nx}]_0^1$.

2- Soit $f_n(x) = nxe^{-nx}$. Cette fonction est continue sur le compact $[0,1]$ donc est intégrable.

21- Pour $x \in \mathbb{R}_+^*$ fixé la suite des réels $(f_n(x))_{n\in\mathbb{N}}$ converge vers 0, et donc la suite (f_n) converge simplement presque partout sur $\mathbb{R}_+$ vers la fonction nulle.

22- $f_n'(x) = (n-n^2x)e^{-nx}$, et f_n atteint son maximum sur $[0,1]$ en $x = \frac{1}{n}$ avec $f_n(\frac{1}{n}) = e^{-1}$. Donc $f_n(x) \leq e^{-1}$ pour tout n. Soit h la fonction constante $h = e^{-1}1_{[0,1]}$. Cette fonction est intégrable sur $[0,1]$, elle majore la suite f_n. On peut appliquer le théorème : $\lim_{n\to\infty}\int_0^1 f_n(x)\,dx = \int_0^1 \lim_{n\to\infty}(f_n(x))\,dx = \int_0^1 0\,dx = 0$.

Même démarche pour $\int_0^1 nx^k e^{-nx}\,dx$ puisque $nx^k e^{-nx} \leq nxe^{-nx} \leq e^{-1}$ sur $[0,1]$ pour $k \geq 1$, le calcul direct par intégration par parties étant plus long (récurrence). ▪

Exercice 5.33 Montrer 1- directement, puis 2- en appliquant le théorème de convergence dominée que :

$$\int_1^\infty nxe^{-nx}\,dx \underset{n\to\infty}{\longrightarrow} 0. \tag{5.36}$$

Réponse. 1- Par intégration par parties : $\int_1^\infty nxe^{-nx}\,dx = [nx\frac{e^{-nx}}{-n}]_1^\infty - \int_1^\infty n\frac{e^{-nx}}{-n}\,dx = -0 + e^{-n} + [\frac{e^{-nx}}{-n}]_1^\infty = e^{-n} + \frac{e^{-n}}{n} \longrightarrow_{n\to\infty} 0$.

2- Soit $f(n,x) = nxe^{-nx}$.

21- Pour $x = x_0 \in [1,\infty[$ fixé, la suite $f_{x_0}(n) = f(n,x_0) = nx_0e^{-nx_0}$ tend vers 0 quand $n\to\infty$ car $y \to ye^{-y}$ converge vers 0 quand $y\to\infty$. Donc f_n converge simplement vers la fonction nulle sur $[1,\infty[$.

22- Montrons que la suite $(f_n)_{n\geq 1}$ est décroissante sur $[1,\infty[$, i.e. $f_n(x) \geq f_{n+1}(x)$ sur $[1,\infty[$. On a $\frac{f_n(x)}{f_{n+1}(x)} = \frac{n}{n+1}e^x$, avec $\frac{n}{n+1} \geq \frac{1}{2}$ pour $n \geq 1$ et $e^x \geq e^1 > 2$ pour $x \in [1,\infty[$, donc $\frac{f_n(x)}{f_{n+1}(x)} \geq 1$ pour $n \geq 1$ et $x \in [1,\infty[$, cqfd.

Donc $0 \leq f_n \leq f_1$ sur $[1,\infty[$ avec $f_1 \in \mathcal{L}^1(]1,\infty[)$.

Donc (théorème de convergence dominée de Lebesgue) $\int_1^\infty nxe^{-nx}\,dx \longrightarrow_{n\to\infty} = \int_1^\infty 0\,dx$. ▪

Exercice 5.34 Démarche générique – Montrer en appliquant le théorème de convergence dominée que :

$$\int_0^\infty nxe^{-nx}\,dx \underset{n\to\infty}{\longrightarrow} 0. \tag{5.37}$$

Réponse. On a $\int_0^\infty = \int_0^1 + \int_1^\infty$, ce qui permet de séparer le comportement sur un borné et le comportement en l'infini, comportements traités dans les exercices précédents. ▪

Exemple 5.35 En revanche :

$$\int_0^1 nxe^{\frac{-nx^2}{2}}\,dx = 1 - e^{\frac{-n}{2}} \underset{n\to\infty}{\rightarrow} 1 \tag{5.38}$$

alors que, à x fixé, $f_n(x) = nxe^{\frac{-nx^2}{2}} \underset{n\to\infty}{\rightarrow} 0$ (convergence p.p. vers 0 de la suite (f_n)). Mais cette suite f_n n'est pas dominée par une fonction intégrable : sinon on pourrait appliquer le théorème de convergence dominée et on aurait $1 = 0$.

Remarque : on vérifiera par exemple que f_n est maximum pour $x_n = \frac{1}{\sqrt{n}}$ et vaut en ce point $f_n(\frac{1}{\sqrt{n}}) = \sqrt{n}$, soit $f_n(x) = \frac{1}{x}$ pour $x = \frac{1}{\sqrt{n}}$. Or $\frac{1}{x}$ n'est pas intégrable au voisinage de 0. ▪

Exemple 5.36 Calculer $\displaystyle\lim_{n\to\infty}\int_{[0,\infty[}\frac{e^{-nx}}{\sqrt{x}}\,dx$.

Réponse. Soit $f(n,x) = \frac{e^{-nx}}{\sqrt{x}}$, pour $n \geq 1$ et $x \geq 0$.

1- Pour $x_0 > 0$ fixé, $f(n,x_0) \underset{n\to\infty}{\longrightarrow} 0$, donc la suite $(f_n)_{\mathbb{N}^*}$ converge p.p. vers $f = 0$ sur $\mathbb{R}_+$.

2- Soit $n \geq 1$. Sur $[0,1]$ on a $|f_n(x)| = f_n(x) = x^{-\frac{1}{2}}e^{-nx} \leq x^{-\frac{1}{2}}$, donc, posant $h(x) = x^{-\frac{1}{2}}$ on a $h \in \mathcal{L}^1(]0,1[)$ et $|f_n| \leq h$, domination indépendante de n. Et sur $[1,\infty]$ on a $|f_n(x)| = f_n(x) \leq e^{-x}$, donc posant $k(x) = e^{-x}$ on a $k \in \mathcal{L}^1(]1,\infty[)$ et $|f_n| \leq k$, domination indépendante de n.

Donc (théorème de convergence dominée de Lebesgue) la limite cherchée est 0. ▪

Exemple 5.37 Montrer que $\displaystyle\lim_{n\to\infty}\int_0^1 \frac{1+nx}{(1+x)^n}\,dx = 0$.

Réponse. On pose $f_n(x) = f_x(n) = f(n,x) = \frac{1+nx}{(1+x)^n}$.

1- À $x = x_0 > 0$ fixé, $n \geq 2$, $(1+x)^n \geq 1+nx_0+\frac{n(n-1)}{2}{x_0}^2$, donc $0 \leq \frac{1+nx_0}{(1+x_0)^n} \leq \frac{1+nx_0}{1+nx_0+\frac{n(n-1)}{2}{x_0}^2} \longrightarrow_{n\to\infty} 0$, donc f_n tend presque partout vers $f = 0$ sur $[0,1]$ (partout sur $]0,1]$). (Ou encore poser $y_0 = 1+x_0$, donc $\frac{1+nx_0}{(1+x_0)^n} = \frac{1+n(y_0-1)}{y_0^n} = \frac{1}{y_0^n} + (y_0-1)ne^{-n\log(y_0)} \to_{n\to\infty} 0$ quand $y_0 > 1$.)

2- $|f(n,x)| \leq 1$ pour tout $n \in \mathbb{N}$ et tout $x \in [0,1]$ (cf. calcul précédent), et donc $h = 1_{[0,1]} \in \mathcal{L}^1(]0,1[)$ domine $(|f_n|_{\mathbb{N}})$. D'où l'application du théorème de convergence dominée et sa conclusion. ∎

Exemple 5.38 Soit $A > 0$. Soit $\displaystyle F_n(A) = \int_0^A \big(1-\frac{x}{n}\big)^n \cos(x)\,dx$.

(i) Calculer $F(A) = \lim\limits_{n\to\infty} F_n(A)$. (On rappelle que $e^x = \lim_{n\to\infty}(1+\frac{x}{n})^n$.)

(ii) Montrer que $\lim_{n\to\infty}\int_0^n \big(1-\frac{x}{n}\big)^n \cos(x)\,dx = \frac{1}{2}$.

Réponse. (i) On pose $f_n(x) = \big(1-\dfrac{x}{n}\big)^n \cos(x)$.

1- À $x = x_0 \geq 0$ fixé, $\big(1-\frac{x_0}{n}\big)^n \longrightarrow_{n\to\infty} e^{-x_0}$, voir annexe B, donc la suite $(f_n)_{\mathbb{N}^*}$ converge simplement sur $\mathbb{R}_+^*$ vers la fonction f donnée par $f(x) = e^{-x}\cos(x)$.

2- Pour $x \in \mathbb{R}$ fixé et pour $n > |x|$, on a $0 \leq \big(1-\frac{x}{n}\big)^n \leq e^{-x}$, voir annexe B. Donc $|f_n(x)| \leq h(x)$ pour $x \in]0,A[$ et $n > A$, où $h(x) = e^{-x}$, avec $h \in \mathcal{L}^1(]0,A[)$.

Donc (convergence dominée) $F(A) = \lim_{n\to\infty} F_n(A) = \int_0^A(\lim_{n\to\infty} f_n(x))\,dx = \int_0^A e^{-x}\cos(x)\,dx$.

Calcul : $F(A) = [e^{-x}\sin x]_0^A + \int_0^A e^{-x}\sin(x)\,dx = e^{-A} + [e^{-x}(-\cos(x))]_0^A - \int_0^A e^{-x}\cos(x)\,dx = e^{-A} - e^{-A}\cos(A) + 1 - F(A)$, d'où $F(A) = \frac{1}{2}(e^{-A} - e^{-A}\cos(A)+1)$.

(ii) On a $G_n = \int_0^n \big(1-\frac{x}{n}\big)^n \cos(x)\,dx = \int_0^\infty g(n,x)\,dx$, où $g_n(x) = \big(1-\frac{x}{n}\big)^n \cos(x)1_{]0,n[}(x)$: ici on a rendu les bornes de l'intégrale indépendantes de n. On procède comme au (i) :

1- à $x = x_0 \geq 0$ fixé, $\big(1-\dfrac{x_0}{n}\big)^n \underset{n\to\infty}{\longrightarrow} e^{-x_0}$. Avec $1_{]0,n[}(x_0) \underset{n\to\infty}{\longrightarrow} 1$. Donc la suite $(g_n)_{\mathbb{N}^*}$ converge simplement vers la fonction f donnée par $f(x) = e^{-x}\cos(x)$ sur $\mathbb{R}_+$.

2- Pour $0 < x < n$, donc $0 < \big|\frac{x}{n}\big| < 1$, on a $0 < \big(1-\frac{x}{n}\big)^n \leq e^{-x}$, voir annexe B, donc $|g(n,x)| \leq e^{-x}$. Et pour $x \geq n$ on a $g(n,x) = 0$ donc en particulier $|g(n,x)| \leq e^{-x}$. Donc $|g(n,x)| \leq e^{-x}$ pour tout $n \in \mathbb{N}^*$ et tout $x > 0$.

Donc (convergence dominée) $\lim_{n\to\infty} G_n(x) = \int_0^\infty(\lim_{n\to\infty} g(n,x))\,dx = \int_0^\infty e^{-x}\cos(x)\,dx = \frac{1}{2}$ (faire une double intégration par parties, ou calculer $\int_0^\infty e^{-x}e^{ix}\,dx$ qui donne aussi $\int_0^\infty e^{-x}\sin(x)\,dx = \frac{1}{2}$). ∎

Exemple 5.39 Montrer que $\displaystyle\lim_{n\to\infty}\int_0^\infty \frac{\sin(\pi x)}{1+x^n}\,dx = \frac{2}{\pi}$.

Réponse. On se contente de $n \geq 2$ (on s'intéresse à $n \to \infty$). On pose $f(n,x) = \frac{\sin(\pi x)}{1+x^n}$.

1- si $0 < x < 1$, $f(n,x) \longrightarrow_{n\to\infty} \sin(\pi x)$, et si $1 < x < \infty$, $x^n \longrightarrow_{n\to\infty} \infty$. Donc f_n converge simplement p.p vers $f(x) = \sin(\pi x)1_{[0,1]}(x)$.

2- Une fonction dominante est donnée par $h(x) = \sin(\pi x)1_{[0,1]}(x) + \frac{1}{1+x^2}1_{[1,\infty[}(x)$ qui est bien intégrable sur $\mathbb{R}_+$. On peut donc appliquer le théorème de convergence dominée : la limite cherchée vaut $\int_0^\infty \sin(\pi x)1_{[0,1]}\,dx = [-\frac{\cos(\pi x)}{\pi}]_0^1 = \frac{2}{\pi}$. ∎

5.4.3 La primitive d'une fonction $\mathcal{L}^1$ est continue

Proposition 5.40 *Si $f \in \mathcal{L}^1([a,b])$ alors sa primitive F définie sur $[a,b]$ par $F(x) = \int_a^x f(t)\,dt$ est continue. (Une primitive quelconque est la fonction F à une constante près.)*

Et on note alors $F' = f$ dans $\mathcal{L}^1$ (même si F n'est pas dérivable au sens classique), voir ci-dessous exercice 5.42.

Preuve. Comme $f \in \mathcal{L}^1([a,b])$, pour $x \in [a,b]$ on a $f \in \mathcal{L}^1([a,x])$, car $\int_a^x |f(t)|\,dt \leq \int_a^b |f(t)|\,dt = ||f||_{\mathcal{L}^1} < \infty$. Donc F est bien définie sur $[a,b]$.

Soit $(x_n)_{\mathbb{N}^*}$ une suite de points dans $[a,b]$ qui converge vers $x \in [a,b]$. Il s'agit de montrer que $F(x_n) \longrightarrow_{n\to\infty} F(x)$. On a $F(x_n) = \int_a^b g_n(t)\,dt$ où on a posé $g_n(t) = 1_{[a,x_n]}(t)f(t)$. Appliquons le théorème de convergence dominée. À t fixé on a $g_n(t) \longrightarrow_{n\to\infty} 1_{[a,x]}(t)f(t)$ puisqu'ici $f(t)$ est une constante. On a $|g_n| \leq |f|$ pour tout n avec $\mu(|f|) < \infty$ donc $g_n \in \mathcal{L}^1([a,b])$ et (g_n) est dominée par $|f| \in \mathcal{L}^1$. Donc $\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty}\int_a^b g_n(t)\,dt = \int_a^b \lim_{n\to\infty}(g_n(t))\,dt = \int_a^b 1_{[a,x]}(t)f(t)\,dt = \int_a^x f(t)\,dt = F(x)$, et F est continue. ∎

Remarque 5.41 Dans la proposition 5.40 la notation $F' = f$ est abusive au sens classique, car la fonction F n'est pas nécessairement dérivable dans tout $[a,b]$, voir l'exemple de la fonction de Cantor, remarque 1.33. ∎

Exercice 5.42 Nécessite le cours de distribution. Soit $f \in \mathcal{L}^1(]0,1[;\mathbb{R})$. Montrer, pour presque tout $x \in]0,1[$:

$$(\int_0^x f(t)\,dt)' = f(x). \tag{5.39}$$

Indication : se servir de $\mathcal{D}(]0,1[)$ dense dans $\mathcal{L}^1(]0,1[)$ et raisonner au sens des distributions.

Réponse. Le résultat est connu pour $f \in C^0([0,1])$, puisque $F(x) = \int_0^x f(t)\,dt$ est la primitive d'une fonction continue, donc F est dérivable, et $F' = f$.

Soit $f \in \mathcal{L}^1$ et soit (f_n) une suite de fonctions de $\mathcal{D}(]0,1[)$ (C^∞ à support compact dans $]0,1[$) qui tend vers $f \in \mathcal{L}^1$: $||f - f_n||_1 \to 0$. Quitte à considérer $(f_n)_{n\geq N}$ pour N assez grand, on peut supposer que $||f_n||_1 \leq 2||f||_1$ pour tout n, et la suite (f_n) est dominée par une fonction intégrable.

Montrons (5.39) au sens des distributions, i.e. que pour tout $\varphi \in \mathcal{D}(]0,1[)$ on a $\langle(\int_0^x f(t)\,dt)', \varphi(x)\rangle = \langle f(x), \varphi(x)\rangle_{dx}$. Les f_n étant dans $\mathcal{D}(]0,1[)$, les f_n sont C^0 et donc on a :

$$\langle(\int_0^x f_n(t)\,dt)', \varphi(x)\rangle = \langle f_n(x), \varphi(x)\rangle_{dx} = \int_0^1 f_n(x)\varphi(x)\,dx.$$

Le membre de droite $\int_0^1 f_n(x)\varphi(x)\,dx$ tend vers $\int_0^1 f(x)\varphi(x)\,dx$: appliquer le théorème de convergence dominée à la suite de fonctions $g_n = f_n\varphi$.

Le membre de gauche vaut $\int_0^1(\int_0^x f_n(t)\,dt)'\varphi(x)\,dx = -\int_0^1(\int_0^x f_n(t)\,dt)\varphi'(x)\,dx$ et montrons qu'il tend vers $-\int_0^1(\int_0^x f(t)\,dt)\varphi'(x)\,dx$: appliquons le théorème de convergence dominée à la suite (F_n) donnée par $F_n(x) = (\int_0^x f_n(t)\,dt)\varphi'(x)$ pour $x \in [0,1]$: 1- les F_n sont intégrables car continues sur le compact $[0,1]$; 2- à x fixé, $F_n(x) \underset{n\to\infty}{\longrightarrow} F(x) = (\displaystyle\int_0^x f(t)\,dt)\varphi'(x)$, cf. (5.28) avec μ mesure de Lebesgue sur $[0,x]$; 3- $|F_n(x)| \leq ||f_n||_{\mathcal{L}^1}|\varphi'(x)| \leq 2||f||_{\mathcal{L}^1}|\varphi'(x)|$ intégrable car φ' continue sur $[0,1]$ compact. D'où $\int_0^1 F_n(x)\,dx \to \int_0^1 F(x)\,dx$, soit $\int_0^1(\int_0^x f_n(t)\,dt)\varphi'(x)\,dx \to \int_0^1(\int_0^x f(t)\,dt)\varphi'(x)\,dx$, ce pour tout $\varphi \in \mathcal{D}(]0,1[)$. ▪

5.5 Continuité d'une intégrale dépendant d'un paramètre

Ci-dessus on a considéré le cas des fonctions

$$f : \begin{cases} \mathbb{N}^* \times E & \to \mathbb{R}, \\ (n,x) & \to f(n,x) = f_n(x), \end{cases} \quad \text{puis} \quad F : \begin{cases} \mathbb{N}^* & \to \mathbb{R} \\ n & \to F(n) = \displaystyle\int_{x\in E} f(n,x)\,d\mu(x) \quad (= \mu(f_n)). \end{cases} \tag{5.40}$$

On se donne maintenant les fonctions

$$f : \begin{cases} I \times E & \to \mathbb{R} \\ (p,x) & \to f(p,x), \end{cases} \quad \text{puis} \quad F : \begin{cases} I & \to \mathbb{R} \\ p & \to F(p) = \displaystyle\int_{x\in E} f(p,x)\,d\mu(x) \quad (= \mu(f_p)), \end{cases} \tag{5.41}$$

où I est un intervalle ouvert de $\mathbb{R}$ (ouvert, fermé, semi-ouvert, borné ou non), où donc p est le paramètre de "l'intégrale $F(p)$ qui dépend du paramètre p", et où x est la variable d'intégration. On notera :

$$f_p : \begin{cases} E \to \mathbb{R} \\ x \mapsto f_p(x) := f(p,x), \end{cases} \quad \text{et} \quad f_x : \begin{cases} I \to \mathbb{R} \\ p \mapsto f_x(p) := f(p,x). \end{cases} \tag{5.42}$$

But : supposant f continue en p_0, on veut savoir si F est continue en p_0 et surtout si

$$\lim_{p\to p_0} F(p) = F(p_0), \quad \text{i.e.} \quad \lim_{p\to p_0}\Big(\int_{x\in E} f(p,x)\,d\mu(x)\Big) = \int_{x\in E} f(p_0,x)\,d\mu(x), \tag{5.43}$$

i.e. si on peut passer à la limite sous le signe $\int$, i.e. si

$$\lim_{p\to p_0}\Big(\int_{x\in E} f(p,x)\,d\mu(x)\Big) = \int_{x\in E}(\lim_{p\to p_0} f(p,x))\,d\mu(x).$$

I.e. si pour toute mesure μ :

$$\lim_{p\to p_0}(\mu(f_p)) = \mu(\lim_{p\to p_0} f_p) = \mu(f_{p_0})). \tag{5.44}$$

Théorème 5.43 *(Continuité d'une intégrale dépendant d'un paramètre.) Soit $(E,\mathcal{A},\mu)$ un ensemble mesuré et $\mathbb{R}$ muni de sa topologie usuelle. Soit $I =]a,b[$ avec $a < b$ et $p_0 \in I$. Si $I = [p_0,b[$, avec $p_0 < b$, on sous-entendra la continuité à droite en p_0, et si $I =]a,p_0]$, avec $a < p_0$, on sous-entendra la continuité à gauche en p_0. Si $p_0 = +\infty$ et $I =]a,p_0[$ avec $a \in \mathbb{R}$, ou si $p_0 = -\infty$ et $I =]p_0,b$ avec $b \in \mathbb{R}$, on sous-entendra la limite en $\pm\infty$.*

Soit f et F donnés en (5.41) t.q., avec (5.42), f_p est μ-mesurable pour tout $p \in I$.

Hypothèses :

1. *pour μ-presque tout $x \in E$, la fonction f_x est continue en p_0 (l'intégrant est continu en p_0), i.e.*

$$\lim_{p\to p_0} f(p,x) = f(p_0,x) \ \in \mathbb{R}, \tag{5.45}$$

2. *domination indépendante du paramètre p par une fonction h μ-intégrable :*

$$\exists h \in \mathcal{L}^1(E), \quad \text{t.q.} \quad \forall p \in I, \quad |f_p| \le h, \tag{5.46}$$

i.e., $\exists h \in \mathcal{L}^1(E)$ t.q. $\forall p \in I$ et pour presque tout $x \in E$ on a $|f(p,x)| \le h(x)$.

Conclusion : *la fonction F donnée en (5.41) est continue en p_0 :*

$$\lim_{p\to p_0} F(p) = F(p_0), \quad \text{i.e.} \quad \lim_{p\to p_0} \mu(f_p) = \mu(f_{p_0}) \quad (= \mu(\lim_{p\to p_0} f_p)), \tag{5.47}$$

noté :

$$\lim_{p\to p_0} \int_E f(p,x)\,d\mu(x) = \int_E f(p_0,x)\,d\mu(x) \quad (= \int_E \big(\lim_{p\to p_0} f(p,x)\big)\,d\mu(x)) \tag{5.48}$$

(dérivation sous le signe $\int$ sous l'hypothèse de domination).

Preuve. (corollaire du théorème de convergence dominée.) Soit $(p_n)_{\mathbb{N}^*}$ une suite qui converge vers p_0, et soit $g_n(x) = f(p_n,x)$. On a $(g_n(x))_{n\in\mathbb{N}^*}$ qui converge vers $f(p_0,x) = f_{p_0}(x)$ cf. hyp 1, pour presque tout x ; et $|g_n| \le h$ pour tout n avec $h \in L^1(E)$ cf. hyp 2. Donc $\lim_{n\to\infty} \mu(g_n) = \mu(f_{p_0})$ cf. (5.33) (thm de convergence dominée), i.e. $\lim_{n\to\infty} F(p_n) = F(p_0)$. Vrai pour toute suite $(p_n)_{\mathbb{N}^*}$ qui converge vers p_0, d'où (5.47). ∎

Exemple 5.44 Montrer que, pour $f \in \mathcal{L}^1(\mathbb{R};\mathbb{R})$, sa transformée de Fourier $\widehat{f}\mathcal{L}^1(\mathbb{R};\mathbb{C})$ définie par :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)\,e^{-i\xi x}\,dx \tag{5.49}$$

est continue sur $\mathbb{R}$ (et est borné).

Réponse. Posons $g(\xi,x) = f(x)\,e^{-i\xi x}$. 1- À x fixé (quelconque dans $\mathbb{R}$), la fonction $g_x : \xi \to f(x)\,e^{-i\xi x}$ est continue car l'exponentielle l'est. 2- On a (module) $|g(\xi,x)| \le |f(x)|$ pour tout x,ξ réels avec $|f| \in \mathcal{L}^1(\mathbb{R})$ (hypothèse) (majoration indépendante du paramètre ξ). Donc on peut appliquer le corollaire précédent : $\widehat{f} \in C^0(\mathbb{R})$. (Et $|\widehat{f}(\xi)| \le ||f||_{\mathcal{L}^1}$ donc $||f||_{\mathcal{L}^\infty} \le ||f||_{\mathcal{L}^1}$.) ∎

Exemple 5.45 Montrer que, pour $t \in [0,\infty[$, posant $F(t) = \displaystyle\int_0^\infty \frac{\sqrt{1+x^2}-1}{(1+x^2)\sqrt{x^2+t}}\,dx$, on a $F(t) \underset{t\to 0+}{\longrightarrow} F(0+)$ $(= \ln 2)$, et que $F(t) \underset{t\to\infty}{\longrightarrow} 0$.

Réponse. Notons f l'intégrant : $f(t,x) = \dfrac{\sqrt{1+x^2}-1}{(1+x^2)\sqrt{x^2+t}}$. Ici le paramètre est t.

1- À x fixé dans $]0,\infty[$, la fonction f_x est continue pour $t \in [0,\infty[$ (on a exclu $x = 0$ mais on a besoin de la continuité des f_x uniquement pour presque tout x).

2- Pour $t \in \mathbb{R}_+$ et $x \in \mathbb{R}_+^*$, on a $\sqrt{1+x^2} - 1 \le x$ (car $1+x^2 \le (1+x)^2$) ; et on a $\sqrt{x^2+t} \ge x$, donc $\frac{1}{\sqrt{x^2+t}} \le \frac{1}{x}$; donc $0 \le f(x,t) \le \frac{1}{1+x^2}$. Et la fonction $h(x) = \frac{1}{1+x^2}$ est indépendante de t et intégrable sur $\mathbb{R}_+$. On peut donc passer à la limite sous le signe $\int$, i.e., on a la continuité de la fonction $F :]0,\infty[\to \mathbb{R}$.

Donc $F(t) \underset{t\to 0+}{\longrightarrow} \displaystyle\int_0^\infty (\lim_{t\to 0+} f(t,x))\,dx = \int_0^\infty f(0,x)\,dx = \int_0^\infty \frac{\sqrt{1+x^2}-1}{(1+x^2)x}\,dx.$

Et $F(t) \underset{t\to\infty}{\longrightarrow} \displaystyle\int_0^\infty (\lim_{t\to\infty} f(t,x))\,dx = \int_0^\infty 0\,dx = 0.$

Reste à voir que $F(0+) = \ln(2)$. On fait le changement de variable $y = \sqrt{1+x^2}$, d'où $ydy = xdx$, d'où $F(0) = \int_1^\infty \frac{y-1}{y^2\sqrt{y^2-1}}\,\frac{y}{\sqrt{y^2-1}}dy = \int_1^\infty \frac{1}{y(y+1)}dy = \int_1^N (\frac{1}{y} - \frac{1}{y+1})\,dy + \int_N^\infty \frac{1}{y(y+1)}dy$. La dernière intégrale est le reste d'une intégrale convergente, et la première vaut $\ln(2) + \ln(\frac{N}{N+1})$ qui tend vers $\ln(2)$ quand $N \to \infty$. ∎

Exemple 5.46 Montrer que, pour $t \in]0,1[$, posant $F(t) = \displaystyle\int_0^1 \frac{1}{\sqrt{x(1-x)(1-tx)}}\,dx$, on a F continue sur $]0,1[$, et $F(0+) = \pi$, et $F(1-) = +\infty$.

Réponse. On pose $g(t,x) = \frac{1}{\sqrt{x(1-x)(1-tx)}}$ pour $x \in]0,1[$ et $t \in [0,1[$.

1- À x fixé dans $]0,1[$, la fonction $g_x : t \in |0,1[\to g_x(t)$ est continue (sauf éventuellement en $t = \frac{1}{x}$ mais $\frac{1}{x} \notin]0,1[$).

2- Pour $t \in [0,\alpha]$ avec $\alpha < 1$, la fonction $g(t,x)$ est dominée par la fonction $g(\alpha,x)$ qui est intégrable (car la fonction $y \to y^{-\frac{1}{2}}$ est intégrable au voisinage de $0+$). Donc $F : t \to F(t)$ est continue sur $[0,\alpha]$, ce pour tout $\alpha < 1$, dont F est continue sur $[0,1[$. Donc $F(0+) = \lim_{t\to 0+} F(t) = \int_0^1 (\lim_{t\to 0+} g_x(t))\,dx = \int_0^1 \frac{dx}{\sqrt{x(1-x)}} = \int_0^1 \frac{2y\,dy}{y\sqrt{1-y^2}} = 2[\arcsin(y)]_0^1 = \pi$.

Et pour $t \in]0,1[$ et $x \geq 0$ on a $1-x \leq 1-tx$, donc $\frac{1}{1-x} \geq \frac{1}{1-tx}$, avec si $x \leq 1$, alors $\frac{1}{x} \geq 1$. Donc $\int_0^1 g(t,x)\,dx \geq \int_0^1 \frac{dx}{\sqrt{x(1-x)(1-tx)}} \geq \int_0^1 \frac{dx}{\sqrt{(1-tx)^2}} = \int_0^1 \frac{dx}{1-tx} = [\frac{-1}{t}\ln(1-tx)]_{x=0}^1 = \frac{-1}{t}\ln(1-t) \longrightarrow_{t\to 1-} +\infty$. Donc $F(t) \underset{t\to 1-}{\longrightarrow} +\infty$. ▟

Exercice 5.47 Soit $f \in \mathcal{L}^1(\mathbb{R})$ et $g \in C^0(\mathbb{R})$ t.q. $||g||_\infty < \infty$ (i.e. g borné). Soit $\varphi := f * g$, i.e. $\varphi(x) = \int_{t=-\infty}^{\infty} f(t)g(x-t)\,dt$. Montrer que φ est continue sur $\mathbb{R}$.

Réponse. On pose $h(x,t) = f(t)g(x-t)$ pour $x,t \in \mathbb{R}$ (l'intégrant).

11- Soit t fixé ; $h_t : x \to h_t(x) := h(x,t)$ est continue sur $\mathbb{R}$ (continuité en x) car g l'est.

12- $|h(x,t)| = |f(t)|\,||g||_\infty$ (domination indépendante du paramètre x), avec $f \in \mathcal{L}^1(\mathbb{R})$: le thm de cvgce dominée donne le résultat. ▟

Exercice 5.48 Soit $f \in \mathcal{L}^1_{\text{loc}}(\mathbb{R})$ et $g \in C^0(\mathbb{R})$ à support compact. Soit $\varphi := f * g$, i.e. $\varphi(x) = \int_{t=-\infty}^{\infty} f(t)g(x-t)\,dt$. Montrer que φ est continue sur $\mathbb{R}$.

Réponse. Soit $h(x,t) = f(t)g(x-t)$ (l'intégrant).

11- Soit t fixé ; $h_t : x \to h_t(x) := h(x,t)$ est continue sur $\mathbb{R}$ (continuité en x) car g l'est.

12- Soit $x_0 \in \mathbb{R}$. Montrons que φ est continue en x_0. Soit $I =]x_0-1, x_0+1[$; g à support compact donc $\exists a,b \in \mathbb{R}$ t.q. $g = 0$ sur $\mathbb{R} - [a,b]$. Donc $g(x-t) = 0$ quand $x-t \notin [a,b]$, i.e. quand $t-x \notin [-b,-a]$, i.e. quand $t \notin [x-b, x-a]$, de pour tout $x \in]x_0-1, x_0+1[$; donc $g(x-t) = 0$ quand $t \notin [x_0-1-b, x_0+1-a] =^{\text{noté}} J$ (borné). Et g est C^0 sur le compact $[a,b]$ donc bornée. Donc $|h(t,x)| = |f(t)g(x-t)1_J(t)| \leq ||g||_\infty |f(t)1_J(t)|$ domination indépendante de x par la fonction $f1_J \in L^1(\mathbb{R})$. Et $\varphi(x) = \int_{t\in J} h(t,x)\,dt$, donc φ est continue en x_0. Vrai pour tout $x_0 \in \mathbb{R}$, donc $\varphi = f * g$ est continue dans $\mathbb{R}$. ▟

Exercice 5.49 Montrer que le théorème de convergence dominée ne s'applique pas au calcul de $\lim_{h\to 0+} F(h)$ (limite à droite en 0) pour :

$$F(h) = \int_0^1 \frac{x}{h} e^{\frac{-x^2}{2h}}\,dx, \tag{5.50}$$

sous peine de conclure que $1 = 0$.

Réponse. Calcul direct : $F(h) = \displaystyle\int_0^1 \frac{x}{h} e^{\frac{-x^2}{2h}}\,dx = [-e^{\frac{-x^2}{2h}}]_0^1 = 1 - e^{\frac{-1}{2h}} \underset{h\to 0+}{\longrightarrow} 1$ (ainsi on peut prolonger F par continuité à droite en 0 en posant $F(0) = F(0+) = 0$).

Et $\displaystyle\int_0^1 \lim_{h\to 0+} (\frac{x}{h} e^{\frac{-x^2}{2h}})\,dx = \int_0^1 0\,dx = 0$ (car l'exponentielle "gagne" : $nxe^{\frac{-nx^2}{2}} \longrightarrow_{n\to\infty} 0$). Comme $1 \neq 0$, on ne peut pas passer à la limite sour $\int$. Ici la fonction intégrée n'est pas dominée, dans un voisinage de $h = 0$, par une fonction intégrable indépendante de h (voir exemple 5.35). ▟

Exercice 5.50 Soit $e = \exp(1) = e^1$, et soit $I(p) = \int_p^{ep} \frac{1}{x}\,dx$ pour $p \geq 0$. Montrer que le théorème de convergence dominée ne s'applique pas pour calculer $\lim_{p\to 0+} I(p)$.

Réponse. On a $I(p) = \int_p^{ep} \frac{1}{x}\,dx = \ln(\frac{ep}{p}) = 1$ constante indépendante de $p > 0$. En particulier $\lim_{p\to 0+} I(p) = 1$. (Et on peut prolonger I par continuité en 0 en posant $I(0) = 1$, ce qui peut surprendre puisque $I(0)$ ressemble à un intégrale sur l'intervalle $[0,0] = \{0\}$ de mesure nulle : cet exemple s'inspire de l'exemple 5.8 où $f(n,x) = n1_{[0,\frac{1}{n}]}(x)$, faire un dessin).

Essayons d'appliquer le théorème de convergence dominée. L'intégrant vaut $f(p,x) = \frac{1}{x}1_{[p,ep]}(x)$ pour $p \geq 0$ et $x > 0$, et $f(0,x) = 0$ et $f(p,x) = 0$ dès que $p < x$; donc, à $x > 0$ fixé, la fonction $f_x : p \to f(x,p)$ est continue en $p = 0$.

Mais la fonction $f : (p,x) \in [0,1]\times]0,1] \to \frac{1}{x}1_{[p,ep]}(x)$ n'est pas dominée indépendamment de $p \in I = [0,\varepsilon]$, $\varepsilon > 0$, par une fonction intégrable au voisinage de $x = 0$ (une fonction dominante h vérifie $h(x) \geq \frac{1}{x}$). ▟

5.6 Dérivation d'une intégrale dépendant d'un paramètre

Notation du § 5.5. Et on s'intéresse encore à l'intégrale dépendant du paramètre p :

$$F : I \to \mathbb{R}, \quad F(p) = \mu(f_p) \quad (= \int_{x\in E} f(p,x)\,d\mu(x)).$$

But : supposant f dérivable dans un voisinage de p_0, on veut savoir si F est dérivable en p_0 et surtout si

$$F'(p_0) = \int_{x\in E} \frac{\partial f}{\partial p}(p_0,x)\,d\mu(x),$$

i.e. si on peut dériver sous le signe $\int$:

$$\frac{d}{dp}\Big(\int_{x\in E} f(p,x)\,d\mu(x)\Big)_{|p=p_0} = \int_{x\in E} \frac{\partial f}{\partial p}(p_0,x)\,d\mu(x),$$

i.e. si

$$\lim_{h\to 0} \frac{F(p_0+h)-F(p_0)}{h} = \int_{x\in E} \lim_{h\to 0} \frac{f(p_0+h,x)-f(p_0,x)}{h}\,d\mu(x). \tag{5.51}$$

Théorème 5.51 *(Dérivabilité d'une intégrale dépendant d'un paramètre.) Soit $(E,\mathcal{A},\mu)$ un ensemble mesuré et $\mathbb{R}$ muni de sa topologie usuelle. Soit $I=]a,b[$ avec $a<b$ et $p_0\in I$. Si $I=[p_0,b[$, avec $p_0<b$, on sous-entendra la dérivabilité à droite en p_0, et si $I=]a,p_0]$, avec $a<p_0$, on sous-entendra la dérivabilité à gauche en p_0. Si $p_0=+\infty$ et $I=]a,p_0[$ avec $a\in\mathbb{R}$, ou si $p_0=-\infty$ et $I=]p_0,b$ avec $b\in\mathbb{R}$, on sous-entendra la limite en $\pm\infty$.*

Soit f et F donnés en (5.41) t.q., avec (5.42), f_p est μ-mesurable pour tout $p\in I$.

Hypothèses :

1. *pour μ-presque tout $x\in E$, la fonction f_x est dérivable en tout $p\in I$, i.e., pour tout $p\in I$,*

$$f_x'(p) = \frac{\partial f}{\partial p}(p,x) \ \in\mathbb{R}, \tag{5.52}$$

2. *domination indépendante du paramètre p par une fonction h μ-intégrable :*

$$\exists h\in\mathcal{L}^1(E),\quad \forall p\in I,\quad |\frac{\partial f}{\partial p}(p,\cdot)|\le h, \tag{5.53}$$

i.e., $\exists h\in\mathcal{L}^1(E)$, $\forall p\in I$, et pour presque tout $x\in E$, on a $|\frac{\partial f}{\partial p}(p,x)|\le h(x)$.

Conclusion : *alors la fonction F est dérivable en p_0, et*

$$(F'(p_0)=)\quad \frac{dF}{dp}(p_0) = \int_E \frac{\partial f}{\partial p}(p_0,x)\,d\mu(x), \tag{5.54}$$

noté :

$$\frac{d}{dp}\Big(\int_E f(p,x)\,d\mu(x)\Big)_{|p=p_0} = \int_E \frac{\partial f}{\partial p}(p_0,x)\,d\mu(x), \tag{5.55}$$

(dérivation sous le signe $\int$ sous l'hypothèse de domination).

Preuve. Avec (5.51), il s'agit de montrer que :

$$\lim_{h\to 0}\Big(\frac{\int_E f(p_0{+}h,x)\,d\mu(x) - \int_E f(p_0,x)\,d\mu(x)}{h}\Big) = \int_E \lim_{h\to 0}\Big(\frac{f(p_0{+}h,x)-f(p_0,x)}{h}\Big)\,d\mu.$$

Soit $g_n(x) = \frac{f(p_0+h_n,x)-f(p_0,x)}{h_n}$ où $(h_n)_{n\in\mathbb{N}^*}$ converge vers 0 et t.q. $p_n=p_0{+}h_n\in I$ pour tout n. Le théorème des accroissements finis donne $g_n(x) = \frac{f_x(p_n)-f_x(p_0)}{h_n} = f_x'(p_k) = \frac{\partial f}{\partial p}(p_k,x)$ avec $p_k\in[p_0,p_0{+}h_n]$ si $h_n>0$ ou $p_k\in[p_0{+}h_n,p_0]$ si $h_n<0$; donc $|g_n(x)|\le h(x)$ p.p., cf. (5.53) ; donc $\lim_{n\to\infty}(\int_E g_n(x)\,d\mu(x)) = \int_E \lim_{n\to\infty} g_n(x)\,d\mu = \int_E \frac{\partial f}{\partial p}(p_0,x)\,d\mu$. Vrai pour toute suite $(p_n)_{\mathbb{N}^*} = (p_0{+}h_n)_{\mathbb{N}^*}$ qui converge vers p_0. ▪

Exemple 5.52 Pour $f\in\mathcal{L}^1(\mathbb{R})$ t.q. $x\to xf(x)$ est dans $\mathcal{L}^1(\mathbb{R})$ (i.e. $\int_\mathbb{R}|x|\,|f(x)|\,dx<\infty$), et sa transformée de Fourier donnée par $\widehat{f}(\xi)=\int_\mathbb{R} f(x)\,e^{-i\xi x}\,dx$, montrer $\widehat{f}\in C^1(\mathbb{R})$ et, pour $\xi\in\mathbb{R}$,

$$(\widehat{f}\,)'(\xi) = \int_\mathbb{R} -ix\,f(x)\,e^{-i\xi x}\,dx. \tag{5.56}$$

Réponse. Posons $g(\xi,x)=f(x)\,e^{-i\xi x}$. À x fixé, la fonction $g_x:\xi\to f(x)\,e^{-i\xi x}$ est dérivable car l'exponentielle l'est, et $g_x'(\xi)=\frac{\partial g}{\partial \xi}(\xi,x) = -ix\,f(x)\,e^{-i\xi x}$. Et donc $|\frac{\partial g}{\partial\xi}(\xi,x)|\le|x|\,|f(x)|$, ce pour tout $x,\xi\in\mathbb{R}$. D'où : la fonction $\widehat{f}$ est dérivable sur $\mathbb{R}$ et sa dérivée est donnée par dérivation sous $\int$: on a (5.56).

Il reste à montrer que $(\widehat{f}\,)'$ est continue. On procède comme dans l'exemple 5.44. ▪

Exemple 5.53 Calculer la dérivée de $F:]0,\infty[\to\mathbb{R}$ définie par $F(t)=\displaystyle\int_{x=0}^\infty x^2\,e^{-tx}\,dx$.

Réponse. Soit $f:(t,x)\in I\times\mathbb{R}_+^*\to f(t,x)=x^2\,e^{-tx}\in\mathbb{R}$ (l'intégrant) où $I=\mathbb{R}_+^*$. 1- Dérivation : à $x\in\mathbb{R}_+^*$ fixé, la fonction $f_x:t\to f_x(t)=f(t,x)$ est dérivable dans I, et $f_x'(t)=\frac{\partial f}{\partial t}(t,x)=-x^3e^{-tx}$. 2- Domination : $|\frac{\partial f}{\partial t}(t,x)|\le|x^3|$ donne une domination indépendante du paramètre $t\in I$... mais $x\to x^3$ n'est pas intégrable sur $\mathbb{R}_+$ (ici l'intervalle I est trop grand). En fait on veut montrer que F est dérivable en tout $t_0\in\mathbb{R}_+^*$, à l'aide du thm de cvgce dominée : on a donc besoin d'un intervalle ouvert I_{t_0} contenant t_0 ; Ici on prend simplement $I_{t_0}=]\frac{t_0}{2},\infty[$ qui donne $|\frac{\partial f}{\partial t}(t,x)\le -x^3e^{-\frac{t_0}{2}x}$ fonction $\mathcal{L}^1(\mathbb{R})$ indépendante de $t\in I_{t_0}$. 3- Le théorème 5.51 donne : F est dérivable en tout $t\in I_{t_0}$, donc en t_0, et $F'(t_0)=-\int_0^\infty x^3e^{-t_0x}\,dx$. Vrai pour tout $t_0>0$, et F est dérivable dans $\mathbb{R}_+^*$. ▪

Exemple 5.54 Cas pour lequel l'intégrale de Riemann ne permet pas de répondre. Soit $g \in C^0([0,1])$, $g \geq 0$, $G_0 = g^{-1}(\{0\}) = \{x \in [0,1] : g(x) = 0\}$ (dessin), $G_0^c = [0,1] - G_0$ (complémentaire),

$$\varphi : \left\{ \begin{array}{ll} I = [0,1] & \to \mathbb{R} \\ \quad t & \to \varphi(t) = \displaystyle\int_0^1 \sqrt{g(x) + t^2}\, dx \end{array} \right\}. \quad \text{Montrer :} \quad \varphi'(0+) = \mu_l(G_0), \tag{5.57}$$

i.e. la dérivée à droite en 0 est la longueur en "contact avec le sol".

Réponse. Soit $f(t,x) = \sqrt{g(x)+t^2}$. 1- À $x \in]0,1|$ fixé, $f_x : t \to f_x(t) := f(t,x)$ est continue dans I (donc continue à droite en 0 et continue à gauche en 1). 2- $|f(t,x)| \leq \sqrt{g(x)+1}$ (domination indépendante de t) avec $x \to \sqrt{g(x)+1}$ continue sur le compact $[0,1]$ (car g et $\sqrt{\ }$ sont continues), donc intégrable sur $[0,1]$. Donc (thm de cvgce dominée) $\varphi \in C^0(I)$.

On a $\left\{ \begin{array}{l} \forall (t,x) \in G_0 \times I, \;\; f(t,x) = t, \\ \forall (t,x) \in G_0^c \times I, \;\; f(t,x) = \sqrt{g(x)+t^2}. \end{array} \right\}$ 1- À $x \in]0,1[$ fixé, la fonction f_x est dérivable dans I (composée de fonctions dérivables),

$$\left\{ \begin{array}{ll} \forall (t,x) \in G_0 \times I, \;\; \dfrac{\partial f}{\partial t}(t,x) = 1, & \text{donc} \;\; \dfrac{\partial f}{\partial t}(0,x) = 1, \\ \forall (t,x) \in G_0^c \times I, \;\; \dfrac{\partial f}{\partial t}(t,x) = \dfrac{t}{\sqrt{g(x)+t^2}}, & \text{donc} \;\; \dfrac{\partial f}{\partial t}(0,x) = 0 \end{array} \right. \tag{5.58}$$

(au sens de la dérivée à droite en 0). 2- $|\frac{\partial f}{\partial t}(t,x)| \leq 1 = 1_{]0,1[}(x)$ pour tout $(t,x) \in I \times [0,1]$, avec $1_{]0,1[} \in \mathcal{L}^1(]0,1[)$ (fonction indépendante de t). Donc φ est dérivable dans I et, pour tout $t \in I$,

$$\varphi'(t) = \int_0^1 1_{G_0}(x)\, dx + \int_0^1 \frac{t}{\sqrt{g(x)+t^2}} 1_{G_0^c}(x)\, dx = \mu_\ell(G_0) + \int_0^1 \frac{t}{\sqrt{g(x)+t^2}} 1_{G_0^c}(x)\, dx. \tag{5.59}$$

En particulier $\varphi'(0+) = \mu_\ell(G_0) + \int_0^1 0\, dx$, i.e. (5.57). ∎

Exercice 5.55 Soit $f \in \mathcal{L}^1_{\rm loc}(\mathbb{R})$ et $g \in C^1(\mathbb{R})$ à support compact, soit $\varphi(x) := (f*g)(x) = \int_{t=-\infty}^{\infty} f(t)g(x-t)\, dt$. Montrer que φ est C^1 sur $\mathbb{R}$.

Réponse. Appliquer la démarche de l'exercice 5.48. ∎

N.B. : on peut relaxer la condition de "dérivée définie et dominée dans tout I" : en supposant la dérivabilité en p_0 sous l'hypothèse de "domination Lipschitzienne" dans un voisinage de p_0 :

Théorème 5.56 *Soit f comme en (5.41) telle que f_p est mesurable pour tout $p \in I$. Soit $p_0 \in I$. On suppose :*

1. *pour presque tout $x \in E$, la fonction f_x est dérivable en p_0,*
2. *et domination lipschitzienne :*

$$\exists h \in \mathcal{L}^1, \quad \forall p \in I,\ p \neq p_0, \quad \frac{|f(p,.) - f(p_0,.)|}{|p - p_0|} \leq h, \tag{5.60}$$

i.e. $|f(p,x) - f(p_0,x)| \leq h(x)\, |p - p_0|$ pour tout $p \in I$ et presque tout $x \in E$.

Alors la fonction F est dérivable en p_0, et sa dérivée est donnée par :

$$F'(p_0) = \int_E \frac{\partial f}{\partial p}(p_0,x)\, d\mu(x), \tag{5.61}$$

(dérivation sous le signe $\int$ sous l'hypothèse de domination lipschitzienne).

Preuve. Il s'agit de montrer que (5.51) :

$$\lim_{\varepsilon \to 0}\Big(\frac{\int_E f(p_0+\varepsilon, x)\, d\mu(x) - \int_E f(p_0,x)\, d\mu(x)}{\varepsilon}\Big) = \int \lim_{\varepsilon \to 0}\Big(\frac{f(p_0+\varepsilon,x) - f(p_0,x)}{\varepsilon}\Big)\, d\mu,$$

i.e.$\lim_{\varepsilon\to 0}(\int_E \frac{f(p_0+\varepsilon,x)-f(p_0,x)}{\varepsilon}\, d\mu(x)) = \int_E (\lim_{\varepsilon\to 0} \frac{f(p_0+\varepsilon,x)-f(p_0,x)}{\varepsilon})\, d\mu$. On applique le théorème de convergence dominée en posant $g_n(x) = \frac{f(p_0+\varepsilon_n,x)-f(p_0,x)}{\varepsilon_n}$ où $(\varepsilon_n)_{n\in\mathbb{N}}$ est une suite quelconque dans I qui converge vers 0. Et ici $|g_n(x)| = |\frac{f_x(p_0+\varepsilon_n)-f_x(p_0)}{\varepsilon_n}| \leq h(x)$ p.p., on peut donc appliquer le théorème de convergence dominée. ∎

Exercice 5.57 $F(p) = \int_{x\in\mathbb{R}} f(x) e^{-|p-x|}\, dx$ pour $f \in L^1(\mathbb{R})$ et $p \in \mathbb{R}$.

Montrer : $F'(p) = \int_{x\in\mathbb{R}} f(x)\, \mathrm{sgn}(x-p)\, e^{-|p-x|}\, dx$.

Réponse. Intégrant $g(p,x) = f(x)e^{-|p-x|} \left\{ \begin{array}{l} = f(x)e^{x-p} \text{ si } p \geq x \\ = f(x)e^{p-x} \text{ si } p \leq x \end{array} \right\}$. Donc, à x fixé, pour tout $p \neq x$ la fonction f est dérivable en p avec $\frac{\partial g}{\partial p}(p,x) \left\{ \begin{array}{l} = -f(x)e^{x-p} \text{ si } p > x \\ = f(x)e^{p-x} \text{ si } p < x \end{array} \right\} = \mathrm{sgn}(x-p)\, e^{-|p-x|}$, p.p. x, avec $|\frac{g(q,x)-g(p,x)}{q-p}| \leq f(x)|$, pour tout $q \neq p$ p.p. x (caractère Lipschitzien). D'où le résultat. ∎

5.7 Application aux lois de conservation en mécanique

Toujours avec une fonction comme en (5.41), mais en adoptant les notations t = temps et x = espace :

$$f : \begin{cases} \mathbb{R} \times \mathbb{R}^n & \to \mathbb{R}, \\ (t, x) & \to f(t, x), \end{cases} \tag{5.62}$$

où $n = 1$, 2 ou 3. Et, pour $t_0 \in \mathbb{R}$ appelé instant initial, et pour Ω_{t_0} un ouvert dans $\mathbb{R}^n$, on se donne "les trajectoires des particules" se trouvant dans Ω_{t_0} à t_0 :

$$\Phi : \begin{cases} \mathbb{R} \times \Omega_{t_0} & \to \mathbb{R}^n, \\ (t, X) & \to x = \Phi(t, X), \end{cases} \tag{5.63}$$

À $X \in \Omega_{t_0}$ fixé, la trajectoire de (la particule qui se trouve en) X est :

$$\Phi_X : \left\{ \begin{array}{ll} \mathbb{R} & \to \mathbb{R}^n, \\ t & \to x_t = \Phi_X(t) \stackrel{\text{déf}}{=} \Phi(t, X), \end{array} \right\} \quad \text{et} \quad {\Phi_X}'(t) \stackrel{\text{noté}}{=} \vec{v}(t, x_t) \quad \text{quand} \quad x_t = \Phi(t, X), \tag{5.64}$$

où on a supposé Φ_X dérivable, et où $\vec{v}(t, x_t)$ = la vitesse eulérienne (de la particule qui se trouve en x_t à t).

À $t \in \mathbb{R}$ fixé, on note :

$$\Phi_t : \left\{ \begin{array}{ll} \Omega_{t_0} & \to \mathbb{R}^n, \\ X & \to x = \Phi_t(X) \stackrel{\text{déf}}{=} \Phi(t, X), \end{array} \right\} \quad \text{et on note} \quad \Omega_t = \Phi_t(\Omega_{t_0}). \tag{5.65}$$

À $X \in \Omega_{t_0}$ fixé on considère la fonction :

$$g(t) = f(t, \Phi_X(t)). \tag{5.66}$$

Supposant f et Φ C^1, la dérivée particulaire de f est la fonction (eulérienne) définie sur $\bigcup_{t\in\mathbb{R}}(\{t\} \times \Omega_t)$ par :

$$\frac{Df}{Dt}(t, \Phi_X(t)) \stackrel{\text{déf}}{=} g'(t), \tag{5.67}$$

donc avec la vitesse eulérienne définie en (5.64) :

$$\begin{aligned} \frac{Df}{Dt}(t, \Phi_X(t)) &= \frac{\partial f}{\partial t}(t, \Phi_X(t)) + df(t, \Phi_X(t)).{\Phi_X}'(t) \\ &= (\frac{\partial f}{\partial t} + df.\vec{v})(t, \Phi_X(t)), \end{aligned} \tag{5.68}$$

où $df(t, x) =^{\text{déf}} df_t(x)$ est la dérivée "en espace" de f, i.e. la différentielle à t fixé de la fonction f_t définie par $f_t(x) =^{\text{déf}} f(t, x)$ pour tout $x \in \mathbb{R}^n$. Autrement dit :

$$\frac{Df}{Dt} \stackrel{\text{déf}}{=} \frac{\partial f}{\partial t} + df.\vec{v}. \tag{5.69}$$

Soit :

$$F(t) = \int_{x_t\in\Omega_t} f(t, x_t)\, d\Omega_t, \tag{5.70}$$

But : Calculer $F'(t) =^{\text{noté}} \frac{d}{dt}(\int_{x_t\in\Omega_t} f(t, x_t)\, d\Omega_t)$, sachant qu'ici le domaine Ω_t d'intégration varie avec t.

Proposition 5.58 *Soit $t_0, T \in \mathbb{R}$ avec $t_0 < T$. Soit U un ouvert borné de $\mathbb{R}^n$ t.q. $U \supset \Omega_t$ pour tout $t \in [t_0, T]$.*

Hypothèses (pour la mécanique) : Ω_{t_0} est un ouvert borné, et

Hyp 1- 11- La fonction $\Phi : [t_0, T] \times \Omega_{t_0} \to \mathbb{R}^n$ est C^1 telle que 12- $d\Phi_t : \Omega_{t_0} \to \Omega_t$ est un difféomorphisme C^1 pour tout $t \in [t_0, T]$, 13- et, notant $J(t, X) = \det(d\Phi_t(x_t))$ le déterminant de la matrice jacobienne (relativement à la base canonique de $\mathbb{R}^n$) :

$$\exists C > 0,\ \forall t \in [t_0, T],\ \forall X \in \Omega_{t_0}, \quad |J(t, X)| \le C \quad \text{et} \quad |\frac{\partial J}{\partial t}(t, X)| \le C. \tag{5.71}$$

Hyp 2- La fonction f est $C^1([t_0, T] \times \overline{U})$.

Conclusion : F est dérivable et on a :

$$(F'(t) =) \quad \frac{d}{dt}\Big(\int_{x_t\in\Omega_t} f(t, x_t)\, d\Omega_t\Big) = \int_{x_t\in\Omega_t} \frac{Df}{Dt}(t, x_t) + f(t, x_t)\mathrm{div}\vec{v}(t, x_t)\, d\Omega_t. \tag{5.72}$$

Preuve. Grâce au changement de variable dans les intégrales, on se ramène au domaine Ω_{t_0} (qui ne varie pas avec t) :

$$F(t) = \int_{X\in\Omega_{t_0}} f(t,\Phi(t,X))|J(t,X)|\,d\Omega_{t_0}. \tag{5.73}$$

Ici $J(t,X)$ est >0 car $J(t_0,X)=1$, J est continue en t et J ne s'annule pas (Φ_t est un difféomorphisme pour tout t). Notons G l'intégrant (fonction lagrangienne) :

$$G(t,X) = f(t,\Phi(t,X))J(t,X). \tag{5.74}$$

1- G est dérivable en t avec $\frac{\partial G}{\partial t}(t,X) = \frac{Df}{Dt}(t,\Phi(t,X))J(t,X) + f(t,\Phi(t,X))\frac{\partial J}{\partial t}(t,X)$, et

2- $|\frac{\partial G}{\partial t}(t,X)| \le C$ sur $[t_0,T]\times U$, avec C constante = le sup du membre de droite sur $[t_0,T]\times U$, car U est supposé borné et f est C^1 sur le compact $[t_0,T]\times\overline{U}$.

Donc on peut appliquer le théorème de convergence dominée :

$$\frac{d}{dt}\Big(\int_{\Omega_{t_0}} G(t,X)\,d\Omega_{t_0}\Big) = \int_{\Omega_{t_0}} \frac{\partial G}{\partial t}(t,X)\,d\Omega_{t_0}. \tag{5.75}$$

Soit :

$$F'(t) = \int_{\Omega_{t_0}} \frac{Df}{Dt}(t,\Phi_X(t))J(t,X) + f(t,\Phi_X(t))\frac{\partial J}{\partial t}(t,X)\,d\Omega_{t_0}. \tag{5.76}$$

Et un calcul direct donne (voir polycopié “Objectivity in classical mechanics : Motions, flows, Eulerian and Lagrangian functions...”)

$$\frac{\partial J}{\partial t}(t,X) = J(t,X)\,\mathrm{div}\vec{v}(t,x_t) \quad \text{quand} \quad x_t = \Phi(t,X). \tag{5.77}$$

D'où le résultat par changement de variable pour revenir à Ω_t. ∴

6 Théorème de Fubini

6.1 Le théorème de Fubini

On se donne deux ensembles E et F munis de tribus $\mathcal{A}_1$ et $\mathcal{A}_2$, et on considère l'ensemble produit $E\times F = \{(x,y)\in E\times F\}$.

Définition 6.1 On appelle tribu produit de $\mathcal{A}_1$ et $\mathcal{A}_2$ la tribu $\mathcal{A}_1\otimes\mathcal{A}_2$ engendrée par les éléments de la forme $A_1\times A_2$ pour $A_1\in\mathcal{A}_1$ et $A_2\in\mathcal{A}_2$.

Exemple 6.2 La tribu borélienne de $\mathbb{R}^2$ est engendrée par les pavés $]a_1,b_1[\times]a_2,b_2[$ (elle contient donc tous les ouverts de $\mathbb{R}^2$). ∴

On a le théorème (admis) :

Théorème 6.3 *Étant données deux mesures μ_1 et μ_2 définies sur $\mathcal{A}_1$ et $\mathcal{A}_2$, il existe une unique mesure μ, notée $\mu_1\otimes\mu_2$, définie sur $\mathcal{A}_1\otimes\mathcal{A}_2$ vérifiant :*

$$\mu(A_1\times A_2) = \mu_1(A_1)\mu_2(A_2) \overset{\text{noté}}{=} (\mu_1\otimes\mu_2)(A_1\times A_2). \tag{6.1}$$

Pour f_1 fonction sur E_1 et f_2 fonction sur E_2, on note $f_1\otimes f_2$ la fonction (à variables séparées) définie sur $E_1\times E_2$ par le produit :

$$(f_1\otimes f_2)(x,y) \overset{\text{déf}}{=} f_1(x)f_2(y). \tag{6.2}$$

Du théorème précédent on déduit (pour les fonctions à variables séparées) :

Théorème 6.4 *Si f_1 est μ_1-intégrable, et si f_2 est μ_2-intégrable, alors la fonction $f_1\otimes f_2$ est $\mu_1\otimes\mu_2$ intégrable et (intégration de fonctions à variables séparées) :*

$$(\mu_1\otimes\mu_2)(f_1\otimes f_2) = \mu_1(f_1)\mu_2(f_2). \tag{6.3}$$

Noté $\int_{E\times F} f_1\otimes f_2\,d(\mu_1\otimes\mu_2) = (\int_E f_1\,d\mu_1)(\int_F f_2\,d\mu_2)$.

Pour les fonctions non à variables séparées, on a le théorème de Fubini (admis) :

Théorème 6.5 *(Fubini.) (i) (Cas des fonctions positives) Si $f : E \times F \to [0,\infty]$ est mesurable positive alors, dans $\overline{\mathbb{R}_+}$,*

$$(\mu_1 \otimes \mu_2)(f) = \mu_1(\mu_2(f)) = \mu_2(\mu_1(f)), \tag{6.4}$$

i.e.

$$\begin{aligned}\int_{(x,y)\in E_1\times E_2} f(x,y)\, d(\mu_1\otimes\mu_2)(x,y) &= \int_{x\in E_1} \Big(\int_{y\in E_2} f(x,y)\, d\mu_2(y) \Big) d\mu_1(x) \\ &= \int_{y\in E_2} \Big(\int_{x\in E_1} f(x,y)\, d\mu_1(x) \Big) d\mu_2(y).\end{aligned} \tag{6.5}$$

(ii) (Cas des fonctions intégrables) Si $f \in L^1(E_1 \times E_2)$ est t.q.

$$(\mu_1\otimes\mu_2)(|f|) < \infty, \quad \text{i.e.} \quad \int_{E_1\times E_2} |f(x,y)|\, d(\mu_1\otimes\mu_2)(x,y) < \infty, \tag{6.6}$$

alors : pour presque tout $x \in E_1$ la fonction $y \to f(x,y)$ est μ_2-intégrable, et pour presque tout $y \in E_2$ la fonction $x \to f(x,y)$ est μ_1-intégrable, et on a (6.4) et (6.5).

Remarque 6.6 Attention, dans (i) si f n'est pas positive, ou dans (ii) si f n'est pas intégrable, on ne peut pas appliquer le théorème de Fubini. Par exemple, on a :

$$\int_{x=0}^1 \Big(\int_{y=0}^1 \frac{x^2-y^2}{(x^2+y^2)^2}\, dy \Big) dx = \int_{x=0}^1 [\frac{y}{x^2+y^2}]_{y=0}^1\, dx = \int_{x=0}^1 \frac{1}{x^2+1}\, dx = \arctan 1 = \frac{\pi}{4}, \tag{6.7}$$

alors que :

$$\int_{y=0}^1 \Big(\int_{x=0}^1 \frac{x^2-y^2}{(x^2+y^2)^2}\, dx \Big) dy = \int_{y=0}^1 [\frac{-x}{x^2+y^2}]_{x=0}^1\, dy = -\int_{y=0}^1 \frac{1}{1+y^2}\, dy = -\arctan 1 = -\frac{\pi}{4}. \tag{6.8}$$

Les deux intégrales existent mais ne sont pas égales : on ne peut pas inverser l'ordre d'intégration : en effet (i) ne peut s'appliquer, et dans (ii), la seule hypothèse à vérifier (à savoir f est $\mu_1 \otimes \mu_2$-intégrable) n'est pas vérifier ! En effet, l'intégrant se comporte au mieux comme $\frac{1}{x^2+y^2}$ au voisinage de $(0,0)$. Et si f était intégrable, on aurait $\frac{\pi}{4} = -\frac{\pi}{4}$. (Noter également qu'on ne peut pas intégrer en séparant les termes $\frac{x^2-y^2}{(x^2+y^2)^2} = \frac{1}{x^2+y^2} - \frac{2y^2}{(x^2+y^2)^2}$ puisque par exemple $\frac{1}{x^2+y^2}$ n'est pas intégrable sur $]0,1[^2$.) ▲

Rappel : On vérifie que dans $\mathbb{R}^n$, si $r = \sqrt{x_1^2 + ... + x_n^2}$, alors

$$\begin{aligned}&\int_{r\le 1} r^\alpha\, dx_1...dx_n < \infty \text{ ssi } \alpha > -n, \\ &\int_{r\ge 1} r^\alpha\, dx_1...dx_n < \infty \text{ ssi } \alpha < -n.\end{aligned} \tag{6.9}$$

En effet, on passe en coordonnées sphériques : $dx_1...dx_n = r^{n-1} dr dS_1$ où dS_1 est l'élément de surface de la sphère unité (et $dS_r = r^{n-1}dS$ celui de la sphère de rayon r), et on applique le théorème 6.4 sur la boule unité :

$$\int_{r\le 1} r^\alpha\, dx_1...dx_n = S_1 \int_0^1 r^{\alpha+n-1}\, dr. \tag{6.10}$$

Et l'intégrale n'est réelle (n'est finie) que si $\alpha + n - 1 > -1$.

6.2 Le théorème de Tonelli

Et on a également le théorème d'intégration de Tonelli qui donne des conditions suffisantes, faciles à vérifier, pour qu'une fonction soit dans $L^1(\Omega_1 \times \Omega_2)$:

Théorème 6.7 *(Tonelli) Si la fonction $f : \Omega_1 \times \Omega_2 \subset \mathbb{R}^2 \to \mathbb{R}$ satisfait aux deux hypothèses :*

$$\int_{\Omega_2} |f(x,y)|\, dy < \infty \quad p.p.\ x, \quad \text{et} \quad \int_{\Omega_1} \Big(\int_{\Omega_2} |f(x,y)|\, dy \Big) dx < \infty, \tag{6.11}$$

alors $f \in L^1(\Omega_1 \times \Omega_2)$, et on peut appliquer le théorème de Fubini.

Remarque 6.8 Si on reprend l'exemple de la remarque 6.6, l'intégrant est $|f(x,y)| = \frac{x^2-y^2}{(x^2+y^2)^2}$ si $y \le x$ et $= -\frac{x^2-y^2}{(x^2+y^2)^2}$ si $y \ge x$. Et on a $\int_{y=0}^1 |f(x,y)|\, dy = \int_{y=0}^x \frac{x^2-y^2}{(x^2+y^2)^2}\, dx + \int_{y=x}^1 -\frac{x^2-y^2}{(x^2+y^2)^2}\, dx$. Et $\int_{y=0}^x \frac{x^2-y^2}{(x^2+y^2)^2}\, dx = [\frac{y}{x^2+y^2}]_{y=0}^x = \frac{1}{2x}$ qui n'est pas intégrable sur l'intervalle $[0,1] \ni x$. Et les hypothèses du théorème de Tonelli ne sont pas vérifiées. Il n'est donc pas étonnant que le théorème de Fubini ne soit pas applicable. ▲

6.3 Fourier inverse

Soit la transformée de Fourier $\mathcal{F} : \left\{ \begin{array}{rl} \mathcal{L}^1(\mathbb{R}) & \to C^0(\mathbb{R}) \\ f & \to \mathcal{F}(f) = \widehat{f} \end{array} \right\}$ définie par (cf. ex. 5.44 au facteur $\frac{1}{\sqrt{2\pi}}$ près) :

$$(\mathcal{F}(f)(\omega) =) \quad \widehat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{t\in\mathbb{R}} f(t) e^{-i\omega t}\, dt. \tag{6.12}$$

Alors on a, pour $f \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^\infty(\mathbb{R}) \cap C^0(\mathbb{R})$ t.q. $\widehat{f} \in \mathcal{L}^1(\mathbb{R})$:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{\omega\in\mathbb{R}} \widehat{f}(\omega) e^{+i\omega t}\, dt = \widehat{\widehat{f}}(-t), \quad \text{donc} \quad \mathcal{F}^{-1} = \overline{\mathcal{F}}, \tag{6.13}$$

où $\overline{\mathcal{F}}(f)(\omega) = \frac{1}{\sqrt{2\pi}} \int_{t\in\mathbb{R}} f(t) e^{+i\omega t}\, dt$.

N.B. : plus généralement ce résultat tient pour "$f \in \mathcal{L}^1(\mathbb{R})$ t.q. $\widehat{f} \in \mathcal{L}^1(\mathbb{R})$", voir cours de distribution.

N.B. : ce résultat tient pour $\mathcal{F} : \mathcal{L}^2(\mathbb{R}) \to \mathcal{L}^2(\mathbb{R})$, et $\mathcal{F}$ est un isomorphisme $\mathcal{L}^2(\mathbb{R}) \to \mathcal{L}^2(\mathbb{R})$, voir cours de distribution.

Exercice 6.9 Démonstration de (6.13).

1. Soit $f, h \in \mathcal{L}^1(\mathbb{R})$ t.q. $\widehat{f}, \widehat{h} \in \mathcal{L}^1(\mathbb{R})$. Montrer (avec Fubini) :

$$\int_{\omega\in\mathbb{R}} \widehat{f}(\omega)\, h(\omega)\, d\omega = \int_{t\in\mathbb{R}} f(t)\, \widehat{h}(t)\, dt. \tag{6.14}$$

2. Montrer : si $f \in \mathcal{L}^1(\mathbb{R})$ alors $\widehat{f} \in \mathcal{L}^\infty(\mathbb{R}) \cap C^0(\mathbb{R})$, mais $\widehat{f} \notin \mathcal{L}^1(\mathbb{R})$ en général.
3. Montrer : si f **<u>et</u>** f' sont dans $L^1(\mathbb{R})$, alors $f(t) \underset{t\to\pm\infty}{\longrightarrow} 0$.
4. Montrer : • si $f, f' \in L^1(\mathbb{R})$ (donc f s'annule à l'infini) alors $\widehat{f'}(\omega) = i\omega\widehat{f}$;

   • si $f, tf \in \mathcal{L}^1(\mathbb{R})$ alors $\widehat{tf}(\omega) = i(\widehat{f}\,)'(\omega)$ où on a noté tf la fonction $t \to tf(t)$.
5. Pour $a > 0$, on note $g_a : \mathbb{R} \to \mathbb{R}$ la gaussienne centrée donnée par :

$$g_a(t) = e^{-\frac{at^2}{2}} \quad (\text{et } g_1 : t \to g_1(t) = e^{-\frac{t^2}{2}} \text{ est la gaussienne centrée réduite}). \tag{6.15}$$

   Montrer :

$$\int_{\mathbb{R}} g_a(t)\, dt = \int_{\mathbb{R}} e^{-\frac{at^2}{2}}\, dt = \sqrt{\frac{2\pi}{a}} \quad (= \widehat{g_a}(0)). \tag{6.16}$$

   et

$$\widehat{g_a} = \frac{1}{\sqrt{a}} g_{\frac{1}{a}}, \quad \text{i.e.} \quad \widehat{g_a}(\omega) = \frac{1}{\sqrt{a}} e^{-\frac{\omega^2}{2a}} = \frac{1}{\sqrt{a}} g_{\frac{1}{a}}(\omega), \quad \text{et} \quad \int_{\mathbb{R}} \widehat{g_a}(\omega)\, d\omega = \sqrt{2\pi}. \tag{6.17}$$

   En particulier $\widehat{g_1} = g_1$: la gaussienne centrée réduite est "conservée" par Fourier. (Et $\widehat{g_a} \in L^1(\mathbb{R})$.)
6. Montrer : si $f \in \mathcal{L}^1(\mathbb{R})$ alors :

$$\int_{t\in\mathbb{R}} f(t) g_\varepsilon(t)\, dt \underset{\varepsilon\to 0}{\longrightarrow} \int_{t\in\mathbb{R}} f(t)\, dt, \tag{6.18}$$

   étalement à la limite de la fonction g_ε vers la fonction constante $1_{\mathbb{R}}$, soit $g_\varepsilon \underset{\varepsilon\to 0}{\rightharpoonup} 1_{\mathbb{R}}$ au sens des distributions.
7. Montrer : si $f \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^\infty(\mathbb{R})$ et f continue dans un voisinage de 0 alors

$$\frac{1}{\sqrt{2\pi}} \int_{\omega\in\mathbb{R}} f(\omega) \widehat{g_\varepsilon}(\omega)\, d\omega \underset{\varepsilon\to 0}{\longrightarrow} f(0), \tag{6.19}$$

   concentration à la limite de $\frac{\widehat{g_\varepsilon}}{\sqrt{2\pi}}$ vers la masse de Dirac δ_0. (Donc $\frac{\widehat{g_\varepsilon}}{\sqrt{2\pi}} \rightharpoonup_{\varepsilon\to 0} \delta_0$ au sens des distributions.)
8. Soit $f \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^\infty(\mathbb{R}) \cap C^0(\mathbb{R})$ t.q. $\widehat{f} \in \mathcal{L}^1(\mathbb{R})$, soit $H_t(\varepsilon) = \frac{1}{\sqrt{2\pi}} \int_{\omega\in\mathbb{R}} \widehat{f}(\omega)\, g_\varepsilon(\omega)\, e^{it\omega}\, d\omega$.

   8.1. Montrer directement que $H_t(\varepsilon) \longrightarrow_{\varepsilon\to 0} \widehat{\widehat{f}}(-t)$.

   8.2. Montrer en utilisant Fubini que $H_t(\varepsilon) \longrightarrow_{\varepsilon\to 0} f(t)$.

   8.3. Conclure que $\widehat{\widehat{f}}(-t) = f(t)$, donc (6.13).

Réponse.

1. $\int_{\omega\in\mathbb{R}} \widehat{f}(\omega)\,h(\omega)\,d\omega = \frac{1}{\sqrt{2\pi}}\int_{\omega\in\mathbb{R}}(\int_{t\in\mathbb{R}} f(t)e^{-it\omega}\,dt)\,h(\omega)\,d\omega$. Soit $z(t,\omega)=f(t)h(\omega)\,e^{-it\omega}$. On a $|z(t,\omega)|\le|f(t)h(\omega)|=|(f\otimes h)(t,\omega)|$ avec $f,h\in\mathcal{L}^1(\mathbb{R})$, donc $f\otimes h\in\mathcal{L}^1(\mathbb{R}^2)$, donc (Fubini) :
$$\int_{\omega\in\mathbb{R}} \widehat{f}(\omega)\,g(\omega)\,d\omega = \frac{1}{\sqrt{2\pi}}\int_{t\in\mathbb{R}} f(t)\Big(\int_{\omega\in\mathbb{R}} h(\omega)\,e^{-it\omega}\,d\omega\Big)\,dt = \int_{u\in\mathbb{R}} f(t)\,\widehat{h}(t)\,dt. \tag{6.20}$$
2. $\widehat{f}\in\mathcal{L}^\infty(\mathbb{R})\cap C^0(\mathbb{R})$: voir ex. 5.44. $f(t)=e^{-at}1_{\mathbb{R}_+}(t)$ vérifie $f\in\mathcal{L}^1(\mathbb{R})$, et $\widehat{f}(\omega)=\frac{1}{\sqrt{2\pi}}\frac{1}{a+i\omega}$ (immédiat), donc $\widehat{f}\notin\mathcal{L}^1(\mathbb{R})$.
3. $f'\in L^1(\mathbb{R})$, donc : $\forall\varepsilon>0,\ \exists A_\varepsilon>0,\ \int_{A_\varepsilon}^\infty|f'(t)|\,dt<\varepsilon$, donc, ayant $|f'|\ge0$,
$$\forall\varepsilon>0,\ \exists A_\varepsilon>0,\ \forall t>A_\varepsilon,\ \int_{A_\varepsilon}^t |f'(u)|\,du<\varepsilon\ ;$$
donc $|\int_{A_\varepsilon}^t f'(u)\,du|<\varepsilon$, donc $|f(t)-f(A_\varepsilon)|<\varepsilon$, donc $\big||f(t)|-|f(A_\varepsilon)|\big|<\varepsilon$, donc $|f(t)|-|f(A_\varepsilon)|<\varepsilon$ et $|f(A_\varepsilon)|-|f(t)|<\varepsilon$, donc $|f(t)>|f(A_\varepsilon)|-\varepsilon$. Si $|f(A_\varepsilon)|-\varepsilon=c>0$ alors $||f||_{L^1(\mathbb{R})}\ge\int_{A_\varepsilon}^\infty|f(t)|\,dt\ge\int_A^\infty(|f(A_\varepsilon)|-\varepsilon)\,dx\ge\int_A^\infty c\,dx=\infty$, faux pour $f\in L^1(\mathbb{R})$. Donc $|f(A_\varepsilon)|\le\varepsilon$. Donc $|f(t)|<2\varepsilon$, $\forall t>A_\varepsilon$, vrai $\forall\varepsilon>0$. Donc $f(t)\longrightarrow_{t\to\infty}0$. Idem en $-\infty$.
4. $\widehat{f'}(\omega)=\frac{1}{\sqrt{2\pi}}\int_{t\in\mathbb{R}}f'(t)e^{-i\omega t}\,dt=\frac{1}{\sqrt{2\pi}}\int_{t\in\mathbb{R}}(i\omega)f(t)e^{-i\omega t}\,dt+\frac{1}{\sqrt{2\pi}}[f(t)e^{-i\omega t}]_{-\infty}^{\infty}=i\omega\widehat{f}(\omega)+0$.
   Et $(\widehat{f}\,)'(\omega)=^{(5.56)}\frac{1}{\sqrt{2\pi}}\int_{t\in\mathbb{R}}(-it)f(t)e^{-i\omega t}\,dt=-i\,\widehat{tf}\,(\omega)$.
5. $\big(\int_{\mathbb{R}}e^{-\frac{ax^2}{2}}\,dx\big)^2=\iint_{\mathbb{R}^2}e^{-\frac{a(x^2+y^2)}{2}}\,dxdy=\int_{r=0}^\infty\int_{\theta=0}^{2\pi}e^{-a\frac{r^2}{2}}r\,dr\,d\theta=2\pi[-\frac1a e^{-a\frac{r^2}{2}}]_0^\infty=\frac{2\pi}{a}$, et $g_a\ge0$, d'où (6.16). Et g_a est solution (trivial) de l'E.D.O.
$$g_a'(t)+a\,t\,g_a(t)=0,\quad g_a(0)=1. \tag{6.21}$$
   $g_a\in\mathcal{L}^1(\mathbb{R})$ et $||g_a||_\infty\le1$ (immédiat), et $\mathcal{F}$ est linéaire sur $\mathcal{L}^1(\mathbb{R})$ (immédiat), d'où, en appliquant $\mathcal{F}$ à (6.21) :
$$i\omega\,\widehat{g_a}(\omega)+ai\,(\widehat{g_a})'\,(\omega)=0,\quad\text{i.e.}\quad(\widehat{g_a})'\,(\omega)+\frac1a\,\omega\,\widehat{g_a}(\omega)=0,\quad\text{avec}\quad\widehat{g_a}(0)\overset{(6.16)}{=}\frac{1}{\sqrt{a}}. \tag{6.22}$$
   Donc $\widehat{g_a}$ satisfait le même type d'équations que g_a, cf. (6.21), d'où (6.17) (avec le théorème de Cauchy–Lipschitz).
6. Soit $G(\varepsilon)=\int_{\mathbb{R}}f(t)g_\varepsilon(t)\,dt$ (intégrale qui dépend du paramètre $\varepsilon\ge0$). L'intégrant $h(\varepsilon,t)=e^{-\frac{\varepsilon t^2}{2}}f(t)$ est continu en ε pour tout ε, et est borné par $|f(t)|$ indépendamment de ε, avec $f\in\mathcal{L}^1(\mathbb{R})$: on peut passer à la limite sous le signe $\int$, et $G(\varepsilon)\longrightarrow_{\varepsilon\to0}G(0)=\int_{\mathbb{R}}f(t)\,dt$.
7. Soit $Z(\varepsilon)=\displaystyle\int_{\omega\in\mathbb{R}}f(\omega)\widehat{g_\varepsilon}(\omega)\,d\omega=\int_{\omega\in\mathbb{R}}\frac{1}{\sqrt{\varepsilon}}e^{-\frac{\omega^2}{2\varepsilon}}f(\omega)\,d\omega$. Soit $Y(\varepsilon)=\displaystyle\int_{u\in\mathbb{R}}e^{-\frac{u^2}{2}}f(\sqrt{\varepsilon}u)\,du$, défini pour tout $\varepsilon\ge0$. l'intégrant $h(\varepsilon,u)=e^{-\frac{u^2}{2}}f(\sqrt{\varepsilon}u)$ vérifie $h(\varepsilon,u)=e^{-\frac{u^2}{2}}||f||_\infty$, donc h est borné par une fonction $\mathcal{L}^1(\mathbb{R})$ indépendante de ε ; et à $u\ne0$ fixé, f étant continue dans un ouvert $]-\eta,\eta[$ (voisinage de 0) h est continu en $\varepsilon\in[0,\frac{\eta}{u^2}[$, donc h est continu en $\varepsilon=0$. Donc (thm 5.43) Y est continu en 0 et $Y(0)=\int_{u\in\mathbb{R}}e^{-\frac{u^2}{2}}f(0)\,du=\sqrt{2\pi}f(0)$. Et $Z(\varepsilon)=Y(\varepsilon)$ (changement de variable $\omega\to\sqrt{\varepsilon}\,u$ pour $\varepsilon>0$), donc $Z(\varepsilon)\longrightarrow_{\varepsilon\to0}\sqrt{2\pi}f(0)$. D'où (6.19).
8. 8.1. Soit $h(\varepsilon,\omega)=\frac{1}{\sqrt{2\pi}}\widehat{f}(\omega)\,g_\varepsilon(\omega)\,e^{it\omega}=\frac{1}{\sqrt{2\pi}}\widehat{f}(\omega)\,e^{\frac{-\varepsilon\omega^2}{2}}\,e^{it\omega}$ l'intégrant de $I_t(\varepsilon)$. h est continue en $\varepsilon\in\mathbb{R}$ (car exp l'est), et $|h(\varepsilon,\omega)|\le\frac{1}{\sqrt{2\pi}}|\widehat{f}(\omega)|$ avec $\widehat{f}\in\mathcal{L}^1(\mathbb{R})$ (domination indépendante de ε). Donc
$$\lim_{\varepsilon\to0}I_t(\varepsilon)=I_t(0)=\frac{1}{\sqrt{2\pi}}\int_{\omega\in\mathbb{R}}\widehat{f}(\omega)\,e^{+it\omega}\,d\omega=\widehat{\widehat{f}}(-t). \tag{6.23}$$
   8.2. $H_t(\varepsilon)=\frac{1}{2\pi}\int_{\omega\in\mathbb{R}}(\int_{u\in\mathbb{R}}f(u)e^{-iu\omega}\,du)\,g_\varepsilon(\omega)\,e^{it\omega}\,d\omega$. Soit $z_\varepsilon(u,\omega)=f(u)\,e^{-iu\omega}g_\varepsilon(\omega)\,e^{+it\omega}$. On a $|z_\varepsilon(u,\omega)|\le|f(u)|\,|g_\varepsilon(\omega)|$ avec $f,g_\varepsilon\in\mathcal{L}^1(\mathbb{R})$, donc $f\otimes g_\varepsilon\in\mathcal{L}^1(\mathbb{R}^2)$, donc on peut appliquer le théorème de Fubini :
$$\begin{aligned}H_t(\varepsilon)&=\frac{1}{2\pi}\int_{u\in\mathbb{R}}f(u)\Big(\int_{\omega\in\mathbb{R}}g_\varepsilon(\omega)\,e^{-i(u-t)\omega}\,d\omega\Big)\,du\\&=\frac{1}{\sqrt{2\pi}}\int_{u\in\mathbb{R}}f(u)\,\widehat{g_\varepsilon}(u-t)\,du=\frac{1}{\sqrt{2\pi}}\int_{z\in\mathbb{R}}\widehat{g_\varepsilon}(z)\,f(z+t)\,dz\overset{(6.19)}{\underset{\varepsilon\to0}{\longrightarrow}}f(t).\end{aligned} \tag{6.24}$$
   8.3. D'où, quand $f\in\mathcal{L}^1(\mathbb{R})\cap\mathcal{L}^\infty(\mathbb{R})\cap C^0(\mathbb{R})$ et $\widehat{f}\in\mathcal{L}^1(\mathbb{R})$, $\widehat{\widehat{f}}(-t)=f(t)$ $(=\frac{1}{\sqrt{2\pi}}\int_{\omega\in\mathbb{R}}\widehat{f}(\omega)e^{+i\omega t}\,dt)$.

7 Espaces $\mathcal{L}^2$ et L^2

Définition 7.1 Une fonction $f:E\to\mathbb{R}$ (ou $f:E\to\mathbb{C}$) est dite de carré intégrable (ou d'énergie finie) si :
1. f est μ-mesurable, et
2. $|f|^2$ est μ-intégrable, i.e., $\mu(|f|^2)<\infty$.

Remarque 7.2 Une fonction ne peut être de carré intégrable que si elle est μ-mesurable (par définition). Eg., la fonction 1_C, avec $C\subset I=[0,1]$, et C non μ-mesurable n'est pas de carré intégrable. Mais c'est un cas qu'on ne rencontre pas (encore) dans les applications ingénieur.

On note $\mathcal{L}^2(E,\mathcal{A},\mu;\mathbb{R})$ (ou $\mathcal{L}^2(E,\mathcal{A},\mu;\mathbb{C})$) l'ensemble des fonctions de carré intégrable, ou plus simplement $\mathcal{L}^2(E;\mathbb{R}) = \mathcal{L}^2$ (ou $\mathcal{L}^2(E;\mathbb{C}) = \mathcal{L}^2$) si le contexte est clair.

Proposition 7.3 *L'application : $(f,g) \in \mathcal{L}^2(E;\mathbb{R}) \times \mathcal{L}^2(E;\mathbb{R}) \to \mu(fg) =^{\text{noté}} (f,g)_{\mathcal{L}^2} \in \mathbb{R}$ définie un semi-produit scalaire (une forme bilinéaire symétrique positive). Et $(f,g) \in \mathcal{L}^2(E;\mathbb{C})^2 \to \mu(f\bar{g}) =^{\text{noté}} (f,g)_{\mathcal{L}^2} \in \mathbb{C}$ définie une forme sesqui-linéaire hermitienne positive (f et g à valeurs complexes.)*

On note $||f||_{\mathcal{L}^2} = \sqrt{\mu(|f|^2)}$ la semi-norme associée ; Et donc, pour tout $f,g \in \mathcal{L}^2$ (Cauchy–Schwarz),

$$|(f,g)_{\mathcal{L}^2}| \le ||f||_{\mathcal{L}^2}||g||_{\mathcal{L}^2}. \tag{7.1}$$

Notation intégrale : $|\int_E f\bar{g}\,d\mu| \le \left(\int_E |f|^2\,d\mu\right)^{\frac{1}{2}} \left(\int_E |g|^2\,d\mu\right)^{\frac{1}{2}}$.

Preuve. Immédiat. (Voir polycopié d'éléments finis pour Cauchy–Schwarz.) ∎

La semi-norme $||.||_{\mathcal{L}^2}$ induit sur $\mathcal{L}^2$ la "topologie de la convergence en moyenne quadratique".

On note $L^2 = \mathcal{L}^2/\mathcal{N}$ l'espace quotient, pour lequel la semi-norme induite $||.||_{L^2}$ est une norme : si $||\dot{f}||_{L^2} = 0$ alors $|\dot{f}|^2 = 0$ μ-p.p., donc $|\dot{f}| = 0$ μ-p.p., donc $\dot{f} \in \dot{0} = \mathcal{N}$. Et L^2 est muni du produit scalaire induit : pour $\dot{f},\dot{g} \in L^2$ avec $f \in \dot{f}$ et $g \in \dot{g}$:

$$(\dot{f},\dot{g})_{L^2} \stackrel{\text{déf}}{=} \int_E f\bar{g}\,d\mu \quad (= \mu(f\bar{g})). \tag{7.2}$$

On a alors :

Théorème 7.4 *(Fischer–Riesz) L'espace vectoriel L^2 muni de la norme $||.||_2$ est complet. Muni du produit scalaire $(\cdot,\cdot)_{L^2}$ c'est donc un espace de Hilbert.*

Preuve. Voir paragraphe 8.5. ∎

8 Espaces $\mathcal{L}^p(I)$ et $L^p(I)$

8.1 Définition

Définition 8.1 Pour $1 \le p < \infty$, on appelle $(\mathcal{L}^p(E),\mu) =^{\text{noté}} \mathcal{L}^p(E)$ l'espace des fonctions f mesurables sur E telles que $|f|^p$ soit intégrable :

$$\begin{aligned} \mathcal{L}^p(E) &= \{f : E \to \mathbb{R}\ \mu\text{-mesurable t.q. } |f|^p \in \mathcal{L}^1(E)\} \\ &\stackrel{\text{noté}}{=} \{f : E \to \mathbb{R}\ \mu\text{-mesurable t.q. } \int_E |f|^p d\mu < \infty\}. \end{aligned} \tag{8.1}$$

Définition 8.2 Et pour $p = \infty$, on note $(\mathcal{L}^\infty(E),\mu) =^{\text{noté}} \mathcal{L}^\infty(E)$ l'ensemble des fonctions mesurables et essentiellement bornées (bornées à un ensemble négligeable près, cf. (3.31)) :

$$\begin{aligned} \mathcal{L}^\infty(E) &= \{f : E \to \mathbb{R}\ \mu\text{-mesurable t.q. } \operatorname*{supess}_{x\in E} |f(x)| < \infty\} \\ &\stackrel{\text{noté}}{=} \{f : E \to \mathbb{R}\ \mu\text{-mesurable t.q. } \sup_{x\in E} |f(x)| < \infty\}. \end{aligned} \tag{8.2}$$

Notation. Pour $f \in \mathcal{L}^p(E)$ on pose :

$$||f||_p \stackrel{\text{déf}}{=} \left(\int_E |f(x)|^p\,dx\right)^{\frac{1}{p}} \quad (= ||\,|f|^p||_1^{\frac{1}{p}}). \tag{8.3}$$

L'inégalité de Minkowski montrera que $||.||_p$ définit une semi-norme sur $\mathcal{L}^p(E)$.

Notation. Et pour $f \in \mathcal{L}^\infty(E)$ on note :

$$||f||_\infty = \operatorname*{supess}_{x\in E} |f(x)| \quad (\stackrel{\text{noté}}{=} \sup_{x\in E} |f(x)|), \tag{8.4}$$

C'est trivialement une semi-norme sur $\mathcal{L}^\infty(E)$.

Enfin, pour $p \ge 1$, on note $L^p(E)$ l'ensemble $\mathcal{L}^p(E)/\mathcal{N}$ des classes de fonctions de $\mathcal{L}^p(E)$. Donc pour $\dot{f} \in L^p(E)$ et $f \in \dot{f}$, $||\dot{f}||_p := ||f||_p$ défini une norme sur $\mathcal{L}^p(E)$.

8.2 Inégalité d'Young

Définition 8.3 Si $p \geq 1$ et $q \geq 1$ sont des réels tels que :

$$\frac{1}{p} + \frac{1}{q} = 1, \tag{8.5}$$

on dit que p et q sont conjugués, ou que q est l'exposant conjugué de p. On dit également que p et q sont conjugués lorsque $p = 1$ (resp. q) et $q = \infty$ (resp. p).

Ainsi p et q sont conjugués ssi $\lambda_1 = \frac{1}{p}$ et $\lambda_2 = \frac{1}{q}$ sont des coordonnées barycentriques positives.

Donc p et q sont conjugués ssi $p \geq 1$ et $q \geq 1$, et dans le cas $p > 1$ et $q > 1$:

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \Leftrightarrow \quad p + q = pq \quad \Leftrightarrow \quad q = \frac{p}{p-1} \quad \Leftrightarrow \quad p = \frac{q}{q-1} \quad \Leftrightarrow \quad (p-1)(q-1) = 1 \tag{8.6}$$

Proposition 8.4 *(Inégalité d'Young) Étant donnés deux réels $a \geq 0$ et $b \geq 0$, si $p > 1$ et $q > 1$ sont conjugués (i.e. $\frac{1}{p} + \frac{1}{q} = 1$) alors :*

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q \tag{8.7}$$

C'est une généralisation du cas $p = q = 2$ (isobarycentre) où $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$.

Preuve. $\frac{1}{p} + \frac{1}{q} = 1$ et donc $\frac{1}{p}a^p + \frac{1}{q}b^q \in [a^p, b^q]$ est barycentre de a^p et b^q. La fonction Log étant concave on en déduit :

$$\mathrm{Log}(\frac{1}{p}a^p + \frac{1}{q}b^q) \geq \frac{1}{p}\mathrm{Log}(a^p) + \frac{1}{q}\mathrm{Log}(b^q) = \mathrm{Log}(ab) \tag{8.8}$$

Et la fonction Log étant croissante, on en déduit le résultat. ∎

8.3 Inégalité d'Hölder

Proposition 8.5 *(Inégalité d'Hölder) Si $f \in \mathcal{L}^p(E)$ et $g \in \mathcal{L}^q(E)$ où p et q sont conjugués, alors $fg \in \mathcal{L}^1(E)$ (le produit est intégrable), et :*

$$||fg||_1 \leq ||f||_p \, ||g||_q. \tag{8.9}$$

En particulier, pour $p = q = 2$, si f et g sont dans $\mathcal{L}^2(E)$, alors $fg \in \mathcal{L}^1(E)$ et on retrouve l'inégalité de Cauchy–Schwarz $||fg||_1 \leq ||f||_2 \, ||g||_2$.

Preuve. Si $p = 1$ et $q = \infty$ (resp. $p = \infty$ et $q = 1$) c'est l'inégalité de la moyenne (4.7).

Sinon, avec l'inégalité d'Young on a, pour tout $x \in E$:

$$|f(x)| \, |g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q. \tag{8.10}$$

f et g étant mesurables, on a fg et $|fg|$ mesurables, avec $|fg| = |f| \, |g|$ positive, d'où :

$$\int_E |f(x)| \, |g(x)| \, dx \leq \frac{1}{p} \int_E |f(x)|^p \, dx + \frac{1}{q} \int_E |g(x)|^q \, dx = \frac{1}{p}||f||_p^p + \frac{1}{q}||g||_q^q < \infty, \tag{8.11}$$

et le produit fg est bien dans $\mathcal{L}^1(\mathbb{R})$.

Et remplaçant f par λf pour $\lambda > 0$, on a :

$$\lambda \int_E |f(x)| \, |g(x)| \, dx \leq \frac{\lambda^p}{p}||f||_p^p + \frac{1}{q}||g||_q^q, \tag{8.12}$$

soit :

$$\int_E |f(x)| \, |g(x)| \, dx \leq \frac{\lambda^{p-1}}{p}||f||_p^p + \frac{1}{q\lambda}||g||_q^q. \tag{8.13}$$

On prend alors le λ qui réalise le minimum du membre de droite :

$$\lambda_{min} = \frac{1}{||f||_p}||g||_q^{\frac{q}{p}} = \frac{1}{||f||_p}||g||_q^{q-1}, \tag{8.14}$$

puisque $\frac{1}{p} + \frac{1}{q} = 1$ donne $\frac{q}{p} = q - 1$. D'où avec cette valeur $\lambda = \lambda_{min}$:

$$||fg||_1 = \int_E |f(x)| \, |g(x)| \, dx \leq ||f||_p ||g||_q (\frac{1}{p} + \frac{1}{q}) = ||f||_p ||g||_q, \tag{8.15}$$

ce qu'il fallait démontrer. ∎

8.4 Inégalité de Minkowski

Proposition 8.6 *(Inégalité de Minkowski) Si $p \geq 1$ et si $f, g \in \mathcal{L}^p(E)$, alors $f+g \in \mathcal{L}^p(E)$ et :*

$$||f+g||_p \leq ||f||_p + ||g||_p, \tag{8.16}$$

et $||.||_p$ définit une norme sur $L^p(E)$.

Preuve. Pour $p=1$ c'est trivial, pour $p=2$ c'est une application de l'inégalité de Cauchy–Schwarz (avec $||f+g||_2^2 \leq (||f||_2 + ||g||_2)^2$), pour $p=\infty$ c'est trivial. Il reste à montrer que c'est encore vrai pour tout les cas intermédiaires :

Soit $p \in]1, \infty[$. On a pour tout $x \in E$:

$$|f(x)+g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p), \tag{8.17}$$

puisque $|a+b|^p \leq (|a|+|b|)^p \leq 2^p \max(|a|^p, |b|^p) \leq 2^p(|a|^p + |b|^p)$ pour a et b réels. D'où $(|f+g|^p \in \mathcal{L}^1(\mathbb{R})$ ou encore $f+g \in \mathcal{L}^p(E)$.

Puis, pour tout $x \in E$:

$$|f(x)+g(x)|^p \leq |f(x)+g(x)|^{p-1}(|f(x)| + |g(x)|), \tag{8.18}$$

d'où, en introduisant l'exposant conjugué q de p, et sachant $p-1 = \frac{p}{q}$:

$$\int_E |f(x)+g(x)|^p \, dx \leq \int_E |f(x)+g(x)|^{\frac{p}{q}} \, |f(x)| \, dx + \int_E |f(x)+g(x)|^{\frac{p}{q}} \, |g(x)| \, dx. \tag{8.19}$$

On a $|f+g|^{\frac{p}{q}} \in \mathcal{L}^q(E)$ puisque $|f+g|^p \in \mathcal{L}^1(\mathbb{R})$, et on a $f \in \mathcal{L}^p(E)$ par hypothèse. On en déduit avec l'inégalité d'Hölder (8.9) :

$$||f+g||_p^p \leq || \, |f+g|^{\frac{p}{q}} \, ||_q ||f||_p + || \, |f+g|^{\frac{p}{q}} \, ||_q ||g||_p. \tag{8.20}$$

Et :

$$|| \, |f+g|^{\frac{p}{q}} \, ||_q = \Big(\int_E |f+g|^p \, dx \Big)^{\frac{1}{q}} = \Big(\int_E |f+g|^p \, dx \Big)^{\frac{p-1}{p}} = ||f+g||_p^{p-1}, \tag{8.21}$$

donc :

$$||f+g||_p^p \leq ||f+g||_p^{p-1}(||f||_p + ||g||_p) \tag{8.22}$$

d'où l'inégalité de Minkowski (8.16). ∎

C'est l'inégalité triangulaire pour la fonction $||.||_p$, qui définit donc une semi-norme sur $\mathcal{L}^p$ et donc une norme sur $L^p = \mathcal{L}^p / \mathcal{N}$. On admet alors :

Théorème 8.7 *(Théorème de Fischer–Riesz) Pour $1 \leq p \leq \infty$, l'espace vectoriel L^p muni de la norme $||.||_p$ est complet : c'est donc un espace de Banach.*

Preuve. Voir paragraphe 8.5. ∎

8.5 * Complément : complétude et séparabilité

Théorème 8.8 *(Fischer–Riesz) Cas $\Omega \subset \mathbb{R}^n$, Ω ouvert, et μ la mesure de Lebesgue. L'espace $(L^1(\Omega), ||.||_1)$ est un espace vectoriel normé qui est complet : c'est un espace de Banach.*

D'où, pour $1 \leq p < \infty$, l'espace $(L^p(\Omega), ||.||_{L^p})$ est un Banach.

Preuve. Soit $(\dot{f}_n)_{\mathbb{N}^*}$ une suite de Cauchy dans L^1. Prenant $f_n \in \dot{f}_n$, la suite $(f_n)_{\mathbb{N}^*}$ une suite de Cauchy dans $\mathcal{L}^1$. Quitte à extraire une sous-suite, on peut supposer que :

$$||f_{k+1} - f_k||_{\mathcal{L}^1} \leq \frac{1}{2^k}, \quad \forall k \geq 1. \tag{8.23}$$

Montrons que (f_k) converge dans $\mathcal{L}^1$ p.p.. On pose, pour $n \geq 1$:

$$g_n(x) = \sum_{k=1}^{n} |f_{k+1}(x) - f_k(x)| \quad \text{p.p..}$$

La suite $(g_n)_{\mathbb{N}^*}$ est positive croissante (immédiat) et bornée dans $\mathcal{L}^1$ car, avec (8.23) :

$$||g_n||_{\mathcal{L}^1} \leq 1, \quad \forall n \geq 1.$$

Et le théorème de convergence monotone implique que $(g_n)_{\mathbb{N}^*}$ est convergente vers une fonction $g \in \mathcal{L}^1$. Avec

$m \geq n \geq 2$ on obtient, pour presque tout x :

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f_{m-1}(x)| + ... + |f_{n+1}(x) - f_n(x)| \leq g_m(x) - g_{n-1}(x) \leq g(x) - g_{n-1}(x),$$

et donc $(f_n(x))_n$ est une suite de Cauchy dans $\mathbb{R}$, ce pour presque tout x. Donc $(f_n(x))$ est convergente vers un réel qu'on note $f(x)$, et on a (g_{n-1} étant positive) :

$$|f(x) - f_n(x)| \leq g(x) \quad p.p., \quad \forall n \geq 2,$$

d'où $f \in \mathcal{L}^1$ (car $|f| \leq g + |f_2|$ p.p.). Il reste à montrer que $||f - f_n||_{\mathcal{L}^1} \to 0$. On applique le théorème de convergence dominée : on a $f_n - f \to 0$ p.p. (définition de f) et $|f - f_n| \leq g$ p.p. avec $g \in L^1$ (majorante indépendante de n), d'où $||f - f_n||_{\mathcal{L}^1} \to ||0||_{\mathcal{L}^1} = 0$. D'où $||\dot{f} - \dot{f}_n||_{L^1} \to 0$ où $\dot{f}$ est la classe de f. D'où L^1 est complet.

Pour $1 \leq p < \infty$, se servir de l'inégalité de Minkowski, voir (8.16), et se ramener au cas L^1 : dans la démonstration précédente remplacer $||.||_{\mathcal{L}^1}$ par $||.||_{\mathcal{L}^p}$. ∎

Théorème 8.9 *Pour $1 \leq p < \infty$, l'espace $(\mathcal{L}^p(\Omega), ||.||_p)$ est séparable.*

Preuve. Il s'agit de montrer qu'il existe une famille dénombrable dense dans $\mathcal{L}^p(\Omega)$. Soit $(Q_i)_{i\in I}$ la famille dénombrable des pavés de la forme $Q = \Pi_{k=1}^n]a_k, b_k[\subset \Omega$ où $a_k, b_k \in \mathbb{Q}$. Ces pavés forment une base de voisinage pour la topologie usuelle. Soit E l'espace des fonctions en escalier engendrées par les combinaisons linéaires à coefficients rationnels des 1_{Q_i} : E est donc dénombrable. Montrons que E est dense dans $\mathcal{L}^p$. Soit $f \in \mathcal{L}^p$ et soit $\varepsilon > 0$. L'espace C_c^0 des fonctions C^0 à support compact étant dense dans $\mathcal{L}^p$, voir cours de distributions (régularisation par convolution), soit $\varphi \in C_c^0$ t.q. $||f - \varphi||_{\mathcal{L}^p} < \varepsilon$. Soit $K = \text{supp}\varphi =$ le support de φ, et soit $\Omega' \subset \Omega$ tel que Ω' est un ouvert borné $\supset K$ (si Ω est borné, on peut prendre $\Omega' = \Omega$). En particulier φ est uniformément continue, et, K étant compact, il existe $\bigcup_{i=1}^N Q_i \supset K$ un recouvrement fini, où les Q_i appartiennent à la famille dénombrable $(Q_i)_{i\in I}$, et où $\bigcup_{i=1}^N Q_i \subset \Omega'$, et pour tout $i = 1, ..., N$ et pour tout $x, y \in Q_i$ on a $|\varphi(x) - \varphi(y)| \leq \frac{\varepsilon}{|\Omega'|^p}$ (où $|\Omega'|$ est le volume de Ω'). On choisit alors $g \in E$, $g = \sum_{i=1}^N \alpha_i 1_{Q_i}$ où $\alpha_i = \varphi(x_i)$ pour un $x_i \in Q_i$. Et on a $|g(x) - \varphi(x)| \leq \frac{\varepsilon}{|\Omega'|^p}$, d'où $||g - \varphi||_{\mathcal{L}^p} \leq \varepsilon$, d'où $||f - g||_{\mathcal{L}^p} \leq 2\varepsilon$. ∎

Théorème 8.10 *L'espace $(\mathcal{L}^\infty, ||.||_\infty)$ est un Banach.*

Preuve. Soit $(\dot{f}_n)_{\mathbb{N}^*}$ une suite de Cauchy dans L^∞, et soit $f_n \in \dot{f}_n$. Donc $(f_n)_{\mathbb{N}^*}$ une suite de Cauchy dans $\mathcal{L}^\infty$: $||f_n - f_m||_\infty \longrightarrow_{n,m\to\infty} 0$. Donc

$\forall k \in \mathbb{N}$, $\exists N_k \in \mathbb{N}^*$, $\forall m, n \geq N_k$, $||f_n - f_m||_\infty < \frac{1}{k}$, soit

$\forall k \in \mathbb{N}$, $\exists N_k \in \mathbb{N}^*$, $\forall m, n \geq N_k$, $\exists E_k \in \mathcal{A}_E$ t.q. $\mu(E_k) = 0$, $\forall x \in E - E_k$, $|f_n(x) - f_m(x)| < \frac{1}{k}$.

On pose $F = \bigcup_{k\in\mathbb{N}^*} E_k$ union dénombrable d'ensembles négligeables, donc $\mu(F) = 0$, et pour tout $x \in E - F$, la suite $(f_n(x))_{\mathbb{N}^*}$ est de Cauchy dans $\mathbb{R}$, donc converge vers un réel qu'on note $f(x)$. Et $|f(x) - f_{N_k}(x)| \leq \frac{1}{k}$ pour tout $x \in E - F$. D'où $f \in \mathcal{L}^\infty$ et $||f - f_{N_k}||_\infty \leq \frac{1}{k} \longrightarrow_{k\to\infty} 0$ p.p.. D'où $(\dot{f}_n)_{\mathbb{N}^*}$ est convergente dans L^∞ vers $\dot{f}$ la classe de f. ∎

Théorème 8.11 *L'espace $\mathcal{L}^\infty$ n'est pas séparable.*

Preuve. On considère la famille non dénombrable des ouverts pour $a \in \Omega$:

$$O_a = \{f \in L^\infty : ||f - 1_{B(a,r_a)}||_{L^\infty} < \frac{1}{2}\},$$

où $r_a = d(a, \mathbb{R}^n - \Omega)$ est la plus petite distance de a au bord de Ω. On a $O_a \bigcap O_b = \emptyset$ quand $a \neq b$. On en déduit que $\mathcal{L}^\infty(\Omega)$ n'est pas séparable à l'aide du lemme suivant : ∎

Lemme 8.12 *Soit E un espace de Banach. On suppose qu'il existe une famille non dénombrable $(O_i)_{i\in I}$ d'ouverts non vides telle que $O_i \bigcap O_j = \emptyset$ si $i \neq j$. Alors E n'est pas séparable.*

Preuve. Par l'absurde : sinon, il existerait une suite $(u_n)_\mathbb{N}$ dense dans E, et on aurait donc : $\forall i \in I$, $\exists u_i$ t.q. $u_i \in O_i$. Et l'application $i \to u_i$ est injective (trivial) et donc card$I \leq$ card$\mathbb{N}$, et donc I est au plus dénombrable. Absurde. ∎

8.6 * Fonctions Lebesgue-mesurables et fonctions en escalier

Cadre : $(\mathbb{R}^n, \mathcal{A}_{\mathbb{R}^n}, \mu_\ell)$ l'espace $\mathbb{R}^n$ muni de sa tribu borélienne et de la mesure de Lebesgue. Pour alléger la présentation on prend n=1. Dans $\mathbb{R}$ on notera (a,b) un intervalle quelconque (ouvert, fermé, ou semi-ouvert...).

Le but ici est de montrer que toute fonction $(L^1(\mathbb{R}^n), \mathcal{A}_{\mathbb{R}^n}, \mu_\ell)$ est presque partout limite d'une suite de fonction en escaliers (combinaison linéaire de fonctions caractéristiques d'intervalles).

On va utiliser une "presque mesure", la mesure extérieure (qui permet également de prouver l'existence de la mesure de Lebesgue sur $\mathcal{A}_{\mathbb{R}^n}$.)

8.6.1 Mesure extérieure

Soit $(E, \mathcal{O}_E)$ un ensemble topologique. Pour $F \subset E$ on note F^o l'intérieur de F.

Définition 8.13 soit I un ensemble quelconque, et soit $(R_i)_{i\in I} \subset \mathcal{A}_{\mathbb{R}}^I$ une famille d'ensembles. L'union $R = \bigcup_{i\in I} R_i$ est un recouvrement de E ssi $R \supset E$, i.e. ssi $\forall x \in E$, $\exists i \in I$, $x \in R_i$.
Le recouvrement est fini ou dénombrable ssi I est fini ou dénombrable.
Le recouvrement est ouvert (resp. fermé) si tous les R_i sont ouverts (resp. fermés), dans le cas sous-entendu où E est muni d'une topologie.

Dans $(\mathbb{R}, \mathcal{O}_{\mathbb{R}})$: tout ensemble $E \subset \mathbb{R}$ admet un recouvrement dénombrable par des intervalles ouverts ou fermés t.q. $0 < \mu_\ell(R_i) < \infty$. En effet $E \subset \mathbb{R} = \bigcup_{\mathbb{N}}]-n, n[= \bigcup_{\mathbb{Z}} [-n, n]$.

Remarque : le cas $E = \bigcup_{i=1}^n (a_i, b_i)$, union finie d'intervalles (a_i, b_i) quelconques, donne l'intégrale de Riemann. On généralise ce cas pour retrouver l'intégrale de Lebesgue.

Définition 8.14 Soit $E \subset \mathbb{R}$. Soit $\mathcal{R}(E)$ l'ensemble des recouvrements ouverts dénombrables de E par des intervalles ouverts. Donc $\bigcup_{i\in\mathbb{N}^*} U_i \in \mathcal{R}(E)$ ssi les U_i sont tous des intervalles ouverts et $E \subset \bigcup_{i\in\mathbb{N}^*} U_i$. On note :

$$\mu^*(E) \stackrel{\text{déf}}{=} \inf_{\bigcup_{i\in\mathbb{N}^*} U_i \in \mathcal{R}(E)} \sum_{i=1}^{\infty} \mu_\ell(U_i) \quad \in \overline{\mathbb{R}_+}. \tag{8.24}$$

Ayant $\mu_\ell(\bigcup_{i\in\mathbb{N}^*} U_i) \le \sum_{i\in\mathbb{N}^*} \mu_\ell(U_i) \in \overline{\mathbb{R}_+}$, le réel $\mu^*(E)$ existe dans $\overline{\mathbb{R}_+}$: borne inférieure d'un sous-ensemble de $\overline{\mathbb{R}_+}$.

Définition 8.15 $\mu^*(E)$ est appelée la mesure extérieure de E.
(Ce n'est malheureusement pas une mesure : elle n'est pas additive, voir lemme 8.18 suivant.)

Exemple 8.16 Si $x \in \mathbb{R}$ alors $\mu^*(\{x\}) = 0$, car $\{x\} \le]x-\varepsilon, x+\varepsilon[$ donne $\mu^*(\{x\}) \le 2\varepsilon$ pour tout $\varepsilon > 0$. ▪

Exercice 8.17 Montrer : dans (8.24) on peut remplacer $\mathcal{R}(E)$ par l'ensemble des recouvrements par des intervalles fermés (ou quelconques).

Réponse. Notons $\mathcal{R}_F(E)$ l'ensemble des recouvrements dénombrables de E par des intervalles fermés.
1- Adoptons la définition (8.24). Si $\bigcup_{i\in\mathbb{N}^*} U_i \supset \mathcal{R}(E)$, i.e. si $E \subset \bigcup_{i\in\mathbb{N}^*} U_i$ où $U_i =]a_i, b_i[$, alors $E \subset \bigcup_{i\in\mathbb{N}^*} \overline{U_i}$ où $\overline{U_i} = [a_i, b_i]$, donc $\bigcup_{i\in\mathbb{N}^*} \overline{U_i} \in \mathcal{R}_F(E)$. Et $\mu_\ell(U_i) = \mu_\ell(\overline{U_i}) = b_i - a_i$, donc $\mu^*(E) = \inf_{\bigcup_{i\in\mathbb{N}^*} F_i \in \mathcal{R}_F(E)} \sum_{i=1}^\infty \mu_\ell(F_i)$.
2- Adoptons la définition : $\mu^*(E) \stackrel{\text{déf}}{=} \inf_{\bigcup_{i\in\mathbb{N}^*} F_i \in \mathcal{R}_F(E)} \sum_{i=1}^\infty \mu_\ell(F_i)$. Donc $\forall \varepsilon > 0$, $\exists F_i = [a_i, b_i]$ pour $i \in \mathbb{N}^*$ t.q. $\mu^*(E) \le \sum_{i=1}^\infty \mu_\ell(F_i) + \varepsilon$. Notons $U_i = F_i +]-\frac{\varepsilon}{2^{i+1}}, \frac{\varepsilon}{2^{i+1}}[=]a_i - \frac{\varepsilon}{2^{i+1}}, b_i + \frac{\varepsilon}{2^{i+1}}[$. On a $\mu_\ell(U_i) = \mu_\ell(F_i) + \frac{\varepsilon}{2^i}$, donc $E \subset \bigcup_{i\in\mathbb{N}^*} U_i$ avec $\mu^*(E) \le \sum_{i=1}^\infty \mu_\ell(U_i) + \varepsilon = 2\varepsilon$. Vrai pour tout ε, donc $\mu^*(E) = \inf_{\bigcup_{i\in\mathbb{N}^*} U_i \in \mathcal{R}(E)} \sum_{i=1}^\infty \mu_\ell(U_i)$. ▪

8.6.2 Sous-additivité

Lemme 8.18 *La fonction $\mu^* : \mathcal{P}(\mathbb{R}) \to \overline{\mathbb{R}_+}$ est telle que :*
1- $\mu^*(\emptyset) = 0$.
2- Pour $c \in \mathbb{R}$ et $E \subset \mathbb{R}$, le translaté $c + E$ vérifie $\mu^*(c + E) = \mu^*(E)$.
3- Si $E \subset F$, alors $\mu^*(E) \le \mu^*(F)$ (monotonie).
4- Propriété de sous-additivité dénombrable : si $(E_i)_{i\in\mathbb{N}^*} \in \mathcal{P}(\mathbb{R})^{\mathbb{N}^*}$, et si $E = \bigcup_{i\in\mathbb{N}^*} E_i$, alors :

$$\mu^*(\bigcup_{i=1}^\infty E_i) \le \sum_{i=1}^\infty \mu^*(E_i). \tag{8.25}$$

5- Mais μ^* ne vérifie pas la propriété de σ-additivité (additivité dénombrable), cf (1.1) : elle ne vérifie même pas la propriété d'additivité finie :

$$\exists N \in \mathbb{N}^*,\ \exists (E_i)_{i=1,\dots,N} \in E^N, \quad \text{t.q.} \quad \forall i \ne j,\ E_i \bigcap E_j = \emptyset \quad \text{et} \quad \mu^*(\bigcup_{i=1}^N E_i) < \sum_{i=1}^N \mu^*(E_i). \tag{8.26}$$

Ce n'est donc pas une mesure. (L'ensemble $\mathcal{P}(\mathbb{R})$ est trop grand : il faudra se contenter d'un sous-ensemble de $\mathcal{P}(\mathbb{R})$ pour définir une mesure, à savoir la tribu borélienne étendue $\mathcal{A}_\mu$.)

Preuve. 1- On a $\emptyset \subset]0,\varepsilon[$ pour tout ε, donc $\mu^*(\emptyset) \le \mu_\ell(]0,\varepsilon[) = \varepsilon$, donc $\mu^*(\emptyset) = 0$.

2- Si U_i est un intervalle, alors $\mu_\ell(U_i) = \mu_\ell(c+U_i)$, et $\mu_\ell(U) = \mu_\ell(c+U)$, donc avec les notations de (8.24), $\mu^*(c+E) = \mu^*(E)$.

3- Si $E \subset F$ et si U recouvre F alors U recouvre F, donc $\mu^*(E) \le \mu^*(F)$, cf. (8.24).

4- Si un des E_i vérifie $\mu^*(E_i) = \infty$, alors (8.25) est trivial grâce à la monotonie. Supposons $\mu^*(E_i) < \infty$ pour tout i. Avec (8.24), on a : $\forall i \in \mathbb{N}^*$, $\forall \varepsilon > 0$, $\exists (\bigcup_{\mathbb{N}^*} U_i^j) \in \mathcal{R}(E_i)$ t.q. :

$$\sum_{j=1}^{\infty} \mu_\ell(U_i^j) \le \mu^*(E_i) + \frac{\varepsilon}{2^i}.$$

Comme $E = \bigcup_{\mathbb{N}^*} E_i$, on a $E \subset \bigcup_{i,j\in\mathbb{N}^*} U_i^j$, recouvrement dénombrable par des ouverts. Donc

$$\mu^*(E) \le \sum_{i,j=1}^{\infty} \mu_\ell(U_j^i) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \mu_\ell(U_j^i) \le \sum_{i=1}^{\infty} \mu^*(E_i) + \sum_{i=1}^{\infty} + \frac{\varepsilon}{2^i} = \sum_{i=1}^{\infty} \mu^*(E_i) + \varepsilon.$$

(Fubini car tous les termes de la série sont positifs.) Vrai pour tout ε, d'où (8.25).

5- On reprend la remarque 1.37 : relation d'équivalence $x\mathcal{R}y$ ssi $y-x \in \mathbb{Q}$, et $\dot{x}$ classe d'équivalence de x. L'axiome du choix permet de créer le sous-ensemble A de $\mathbb{R}$ en prenant un et un seul élément de chaque classe. Quitte à choisir un élément dans la classe on le prend dans $[0,1[$ (possible car si $x \in A$, alors $\exists k \in \mathbb{Z} \subset \mathbb{Q}$ t.q. $x \in [k, k+1[$), notons B l'ensemble obtenu, sous-ensemble de $[0,1[$. Donc $\mu^*(B) \le 1$. Et on a $\mathbb{R} = \bigcup_{q\in\mathbb{Q}} q + B$, cf. remarque 1.37. Donc ; avec 4- et 2- :

$$\infty = \mu^*(\mathbb{R}) \le \sum_{q\in\mathbb{Q}} \mu^*(q+B) = \sum_{q\in\mathbb{Q}} \mu^*(B), \quad \text{donc} \quad \mu^*(B) > 0.$$

Et l'ensemble $P = \bigcup_{q\in[0,1]_\mathbb{Q}} q + B$ est une partition dénombrable de $[0,2[$, reprendre la démarche de la remarque 1.37. Avec $P \subset [0,2[$, donc $\mu^*(P) \le 2$.

Soit le réel $r = \mu^*(B) > 0$. Soit $N > \frac{2}{r}$. Soit les rationnels $q_i = \frac{1}{N}$. Si on avait l'additivité, on aurait $\sum_{i=1}^N \mu(q_i + B) = \sum_{i=1}^N \mu(B)$, cf. 2-, donc avec 4- on aurait $2 \ge \mu(P) \ge \sum_{i=1}^N r = Nr > 2$: absurde. ∴

Exemple 8.19 $\mu^*(\mathbb{R}) = \infty$ car $\mathbb{R} \supset]-n,n[$ donne (sous-additivité) $\mu^*(\mathbb{R}) \ge \mu^*(]-n,n[) = 2n$, pour tout n. ∴

Exemple 8.20 Soit $f : \mathbb{N}^* \to \mathbb{Q}$ bijection ($\mathbb{Q}$ est dénombrable), donc $\mathbb{Q} = \bigcup_{\mathbb{N}^*}\{q_i\}$. Donc $\mathbb{Q} \subset \bigcup_{\mathbb{N}^*}\{q_i\}$. Donc $\mu^*(\mathbb{Q}) \le \sum_{i=1}^{\infty} \mu^*(\{q_i\}) = 0$, donc $\mu^*(\mathbb{Q}) = 0$. ∴

Exemple 8.21 $\mu^*(\mathbb{R} - \mathbb{Q}) = \infty$. Sinon $\mu^*(\mathbb{R}) = \mu^*(\mathbb{Q} + (\mathbb{R}-\mathbb{Q})) \le \mu^*(\mathbb{Q}) + \mu^*(\mathbb{R}-\mathbb{Q})$, cf. (8.25), donnerait $\mu^*(\mathbb{R}) < \infty$. ∴

Exemple 8.22 Soit C l'ensemble de Cantor, cf. (1.4.2). C est non dénombrable et $\mu^*(C) = 0$ car $\mu^*(C) \le \mu^*(A_i) = \mu_\ell(A_i) = (\frac{2}{3})^n$ pour tout n. ∴

8.6.3 Outer Regularity

Lemme 8.23 *(Outer Regularity). 1- Si $E \subset \mathbb{R}$ alors :*

$$\mu^*(E) = \inf_{U\in\mathcal{O}_\mathbb{R},\, U\supset E} \mu^*(U), \tag{8.27}$$

où $\mathcal{O}_\mathbb{R}$ est l'ensemble des ouverts (topologie usuelle de $\mathbb{R}$).

2- Si U est un ouvert alors $\mu^*(U) = \mu_\ell(U) = \mu^*(\overline{U}) = \mu_\ell(\overline{U})$. Et donc :

$$\mu^*(E) = \inf_{U\in\mathcal{O}_\mathbb{R},\, U\supset E} \mu_\ell(U), \tag{8.28}$$

Preuve. 1- Il s'agit de montrer : $\forall\varepsilon > 0$, $\exists U \supset E$, U ouvert, $\mu^*(E) \le \mu^*(U) + \varepsilon$.

(8.24) donne : pour $\varepsilon > 0$, $\exists (U_i)_{\mathbb{N}^*}$ famille d'intervalles ouverts t.q. $\bigcup_{i\in\mathbb{N}^*} U_i \supset E$ et $\sum_{i=1}^{\infty} \mu_\ell(U_i) \le \mu^*(E) + \varepsilon$. On pose $U = \bigcup_{\mathbb{N}^*} U_i$, et U est ouvert car réunion d'ouverts.

2- Puis soit U un ouvert. Donc $U = \bigcup_{\mathbb{N}^*}[c_i, d_i]$ où les $([c_i,d_i])_{\mathbb{N}^*}$ est une famille presque disjointe d'intervalles fermés, cf. exercice 0.20. Et c'est le plus petit recouvrement fermé, donc $\mu^*(U) = \sum_{\mathbb{N}^*} \mu_\ell([c_i,d_i]) = \mu_\ell(U)$, cf. (8.24) et exercice 8.17. ∴

Remarque 8.24 On montre qu'un borélien (un $A \in \mathcal{A}_E$ où $\mathcal{A}_E$ est la tribu borélienne d'un espace topologique $(E, \mathcal{O}_E)$) de mesure borélienne finie est de la forme $A = F \bigcup N$ où N négligeable et $F \in \mathcal{F}_\sigma = \{$unions dénombrables de fermés$\}$. Ou encore de la forme $A = G - N'$ où N' négligeable et et $\mathcal{G}_\delta = \{$intersections dénombrables d'ouverts$\}$.

Et pour qu'un $E \in \mathcal{P}(\mathbb{R})$ soit dans $\mathcal{A}_\mathbb{R}$ il faudrait de plus que

$\mu^*(E) = \sup\{\mu_\ell(K) : K \subset E,\ K \text{ compact}\}$.

Et alors $\mu^* = \mu_\ell$ sur $\mathcal{A}_\mathbb{R}$. (Voir Villani [18] par exemple.) ∴

8.6.4 Littlewood First Principle

(Voir Muthukumar [10] par exemple.)

Lemme 8.25 *(Littlewood First Principle : un ensemble mesurable de mesure finie est "presque" une union finie d'intervalles.) Soit $A \in \mathcal{A}_\mathbb{R}$ (un ensemble Lebesgue mesurable dans $\mathbb{R}$) tel que $\mu_\ell(A) < \infty$. Alors pour tout $\varepsilon > 0$, il existe $n_\varepsilon \in \mathbb{N}^*$, il existe une famille finie $(F_i)_{i=1,...,n_\varepsilon}$ d'intervalles fermés t.q., notant $F = \bigcup_{i=1}^{n_\varepsilon} F_i$, on a $\mu_\ell(A\Delta F) \le \varepsilon$.*

Preuve. On applique (8.28) et l'exercice 0.20. Soit $A \in \mathcal{A}_\mathbb{R}$. Soit $\varepsilon > 0$. On a : $\exists (F_i)_{\mathbb{N}^*}$ suite d'intervalles fermés t.q. $A \subset \bigcup_{\mathbb{N}^*} F_i$ et $\sum_{i=1}^\infty \mu_\ell(F_i) \le \mu(A) + \varepsilon$. Donc la série $\sum_{i=1}^\infty \mu_\ell(F_i)$ est convergente : $\exists n_\varepsilon$ t.q. $\sum_{i>n_\varepsilon} \mu_\ell(F_i) < \varepsilon$. On pose $F = \bigcup_{i\le n_\varepsilon} F_i$, union finie d'intervalles fermés.

On a $A - F \subset \bigcup_{\mathbb{N}^*} F_i - \bigcup_{i\le n_\varepsilon} F_i \subset \bigcup_{i>n_\varepsilon} F_i$. D'où :

$$\mu(A-F) \le \mu(\bigcup_{i>n_\varepsilon} F_i) \le \sum_{i>n_\varepsilon} \mu(F_i) < \varepsilon.$$

On a $F - A \subset \bigcup_{\mathbb{N}^*} F_i - A$, avec $A \subset \bigcup_{\mathbb{N}^*} F_i$, donc $\mu(A) + \mu(A - \bigcup_{\mathbb{N}^*} F_i) = \mu(A)$, d'où :

$$\mu(F-A) \le \mu(\bigcup_{i=1}^\infty F_i - A) \le \mu(\bigcup_{i=1}^\infty F_i) - \mu(A) \le \sum_{i=1}^\infty \mu(F_i) - \mu(A) \le \varepsilon,$$

D'où $\mu(A \,\Delta\, U) \le 2\varepsilon$. ∎

Lemme 8.26 *Dans le lemme précédent 8.25, on peut remplacer la famille finie $(F_i)_{i=1,...,n_\varepsilon}$ d'intervalles fermés par une famille finie $(B_i)_{i=1,...,m_\varepsilon}$ presque disjointe d'intervalles fermés, $m_\varepsilon \in \mathbb{N}^*$.*

Preuve. Si les F_i° sont deux à deux disjoints, la famille $(B_i) = (F_i)$ convient. Sinon $m_\varepsilon < n_\varepsilon$: si $i < j$ et $F_i^\circ \cap F_j^\circ \ne \emptyset$ on forme l'intervalle $F_i \bigcup F_j$ qu'on rebaptise F_i (qui reste fermé). Et on recommence jusqu'à n'avoir que des intervalles fermés presque disjoints (ici n_ε est fini, on recommence au plus un nombre fini de fois). ∎

8.6.5 Approximation d'une fonction $L^1(\mathbb{R})$ par une fonction en escalier

Corollaire 8.27 *Si $f \in \mathcal{L}^1(\mathbb{R})$ et $\text{supp}(f)$ borné, alors il existe une suite $(f_n)_{\mathbb{N}^*}$ de fonctions en escalier qui converge simplement p.p. vers f. En particulier, $\forall \varepsilon > 0$, $\exists g$ en escalier, p.p. $x \in \mathbb{R}$, $|g(x) - f(x)| < \varepsilon$.*

Preuve. Il suffit de le prouver pour $f = 1_A$ où A est mesurable. Avec les lemmes 8.25 et (8.6.4), $\forall k \in \mathbb{N}^*$, $\exists n_k \in \mathbb{N}^*$, $\exists (B_i)_{i=1,...,n_k}$, t.q. $\mu_\ell(A\Delta(\bigcup_{i=1}^{n_k} B_i)) \le \frac{1}{2^k}$, où $(B_i)_{i=1,...,n_k}$ est une famille presque disjointe d'intervalles fermés. On pose :

$$C_k = \bigcup_{i=1}^{n_k} B_i^\circ, \quad \text{où donc} \quad 1_{C_k} = \sum_{i=1}^{n_k} 1_{B_i^\circ}, \quad \text{avec} \quad \mu_\ell(A\Delta C_k) \le \frac{1}{2^k}. \tag{8.29}$$

Montrons que la suite de fonctions en escaliers $(f_k = 1_{C_k})_{k\in\mathbb{N}^*}$ tend p.p. vers $f = 1_A$.

Posons $D_k =^{\text{déf}} A\Delta C_k$, avec donc $\mu_\ell(D_k) < \frac{1}{2^k}$.

On a $1_A(x) = 1_{C_k}(x) = 1$ pour $x \in (A \cap C_k)$ et $1_A(x) = 1_{C_k}(x) = 0$ pour $x \in (A \cup C_k)^C$. Donc $1_A = 1_{C_k}$ sur $\mathbb{R} - D_k$.

On pose $E_n = \bigcup_{k\ge n} D_k$ (l'union à partir de D_n). En particulier $1_A = 1_{C_k}$ sur $\mathbb{R} - E_k$ avec la suite $(E_n)_{n\in\mathbb{N}^*}$ qui est décroissante t.q. $\mu_\ell(E_n) \le \sum_{k=n}^\infty \frac{1}{2^k} = \frac{1}{2^{n-1}}$. Donc notant $E = \bigcap_{n\in\mathbb{N}^*} E_n$ on a $\mu_\ell(E) = 0$. Et pour tout $x \in E^C$ on a $f_n(x) \longrightarrow_{n\to\infty} f(x)$.

En effet $E^C = \bigcup_{n\in\mathbb{N}^*} E_n^C = \bigcup_{n\in\mathbb{N}^*} \bigcap_{k\ge n} D_k^C$, donc si $x \notin E$ alors $\exists n$ t.q. $x \in \bigcap_{k\ge n} D_k^C$, donc $x \in D_k^C$ pour tout $k \ge n$, donc $1_A(x) = 1_{C_k}(x)$ pour tout $k \ge n$. ∎

9 Changement de variables dans les intégrales dans $\mathbb{R}^n$

9.1 Volume d'un ouvert

On se place dans l'espace affine $\mathbb{R}^n$ (ensemble de points).

On note O un point dans $\mathbb{R}^n$ qu'on appellera l'origine.

On note $\overrightarrow{\mathbb{R}^n}$ l'espace vectoriel associé (ensemble de vecteurs).

On note $(\vec{e}_i)_{i=1,...,n} =^{\text{noté}} (\vec{e}_i)$ la base canonique de $\overrightarrow{\mathbb{R}^n}$.

Un point $p \in \mathbb{R}^n$ sera repéré par le vecteur "bi-point" $\overrightarrow{0p} =^{\text{noté}} \vec{x}$, et ses coordonnées sont les x_i données par $\vec{x} = \sum_{i=1}^n x_i\vec{e}_i$, soit $p = O + \sum_{i=1}^n x_i\vec{e}_i$. Et on dit abusivement que $\vec{x}$ est un point.

Pour $\alpha < \beta$, on note (α, β) un intervalle de type $[\alpha, \beta]$, $]\alpha, \beta]$, $[\alpha, \beta[$ ou $]\alpha, \beta[$.

Définition 9.1 Un pavé P dans $\mathbb{R}^n$ est un ensemble de type : il existe $\vec{x} \in \mathbb{R}^n$ et il existe des $a_i, b_i \in \mathbb{R}$ avec $a_i < b_i$ t.q. :

$$\begin{aligned} P = \vec{x} + \prod_{i=1}^{n}(a_i, b_i) &= \{\vec{x} = \sum_{i=1}^{n} x_i\vec{e}_i \in \mathbb{R}^n : \forall i = 1, ..., n,\ x_i \in (a_i, b_i)\} \\ &= \{\vec{y} \in \mathbb{R}^n : \forall i = 1, ..., n,\ \exists\lambda_i \in (0, b_i - a_i),\ \vec{y} = \vec{x} + \sum_{i=1}^{n} \lambda_i\vec{e}_i\}, \end{aligned} \tag{9.1}$$

et $\vec{x}$ est un sommet du pavé.

Proposition 9.2 *Les pavés ouverts de $\mathbb{R}^n$ forment une base de voisinage de $\mathbb{R}^n$.*

Preuve. Voir polycopié "Topologie". ∎

Définition 9.3 La mesure de Lebesgue (le volume) d'un pavé donné en (9.1) est définie par :

$$\mu_\ell(P) = \prod_{i=1}^{n}(b_i - a_i) \overset{\text{noté}}{=} \int_{\vec{x}\in P} dx, \tag{9.2}$$

produit des longueurs des côtés.

De la proposition 9.2, on déduit, pour tout U ouvert, la mesure de 1_U (la fonction fonction indicatrice de U) :

$$\mu_\ell(1_U) \overset{\text{déf}}{=} \mu_\ell(U) \overset{\text{noté}}{=} \int_{\mathbb{R}^n} 1_U(\vec{x})\, dx. \tag{9.3}$$

Et donc la mesure de toute fonction f mesurable est donnée par :

$$\mu_\ell(f) \overset{\text{noté}}{=} \int_{\mathbb{R}^n} f(\vec{x})\, dx. \tag{9.4}$$

9.2 Formule de changement de variables

On décompose un ouvert $U \subset \mathbb{R}^n$ à l'aide de la proposition 9.2 : $U = \bigcup_{i\in\mathbb{N}^*} P_i$, où les P_i sont des pavés deux à deux distincts.

Soit $\Phi : \left\{ \begin{array}{ll} \mathbb{R}^n & \to \mathbb{R}^n \\ \vec{x} & \to \vec{y} = \Phi(\vec{x}) = \begin{pmatrix} \Phi_1(\vec{x}) \\ \vdots \\ \Phi_n(\vec{x}) \end{pmatrix} \end{array} \right\}$ une application différentiable de l'espace affine $\mathbb{R}^n$ dans lui-même : pour tout point $\vec{x}_0 \in \mathbb{R}^n$. Soit $d\Phi(\vec{x}_0)$ donnée par :

$$L_{\vec{x}_0} = d\Phi(\vec{x}_0) : \left\{ \begin{array}{ll} \overrightarrow{\mathbb{R}^n} & \to \overrightarrow{\mathbb{R}^n} \\ \vec{u} & \to \vec{v} = d\Phi(\vec{x}_0).\vec{u}, \end{array} \right\} \tag{9.5}$$

son application linéaire tangente en $\vec{x}_0$. Ainsi au voisinage d'un point $\vec{x}_0 \in \mathbb{R}^n$:

$$(\vec{y} =) \quad \Phi(\vec{x}) = \vec{x}_0 + d\Phi(\vec{x}_0).(\vec{x} - \vec{x}_0) + o(\vec{x} - \vec{x}_0), \tag{9.6}$$

Définition 9.4 La matrice jacobienne de Φ en $\vec{x}_0$ est la matrice $[d\Phi(\vec{x}_0)] = [\frac{\partial\Phi_i}{\partial x_j}(\vec{x}_0)]$ de l'endomorphisme $d\Phi(\vec{x}_0)$ représenté dans la base canonique de $\overrightarrow{\mathbb{R}^n}$.

La jacobienne $J_\Phi(\vec{x}_0)$ de Φ en $\vec{x}_0$ est le déterminant de la matrice jacobienne $[d\Phi(\vec{x}_0)]$:

$$J_\Phi(\vec{x}_0) \overset{\text{déf}}{=} \det([d\Phi(\vec{x}_0)]). \tag{9.7}$$

On s'intéresse au volume de $\Phi(U)$:

Proposition 9.5 *Si Φ est un difféomorphisme (application bijective C^1 d'inverse C^1) on a :*

$$\mu_\ell(\Phi(U)) = \mu_\ell(1_U J_\Phi), \tag{9.8}$$

soit :

$$\int_{\vec{y}\in\Phi(U)} d\Omega_y = \int_{\vec{x}\in U} |J_\Phi(\vec{x})|\, d\Omega_x \tag{9.9}$$

(l'élément de volume $d\Omega_y$ est envoyé par Φ^{-1} sur l'élément de volume $|J_\Phi(\vec{x})|\, d\Omega_x$). Plus généralement, pour tout fonction g mesurable :

$$\int_{\vec{y}\in\Phi(U)} g(\vec{y})\, dy = \int_{\vec{x}\in U} g(\Phi(\vec{x}))|J_\Phi(\vec{x})|\, dx. \tag{9.10}$$

Preuve. 1- Cas Φ affine : un pavé P est transformé en parallélogramme $\Phi(P)$ de volume $|J_\Phi(\vec{x})| = |\det([d\Phi(\vec{x})])|$ qui est indépendant de $\vec{x}$ (car Φ est affine). Et on obtient la formule $\int_{\Phi(P)} dy = \int_P 1_P(\vec{x})|J_\Phi(\vec{x})|\,dx$, soit (9.9) pour les pavés.

Avec la proposition 9.2, toujours avec Φ affine, on obtient (9.9).

2- Cas Φ difféomorphisme : exercice (construction de Riemann). ∴

Et on a donc également, pour tout ouvert V :

$$\int_{\vec{x}\in\Phi^{-1}(V)} dx = \int_{\vec{y}\in V} |J_{\Phi^{-1}}(\vec{y})|\,dy, \tag{9.11}$$

et pour toute fonction f mesurable :

$$\int_{\vec{x}\in\Phi^{-1}(V)} f(\vec{x})\,dx = \int_{\vec{y}\in V} f(\Phi^{-1}(\vec{y}))\,|J_{\Phi^{-1}}(\vec{y})|\,dy. \tag{9.12}$$

Exemple 9.6 Cas 1-D et $\Phi : x \in [0,1] \to y = \Phi(x) = -2x \in [-2,0]$ (changement de variable ici affine) de différentielle $\Phi'(x) = -2 = J_\Phi(x)$ sur $[0,1]$; vérifions (9.10) avec ici $g = 1_B$:

$$\mu_\ell(\Phi([0,1])) = \mu_\ell([-2,0]) = \int_{x\in[-2,0]} dx = 2,$$

de même que :

$$|J_\Phi(x)|\,\mu_\ell([0,1]) = |-2| \int_{x\in[0,1]} dx = 2,$$

les intervalles étant toujours exprimés sous la forme $[a,b]$ avec $a \le b$. On retrouve les résultats du calcul de changement de variable usuel de l'intégrale de Riemann avec $y = -2x$:

$$\int_{y\in[-2,0]} dy = \int_{y=-2}^{0} dy = \int_{x=1}^{0} -2\,dx = -\int_{x=0}^{1} -2\,dx = 2\int_{x=0}^{1} dx = \int_{x\in[0,1]} |-2|\,dx,$$

voir remarque 4.11. ∴

Exercice 9.7 Un parallélogramme dans $\mathbb{R}^n$ est un ensemble Ω de points de type :

$$\Omega = \{\vec{y} \in \mathbb{R}^n : \exists \vec{x} \in \mathbb{R}^n,\ \exists \lambda_1, ..., \lambda_n \in (0,1),\ \vec{y} = \vec{x} + \lambda_1 \vec{b}_1 + ... + \lambda_n \vec{b}_n\}, \tag{9.13}$$

où $(\vec{b}_i)$ est une base de $\overrightarrow{\mathbb{R}^n}$. Le point $\vec{x}$ est un sommet d'un parallélogramme. (Un pavé est un parallélogramme de côtés parallèles aux axes.)

Montrer que le volume (la mesure de Lebesgue) d'un parallélogramme donné en (9.13) est :

$$\mu_\ell(\Omega) = |\det(\vec{b}_1, ..., \vec{b}_n)| \overset{\text{noté}}{=} \int_{\vec{\omega}\in\Omega} d\omega, \tag{9.14}$$

(Ne pas oublier la valeur absolue : une mesure est positive.)

Réponse. Prendre Φ application affine qui transforme l'hyper-cube $(0, \vec{e}_1, ..., \vec{e}_n)$ en Ω (faire un dessin). ∴

9.3 Vers la mesure image

(En vue des applications aux probabilités et à la loi image.)

(9.9) se lit, avec $V = \Phi(U)$:

$$\mu_\Phi(V) = \mu(U), \quad \text{soit} \quad \int_{\vec{y}\in V} d\mu_\Phi(\vec{y}) = \int_{\vec{x}\in\Phi^{-1}(V)} d\mu(\vec{x}), \tag{9.15}$$

quand on pose :

$$\begin{cases} \mu_\Phi(\vec{y}) \overset{\text{déf}}{=} \mu_\ell(\vec{y}) \quad (= dy) \quad (\text{dans l'espace d'arrivée : } \vec{y} = \Phi(\vec{x})), \\ \mu(\vec{x}) \overset{\text{déf}}{=} |J_\Phi(\vec{x})|\,\mu_\ell(\vec{x}) \quad (\text{dans l'espace de départ}), \end{cases} \tag{9.16}$$

μ étant donc la mesure de densité $|J_\Phi(\vec{x})|$ relativement à la mesure de Lebesgue. Ainsi (9.8) se lit (on est ici dans le cas Φ difféomorphisme) :

$$(\mu_\Phi \circ \Phi)(U) = \mu(U), \tag{9.17}$$

soit :

$$\mu_\Phi(V) = (\mu \circ \Phi^{-1})(V). \tag{9.18}$$

Et (9.10) se lit : pour g intégrable sur V ouvert :

$$\int_V g\, d\mu_\Phi = \int_{\Phi^{-1}(V)} g \circ \Phi\, d\mu, \tag{9.19}$$

soit :

$$\int_{\vec{y}\in V} g(\vec{y})\, d\mu_\Phi(y) = \int_{\vec{x}\in\Phi^{-1}(V)} g(\Phi(\vec{x}))\, d\mu(\vec{x}), \tag{9.20}$$

par définition de la mesure μ, cf. (9.16).

9.4 Mesure image

9.4.1 Rappel : définition de f^{-1}

Soit $f : E \to F$ une fonction non nécessairement bijective. On rappelle que f^{-1} est définie par :

$$f^{-1} : \begin{cases} \mathcal{P}(F) & \to \mathcal{P}(E), \\ \quad B & \to A = f^{-1}(B) = \{\omega \in E : f(\omega) \in B\}. \end{cases} \tag{9.21}$$

Exemple 9.8 $E = \mathbb{R} = F$ et $f = 1_A$ où $A \in \mathcal{A}_\mathbb{R}$, la fonction indicatrice de A. On a $1_A(\mathbb{R}) = \{0,1\}$, et pour $V \in \mathcal{A}_\mathbb{R}$:

$$1_A^{-1}(V) = \{\omega \in E : 1_A(\omega) \in V\} = \begin{cases} \emptyset & \text{si } V \cap \{0,1\} = \emptyset, \text{ i.e. } 0 \notin V \text{ et } 1 \notin V, \\ A & \text{si } V \cap \{0,1\} = \{1\}, \text{ i.e. } 0 \notin V \text{ et } 1 \in V, \\ A^c & \text{si } V \cap \{0,1\} = \{0\}, \text{ i.e. } 0 \in V \text{ et } 1 \notin V, \\ \mathbb{R} & \text{si } V \cap \{0,1\} = \{0,1\}, \text{ i.e. } 0 \in V \text{ et } 1 \in V. \end{cases} \tag{9.22}$$

Faire un dessin. En particulier, pour $a < b$ on a :

$$1_A^{-1}([a,b]) = \begin{cases} \emptyset & \text{si } a > 1, \text{ ou } b < 0, \text{ ou } 0 < a < b < 1, \\ A & \text{si } 0 < a < b \geq 1, \\ A^c & \text{si } a \leq 0 < b < 1, \\ \mathbb{R} & \text{si } a < 0 \text{ et } b > 1. \end{cases} \tag{9.23}$$

▲

Exemple 9.9 Soit $f : \mathbb{R} \to \mathbb{R}$ donnée par $f(x) = \sin x$. Alors $f^{-1}(\{0\}) = \pi\mathbb{Z}$. ▲

Cas particulier f bijective, alors f^{-1} est identifiée à une fonction $F \to E$, encore notée f^{-1}, où on note $f^{-1}(x) = f^{-1}(\{x\}) \in E$.

9.4.2 Définition de la mesure image

Soit $(E, \mathcal{A}_E, \mu)$ un espace mesuré, et $(F, \mathcal{A}_F)$ un espace mesurable où F est un espace topologique et $\mathcal{A}_F$ la tribu borélienne de F. Munissons $(F, \mathcal{A}_F)$ d'une mesure.

Définition 9.10 Soit $\Phi : E \to F$ une application mesurable non bijective a priori. On appelle mesure image de μ par Φ la mesure $\mu_\Phi : \mathcal{A}_F \to \mathbb{R}$ (mesure sur F) définie par :

$$\mu_\Phi \stackrel{\text{déf}}{=} \mu \circ \Phi^{-1}. \tag{9.24}$$

(Comme en (9.18), sauf qu'on ne suppose plus Φ bijective.)

Proposition 9.11 *La fonction $\mu_\Phi : \mathcal{T}_F \to \mathbb{R}$ définie par (9.24) est bien une mesure, et $(F, \mathcal{A}_F, \mu_\Phi)$ est donc un espace mesuré.*

Preuve. Vérification immédiate des points 1,2,3 de la définition 1.14 page 8. ▲

Remarque 9.12 Pour Φ non bijective, on ne peut pas remplacer (9.24) par $\mu_\Phi \circ \Phi = \mu$.

Car, pour Φ non bijective, si m est une mesure, alors $m \circ \Phi$ n'est pas une mesure. Exemple : prendre $m = \mu_\ell$ (comme dans (9.16)) et $\Phi = \sin$, avec $A_1 = [0, \frac{\pi}{2}[$ et $A_2 =]\frac{\pi}{2}, \pi]$ qui vérifient $A_1 \cap A_2 = \emptyset$ avec $(m \circ \Phi)(A_1) = \mu_\ell([0,1[) = 1 = (m \circ \Phi)(A_2)$, et avec $(m \circ \Phi)(A_1 \cup A_2) = \mu_\ell([0,1[) = 1 \neq 2 = (m \circ \Phi)(A_1) + (m \circ \Phi)(A_2)$.

(En revanche, si $A_1 \cap A_2 = \emptyset$ alors $\Phi^{-1}(A_1 \cup A_2) = \Phi^{-1}(A_1) + \Phi^{-1}(A_2)$.) ▟

Donc, pour tout $V \in \mathcal{A}_F$:

$$\mu_\Phi(V) = \mu(U) \quad \text{quand} \quad U = \Phi^{-1}(V). \tag{9.25}$$

(En particulier, dans le cas Φ difféomorphisme, on retrouve (9.17).)

Exemple 9.13 Soit $(E, \mathcal{A}_E, \mu)$ un espace mesuré. Soit $(F, \mathcal{A}_F) = (\mathbb{R}, \mathcal{A}_\mathbb{R})$, soit $A \in \mathcal{A}_E$, et soit $\Phi = 1_A$. Ainsi $(\mathbb{R}, \mathcal{A}_\mathbb{R}, \mu_{1_A})$ est un espace mesuré, où avec (9.23) on obtient, pour $V \in \mathcal{A}_\mathbb{R}$:

$$\mu_{1_A}(V) = \mu(1_A^{-1}(V)) = \begin{cases} \mu_(\emptyset) = 0 & \text{si } V \cap \{0,1\} = \emptyset, \\ \mu(A) & \text{si } V \cap \{0,1\} = \{1\}, \\ \mu(A^c) & \text{si } V \cap \{0,1\} = \{0\}, \\ \mu(\mathbb{R}) & \text{si } V \cap \{0,1\} = \{0,1\}. \end{cases}$$ ▟

9.4.3 Propriétés

Lemme 9.14 *Soit $\Phi : E \to F$ mesurable (bijective ou non). Pour tout $B \in \mathcal{T}_F$ on a :*

$$\Phi^{-1}(B \cap \Phi(E)) = \Phi^{-1}(B), \tag{9.26}$$

et donc, pour tout $B \in \mathcal{T}_F$:

$$\mu_\Phi(B \cap \Phi(E)) = \mu_\Phi(B). \tag{9.27}$$

Preuve. On a $\Phi^{-1}(B \cap \Phi(E)) = \{x \in E : \Phi(x) \in B \cap \Phi(E)\} = \{x \in E : \Phi(x) \in B\}$ (puisque $\Phi(x) \in \Phi(E)$ est automatiquement vérifié), d'où (9.26). D'où (9.27), cf. (9.24). ▟

Proposition 9.15 *Si $\Phi : E \to F$ est mesurable, si $g : F \to \mathbb{R}$ est mesurable, et si $g \circ \Phi : E \to \mathbb{R}$ est μ-intégrable alors $g : \Phi(E) \to \mathbb{R}$ est μ_Φ-intégrable et :*

$$\mu_\Phi(1_{\Phi(E)} g) = \mu(g \circ \Phi), \tag{9.28}$$

soit (formule de changement de variable) :

$$\int_{y \in \Phi(E)} g(y)\, d\mu_\Phi(y) = \int_{x \in E} g(\Phi(x))\, d\mu(x). \tag{9.29}$$

Preuve. Vérifions (9.28) quand $g = 1_V$ pour un $V \in \mathcal{T}_F$, i.e. $\mu_\Phi(1_V 1_{\Phi(E)}) = \mu(1_V \circ \Phi)$. On a $1_{\Phi(E)} 1_V = 1_{V \cap \Phi(E)}$, et :

$$(1_V \circ \Phi)(x) = 1_V(\Phi(x)) = \left\{ \begin{matrix} 1 & \text{si } \Phi(x) \in V, \text{ i.e. si } x \in \Phi^{-1}(V) \\ 0 & \text{si } \Phi(x) \notin V, \text{ i.e. si } x \notin \Phi^{-1}(V) \end{matrix} \right\} = 1_{\Phi^{-1}(V)}, \tag{9.30}$$

Et donc on veut vérifier :

$$\mu_\Phi(1_{V \cap \Phi(E)}) = \mu(1_{\Phi^{-1}(V)}), \quad \text{i.e.} \quad \mu_\Phi(V \cap \Phi(E)) = \mu(\Phi^{-1}(V)). \tag{9.31}$$

Avec (9.27), c'est vrai.

Si $g = \sum_{i=1}^n c_i 1_{V_i}$ (fonction en escalier), alors (9.28) est vraie par linéarité.

Si g est une fonction mesurable positive, cela reste vrai, cf. (3.7).

D'où si g est intégrable c'est encore vrai (on écrit $g = g_+ - g_-$). ▟

Corollaire 9.16 *Avec $g(y) = y^n$, on a (moments d'ordre n) :*

$$\int_{y \in \Phi(E)} y^n\, d\mu_\Phi(y) = \int_{x \in E} (\Phi(x))^n\, d\mu(x), \tag{9.32}$$

$=^{noté}$ *$E(\Phi^n)$ dans le cours de probabilité (espérance de la fonction intégrable Φ^n).*

9.4.4 Formule de changement de variables dans les intégrales

Exercice 9.17 Cas d'une mesure de densité $\mu = p(\vec{x})\,dx$ sur $\mathbb{R}^n$ muni de sa tribu borélienne, avec $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ un difféomorphisme. Montrer que la mesure image par Φ de la mesure μ est la mesure de densité $\mu_\Phi(\vec{y}) = p_\Phi(\vec{y})\,dy$ où p_Φ est donnée par :

$$p_\Phi(\vec{y}) = |J_{\Phi^{-1}}(\vec{y})|\, p(\Phi^{-1}(\vec{y})) \quad (= \frac{p(\vec{x})}{|J_\Phi(\vec{x})|} \quad \text{quand} \quad \vec{y} = \Phi(\vec{x})). \tag{9.33}$$

En déduire que, pour toute fonction g intégrable :

$$\int_{\vec{y}\in\Phi(U)} g(\vec{y})\, p_\Phi(\vec{y}))\, dy = \int_{\vec{x}\in U} g(\Phi(\vec{x}))\, p(\vec{x})\, dx, \tag{9.34}$$

et retrouver (9.10).

Réponse. (9.10) donne :

$$\int_{\vec{x}\in U} p(\vec{x})\, dx = \int_{\vec{x}\in U} \frac{p(\vec{x})}{|J_\Phi(\vec{x})|} |J_\Phi(\vec{x})|\, dx = \int_{\vec{y}\in\Phi(U)} g(\vec{y})\, dy, \tag{9.35}$$

où $g(\vec{y}) = (g\circ\Phi)(x) = \frac{p(\vec{x})}{|J_\Phi(\vec{x})|}$, d'où (9.33). D'où (9.10) donne (9.34).

En particulier pour $p(\vec{x}) = |J_\Phi(\vec{x})|\, dx$, on obtient $p_\Phi(\vec{y}) = 1$, donc (9.10).

10 Convolution et régularisation

10.1 Notations $\check{f}$ et $\tau_x f$

Définition 10.1 Pour $f : \mathbb{R} \to \mathbb{R}$, on définit $\check{f} : \mathbb{R} \to \mathbb{R}$ par :

$$\check{f}(x) = f(-x). \tag{10.1}$$

Autrement dit $\check{f} = f \circ g$ où $g(x) = -x$.

(Le graphe de $\check{f}$ est le symétrique du graphe de f par rapport à "l'axe des y".)

Définition 10.2 Pour $f : \mathbb{R} \to \mathbb{R}$, et $c \in \mathbb{R}$ on définit la translatée $\tau_c(f) =^{\text{noté}} \tau_c f$ de f par :

$$\tau_c f(x) = f(x - c). \tag{10.2}$$

Autrement dit $\tau_c f = f \circ h_c$ où $h_c(x) = x - c$.

(Le graphe de $\tau_c f$ est le translaté du graphe de f de c : en particulier $(\tau_c f)(c) = f(0)$.)

Exercice 10.3 Montrer que $\widecheck{\tau_c f}(t) = f(-t-c)$, et que $\tau_c \check{f}(t) = f(c-t)$

Réponse. Soit $g = \tau_c f$. On a $\check{g}(t) = g(-t) = f(-t-c)$. Et $\tau_c\check{f}(t) = \check{f}(t-c) = f(c-t)$.

Exercice 10.4 Montrer que :

$$\widecheck{\tau_c f} = \tau_{-c}\check{f}. \tag{10.3}$$

Autrement dit, les opérateurs $\check{}$ et τ_c ne commutent pas pour $c \neq 0$: $(\check{} \circ \tau_c)(f) = (\tau_{-c} \circ \check{})(f)$.

Réponse. $\widecheck{\tau_c f}(x) = \tau_c f(-x) = f(-x-c) = \check{f}(x+c) = \tau_{-c}\check{f}(x)$.

Proposition 10.5 *Pour $f : \mathbb{R} \to \mathbb{R}$ et pour $c \in \mathbb{R}$, on a :*

$$\mathrm{supp}\check{f} = -\mathrm{supp}f, \qquad \mathrm{supp}(\tau_c f) = \mathrm{supp}f + c, \qquad \mathrm{supp}(\tau_c\check{f}) = -\mathrm{supp}f + c. \tag{10.4}$$

Immédiat sur un dessin. En particulier, si $\mathrm{supp}f \subset [a,b]$ où $a \le b$, alors :

$$\mathrm{supp}(\check{f}) \subset [-b,-a], \qquad \mathrm{supp}(\tau_c f) \subset [a{+}c, b{+}c], \qquad \mathrm{supp}(\tau_c\check{f}) \subset [-b{+}c, -a{+}c]. \tag{10.5}$$

Preuve. Pour $\check{f}$: on a $\{x : \check{f}(x) \neq 0\} = \{x : f(-x) \neq 0\} = \{-y : f(y) \neq 0\}$, d'où, en prenant l'adhérence, $\mathrm{supp}\check{f} = -\mathrm{supp}f$.

Pour $\tau_c f$: on a $\{x : \tau_c f(x) \neq 0\} = \{x : f(x-c) \neq 0\} = \{y + c : f(y) \neq 0\}$, d'où, en prenant l'adhérence, $\mathrm{supp}\tau_c f = \mathrm{supp}f + c$.

10.2 Définition de la convolution

On rappelle que si $f,g:\mathbb{R}^n\to\mathbb{R}$ sont deux fonctions (mesurables), alors $f*g:\mathbb{R}^n\to\mathbb{R}$ est la fonction donnée formellement par :

$$(f*g)(x)=\int_{t\in\mathbb{R}^n} f(t)g(x-t)\,dt=\int_{t\in\mathbb{R}^n} f(t)\tau_x\breve{g}(t)\,dt, \tag{10.6}$$

intégrale qui dépend du paramètre x. En particulier, si $f,g\in\mathcal{L}^2(\mathbb{R})$ alors pour chaque x on a $\tau_x\breve{g}\in\mathcal{L}^2(\mathbb{R})$ (car $\int_{t\in\mathbb{R}}|g(x-t)|^2\,dt=\int_{s\in\mathbb{R}}|g(s)|^2\,ds<\infty$), et donc :

$$(f*g)(x)=(f,\tau_x\breve{g})_{\mathcal{L}^2(\mathbb{R})}.$$

Dans la suite, pour simplifier la présentation, on considèrera essentiellement le cas $\mathbb{R}^n=\mathbb{R}$.

Exemple 10.6 Si $f\in L^1(\mathbb{R})$ et $g=1_\mathbb{R}$, on a $(f*1_\mathbb{R})(x)=\int_\mathbb{R} f(t)\,dt=$ constante, et donc l'application $f\to f*1_\mathbb{R}$ est la fonction "aire sous la courbe f" (indépendante de x). ▪

Exemple 10.7 Si $f\in L^1(\mathbb{R})$ et $g=\Pi_k=k1_{[\frac{1}{2k},\frac{1}{2k}]}$, on a $(f*g)(x)=k\int_{-\frac{1}{2k}}^{\frac{1}{2k}} f(x-t)\,dt=$ la "valeur moyenne de f à travers une fenêtre de largeur $\frac{1}{k}$ centrée en x". Dessin.

La "mesure" d'une fonction f à travers un appareil d'une certaine précision est une application de type $M_k:f\to M_k(f)=f*\Pi_k$, avec donc $M_k(f)(x)$ approchant $f(x)$, approximation d'autant meilleure que k est grand, i.e. que l'appareil est précis. L'appareil idéal (de précision parfaite) est $M_\infty:f\to M_\infty(f)=f(x)$, soit $M_\infty=\delta_0$, voir plus loin. ▪

Proposition 10.8 *Quand elle est définie, l'opération $*$ est distributive et commutative :*

$$g*f=f*g,\qquad f*(g_1+\lambda g_2)=f*g_1+\lambda f*g_2, \tag{10.7}$$

d'où le nom de "produit" (commutatif) de convolution. Et on a :

$$\widetilde{f*g}=\breve{f}*\breve{g},\qquad \text{et}\qquad \tau_a(f*g)=(\tau_a f)*g=f*(\tau_a g). \tag{10.8}$$

Et :

$$\begin{cases}(f \text{ et } g \text{ paires}) \text{ ou } (f \text{ et } g \text{ impaires}) \Rightarrow\ f*g \text{ paire},\\ (f \text{ paire et } g \text{ impaire}) \text{ ou } (f \text{ impaire et } g \text{ paire}) \Rightarrow\ f*g \text{ impaire}.\end{cases} \tag{10.9}$$

Preuve. $(f*g)(x)=\int_\mathbb{R} f(t)g(x-t)\,dt=\int_\mathbb{R} f(x-u)g(u)\,du=(g*f)(x)$ donne la commutativité. Et la distributivité résulte de la distributivité de la multiplication de $\mathbb{R}$ et de la linéarité de l'intégrale.

Puis $(\breve{f}*\breve{g})(x)=\int_\mathbb{R} f(-t)g(-x+t)\,dt=\int_\mathbb{R} f(u)g(-x-u)\,du=(f*g)(-x)$.

Puis $\tau_a(f*g)(x)=(f*g)(x-a)=\int_\mathbb{R} f(t)g(x-a-t)\,dt=\int_\mathbb{R} f(t)\tau_a g(x-t)\,dt=(f*\tau_a g)(x)$ et $f*g=g*f$.

Puis $(f*g)(-x)=\int_\mathbb{R} f(t)g(-x-t)\,dt$: si g est paire, alors $(f*g)(-x)=\int_\mathbb{R} f(t)g(x+t)\,dt=\int_\mathbb{R} f(-u)g(x-u)\,dt$, d'où si f est paire alors $(f*g)(-x)=(f*g)(x)$, et si f est impaire alors $(f*g)(-x)=-(f*g)(x)$; et $f*g=g*f$. ▪

Proposition 10.9 *Pour $f,g:\mathbb{R}\to\mathbb{R}$, la fonction convolée $f*g$ vérifie (quand elle a un sens) :*

$$\mathrm{supp}(f*g)\subset\overline{\mathrm{supp}f+\mathrm{supp}g}. \tag{10.10}$$

Preuve. On a $(f*g)(x)=\displaystyle\int_{t\in\mathbb{R}} f(t)g(x-t)\,dt=\int_{t\in\mathrm{supp}f\bigcap\mathrm{supp}\tau_x\breve{g}} f(t)\tau_x\breve{g}(t)\,dt$.

Cas simple : $\mathrm{supp}f=[a,b]$ et $\mathrm{supp}g=[c,d]$, avec $a\le b$ et $c\le d$, donc $\mathrm{supp}f+\mathrm{supp}g=[a{+}c,b{+}d]$. Et $\mathrm{supp}\tau_x\breve{g}=[-d{+}x,-c{+}x]$, donc $\mathrm{supp}f\bigcap\mathrm{supp}\tau_x\breve{g}=[a,b]\cap[-d{+}x,-c{+}x]$, donne $\mathrm{supp}f\bigcap\mathrm{supp}\tau_x\breve{g}=\emptyset$ ssi soit $a>-c{+}x$ soit $b<-d{+}x$, i.e. ssi $x<a{+}c$ ou $x>b{+}d$. Donc $\mathrm{supp}f\bigcap\mathrm{supp}\tau_x\breve{g}=\emptyset$ dès que $x\notin[a{+}c,b{+}d]$, donc $\mathrm{supp}(f*g)\subset\mathrm{supp}f+\mathrm{supp}g$.

Cas général : on a $(f*g)(x)=0$ dès que $\mathrm{supp}f\bigcap\mathrm{supp}\tau_x\breve{g}=\emptyset$. Et $\exists t\in\mathrm{supp}f\bigcap\mathrm{supp}\tau_x\breve{g}$ ssi $\exists t\in\mathrm{supp}f$ et $t\in x-\mathrm{supp}g$, i.e. ssi $\exists t\in\mathrm{supp}f$ et $x\in t+\mathrm{supp}g$ ($\subset\mathrm{supp}f+\mathrm{supp}g$). Donc si $x\notin\mathrm{supp}f+\mathrm{supp}g$ alors $\mathrm{supp}f\bigcap\mathrm{supp}\tau_x\breve{g}=\emptyset$ donc $(f*g)(x)=0$. Donc $\{x:(f*g)(x)\ne0\}\subset\mathrm{supp}f+\mathrm{supp}g$. D'où (10.10). ▪

Remarque 10.10 Rappel : la somme de deux fermés n'est pas nécessairement un fermé : prendre $F=\mathbb{N}^*$ et $G=\bigcup_{k\in\mathbb{N}^*}\{-k+\frac{1}{k}\}$ qui donne $F+G=\{n-k+\frac{1}{k},\ k,n\in\mathbb{N}^*\}$. Ici $\mathbb{R}-F=\bigcup_{n\in\mathbb{N}^*}]-n,n{+}1[$ et $\mathbb{R}-G=\bigcup_{k\in\mathbb{N}^*}]-k+\frac{1}{k},-k{+}1+\frac{1}{k-1}[$ sont des ouverts (union d'ouverts), donc F et G sont fermés, mais $F+G$ contient la suite $(\frac{1}{k})_{k\in\mathbb{N}^*}$ qui converge vers 0 dans $\mathbb{R}$, avec $0\notin F+G$, donc $F+G$ n'est pas fermé.

Rappel : la somme d'un compact et d'un fermé est un fermé : soit K compact et G fermé, soit (z_n) une suite dans $K+G$ qui converge vers z dans $\mathbb{R}$. Montrons que $z\in K+G$. On a $z_n=k_n+g_n$, et quitte à extraire une sous-suite, on a $k_n\to k$ dans K. Donc $g_n=z_n-k_n\in G$ converge vers $z-k$, avec G fermé, donc $g=^{\text{déf}}z-k\in G$, donc $z=k+g\in K+G$, donc $K+G$ est fermé. ▪

10.3 Stabilité par convolution : $L^1(\mathbb{R}) * L^p(\mathbb{R}) \subset L^p(\mathbb{R})$ pour $1 \le p \le \infty$

Proposition 10.11 *Si $f, g \in L^1(\mathbb{R})$ alors $|f| * |g| \in L^1(\mathbb{R})$ et $f * g \in L^1(\mathbb{R})$, avec :*

$$||f * g||_1 \le ||f||_1 ||g||_1. \tag{10.11}$$

Preuve. On a, si ça a un sens :

$$\begin{aligned} ||f * g||_{L^1} &= \int_{x\in\mathbb{R}} |(f*g)(x)|\,dx = \int_{x\in\mathbb{R}} |\int_{t\in\mathbb{R}} f(x-t)g(t)\,dt|\,dx. \\ &\le \int_{x\in\mathbb{R}} (\int_{t\in\mathbb{R}} |f(x-t)|\,|g(t)|\,dt)dx = \int_{x\in\mathbb{R}} (|f| * |g|)(x)\,dx. \end{aligned} \tag{10.12}$$

Comme f et g sont dans $L^1(\mathbb{R})$ la fonction $f \otimes g : (x,y) \to f(x)g(y)$ (fonction à variables séparées) est dans $L^1(\mathbb{R}^2)$. Et on peut appliquer Fubini :

$$\infty > ||f||_{L^1}||g||_{L^1} = \int_{(y,t)\in\mathbb{R}^2} |f(y)|\,|g(t)|\,dtdy = \int_{(x,t)\in\mathbb{R}^2} |f(x-t)|\,|g(t)|\,dtdx,$$

où on a utilisé le changement de variable $F : (y,t) \in \mathbb{R}^2 \to F(y,t) = \begin{pmatrix} x = F_1(y,t) = y+t \\ t = F_2(y,t) = t \end{pmatrix}$, difféomorphisme de $\mathbb{R}^2$ dans lui-même, de jacobien $\det\begin{pmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial t} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial t} \end{pmatrix}(y,t) = \det\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = 1$ qui donne $(dtdy) = |1|\,(dtdx) = (dtdx)$. D'où $||f*g||_{L^1} \le ||f||_{L^1}||g||_{L^1} < \infty$, i.e. (10.11). ∎

Exercice 10.12 Montrer (10.11) à l'aide du théorème d'intégration de Tonelli (cours d'intégration).

Réponse. Rappel de Tonelli : si la fonction $h : \Omega_1 \times \Omega_2 \to \mathbb{R}$ satisfait aux deux hypothèses :

$$\int_{y\in\Omega_2} |h(x,y)|\,dy < \infty \quad \text{p.p. } x \quad \text{et} \quad \int_{x\in\Omega_1} \Big(\int_{y\in\Omega_2} |h(x,y)|\,dy\Big)\,dx < \infty \tag{10.13}$$

alors $h \in L^1(\Omega_1 \times \Omega_2)$, et alors on peut inverser l'ordre d'intégration (Fubini).

Ici on pose $h(t,x) = |f(t)|\,|g(x-t)|$ et on vérifie les hypothèses : commençant par intégrer en x à t fixé, il vient, à t fixé :

$$\int_{x\in\mathbb{R}} |f(t)|\,|g(x-t)|\,dx = |f(t)|(\int_{x\in\mathbb{R}} |g(x-t)|\,dx) = |f(t)|(\int_{y\in\mathbb{R}} |g(y)|\,dy) \le |f(t)|\,||g||_1 < \infty, \tag{10.14}$$

puis :

$$\int_{t\in\mathbb{R}} |f(t)|\,||g||_1\,dt \le ||g||_1 \int_{t\in\mathbb{R}} |f(t)|\,dt \le ||g||_1\,||f||_1 < \infty. \tag{10.15}$$

D'où le résultat. ∎

Exemple 10.13 $f(t) = g(t) = e^{-t}1_{\mathbb{R}_+}(t)$. On a $\int_{\mathbb{R}} |f(t)|\,dt = \int_0^\infty e^{-t}\,dt = 1 < \infty$, et $f, g \in L^1(\mathbb{R})$. Et $(f*g)(x) = \int_{\mathbb{R}} e^{-t}1_{\mathbb{R}_+}(t)e^{-(x-t)}1_{\mathbb{R}_+}(x-t)\,dt = \int_{t=0}^x e^{-t}e^{-(x-t)}1_{\mathbb{R}_+}(x)\,dt = \int_0^x e^{-x}1_{\mathbb{R}_+}(x)\,dt = xe^{-x}1_{\mathbb{R}_+}(x)$, intégrable sur $\mathbb{R}$, donc on a bien $f * g \in L^1(\mathbb{R})$. ∎

Exercice 10.14 Montrer que si $f, g, h \in L^1(\mathbb{R})$ sont positives et $f \le g$, alors $f * h \le g * h$.

Réponse. On a $(f*h)(x) = \int_{\mathbb{R}} f(t)h(x-t)\,dt \le \int_{\mathbb{R}} g(t)h(x-t)\,dt = (h*g)(x)$. ∎

Remarque 10.15 L'inégalité (10.11) obtenue est une inégalité où à gauche on a de fait une intégrale double, cf. (10.12), alors qu'à droite on a un produit des deux intégrales simples.

En particulier, $||f*g||_1$ (calcul d'une intégrale double) n'a rien à voir avec le produit $||fg||_1$ (calcul d'une intégrale simple) qui en général n'a pas de sens pour f et g dans $L^1(\mathbb{R})$.

Par exemple, f et g données par $f(t) = g(t) = \frac{1}{\sqrt{t}}1_{]0,1]}$ sont dans $L^1(\mathbb{R})$ (car $\int_0^1 |\frac{1}{\sqrt{t}}|\,dt = [2\sqrt{t}]_0^1 = 2 < \infty$), mais $(fg)(t) = \frac{1}{t}1_{]0,1]}$ n'est pas dans $L^1(\mathbb{R})$. Alors que $f*g$ donnée par $(f*g)(x) = \int_t \frac{1}{\sqrt{|t|}}\frac{1}{\sqrt{|x-t|}}\,dt$ est dans $L^1(\mathbb{R})$: cette fonction est définie p.p., et plus précisément pour tout $x \in \mathbb{R}^*$, et n'est pas définie en x=0, mais ce n'est pas gênant puisque, ici, seul le caractère intégrable (au sens de Lebesgue) nous intéresse (notion de presque partout) : autrement dit on a $L^1(\mathbb{R}^*) = L^1(\mathbb{R})$ car $\mathbb{R}^* = \mathbb{R} - \{0\}$ et l'ensemble singleton $\{0\}$ est négligeable pour la mesure de Lebesgue . En particulier, on a vu que $\int_{\mathbb{R}} |f*g|(x)\,dx \le ||f||_1||g||_1 < \infty$. ∎

On rappelle que $g \in L^p(\mathbb{R})$ ssi $|g|^p \in L^1$, et qu'alors $||g||_p = (||\,|g|^p||_{L^1})^{\frac{1}{p}} = (\int_{\mathbb{R}} |g(x)|^p\,dx)^{\frac{1}{p}}$ est une norme dans $L^p(\mathbb{R})$, cf. cours intégrale de Lebesgue.

Proposition 10.16 *Soit $p \in [1,\infty]$. Soit $f \in L^1(\mathbb{R})$ et $g \in L^p(\mathbb{R})$. Alors $f*g \in L^p(\mathbb{R})$, autrement dit $L^1(\mathbb{R}) * L^p(\mathbb{R}) \subset L^p(\mathbb{R})$, et on a :*

$$||f*g||_p \le ||f||_1 ||g||_p. \tag{10.16}$$

Preuve. Le cas $p=1$ vient d'être traité, et le cas $p=\infty$ est immédiat car alors $|(f*g)(x)| \le ||g||_\infty \int_{\mathbb{R}} |f(t)|\,dt$. Supposons donc $1<p<\infty$.

On va utiliser l'inégalité de Hölder : soit q l'exposant conjugué de p, donné par $\frac{1}{p}+\frac{1}{q}=1$; quand $\alpha \in L^q$ et $\beta \in L^p$ alors $\alpha\beta \in L^1(\mathbb{R})$ et $||\alpha\beta||_1 \le ||\alpha||_q ||\beta||_p$ (voir cours d'intégration). On a :

$$\int_{t\in\mathbb{R}} |f|(t)|g|(x-t)\,dt = \int_{t\in\mathbb{R}} |f|^{\frac{1}{q}}(t)|f|^{\frac{1}{p}}(t)|g|(x-t)\,dt \tag{10.17}$$

On pose $\alpha = |f|^{\frac{1}{q}} \in L^q(\mathbb{R})$, donc $\alpha^q = |f| \in L^1(\mathbb{R})$.

À x fixé, on pose $\beta_x(t) = |f(t)|^{\frac{1}{p}}||g|(x-t)$, donc $\beta_x(t)^p = |f(t)||g|^p(x-t)$. Et $\int_{t\in\mathbb{R}} \beta_x(t)^p\,dt = \int_{t\in\mathbb{R}} |f|(t)\,|g|^p(x-t)\,dt = (|f|*|g|^p)(x)$ est bien défini car $|f| \in L^1(\mathbb{R})$, $|g|^p \in L^1(\mathbb{R})$, cf. (10.11). Donc $\beta_x \in L^p(\mathbb{R})$. Donc (Hölder) :

$$\begin{aligned}\int_{t\in\mathbb{R}} |f|^{\frac{1}{q}}(t)|f|^{\frac{1}{p}}(t)|g|(x-t)\,dt &\le (\int_{t\in\mathbb{R}} |f|(t)\,dt)^{\frac{1}{q}}(\int_{t\in\mathbb{R}} |f|(t)|g|^p(x-t)\,dt)^{\frac{1}{p}} \\ &= ||f||_1^{\frac{1}{q}}(|f|*|g|^p)(x)^{\frac{1}{p}}.\end{aligned} \tag{10.18}$$

Donc, avec (10.17) :

$$|(|f|*|g|)|^p(x) \le ||f||_1^{\frac{p}{q}}\; |(|f|*|g|^p)(x)|. \tag{10.19}$$

D'où $|(f*g)|^p$ est dans $L^1(\mathbb{R})$ avec :

$$||\,|f*g|^p||_1 \le ||f||_1^{\frac{p}{q}}\; ||f||_1\; ||\,|g|^p||_1. \tag{10.20}$$

Comme $1+\frac{p}{q}=p$, on a (10.16). (Démonstration similaire dans $\mathbb{R}^n$.) ∎

10.4 Dérivation et convolution

Proposition 10.17 *Si $f \in L^1(\mathbb{R})$, si $g \in L^p(\mathbb{R})$ pour un $p \in [1,\infty]$, si g est dérivable dans $\mathbb{R}$, et si $g' \in L^\infty(\mathbb{R})$ (i.e. g' est bornée), alors $f*g$ est dérivable dans $\mathbb{R}$ et :*

$$(f*g)' = f*g'. \tag{10.21}$$

Preuve. Les hypothèses indiquent que $f*g$ et $f*g'$ ont un sens.

(10.21) signifie $\dfrac{d}{dx}\Big(\displaystyle\int_{t\in\mathbb{R}} f(t)g(x-t)\,dt\Big) = \displaystyle\int_{t\in\mathbb{R}} \frac{\partial}{\partial x}\Big(f(t)g(x-t)\Big)\,dt$. C'est vrai grâce au théorème de convergence dominée : l'intégrant $h(x,t)=f(t)g(x-t)$ est dérivable en x (car g l'est), de dérivée $\frac{\partial h}{\partial x}(x,t) = f(t)g'(x-t)$, et $|\frac{\partial h}{\partial x}(x,t)| \le ||g'||_\infty |f(t)|$, avec $||g'||_\infty |f| \in L^1(\mathbb{R})$ fonction dominante intégrable indépendante de x. ∎

10.5 Stabilité de $L^1_{\text{loc}}(\mathbb{R})$ par convolution "bornée"

Le résultat suivant sera généralisé à la convolution des distributions.

Proposition 10.18 *Soit $f,g \in L^1_{\text{loc}}(\mathbb{R})$.*
1- Si $\text{supp} g$ est compact, alors $f*g \in L^1_{\text{loc}}(\mathbb{R})$.
2- Si $\text{supp} f$ et $\text{supp} g$ sont tous deux limités à gauche (ou tous deux limités à droite), alors $f*g \in L^1_{\text{loc}}(\mathbb{R})$.
3- Les hypothèses $f \in L^1_{\text{loc}}(\mathbb{R})$ et $g \in L^1(\mathbb{R})$ sont insuffisantes.

Preuve. Pour tout $\alpha<\beta$ on veut $f*g \in L^1([\alpha,\beta])$, i.e. $\int_{x=\alpha}^{\beta}(\int_{t=a}^{b} |g(t)f(x-t)|\,dt)\,dx < \infty$. On a :

$$\int_{x=\alpha}^{\beta} |(f*g)(x)|\,dx \le \int_{x=\alpha}^{\beta}(\int_{t=a}^{b} |g(t)f(x-t)|\,dt)\,dx \tag{10.22}$$

1- g à support compact, donc il existe $a,b \in \mathbb{R}$ t.q. $\text{supp}(g) \subset [a,b]$ et $g \in L^1(\mathbb{R})$. Pour inverser l'ordre des intégrations dans (10.22), appliquons le théorème de Fubini–Tonelli.

À t fixé, $\int_{x=\alpha}^{\beta} |g(t)f(x-t)|\,dx = |g(t)|\int_{x=\alpha}^{\beta} |f(x-t)|\,dx = |g(t)|\int_{x=\alpha-t}^{\beta-t} |f(y)|\,dy < \infty$ car $f \in L^1_{\text{loc}}(\mathbb{R})$. Et pour $t \in [a,b]$ on a $\alpha-t \ge \alpha-b$ et $\beta-t \le \beta-a$, donc $\int_{x=\alpha-t}^{\beta-t} |f(y)|\,dy \le \int_{x=\alpha-b}^{\beta-a} |f(y)|\,dy$. Donc

$\int_{t=a}^{b}\int_{x=\alpha}^{\beta}|g(t)f(x-t)|\,dxdt \le \int_{t=a}^{b}|g(t)|\int_{x=\alpha-b}^{\beta-a}|f(y)|\,dy\,dt = \int_{t=a}^{b}|g(t)|\,dt\int_{x=\alpha-b}^{\beta-a}|f(y)|\,dy < \infty$ car $f,g \in L^1_{\text{loc}}(\mathbb{R})$. Donc on peut échanger l'ordre des intégrations dans (10.22) :

$$\begin{aligned}\int_{x=\alpha}^{\beta}|(f*g)(x)|\,dx &\le \int_{t=a}^{b}|g(t)|(\int_{x=\alpha}^{\beta}|f(x-t)|\,dx)\,dt = \int_{t=a}^{b}|g(t)|\int_{y=\alpha-t}^{\beta-t}|f(y)|\,dy\,dt\\ &\le \int_{t=a}^{b}|g(t)|\int_{y=\alpha-b}^{\beta-a}|f(y)|\,dy\,dt \le ||g||_{L^1}\int_{y=\alpha-b}^{\beta-a}|f(y)|\,dy < \infty.\end{aligned}$$

Vrai pour tout α,β, donc $f*g$ est $L^1_{\text{loc}}(\mathbb{R})$.

2- Supports limités à gauche : il existe $a,b \in \mathbb{R}$ t.q. supp$f \subset [a,\infty[$ et supp$g \subset [b,\infty[$. Donc, pour $x \in \mathbb{R}$, on a supp$\tau_x\breve{g} \subset]-\infty,-b+x]$. Donc supp$f \bigcap$ supp$\tau_x\breve{g} \subset [a,-b+x]$, et avec Fubini on a :

$$\begin{aligned}\int_{x=\alpha}^{\beta}|(f*g)(x)|\,dx &\le \int_{x=\alpha}^{\beta}\int_{t=a}^{-b+x}|f(t)\tau_x\breve{g}(t)|\,dt\,dx \le \int_{x=\alpha}^{\beta}\int_{t=a}^{-b+\beta}|f(t)g(x-t)|\,dt\,dx\\ &= \int_{t=a}^{-b+\beta}\int_{x=\alpha}^{\beta}|f(t)g(x-t)|\,dx\,dt = \int_{t=a}^{-b+\beta}\int_{y=\alpha-t}^{\beta-t}|f(t)g(y)|\,dy\,dt\\ &\le \int_{t=a}^{-b+\beta}\int_{y=\alpha+b-\beta}^{\beta-a}|f(t)g(y)|\,dy\,dt \le \int_{t=a}^{-b+\beta}|(f(t)|\,dt\int_{y=\alpha+b-\beta}^{\beta-a}|g(y)|\,dy,\end{aligned}$$

fini car f et g sont $L^1_{\text{loc}}(\mathbb{R})$. Idem pour supports limités à droite.

3- On prend $f(t)=e^{-t}$ sur $\mathbb{R}$ et $g(t)=e^{-t}1_{\mathbb{R}_+}(t)$ donc $g \in L^1(\mathbb{R})$; alors $(f*g)(x)=\int_{\mathbb{R}}g(t)f(x-t)\,dt = \int_0^\infty e^{-t}e^{-(x-t)}\,dt = \int_0^\infty e^{-x}\,dt = \infty$. ▪

10.6 Régularisation par convolution

10.6.1 Régularisation d'une fonction $L^1_{\text{loc}}(\mathbb{R})$

On rappelle que si Ω est un ouvert dans $\mathbb{R}^n$, $\mathcal{D}(\Omega)=\{f \in C^\infty(\Omega) : \text{supp}f \text{ compact}\}$, où le support de f, noté supp(f), est l'adhérence de l'ensemble $\{f \ne 0\} = \{\vec{x} \in \Omega \;:\; f(\vec{x}) \ne 0\}$. En particulier si $f \in \mathcal{D}(\Omega)$ alors $f \in C^\infty(\Omega)$ et f est nulle en dehors d'un compact dans Ω.

Proposition 10.19 *Si $f \in L^1_{\text{loc}}(\mathbb{R})$, si $\varphi \in \mathcal{D}(\mathbb{R})$, alors $f*\varphi \in C^\infty(\mathbb{R})$.*
Si de plus supp$f$ *est compact alors $f*\varphi \in \mathcal{D}(\mathbb{R})$.*

Preuve. On pose $z=f*g$, i.e. $z(x)=\int_{t\in\mathbb{R}}h(x,t)\,dt$ où $h(x,t)=f(t)g(x-t)$ pour $x,t \in \mathbb{R}$ (l'intégrant).

1- Cas simple $f \in L^1(\mathbb{R})$. 11- Soit t fixé ; $h_t : x \to h_t(x) := h(x,t)$ est C^∞ sur $\mathbb{R}$ car g l'est. 12- $|\frac{\partial^k h}{\partial x^k}(x,t)| = |f(t)|\,||g^{(k)}||_\infty$, avec $f \in L^1(\mathbb{R})$, donc h est dominée indépendamment de x par une fonction $L^1(\mathbb{R})$, pour tout $k \in \mathbb{N}$: le thm de cvgce dominée donne le résultat.

2- Cas général $f \in L^1_{\text{loc}}(\mathbb{R})$. Soit $x_0 \in \mathbb{R}$, montrons que z est C^∞ au point x_0.

21- Comme 11-, donc, en particulier, h_t est C^∞ en x_0.

22- Domination indépendante de x dans un voisinage de x_0 : on considère l'intervalle $I=]x_0-1,x_0+1[$. On a $g(x-t)=0$ quand $x-t \notin [a,b]$, donc quand $t-x \notin [-b,-a]$, donc quand $t \notin [x-b,x-a]$. Donc pour tout $x \in I$ on a $g(x-t)=0$ quand $t \notin [x_0-1-b,x_0+1-a] =^{\text{noté}} J$ (borné). Donc $|\frac{\partial^k h}{\partial x^k}(t,x)| = |f(t)g(x-t)1_J(t)| \le ||g^{(k)}||_\infty|f(t)1_J(t)|$. Ayant $f \in L^1_{\text{loc}}(\mathbb{R})$, on a $f1_J \in \mathcal{L}^1(J)$, domination indépendante de x dans I.

Donc, thm de cvgce dominée pour $z : x \in I \to z(x)$: la fonction z est C^∞ dans I, donc en particulier en x_0. Vrai pour tout $x_0 \in \mathbb{R}$, donc $z=f*g$ est C^∞ dans $\mathbb{R}$.

3- Et si supp$f$ est borné, alors supp$(f*\varphi)$ est borné, cf. (10.10), donc $f*\varphi \in \mathcal{D}(\mathbb{R})$. ▪

Exercice 10.20 Montrer que si $f \in L^1_{\text{loc}}(\mathbb{R})$, si $g \in C^\infty(\mathbb{R})$, si supp$f$ et supp$g$ sont tous deux limités à gauche (ou tous deux limités à droite), alors $f*g \in C^\infty(\mathbb{R})$. ▪

10.6.2 Suite régularisante ou approximation de l'identité

Définition 10.21 Une fonction intégrable f est dite de masse unité ssi $\int f(x)\,dx = 1$ (souvent défini avec l'hypothèse supplémentaire $f \ge 0$).

Définition 10.22 On appelle suite régularisante une suite $(\varphi_k)_{\mathbb{N}^*}$ de $\mathcal{D}(\mathbb{R})$ telle que :

$$\begin{cases} \varphi_k(x) \geq 0, \quad \forall x \in \mathbb{R}, \\ \mathrm{supp}(\varphi_k) \subset [-\frac{1}{k}, \frac{1}{k}], \\ \int_{\mathbb{R}} \varphi_k(x)\,dx = 1 \qquad \text{(masse unité)}. \end{cases} \tag{10.23}$$

Définition similaire dans $\mathbb{R}^n$ où $[-\frac{1}{k}, \frac{1}{k}]$ est remplacé par la boule de centre 0 et de rayon $\frac{1}{k}$.

(On verra que $(\varphi_k)_{\mathbb{N}^*}$ approxime la masse de Dirac au sens des distributions, et la masse de Dirac est l'identité du produit de convolution, d'où le nom "approximation de l'identité".)

Soit ζ la fonction de $\mathcal{D}(\mathbb{R})$ définie par :

$$\zeta(x) = \begin{cases} \exp(-\frac{1}{1-x^2}), & \forall x \in]-1,1[, \\ 0, & \forall x \notin]-1,1[. \end{cases} \tag{10.24}$$

On pose :

$$\gamma_1(x) = \frac{\zeta(x)}{||\zeta||_{L^1}}, \qquad \text{puis} \qquad \gamma_k(x) = k\,\gamma_1(kx), \quad k \geq 1. \tag{10.25}$$

Proposition 10.23 *La suite $(\gamma_k)_{k\in\mathbb{N}^*}$ est une suite régularisante.*

Preuve. Comme $\zeta \geq 0$, on a $\gamma_1 \geq 0$.

Comme $\zeta \geq 0$ et ζ non identiquement nulle, on a $||\zeta||_{L^1} = \int_{\mathbb{R}} |\zeta(x)|\,dx = \int_{\mathbb{R}} \zeta(x)\,dx > 0$.

Donc $\int_{\mathbb{R}} \gamma_1(x)\,dx = \frac{1}{||\zeta||_{L^1}} \int_{\mathbb{R}} \zeta(x)\,dx = \frac{1}{||\zeta||_{L^1}} ||\zeta||_{L^1} = 1$.

Donc $\int_{\mathbb{R}} \gamma_k(x)\,dx = \int_{\mathbb{R}} k\,\gamma_1(kx)\,dx = \int_{\mathbb{R}} \gamma_1(y)\,dy = 1$.

Comme $\gamma_1 > 0$ et $k \geq 0$ on a $\gamma_k \geq 0$.

Et $\gamma_1(x) \neq 0$ ssi $k \in]-1,1[$. Donc $\gamma_k(x) \neq 0$ ssi $kx \in]-1,1[$. D'où $\mathrm{supp}(\gamma_k) = [-\frac{1}{k}, \frac{1}{k}]$. ∴

10.6.3 Régularisation C^∞ des fonctions $1_{[a,\infty]}$ et $1_{[a,b]}$

Soit (φ_k) une suite régularisante.

Proposition 10.24 *Soit $a, b \in \mathbb{R}$, $b > a$: dès que k est assez grand, à savoir dès que $\frac{1}{k} \leq \frac{b-a}{2}$, $1_{[a,b]} * \varphi_k \in \mathcal{D}(\mathbb{R})$ et, pour $x \in \mathbb{R}$:*

$$\begin{cases} 0 \leq (1_{[a,b]} * \varphi_k)(x) \leq 1, \\ (1_{[a,b]} * \varphi_k)(x) = 1 \text{ pour } x \in [a+\frac{1}{k}, b-\frac{1}{k}], \\ \mathrm{supp}(1_{[a,b]} * \varphi_k) = [a-\frac{1}{k}, b+\frac{1}{k}]. \end{cases} \tag{10.26}$$

On conserve ce résultat si on ouvre en a ou b. Cas particulier $(\varphi_k) = (\gamma_k)$ cf. (10.25) : de plus $\varphi(a) = \varphi(b) = \frac{1}{2}$.

Et $1_{[a,\infty]} * \varphi_k \in C^\infty(\mathbb{R})$ vérifie, pour $x \in \mathbb{R}$:

$$\begin{cases} 0 \leq (1_{[a,\infty]} * \varphi_k)(x) \leq 1, \\ (1_{[a,\infty]} * \varphi_k)(x) = 1 \text{ pour } x \in [a+\frac{1}{k}, \infty[, \\ \mathrm{supp}(1_{[a,\infty]} * \varphi_k) = [a-\frac{1}{k}, \infty[. \end{cases} \tag{10.27}$$

On conserve ce résultat si on ouvre en a (i.e. sur l'intervalle $]a,\infty[$).

Dans $\mathbb{R}^n$ avec K compact : la fonction $1_K * \varphi_k$ est également dans $\mathcal{D}(\mathbb{R}^n)$, avec $0 \leq \varphi \leq 1$, avec $\mathrm{supp}(\varphi) \subset K + B(\vec{0}, \frac{1}{k})$, et avec $\varphi = 1$ sur $K - B(\vec{0}, \frac{1}{k})$, où $B(\vec{0}, \frac{1}{k})$ est la boule unité de centre $\vec{0}$ et rayon $\frac{1}{k}$.

Preuve. Cas $[a,b]$ (exercice pour $[a,\infty[$ et intervalles ouverts et semi-fermés). On a $\psi_k = 1_{[a,b]} * \varphi_k \in \mathcal{D}(\mathbb{R})$, cf. prop. 10.19. On a $\mathrm{supp}\tau_x(\widetilde{1_{[a,b]}}) = [x-b, x-a]$ et $\mathrm{supp}\varphi_k = [-\frac{1}{k}, \frac{1}{k}]$, donc :

$$\psi_k(x) = \int_{t\in\mathbb{R}} \varphi_k(t)\tau_x(\widetilde{1_{[a,b]}})(t)\,dt = \int_{t\in[x-b,x-a]\bigcap[-\frac{1}{k},\frac{1}{k}]} \varphi_k(t)\,dt, \tag{10.28}$$

et la fonction φ_k est positive et $\int_{\mathbb{R}} \varphi_k = 1$. D'où $0 \leq \psi_k \leq 1$.

Et si $x-b > \frac1k$ ou si $x-a < -\frac1k$, alors $\psi_k(x) = \int_\emptyset ... = 0$, d'où $\mathrm{supp}(\psi_k) \in [a-\frac1k, b+\frac1k]$.

Et si $x-b < -\frac1k$ et si $x-a > \frac1k$, alors $[x-b, x-a] \subset [-\frac1k, \frac1k]$, donc $\psi_k(x) = 1$.

Cas particulier $(\varphi_k) = (\gamma_k)$: on a $\psi_k(a) = \int_{u\in[a-b,0]\bigcap[-\frac1k,\frac1k]} \gamma_k(u)\,du = \int_{u\in[\frac{-1}{k},0]} \gamma_k(u)\,du = \frac12$, dès que $\frac1k < b-a$, car γ_k est paire et $\int_{u\in[\frac{-1}{k},\frac1k]} \gamma_k(u)\,du = 1$. Idem : $\psi_k(b) = \frac12$.

Exercice dans $\mathbb{R}^n$. ∎

Exercice 10.25 Soit $f : x \in]0,\infty[\to f(x) = \frac1x \in \mathbb{R}$. Donner une fonction $g \in C^\infty([0,\infty[\to \mathbb{R}$ telle que $g(0) = 0$ et $g(x) = f(x)$ pour $x \ge 1$. Dessin.

Réponse. Troncature régulière de f : on pose $g = \psi f$ avec $g(0) = 0$ et $\psi = \varphi_{\frac14} * 1_{[\frac12,\infty]}$ (régularisée C^∞ de $1_{[\frac12,\infty]}$, la fonction ψ valant 1 sur $[\frac34,\infty[$). ∎

Exercice 10.26 Donner une fonction $f \in \mathcal{D}(\mathbb{R})$ telle que $f = \exp$ sur $[-1,1]$ et $0 \le f \le \exp$ sur $\mathbb{R}$, où $\exp : x \to e^x \in C^\infty(\mathbb{R})$ est la fonction exponentielle.

Réponse. On "tronque de manière régulière" la fonction exp : on pose $g = 1_{[-2,2]} * \varphi_1$. Avec (10.26) on a $g \in \mathcal{D}(\mathbb{R})$ et $g(x) = 1$ sur $[-1,1]$ et $0 \le g(x) \le 1$. Cette fonction g est notre fonction de "troncature régulière". On pose $f(x) = \exp(x)g(x)$: la fonction f convient, car produit de deux fonctions $C^\infty(\mathbb{R})$, donc est $C^\infty(\mathbb{R})$, et $\mathrm{supp}g$ est borné, et trivialement $\mathrm{supp}f \subset \mathrm{supp}g$, donc $f \in \mathcal{D}(\mathbb{R})$. ∎

Corollaire 10.27 *Soit $a < b \in \overline{\mathbb{R}}$, soit $c < d \in \mathbb{R}$. Si $[c,d] \subset]a,b[$ alors il existe une fonction $\varphi \in \mathcal{D}(]a,b[)$ qui vaut 1 sur $[c,d]$, et telle que $0 \le \varphi \le 1$. (Dessin).*

Dans $\mathbb{R}^n$: si Ω est un ouvert de $\mathbb{R}^n$ et K un compact tel que $K \subset \Omega$, alors il existe une fonction $\varphi \in \mathcal{D}(\Omega)$ qui vaut 1 sur K et telle que $0 \le \varphi \le 1$.

Preuve. Soit (γ_k) une suite régularisante. Soit $\varepsilon = \frac12 \min(c-a, b-d)$ (dessin), soit $e = c-\varepsilon$ et $f = d+\varepsilon$. Soit k t.q. $\frac1k \le \frac{f-e}{2}$ et $\frac1k < \varepsilon$. La fonction $\varphi = 1_{[e,f]} * \varphi_k$ convient, cf. proposition précédente.

Dans $\mathbb{R}^n$: soit $K = \mathrm{supp}\varphi$ et soit $\varepsilon = d(K, \mathbb{R}^n - \Omega) > 0$ la distance de K à $\mathbb{R}^n - \Omega$. Soit $K_\varepsilon = K + \overline{B}(0, \frac\varepsilon2)$... on continue comme précédemment avec la fonction $\varphi = 1_{K_\varepsilon} * \gamma_k$. ∎

Corollaire 10.28 *Soit (φ_k) une suite régularisante. Soit $x_0 \in \mathbb{R}$.*

1- Soit $f \in C^\infty(\mathbb{R})$. Alors pour $r \in \mathbb{R}_+^*$ et $k \in \mathbb{N}^*$ t.q. $\frac1k < r$, la fonction produit :

$$f_{r,k} \stackrel{\text{déf}}{=} f\,(1_{]x_0-r,x_0+r[} * \varphi_k), \tag{10.29}$$

est dans $\mathcal{D}(\mathbb{R})$ et est égale à f dans un voisinage de x_0. Plus précisément on a $f_{r,k} = f$ sur $]x_0-r+\frac1k, x_0+r-\frac1k[$ avec $\mathrm{supp}f_{r,k} \subset]x_0-r-\frac1k, x_0+r+\frac1k[$.

De plus, si f est bornée, alors $||f - f_{r,k}||_\infty \le ||f||_\infty$.

2- Plus généralement, soit $\varepsilon > 0$, et soit $f \in C^\infty([x_0-\varepsilon, x_0+\varepsilon])$. Alors pour $r \in \mathbb{R}$ et $k \in \mathbb{N}^*$ t.q. $0 < r < \frac\varepsilon2$ et $\frac1k < r$, on a $f_{r,k} \in \mathcal{D}(\mathbb{R})$ et $f_{r,k} = f$ dans un voisinage de x_0. Plus précisément on a $f_{r,k} = f$ sur $]x_0-r+\frac1k, x_0+r-\frac1k[$ avec $\mathrm{supp}f_{r,k} \subset]x_0-r-\frac1k, x_0+r+\frac1k[$.

3- De plus, si f est bornée, alors $||f_{r,k}||_{L^\infty(]x_0-\varepsilon,x_0+\varepsilon[)} \le ||f||_{L^\infty(]x_0-\varepsilon,x_0+\varepsilon[)}$.

Ainsi que $||f - f_{r,k}||_{L^\infty(]x_0-\varepsilon,x_0+\varepsilon[)} \le ||f||_{L^\infty(]x_0-\varepsilon,x_0+\varepsilon[)}$

Preuve. 1- Soit $\psi_{r,k} = 1_{]x_0-r,x_0+r[} * \varphi_k$. On a $\psi_{r,k} \in C^\infty(\mathbb{R})$, donc $f_{r,k} \in C^\infty(\mathbb{R})$.. Soit $I_- =]x_0-r+\frac1k, x_0+r-\frac1k[$ et soit $I_+ =]x_0-r-\frac1k, x_0+r+\frac1k[$. On a $\psi_{r,k} = 1$ dans I_-, donc $f_{r,k} = f$ dans I_- et $\psi_{r,k} = 0$ dans I_+, donc $f_{r,k} = 0$ dans $\mathbb{R} - I_+$. D'où 2-.

3- Et $0 \le 1_{]x_0-r,x_0+r[} * \varphi_k \le 1$, donc $0 \le |f_{r,k}(x)| \le |f(x)|$ et $|f(x) - f_{r,k}(x)| \le |f(x)|$. ∎

Exercice 10.29 Soit f en escalier avec $\mathrm{supp}f$ borné. Soit $(\varphi_k)_{\mathbb{N}^*}$ une suite régularisante. Alors pour k assez grand on a $f * \varphi_k \in \mathcal{D}(\mathbb{R})$ avec $||f * \varphi_k||_\infty \le ||f||_\infty$ et $||f - f * \varphi_k||_\infty \le ||f||_\infty$.

Réponse. Ici $\exists n \in \mathbb{N}^*$, $\exists a_1, ..., a_n, c_1, ..., c_n \in \mathbb{R}$ avec $a_1 < ... < a_n$, $f = \sum_{i=1}^{n-1} c_i 1_{[a_i,a_{i+1}]}$. On prend $\frac1k \le \min_i(\frac{a_{i+1}-a_i}{2})$. Et $f * \varphi_k$ vérifie les propriétés demandées (démarche de la prop. 10.24). ∎

10.7 $\mathcal{D}(\mathbb{R}) = C_c^\infty(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$ pour $1 \le p < \infty$

Soit $\mathbb{R}$ muni de sa tribu borélienne $\mathcal{A}_\mathbb{R}$ (engendrée par les intervalles ouverts).

Soit $a, b \in \mathbb{R}$, $a < b$. Soit $(\varphi_k)_{k\in\mathbb{N}}$ une suite régularisante.

10.7.1 $1_{[a,b]}$ et convergence p.p. des régularisées

Proposition 10.30 *On a la convergence simple presque partout :*

$$\varphi_k * 1_{[a,b]} \underset{k\to\infty}{\longrightarrow} 1_{[a,b]} \quad \textit{presque partout.} \tag{10.30}$$

Preuve. Soit $x \in \mathbb{R} - \{a, b\}$, et $d(x) = \min(d(x, a), d(x, b)) > 0$ (dessin). Soit k t.q. $\frac{1}{k} < d(x)$, i.e. $k > \frac{1}{d(x)}$. On a $(\varphi_k * 1_{[a,b]})(x) = 1_{[a,b]}(x)$, cf. prop. 10.24, donc $|\varphi_k * 1_{[a,b]})(x) - 1_{[a,b]}(x)| \longrightarrow_{k\to\infty} 0$. D'où (10.30). ▪

10.7.2 $1_{[a,b]}$ et convergence L^p des régularisées pour $p \in [1, \infty[$

Proposition 10.31 *Pour $p \in [1, \infty[$ on a la convergence dans $L^p(\mathbb{R})$:*

$$\varphi_k * 1_{[a,b]} \underset{k\to\infty}{\longrightarrow} 1_{[a,b]} \quad \textit{dans } L^p(\mathbb{R}), \tag{10.31}$$

i.e. $||1_{[a,b]} - \varphi_k * 1_{[a,b]}||_{L^p(\mathbb{R})} \underset{k\to\infty}{\longrightarrow} 0$. Pour $p = \infty$ (cas $L^\infty(\mathbb{R})$) c'est faux.

On conserve ce résultat pour $g = \sum_{i=1}^n c_i\, 1_{[a_i,b_i]} \in L^p(\mathbb{R})$ fonction en escalier à support compact.

Preuve. Cas $p = 1$. On a $1_{[a,b]} = \varphi_k * 1_{[a,b]}$ sauf sur $K = [a-\frac{1}{k}, a+\frac{1}{k}] \bigcup [b-\frac{1}{k}, b+\frac{1}{k}]$ (pour $k > \frac{1}{2(b-a)}$) ensemble de longueur $|K| = \frac{4}{k}$ sur lequel $|1_{[a,b]}(x) - (\varphi_k * 1_{[a,b]})(x)| \le 1$. Donc $||1_{[a,b]} - \varphi_k * 1_{[a,b]}||_{L^1(\mathbb{R})} \le \displaystyle\int_K dx = \frac{4}{k} \underset{k\to\infty}{\longrightarrow} 0$.

Cas $1 < p < \infty$. Calcul similaire avec $|(1_{[a,b]}(x) - \varphi_k * 1_{[a,b]}(x))^p| \le 1$, et même conclusion.

Cas $p = \infty$. Comme $\varphi_k * g \in C^0(\mathbb{R})$, on a $||\varphi_k * \varphi||_\infty = \sup_{x\in\mathbb{R}} |(\varphi_k * \varphi)(x)|$. Prenons la suite régularisante $\varphi_k = \gamma_k$, cf. (10.25). Soit $k > \frac{1}{2(b-a)}$. On a $(\varphi_k * 1_{[a,b]})(b) = \frac{1}{2}$, cf. (10.28). Donc $|1_{[a,b]}(b) - (\varphi_k * 1_{[a,b]})(b)| = \frac{1}{2}$, donc $||1_{[a,b]} - \varphi_k * 1_{[a,b]}||_\infty$ ne tend pas vers 0 quand $k \to \infty$. On ne converge pas dans $L^\infty(\mathbb{R})$.

Et une fonction en escalier est une somme finie de fonctions indicatrices d'intervalles. ▪

10.7.3 $C^0(K)$ est dense dans $L^p(K)$, pour $1 \le p < \infty$

Soit K un compact dans $\mathbb{R}$. Soit $C^0(K) = \{f \in C^0(\mathbb{R}) : \text{supp}(f) \subset K\}$ (les fonctions continues à support compact dans K).

Proposition 10.32 *$C^0(K)$ est dense dans $L^p(K)$, pour $1 \le p < \infty$.*

Preuve. Soit $\mathcal{A}_K = \mathcal{A}_\mathbb{R} \cap K$ (la restriction de $\mathcal{A}_\mathbb{R}$ à K. Soit $\mathcal{T}$ la sous-tribu de $\mathcal{A}_K$ engendrée par les $A \in \mathcal{A}_K$ t.q. 1_A est limite d'une suite de fonctions $f_n \in C^0(K)$ t.q. $0 \le f_n \le 1$ (où donc $||1_A - f_n||_{L^p} \longrightarrow_{n\to\infty} 0$).

Montrons que $\mathcal{T}$ contient les ouverts de K : soit U un ouvert dans K ; soit $f_n : K \to \mathbb{R}$ définie par $f_n(t) = \min(1, n\, d(t, U^c))$ où $d(t, U^c)$ est la distance de t au complémentaire de U. Les f_n sont continues (immédiat), forment une suite croissante (immédiat), et $0 \le f_n \le 1$ (immédiat). Et $f_n \to 1_U$ dans $L^p(K)$ car $f_n(t) \longrightarrow_{n\to\infty} 1_U(t)$ pour presque tout $t \in K$ (convergence p.p.) et $|1_U(x) - f_n(x)|^p < 1$ (domination indépendante de n) : on peut appliquer le théorème de convergence dominée.

Donc $\mathcal{T} = \mathcal{A}_K$ (= la tribu engendrée par les ouverts). Donc pour tout borélien $B \in \mathcal{A}_K$ la fonction 1_B est limite dans $L^p(K)$ d'une suite de fonctions $f_n \in C^0(K)$ t.q. $0 \le f_n \le 1$. Par linéarité, toute fonction étagée dans K est limite dans $L^p(K)$ d'une suite de fonctions $f_n \in C^0(K)$. Donc, par construction de $L^p(K)$, toute fonction $f \in L^p(K)$ est limite dans $L^p(K)$ d'une suite de fonctions $f_n \in C^0(K)$. ▪

10.7.4 $C_c^0(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$, pour $1 \le p < \infty$

Soit $C_c^0(\mathbb{R}) = \{f \in C^0(\mathbb{R}) : \text{supp}(f) \text{ compact}\}$ (les fonctions continues à support compact dans $\mathbb{R}$).

Proposition 10.33 *$C_c^0(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$, pour $1 \le p < \infty$.*

Preuve. Soit $f \in L^p(\mathbb{R})$, donc $\forall \varepsilon > 0$, $\exists R > 0$, $(\int_{x\notin[-R,R]} |f(x)|^p\, dx)^{\frac{1}{p}} < \varepsilon$. Soit $f_R = 1_{[-R,R]} f$, donc $||f - f_R||_{L^p} = (\int_{x\in\mathbb{R}} |f(x) - f_R(x)|^p\, dx)^{\frac{1}{p}} < \varepsilon$. Avec la proposition 10.32, ayant $f_R \in L^p([-R, R])$, soit $g_R \in C^0([-R, R]) \subset C_c^0(\mathbb{R})$ t.q. $||g_R - f_R||_{L^p} < \varepsilon$: on a $||f - g_R||_{L^p} \le ||f - f_R||_{L^p} + ||f_R - g_R||_{L^p} < 2\varepsilon$. ▪

10.7.5 $\mathcal{D}(\mathbb{R})$ est dense dans $C_c^0(\mathbb{R})$ au sens $L^p(\mathbb{R})$, donc $\mathcal{D}(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$

Proposition 10.34 *$\mathcal{D}(\mathbb{R})$ (contenu dans $C_c^0(\mathbb{R})$) est dense dans $C_c^0(\mathbb{R})$ au sens $L^p(\mathbb{R})$: $\forall f \in C_c^0(\mathbb{R})$, $\forall \varepsilon > 0$, $\exists \varphi \in \mathcal{D}(\mathbb{R})$ t.q. $||f-\varphi||_{L^p} < \varepsilon$. Donc $\mathcal{D}(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$.*

Preuve. Soit $f \in C_c^0(\mathbb{R})$, soit $R > 0$ t.q. $\mathrm{supp}(f) \subset [-R, R]$. Donc f est uniformément continue sur $[-R, R]$, donc f est limite d'une suite de fonctions en escalier : $\forall \varepsilon > 0$, $\exists n \in \mathbb{N}^*$, $\exists c_1, ..., c_n \in \mathbb{R}$, $\exists a_1, ..., a_n, b_1, ..., b_n \in [-R, R]$, $\forall x \in \mathbb{R}$, $|f(x) - \sum_{i=1}^n c_i\, 1_{[a_i,b_i]}(x)| < \varepsilon$. Et pour k assez grand, $|\sum_{i=1}^n c_i\, 1_{[a_i,b_i]}(x) - \sum_{i=1}^n c_i\, (\varphi_k * 1_{[a_i,b_i]})(x)| < \varepsilon$, donc $|f(x) - \sum_{i=1}^n c_i(\varphi_k * 1_{[a_i,b_i]})(x)| < 2\varepsilon$. Donc $\int_{\mathbb{R}} |f(x) - \sum_{i=1}^n c_i(\varphi_k * 1_{[a_i,b_i]})(x)|^p\, dx \le (2\varepsilon)^p (2R)$, donc $||f - \sum_{i=1}^n c_i(\varphi_k * 1_{[a_i,b_i]})||_{L^p} \le 2\varepsilon (2R)^{\frac{1}{p}}$.

Et les propositions 10.33 et 10.34 donnent : $\mathcal{D}(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$. ∎

Exercice 10.35 Rappel : montrer que $(C_b^0(\mathbb{R}), ||.||_\infty)$ est un espace de Banach (résultat classique).

Réponse. C'est un espace vectoriel (immédiat : sous-espace vectoriel de $\mathcal{F}(\mathbb{R};\mathbb{R})$). Soit $(f_n)_{\mathbb{N}}$ une suite de Cauchy dans $C_b^0(\mathbb{R})$, donc $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, $\forall m, n \ge N$, $||f_n - f_m||_\infty \le \varepsilon$. Donc, $|f_n(x) - f_m(x)| \le \varepsilon$ pour tout x, et $\mathbb{R}$ complet, donc $(f_n(x))_{\mathbb{N}}$ est de Cauchy dans $\mathbb{R}$ donc convergente vers un réel qu'on note $f(x)$, ce qui définit la fonction f, avec $|f(x) - f_N(x)| \le \varepsilon$ pour tout x, donc $||f - f_N||_\infty \le \varepsilon$. Et $|f(x)| \le |f(x) - f_N(x)| + |f_N(x)| \le ||f - f_N||_\infty + ||f_N||_\infty + \le \varepsilon + ||f_N||_\infty$, pour tout x, avec $||f_N||_\infty < \infty$, donc f est bornée. Et $|f(y) - f(x)| \le |f(y) - f_N(y)| + |f_N(y) - f_N(x)| + |f_N(x) - f(x)| \le 2||f - f_N||_\infty + |f_N(y) - f_N(x)|$, et on choisit $\eta > 0$ t.q. $|f_N(y) - f_N(x)| \le \varepsilon$ pour tout y t.q. $|y - x| < \eta$, donc $|f(y) - f(x)| \le 3\varepsilon$ pour tout $y \in]x-\varepsilon, x+\varepsilon[$, donc f est continue en x, vrai pour tout x. ∎

10.8 Lemme de Lebesgue

Un résultat de convergence qu'on n'obtient pas avec le théorème de convergence dominée, et qui utilise la densité de $\mathcal{D}(\mathbb{R})$ dans $L^1(\mathbb{R})$:

Lemme 10.36 *Si $f \in L^1(\mathbb{R})$ alors $\displaystyle\lim_{t\to\infty} \int_{x\in\mathbb{R}} f(x)\sin(tx)\, dx = 0$. Interprétation : dès que la fonction sinus "oscille assez vite" (i.e. t "assez grand") l'intégrale (valeur moyenne) est proche de 0 (dessin).*

Preuve. (Ici, à x fixé, $g_x(t) = f(x)\sin(tx)$ ne converge pas quand $t \to \infty$: passer à la limite sous le signe $\int$ n'a pas de sens.)

1- Pour $f = 1_{[a,b]}$, où $a < b$, on a $\displaystyle\int_a^b \sin(tx)\, dx = [-\frac{\cos(tx)}{t}]_{x=a}^b = \frac{\cos(ta) - \cos(tb)}{t} \underset{t\to\infty}{\longrightarrow} 0$.

2- Donc pour g en escalier on a $\displaystyle\int_{x\in\mathbb{R}} g(x)\sin(tx)\, dx \underset{t\to\infty}{\longrightarrow} 0$, comme somme finie d'intégrales convergeant vers 0. Donc $\forall \varepsilon > 0$, $\exists T_\varepsilon > 0$, $\forall t > T_\varepsilon$, $\displaystyle\int_{x\in\mathbb{R}} g(x)\sin(tx)\, dx < \varepsilon$.

3- Soit $f \in \mathcal{D}(\mathbb{R})$ (donc en particulier continue). Donc, pour $\varepsilon > 0$, $\exists g$ en escalier t.q. $||f - g||_{L^1} < \varepsilon$. Le 2- indique qu'il existe T_ε t.q. pour tout $t \ge T_\varepsilon$ on a $|\displaystyle\int_{x\in\mathbb{R}} g(x)\sin(tx)\, dx| < \varepsilon$. D'où, pour tout $t > T_\varepsilon$:

$$\begin{aligned} |\int_{x\in\mathbb{R}} f(x)\sin(tx)\, dx| &\le \int_{x\in\mathbb{R}} |f(x) - g(x)|\, dx + |\int_{x\in\mathbb{R}} g(x)\sin(tx)\, dx| \\ &\le ||f - g||_{L^1} + \varepsilon \le 2\varepsilon. \end{aligned}$$

4- Puis $\mathcal{D}(\mathbb{R})$ est dense dans $L^1(\mathbb{R})$, d'où le résultat en reprenant la démarche du 3-. ∎

10.9 Partition de l'unité

10.9.1 $1_{\mathbb{R}}$ comme somme de régularisées (partition de l'unité de $\mathbb{R}$)

On rappelle que $\tau_c\varphi : x \to \tau_c\varphi(x) = \varphi(x - c)$.

Proposition 10.37 *(Partition de l'unité de $\mathbb{R}$.) Soit $(\gamma_n)_{\mathbb{N}}$ la suite régularisante paire donnée par (10.25). Soit $a, b \in \mathbb{R}$ t.q. $a < b$, et on fixe $n \in \mathbb{N}$ t.q. $\frac{1}{n} < \frac{b-a}{2}$. On pose :*

$$\varphi = \gamma_n * 1_{[a,b]}, \tag{10.32}$$

la régularisée de $1_{[a,b]}$. En particulier $\varphi = 1$ sur $[a+\frac{1}{n}, b-\frac{1}{n}]$ et $\mathrm{supp}\varphi = [a-\frac{1}{n}, b+\frac{1}{n}]$.

Soit $d=b-a$ (distance de a à b = largeur de l'intervalle $[a,b]$). On a :

$$\begin{cases} \varphi+\tau_d\varphi=1 \quad \text{sur} \quad [a+\frac{1}{n},b+d-\frac{1}{n}], \quad \text{et} \quad \mathrm{supp}(\varphi+\tau_d\varphi)=[a-\frac{1}{n},b+d+\frac{1}{n}], \\ \tau_{-d}\varphi+\varphi=1 \quad \text{sur} \quad [a-d+\frac{1}{n},b-\frac{1}{n}], \quad \text{et} \quad \mathrm{supp}(\tau_{-d}\varphi+\varphi)=[a-d-\frac{1}{n},b+\frac{1}{n}], \end{cases} \tag{10.33}$$

Faire un dessin. Et de même, pour tout $k,\ell\in\mathbb{N}$ où $k<\ell$:

$$\tau_{kd}\varphi+\tau_{(k+1)d}\varphi+\ldots+\tau_{(\ell-1)d}\varphi+\tau_{\ell d}\varphi=1 \qquad \text{sur} \qquad [a+kd+\frac{1}{n},b+\ell d-\frac{1}{n}], \tag{10.34}$$

et de support $[a+kd-\frac{1}{n},b+\ell d+\frac{1}{n}]$. Et donc :

$$\sum_{k\in\mathbb{Z}}\tau_{kd}\varphi=1_{\mathbb{R}}, \tag{10.35}$$

formule de partition de l'unité de $\mathbb{R}$ (la fonction constante $1_{\mathbb{R}}$).

Preuve. On reprend le calcul (10.28), avec $\mathrm{supp}(\tau_d 1_{[a,b]})=\mathrm{supp}(1_{[a+d,b+d]})$. En particulier :

$$\varphi(x)=\int_{t\in[x-b,x-a]}\gamma_n(t)\,dt, \qquad \tau_d\varphi(x)=\varphi(x-d)=\int_{t\in[x-b-d,x-a-d]}\gamma_n(t)\,dt.$$

Et $d>0$, donc :

$$\varphi(x)+\tau_d\varphi(x)=\int_{J_x}\gamma_n(t)\,dt, \qquad J_x=[x-b-d,x-a]\bigcap[-\frac{1}{n},\frac{1}{n}].$$

1- Si $x-a\le-\frac{1}{n}$, soit $x\le a-\frac{1}{n}$, alors $\varphi(x)+\tau_d\varphi(x)=0$.
1'- Si $x-b-d\ge\frac{1}{n}$, soit $x\ge b+d+\frac{1}{n}$, alors $\varphi(x)+\tau_d\varphi(x)=0$.
2- Si $[-\frac{1}{n},\frac{1}{n}]\subset[x-b-d,x-a]$, soit $-\frac{1}{n}\ge x-b-d$ et $\frac{1}{n}\le x-a$, soit $x\in[a+\frac{1}{n},b+d-\frac{1}{n}]$ alors $\varphi(x)+\tau_d\varphi(x)=1$
3- Et dans les autres cas $0\le\varphi(x)+\tau_d\varphi(x)\le 1$.
D'où $(10.33)_1$. Puis de même $(10.33)_2$. D'où (10.34) par récurrence, d'où (10.35).

10.9.2 Partition de l'unité dans $\mathbb{R}^n$

Soit Ω un ouvert dans $\mathbb{R}^n$.

Lemme 10.38 *Soit K un compact contenu dans une réunion finie d'ouverts $\bigcup_{j=1}^m\Omega_j$. Alors il existe des compacts $K_j\subset\Omega_j$ tels que $K\subset\bigcup_{j=1}^m\mathring{K}_j$.*

Preuve. Pour $x\in K$, soit $j_x\in[1,m]_{\mathbb{N}}$ t.q. $x\in\Omega_{j_x}$, et soit r_{j_x} t.q. $B(x,2r_{j_x})\subset\Omega_{j_x}$. Comme $K\subset\bigcup_{x\in K}B(x,r_{j_x})$ et K compact, il existe un sous recouvrement fini $K\subset\bigcup_{k=1}^{\ell}B(x_k,r_{j_{x_k}})$. On pose $K_j=\bigcup_{\substack{k=1,\ldots,\ell\\ x_k\in\Omega_j}}\overline{B(x_k,r_{j_{x_k}})}$, réunion finie de compacts donc compact, et $K\subset\bigcup_{j=1}^m\mathring{K}_j$, avec $K_j\subset\bigcup_{\substack{k=1,\ldots,\ell\\ x_k\in\Omega_j}}B(x_k,2r_{j_{x_k}})\subset\Omega_j$.

Lemme 10.39 *Soit Ω un ouvert et soit un compact $K\subset\Omega$. Soit $f\in C^0(\Omega)$ t.q. $f_{|K}=1$. Alors f est strictement positive dans un voisinage ouvert de K : $\exists\varepsilon>0$, $\forall x\in K+B(0,\varepsilon)$, $f(x)>0$.*

Exercice 10.40 Montrer à l'aide des suites que si K est compact dans Ω ouvert, alors il existe $\varepsilon>0$ t.q. $K+B(0,\varepsilon)\subset\Omega$ (donc que K est à plus d'une distance ε du bord de Ω).

Réponse. Sinon, pour tout ε, en particulier $\varepsilon=\frac{1}{m}$, on a $(K+B(0,\frac{1}{m}))\cap(\mathbb{R}^n-\Omega)\neq\emptyset$. Donc il existe $x_m\in K$ et $z_m\in B(0,\frac{1}{m})$ t.q. $x_m+z_m\in(\mathbb{R}^n-\Omega)$. On a construit une suite $(z_m)_{m\in\mathbb{N}^*}$ dans $\mathbb{R}^n$ qui converge vers 0. Et on a construit une suite $(x_m)_{m\in\mathbb{N}^*}$ dans K, et comme K est compact, la suite $(x_m)_{m\in\mathbb{N}^*}$ a une sous-suite convergente $(x_{m_k})_{k\in\mathbb{N}^*}$ dans K ; notons $x_\infty=\lim_{k\to\infty}x_{m_k}\in K$. Donc la suite $(x_{m_k}+z_{m_k})_{\mathbb{N}^*}$ converge vers $x_\infty+0=x_\infty$; et $(x_{m_k}+z_{m_k})_{\mathbb{N}^*}$ est une suite dans le fermé $(\mathbb{R}^n-\Omega)$ (complémentaire d'un ouvert), donc sa limite $x_\infty\in\mathbb{R}^n-\Omega$. Avec $x_\infty\in\Omega$: absurde.
En particulier $x_\infty\in\Omega$ (car $K\subset\Omega$).

Preuve. D'après le lemme précédent, il existe $r>0$ t.q le compact $K+\overline{B(0,r)}=^{\text{noté}}K_r$ est tout entier dans Ω . Soit $N\in\mathbb{N}^*$ t.q. $\frac{1}{N}<r$. Supposons le lemme faux, i.e. $\forall\varepsilon=\frac{1}{n}$ où $n>N$, $\exists x_n\in K+B(0,\frac{1}{n})$ t.q. $f(x_n)=0$. On a construit une suite $(x_n)_{n>N}$ telle que $f(x_n)=0$ pour tout n. Et $(x_n)_{\mathbb{N}^*}$ appartient au compact K_r, donc on peut extraire une sous suite convergente dans K_r, soit $x_\infty\in K_r$ la limite. Mieux, $x_\infty\in K$ car K est fermé : sinon $x_\infty\in\mathbb{R}^n-K$ ouvert, donc $\exists\varepsilon>0$ t.q. $B(x_\infty,\varepsilon)\subset\mathbb{R}^n-K$, donc $d(x_\infty,K)\ge\varepsilon$, absurde par construction de la suite (x_n). Et comme f est continue et $x_n\to x_\infty$, on a $f(x_\infty)=0$. Et comme $x_\infty\in K$ on a $f(x_\infty)=1$. Absurde car $x_\infty\in K\subset\Omega$: donc le lemme est vrai.

Proposition 10.41 *(Partition de l'unité.) Soit K un compact de $\mathbb{R}^n$ dont on considère un recouvrement fini $\bigcup_{j=1}^m \Omega_j \supset K$, les Ω_j étant des ouverts de $\mathbb{R}^n$.*
Il existe alors m fonctions $\chi_j \in \mathcal{D}(\Omega_j)$ telles que $0 \le \chi_j \le 1$ pour tout $j = 1, ..., m$ et :

$$\chi_1(x) + ... + \chi_m(x) = 1 \qquad \textit{dans un voisinage ouvert de } K. \tag{10.36}$$

Preuve. On applique le lemme 10.38 : soit m compacts $K_j \subset \Omega_j$ t.q. $K \subset \bigcup_{j=1}^m \mathring{K}_j$.
Soit alors $\psi_j \in \mathcal{D}(\Omega_j)$ une fonction qui vaut 1 sur K_j (une telle fonction existe d'après le corollaire 10.27). En particulier $\sum_{i=1}^m \psi_i$ est une fonction C^∞ strictement positive dans un voisinage ouvert de de K. On pose dans $\mathbb{R}^n$:

$$\chi_j(x) = \frac{\psi_j(x)}{\sum_{i=1}^m \psi_i(x)}. \tag{10.37}$$

On vérifie immédiatement que les χ_j conviennent. ∎

10.10 $L^p_{loc}(\mathbb{R})$ et résultat de "projection"

Lemme 10.42 *Soit $1 \le p \le \infty$, et soit $f \in L^p_{loc}(\mathbb{R})$. On suppose :*

$$\textit{hypothèse :} \quad \forall \varphi \in \mathcal{D}(\mathbb{R}), \quad \int_{\mathbb{R}} f(x)\varphi(x)\,dx = 0. \tag{10.38}$$

Alors, avec q le conjugué de p défini par $\frac{1}{p} + \frac{1}{q} = 1$ quand $1 < p < \infty$:

$$\textit{conclusion :} \quad \begin{cases} p = 1 : & \forall \psi \in L^\infty(\mathbb{R}) \text{ t.q. } \mathrm{supp}\psi \textit{ compact}, & \int_{\mathbb{R}} f(x)\psi(x)\,dx = 0, \\ p \in]1, \infty[: & \forall \psi \in L^q(\mathbb{R}) \text{ t.q. } \mathrm{supp}\psi \textit{ compact}, & \int_{\mathbb{R}} f(x)\psi(x)\,dx = 0, \\ p = \infty : & \forall \psi \in L^1(\mathbb{R}) \text{ t.q. } \mathrm{supp}\psi \textit{ compact}, & \int_{\mathbb{R}} f(x)\psi(x)\,dx = 0. \end{cases} \tag{10.39}$$

Preuve. Cas $p = 1$. Soit ψ en escalier avec suppψ borné, i.e. $\exists k \in \mathbb{N}$, $\exists a_1, ..., a_k, c_1, ..., c_k \in \mathbb{R}$, $a_1 < a_2 < ... < a_k$, $\psi = \sum_{i=1}^{k-1} c_i 1_{[a_i, a_{i+1}]}$. Soit (γ_n) une suite régularisante et soit $\psi_n \stackrel{\text{déf}}{=} \psi * \gamma_n$. On a $\psi_n \in \mathcal{D}(\mathbb{R})$ et $\psi_n(x) \longrightarrow_{n\to\infty} \psi(x)$ p.p., avec $||\psi - \psi_n||_\infty \le ||\psi||_\infty$, cf. exercice 10.29. Donc :

$$\int_{\mathbb{R}} f(x)(\psi(x) - \psi_n(x))\,dx = \int_{\mathbb{R}} f(x) 1_{[a_1,a_k]}(x)(\psi(x) - \psi_n(x))\,dx \underset{n\to\infty}{\longrightarrow} 0,$$

grâce au théorème de convergence dominée : notant $g(n,x) = (f1_{[a_1,a_k]})(x)(\psi(x) - \psi_n(x))$ l'intégrant, à x fixé $\psi(x) - \psi_n(x) \to 0$ p.p. donne $g(n,x) \to 0$, avec $|g(n,x)| \le ||\psi||_\infty |(f1_{[a_1,a_k]})(x)|$ majoration indépendante de n par une fonction intégrable. Donc $(10.39)_1$ est vraie pour les fonctions en escalier.

Soit $\psi \in L^\infty(\mathbb{R})$ à support borné. Soit $(e_n)_\mathbb{N}$ une suite de fonctions en escalier qui converge p.p. vers ψ, avec $(e_n(x))_\mathbb{N}$ croissante positive si $\psi(x) \ge 0$ et $(e_n(x))_\mathbb{N}$ décroissante négative si $\psi(x) \le 0$ (voir cours d'intégration). Donc on a $||e_n||_\infty \le ||\varphi||_\infty < \infty$ pour tout n.

Comme $\exists a > 0$ t.q. supp$\psi \subset [-a, a]$, quitte à remplacer les e_n par $e_n 1_{[-a,a]}$, on peut considérer les (e_n) toutes à support dans $[-a, a]$. Et on a :

$$\int_{\mathbb{R}} f(x)(\psi(x) - e_n(x))\,dx = \int_{\mathbb{R}} f(x)1_{[-a,a]}(x)(\psi(x) - e_n(x))\,dx \underset{n\to\infty}{\longrightarrow} 0,$$

grâce au théorème de convergence dominée : à x fixé $\psi(x) - e_n(x) \to 0$ p.p., et $|f(x)(\psi(x) - e_n(x))| \le ||\psi||_\infty ||f1_{[-a,a]}||_{L^1(\mathbb{R})}$ majoration indépendante de n par une fonction intégrable. Donc $(10.39)_1$ est vraie pour les fonctions bornées à support borné.

Cas $p \in]1, \infty[$. Soit q t.q. $\frac{1}{p} + \frac{1}{q} = 1$.
Cas ψ en escalier avec suppψ borné : même suite (ψ_n) que précédemment : ici $f1_{[a_1,a_k]} \in L^p(\mathbb{R})$ et $\psi, \psi_n \in L^q(\mathbb{R})$ pour tout n (trivial). Et Hölder : $|\int_{\mathbb{R}} f(x)1_{[a_1,a_k]}(x)(\psi(x) - \psi_n(x))\,dx| \le ||f1_{[a_1,a_k]}||_{L^p} ||\psi - \psi_n||_{L^q} \to 0$.
Soit $\psi \in L^\infty(\mathbb{R})$ à support borné. Même suite (e_n) que précédemment. Et Hölder.

Cas $f \in L^\infty(\mathbb{R})$. Cas ψ en escalier avec suppψ borné : même suite (ψ_n) que précédemment : ici $f1_{[a_1,a_k]} \in L^\infty(\mathbb{R})$ et $\psi, \psi_n \in L^1(\mathbb{R})$ pour tout n (trivial). Et : $|\int_{\mathbb{R}} f(x)1_{[a_1,a_k]}(x)(\psi(x) - \psi_n(x))\,dx| \le ||f1_{[a_1,a_k]}||_\infty ||\psi - \psi_n||_{L^1} \to 0$.
Soit $\psi \in L^1(\mathbb{R})$ à support borné. Même suite (e_n) que précédemment... ∎

Proposition 10.43 *Soit $1 \le p \le \infty$, et soit $f \in L^p_{loc}(\mathbb{R})$. On a :*

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \quad \int_{\mathbb{R}} f(x)\varphi(x)\,dx = 0 \qquad \Longleftrightarrow \qquad f = 0 \text{ p.p.}. \tag{10.40}$$

Preuve. $\Leftarrow$: trivial. C'est $\Rightarrow$ qu'il s'agit d'établir. Avec le lemme 10.42 :

Cas $p = 1$: on prend $\psi(x) = 0$ quand $x \notin]-k,k[$ et quand $f(x) = 0$, et sinon $\psi(x) = 1$ si $f(x) > 0$ et $\psi(x) = -1$ si $f(x) < 0$. Donc $\psi \in L^\infty(\mathbb{R})$ à support borné et $0 = \int_{\mathbb{R}} f(x)\psi(x)\,dx = \int_{\mathbb{R}} |f(x)1_{]-k,k[}(x)|\,dx$. Comme $|f1_{]-k,k[}| \ge 0$, on déduit $|f1_{]-k,k[}| = 0$ p.p., voir cours d'intégration, donc $f1_{]-k,k[} = 0$ donc $f = 0$ sur $]-k,k[$, vrai pour tout k.

Cas $p \in]1,\infty[$: on prend $\psi(x) = 0$ quand $x \notin]-k,k[$ et quand $f(x) = 0$, et sinon $\psi(x) = f(x)^{p-1}$ si $f(x) > 0$ et $\psi(x) = -|f(x)|^{p-1}$ si $f(x) < 0$. Soit q le conjugué de p donné par $\frac{1}{p} + \frac{1}{q} = 1$. On a $|\psi(x)|^q = |f(x)|^{q(p-1)} = |f(x)|$ pour $x \in]-k,k[$ et 0 ailleurs. Donc $\psi \in L^q(\mathbb{R})$. Avec $f1_{[-k,k]} \in L^p(\mathbb{R})$. Donc $(f1_{[-k,k]})\psi \in L^1(\mathbb{R})$ avec $(f1_{[-k,k]})\psi = |f|^p 1_{[-k,k]} \ge 0$ d'intégrale nulle, donc $f = 0$ sur $[-k,k]$, vrai pour tout k.

Cas $p = \infty$: dual du cas $p = 1$. On prend $\psi(x) = 0$ quand $x \notin]-k,k[$ et quand $f(x) = 0$, et sinon $\psi(x) = 1$ si $f(x) > 0$ et $\psi(x) = -1$ si $f(x) < 0$. Comme $[-k,k]$ est borné et ψ borné, $\psi \in L^1(\mathbb{R})$. Donc $\int_{\mathbb{R}} f1_{[-k,k]}\psi = 0$, avec $f1_{[-k,k]}\psi = |f|1_{[-k,k]} \ge 0$, donc $|f|1_{[-k,k]} = 0$ p.p., donc $f = 0$ sur $[-k,k]$, vrai pour tout k. ∎

A Annexe : cardinaux $\aleph_0$ et $\aleph_1$

Proposition A.1 *Soit E un ensemble fini de cardinal $n \ge 0$. Alors $\mathcal{P}(E)$ a pour cardinal 2^n.*

Et pour $n \ge 1$, $\mathcal{P}(E)$ a même cardinal que l'ensemble $\mathcal{F}(E;\{0,1\}) =^{\text{noté}} (\{0,1\})^E$ des fonctions de E à valeurs dans $\{0,1\}$.

Preuve. Par récurrence. C'est vrai pour n=0, car $E = \emptyset$ donne $\mathcal{P}(E) = \{\emptyset\}$. C'est également vrai pour n=1, car si $E = \{x\}$ alors $\mathcal{P}(E) = \{\emptyset, \{x\}\}$.

Supposons que ce soit vrai pour $n \ge 1$. Soit $E_{n+1} = \{x_1, ..., x_{n+1}\}$ de cardinal n+1. Alors $\mathcal{P}(E_{n+1}) = Q \bigcup R$ est l'union disjointe de Q l'ensemble des parties de E_{n+1} qui ne contiennent pas x_{n+1} et de R l'ensemble des parties de E_{n+1} qui contiennent x_{n+1}.

Q est donc l'ensemble $\mathcal{P}(E_n)$ des parties de $E_n = \{x_1, ..., x_n\}$ et donc, par hypothèse de récurrence, Card$Q = 2^n$. Et on a une bijection f immédiate entre Q et R donnée par $f : A \in Q \to A \bigcup \{x_{n+1}\} \in R$. Donc Card$R = 2^n$. Donc Card$\mathcal{P}(E) = 2^n + 2^n = 2^{n+1}$.

Par récurrence montrons que $\mathcal{F}(E;\{0,1\})$ a 2^n éléments. Si E a 1 élément x_1 alors soit $f(x_1) = 0$, soit $f(x_1) = 1$: 2 fonctions en tout. Puis si E_{n+1} a n+1 éléments, les fonctions de $\mathcal{F}(E;\{0,1\})$ vérifient soit $f(x_{n+1}) = 0$, soit $f(x_{n+1}) = 1$, et par hypothèse de récurrence on a 2^n possibilités dans chaque cas, soit $2^n + 2^n = 2^{n+1}$ possibilités en tout. ∎

Définition A.2 Un ensemble E est dit dénombrable s'il existe une bijection entre E et $\mathbb{N}$. On note $\aleph_0$ le cardinal de $\mathbb{N}$ (la lettre $\aleph$ se prononce “aleph”, et c'est la première lettre de l'alphabet hébreu. Le choix de cette lettre $\aleph$ a été fait par Cantor).

On rappelle que l'ensemble des réels est l'ensemble des nombres décimaux (un réel x est un nombre décimal $x = n, a_1 a_2 ...$) où $n \in \mathbb{N}$ et $a_k \in \{0, 1, ..., 9\}$ pour tout $k \in \mathbb{N}^*$), et on pose :

Définition A.3 On note $\aleph_1$ le cardinal de $\mathbb{R}$, dit également puissance du continu.

On donne la proposition suivante, dont la démonstration est connue sous le nom de démonstration de Cantor, ou “procédé d'extraction diagonale de Cantor”.

Proposition A.4 *L'ensemble des nombres décimaux est non dénombrable.*

Preuve. Raisonnons par l'absurde, i.e. supposons que $I = [0,1[$ est dénombrable. On peut donc ordonner I, grâce à la bijection $f : \mathbb{N} \to I$, et I est donc l'ensemble $\{f(i) \overset{\text{noté}}{=} x^{(i)} : i \in \mathbb{N}\}$. Et les $x^{(i)} \in I$ sont de la forme (nombres décimaux), pour $i \in \mathbb{N}$:

$$x^{(i)} = 0, a_1^{(i)} a_2^{(i)} ... \qquad \text{où} \quad a_k^{(i)} \in \{0, 1, ..., 9\} \text{ pour tout } k \in \mathbb{N}^*.$$

On construit alors le nombre $y = 0, b_1 b_2 ...$ où les $b_k \in \{0, 1, ..., 9\}$ sont choisis comme suit : on prend $b_1 \ne a_1^{(1)}$, puis $b_2 \ne a_2^{(2)}$, puis..., $b_n \ne a_n^{(n)}$... (extraction diagonale). Ce nombre y est bien un nombre décimal dans I, mais n'est pas un des $x^{(m)}$ pour $m \in \mathbb{N}$. En effet, s'il existait un m tel que $y = x^{(m)} = 0, a_1^{(m)} a_2^{(m)} ...$ alors $b_m = a_m^{(m)}$, ce qui est faux par construction de b_m. Donc $y \notin f(\mathbb{N})$, donc f non surjectif. Absurde puisqu'on a supposé f bijectif, donc I n'est pas dénombrable. ∎

Proposition A.5 *Si E est infini dénombrable, alors $\mathcal{P}(E)$ est infini non dénombrable. Et donc, si* $\mathrm{Card}E = \aleph_0$, *alors* $\mathrm{Card}(\mathcal{P}(E)) > \aleph_0$.

Preuve. (Analogue à la démonstration de Cantor.) Montrons-le pour $E = \mathbb{N}$.

Raisonnons par l'absurde, i.e. supposons $\mathcal{P}(\mathbb{N})$ dénombrable, i.e. il existe une bijection (qui ordonne $\mathcal{P}(\mathbb{N})$) :

$$f : \begin{cases} \mathbb{N} & \to \mathcal{P}(\mathbb{N}), \\ n & \to f(n) = A_n. \end{cases}$$

Comme $\emptyset \in \mathcal{P}(\mathbb{N})$, on a $f^{-1}(\emptyset) \in \mathbb{N}$. Quitte à appliquer une permutation, on peut supposer $A_0 = f(0) = \emptyset$. Ainsi pour $n \geq 1$, comme f est une bijection, on a $A_n = f(n) \neq \emptyset$. Donc :

$$n \in \mathbb{N}^* \quad \Rightarrow \quad \exists a_n \in \mathbb{N}^*, \quad a_n \in A_n = f(n).$$

En particulier $1_{A_n}(a_n) = 1$ (dès que $n \geq 1$), où 1_{A_n} est la fonction indicatrice de A_n (i.e. $1_{A_n}(m)$=1 si $m \in A_n$, et $= 0$ sinon).

On définit alors la fonction (construction “diagonale”) :

$$g : n \in \mathbb{N} \to g(n) \overset{\text{déf}}{=} 1 - 1_{A_n}(n) \quad \in \{0, 1\}.$$

Autrement dit, $g = 1_B$ (fonction indicatrice d e B) où $B = g^{-1}(\{1\}) = \{n \in \mathbb{N} : g(n) = 1\}$; et comme $g(0) = 1$, on a $0 \in B$, donc $B \neq \emptyset$.

Et comme $B \in \mathcal{P}(\mathbb{N})$ et $B \neq \emptyset$, il existe $n > 0$ t.q. $B = f(n) = A_n$.

Donc il existe $n > 0$ t.q. $g = 1_{A_n}$. Donc $g(a_n) = 1$.

Et comme $a_n \in A_n$ on a $g(a_n) = 1 - 1_{A_n}(a_n) = 0$. Donc $1 = 0$. Absurde.

Donc une bijection $f : \mathbb{N} \to \mathcal{P}(\mathbb{N})$ n'existe pas. ∎

Exercice A.6 Montrer dans la démonstration précédente que, si on restreint g à $\mathbb{N}^*$, alors g n'est pas la fonction nulle.

Réponse. Supposons que $g_{|\mathbb{N}^*}$ est la fonction nulle : $g(n) = 0$ pour tout $n \geq 1$, i.e. $1_{A_n}(n) = 1$ pour tout $n \geq 1$. Donc $n \in A_n$ pour tout $n \geq 1$, i.e. $n \in f(n)$ pour tout $n \geq 1$.

Comme $\{1\} \in \mathcal{P}(\mathbb{N})$, il existe n_1 tel que $f(n_1) = \{1\}$. Comme $n_1 \in f(n_1) = \{1\}$, on a $n_1 = 1$. Donc $f(1) = \{1\}$.

Comme $\{2\} \in \mathcal{P}(\mathbb{N})$, il existe n_2 tel que $f(n_2) = \{2\}$. Comme $n_2 \in f(n_2) = \{2\}$, on a $n_2 = 2$. Donc $f(2) = \{2\}$.

Comme $\{1, 2\} \in \mathcal{P}(\mathbb{N})$, il existe n tel que $f(n) = C$. Comme $n \in f(n) = \{1, 2\}$, on en déduit que $n = 1$ ou $n = 2$. Impossible car $f(1) = \{1\}$ et $f(2) = \{2\}$ sont différents de $\{1, 2\}$. ∎

Corollaire A.7 *Si un ensemble contient au moins $2^{\mathbb{N}}$ éléments, il n'est pas dénombrable.*

Preuve. En effet, $\mathcal{P}(\mathbb{N})$ contient $2^{\mathbb{N}}$ éléments (corollaire de la proposition A.1), et $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable. ∎

Références pour cette annexe :

A. Kartachev, B. Rojdestvenski : Analyse mathématique. Editions Mir, Moscou 1984, traduction française 1988, ISBN 5-03-000160-3.

Gérard Letac, http ://www.les-mathematiques.net/

Gérard Villemin, http ://villemin.gerard.free.fr/index.html/, rubrique arithmétique / infini. Et “références” pour des sites internet relatifs à l'arithmétique. En particulier :

http ://noe-education.com/D111.htm

http ://www.animath.fr :8080/seconde/Ordinaux/Ordinaux.pdf

Définitions : voir http ://www.sciences-en-ligne.com/Frames_dictionary.asp

B Annexe : fonction exponentielle et suite numérique

La fonction $f : (y, z) \in \mathbb{R}^2 \to f(y, z) = z \in \mathbb{R}$ est uniformément 1-lipschitzienne en y, car $|f(y, z_1) - f(y(z_2)| = |z_1 - z_2| \leq 1 \times |z_1 - z_2|$, pour tout y, z_1, z_2. Donc le théorème de Cauchy–Lipschitz indique que l'équation différentielle $u'(x) = f(x, u(x))$ avec la condition initiale $u(0) = 1$ a une unique solution C^1 sur $\mathbb{R}$:

Définition B.1 La fonction exponentielle $\exp : \mathbb{R} \to \mathbb{R}$ est l'unique fonction $u : \mathbb{R} \to \mathbb{R}$ qui vérifie $u'(x) = u(x)$ et $u(0) = 1$ (existence et unicité par application du théorème de Cauchy–Lipschitz).

On a immédiatement $\exp \in C^\infty(\mathbb{R}^n)$, et $\exp^{(n+1)} = \exp^{(n)} = \exp$ pour tout $n \in \mathbb{N}^*$.

Soit $\varphi : x \in \mathbb{R} \to \varphi(x) = \exp(-x)$. Donc $\varphi'(x) = -\exp(-x)$. Donc la fonction $g : x \in \mathbb{R} \to g(x) = \exp(x)\exp(-x)$ vérifie $g'(x) = \exp'(x)\exp(-x) + \exp(x)(-\exp(-x)) = 0$, donc $g(x) = \text{constante} = g(0) = 1$. Donc $\exp(x) \neq 0$, et $\exp(-x) = \frac{1}{\exp(x)}$, et exp est strictement positive (car est continue et $\exp(0) = 1$). Donc $\exp'' = \exp' = \exp > 0$ et exp est strictement croissante concave.

Puis $h(x) = \frac{\exp(x+a)}{\exp(a)}$ vérifie $h'(x) = h(x)$ et $h(0) = 1$, donc $h = \exp$, donc $\exp(a+b) = \exp(a)\exp(b)$ pour tout a, b. D'où la notation $\exp(x) = e^x$ (notation d'une puissance) où $e := \exp(1)$.

Et $\exp(x) \underset{x\to\infty}{\longrightarrow} \infty$ car $\exp(1) > \exp(0) = 1$ et $\exp(n) = \exp(1)^n$. Donc $\exp(x) = \frac{1}{\exp(-x)} \longrightarrow_{x\to-\infty} 0$.

Développement en série, x_0 fixé : $\exp(x_0) = \sum_{n\in\mathbb{N}} \frac{\exp^n(0)}{n!} {x_0}^n = \sum_{n\in\mathbb{N}} \frac{{x_0}^n}{n!}$, série convergente car $\frac{|\frac{{x_0}^{n+1}}{(n+1)!}|}{|\frac{{x_0}^n}{n!}|} = \frac{|x_0|}{n+1} < 1$ pour $n \geq |x_0|$ (règle de d'Alembert). Vrai pour tout $x_0 \in \mathbb{R}$: le rayon de convergence de la série est ∞.

Et l'inverse $\log = \exp^{-1} : \mathbb{R}_+^* \to \mathbb{R}$ (qui vérifie $\log(\exp(x)) = x$) est C^∞ strictement croissante avec $\log(1) = 0$ (car $(\log(\exp(0)) = 0)$.

Et $\log'(x) = \frac{1}{x}$ (car $\log'(\exp(x))\exp'(x) = 1$).

Et $\log(ab) = \log(a) + \log(b)$ pour $a, b > 0$ (car $\exp(\log(a) + \log(b)) = \exp(\log(a))\exp(\log(b)) = ab$).

Et $\log(1+x) \leq x$ pour tout $x > -1$ (faire le tableau de variation de $f(x) = \log(1+x) - x$).

On note :

$$\lambda_n(x) = \big(1 + \frac{x}{n}\big)^n, \quad \text{et} \quad \mu_n(x) = \big(1 - \frac{x}{n}\big)^{-n} = \frac{1}{\lambda_n(-x)}. \tag{B.1}$$

Proposition B.2 *1)* $\forall n \in \mathbb{N}^*,\ \forall x \geq -n,\ \lambda_n(x) \leq e^x$.

2) $\forall n \in \mathbb{N}^*,\ \forall y \geq -1,\ 1 + ny \leq (1+y)^n$ *(inégalité de Bernoulli).*

3) $\forall x \in \mathbb{R}$, *la suite* $(\lambda_n(x))_{n>|x|}$ *est croissante et* $\lambda_n(x) \longrightarrow_{n\to\infty} e^x$, *la suite* $(\mu_n(x))_{n>|x|}$ *est décroissante et* $\mu_n(x) \longrightarrow_{n\to\infty} e^x$ *(convergence simple des suites de fonctions* $(\lambda_n)_{\mathbb{N}^*}$ *et* $(\mu_n)_{\mathbb{N}^*}$ *vers* exp*).*

4) Remarque : $e^{-x} \longrightarrow_{x\to-\infty} 0$, *alors qu'à* $n \in \mathbb{N}^*$ *fixé on a* $\lambda_n(x) \longrightarrow_{x\to-\infty} \pm\infty$ *(polynôme de degré* $n \geq 1$*).*

Preuve. 1) $\log(1+x) \leq x$ pour $x > -1$, donc $\log(1+\frac{x}{n}) \leq \frac{x}{n}$ pour $\frac{x}{n} > -1$, donc $n\log(1+\frac{x}{n}) \leq x$ pour $x > -n$, donc $\log((1+\frac{x}{n})^n) \leq x$ pour $x > -n$, donc $(1+\frac{x}{n})^n \leq e^x$ pour $x > -n$ (l'exponentielle est croissante).

2) C'est trivial pour $n = 0, 1$. Soit $n \geq 2$. Soit $g(x) = (1+x)^n - (1+nx)$, donc $g'(x) = n(1+x)^{n-1} - n$, donc $g''(x) = n(n-1)(1+x)^{n-2}$, donc $g''(x) \geq 0$ pour $x \geq -1$, donc g' croissante sur $[-1, \infty[$, et $g'(0) = 0$, donc g décroissante sur $[-1, 0]$ et croissante sur $[0, \infty[$, donc $g(x) \geq g(0) = 0$, donc $g \geq 0$ sur $[-1, \infty[$, d'où l'inégalité de Bernoulli.

3) Soit $x_0 \in \mathbb{R}$. Pour $n > -x_0$ on a $1 + \frac{x_0}{n} > 0$, donc $\lambda_n(x_0) > 0$; donc la suite $(\lambda_n(x_0))_{n>|x_0|}$ est positive. On a $\log(\lambda_n(x_0)) = n\log(1+\frac{x_0}{n})$. Posons $\varphi(a) = a\log(1+\frac{x_0}{a})$. Montrons que φ est croissante sur $]|x_0|, \infty[$ (pour $a > |x_0|$). On a $\varphi'(a) = \log(1+\frac{x_0}{a}) + a\frac{-\frac{x_0}{a^2}}{1+\frac{x_0}{a}} = \log(1+\frac{x_0}{a}) + \frac{-\frac{x_0}{a}}{1+\frac{x_0}{a}} = \log(1+\frac{x_0}{a}) - 1 + \frac{1}{1+\frac{x_0}{a}} =^{\text{noté}} \psi(z)$, où $z = 1 + \frac{x_0}{a} \in]0, 2[$ et $\psi(z) = \log(z) - 1 + \frac{1}{z}$. On a $\psi'(z) = \frac{1}{z} - \frac{1}{z^2} = \frac{1}{z^2}(z-1)$, donc ψ décroissante sur $]0, 1]$ et croissante sur $[1, 2[$. Donc $\psi(x) \geq \psi(1) = 0$ sur $]0, 2[$: ψ est positive sur $]0, 2[$. Donc φ' est positive sur $]|x_0|, \infty[$, donc φ est croissante sur $]|x_0|, \infty[$. Donc $\varphi(n) \leq \varphi(n+1)$ pour $n > |x_0|$. Et la fonction log est croissante, donc $\lambda_n(x_0) \leq \lambda_{n+1}(x_0)$ pour $n > |x_0|$: la suite $(\lambda_n(x_0))_{n>|x_0|} = (\lambda_n(x_0))_{n>|x_0|}$ est positive et croissante. Donc la suite $(\mu_n(x_0))_{n>|x_0|}$ donnée par $\mu_n(x_0) = \frac{1}{\lambda_n(-x_0)}$ est positive décroissante pour $n > |x_0|$.

Et $\mu_n(x_0) - \lambda_n(x_0) = \big(1 - \frac{x_0}{n}\big)^{-n}\big(1 - \big(1 - \frac{x_0}{n}\big)^n\big(1 + \frac{x_0}{n}\big)^n\big) = \big(1 - \frac{x_0}{n}\big)^{-n}\big(1 - \big(1 - (\frac{x_0}{n})^2\big)^n\big)$, donc, pour $n > |x_0|$ on a $\mu_n(x_0) - \lambda_n(x_0) \geq 0$ ainsi que $\mu_n(x_0) - \lambda_n(x_0) \leq \mu_n(x_0)\frac{{x_0}^2}{n}$ (car Bernoulli pour $y = -(\frac{x_0}{n})^2$ donne $\big(1 - (\frac{x_0}{n})^2\big)^n \geq 1 - n(\frac{x_0}{n})^2$ donc $1 - \big(1 - (\frac{x_0}{n})^2\big)^n \leq \frac{{x_0}^2}{n}$).

Donc $\mu_n(x_0) - \lambda_n(x_0) \longrightarrow_{n\to\infty} 0$. Donc les suites $(\lambda_n(x_0))_{n>|x_0|}$ et $(\mu_n(x_0))_{n>|x_0|}$ sont adjacentes, avec l'une croissante et l'autre décroissante : elles convergent vers un réel noté $\widetilde{\exp}(x_0)$.

Vérifions que $\widetilde{\exp}'(x) = \widetilde{\exp}(x)$, donc que $\widetilde{\exp} = \exp$ (car $\widetilde{\exp}(0) = 1$). (On a ${\lambda_n}'(x) = \big(1 + \frac{x}{n}\big)^{n-1} = \big(1 + \frac{x}{n}\big)^n\big(1 + \frac{x}{n}\big)^{-1}$ semble $\simeq \lambda_n(x)$ quand $n \simeq \infty$...).

Pour cela montrons : $h\,\widetilde{\exp}(x) \leq \widetilde{\exp}(x+h) - \widetilde{\exp}(x) \leq h\,\widetilde{\exp}(x+h)$ pour $h > 0$, h petit. On a ${\lambda_n}'(x) = \big(1 + \frac{x}{n}\big)^{n-1}$ et ${\lambda_n}''(x) = \frac{n-1}{n}\big(1+\frac{x}{n}\big)^{n-2} \geq 0$ pour $n > |x|$, donc ${\lambda_n}'$ est croissante au voisinage de x pour $n > |x|$. Donc le théorème des accroissements finis donne ${\lambda_n}'(x) \leq \frac{\lambda_n(x+h) - \lambda_n(x)}{h} \leq {\lambda_n}'(x+h)$, donc $\lambda_n(x) \leq \frac{\lambda_n(x+h) - \lambda_n(x)}{h(1+\frac{x}{n})} \leq \lambda_n(x+h)$ pour $n > |x[$ et $h > 0$, h petit. Donc, x fixé et $n \to \infty$ donnent $\widetilde{\exp}(x) \leq \frac{\widetilde{\exp}(x+h) - \widetilde{\exp}(x)}{h} \leq \widetilde{\exp}(x+h)$ au voisinage de x, comme annoncé.

Donc $(1+h)\widetilde{\exp}(x) \leq \widetilde{\exp}(x+h) \leq \frac{\widetilde{\exp}(x)}{1-h}$ pour $|h| < 1$. Donc $\widetilde{\exp}$ est continue en x et $\widetilde{\exp}$ est dérivable en x de dérivée elle-même ; et trivialement $\widetilde{\exp}(0) = 1$. Donc $\widetilde{\exp}$ est la fonction exponentielle. ∎

Exercice B.3 À l'aide de la définition de l'exponentielle, montrer : $\forall n \in \mathbb{N}^*,\ \forall t \in \mathbb{R},\ \lambda_n(t^2) \leq e^{t^2}$.

Réponse. $e^x = 1 + x + \frac{x^2}{2} + ... \geq 1 + x$ pour $x \geq 0$. Donc $e^{\frac{x}{n}} \geq 1 + \dfrac{x}{n}$ pour $x \geq 0$ et $n \in \mathbb{N}^*$. Et la fonction $y \to y^n$ est croissante pour $y \geq 0$, donc pour $x \geq 0$ et $n \in \mathbb{N}^*$ on a $e^x \geq (1 + \dfrac{x}{n})^n$.

Pour un autre point de vue / présentation, voir Perrin :
`https://www.math.u-psud.fr/~perrin/CAPES/analyse/fonctions/definition-exponentielle.pdf`

Références

[1] Aubin j.-P. : *Initiation à l'analyse appliquée*. Masson, 1994.

[2] Bartle R. : *The Elements of Integration and Lebesgue Measure*. Wiley Classics Library, 1995.

[3] Bouyssel M. : *Intégrale de Lebesgue, mesure et intégration, exercices avec solutions et rappels de cours*. Cépaduès, 1996.

[4] Brézis H. : *Analyse fonctionnelle, théorie et applications*. Collection mathématiques appliquées pour la maîtrise, Masson (1983).

[5] Deheuvels P. : *L'intégrale*. PUF, collection Que sais-je ?, 1986.

[6] Guichardet A. : *Calcul intégral*. Armand Colin, 1969.

[7] Guichardet A. : *Intégration et analyse hilbertienne*. Ed. de l'Ecole Polytechnique, 1983.

[8] Hennequin P.L. : *Cours de probabilités*. Université de Clermont, Faculté des sciences, maîtrise es sciences mathématiques, 1971.

[9] Kartachev A., Rojdestvenski B. : *Analyse mathématique*. Ed. Mir, 1988.

[10] Muthukumar T. : `home.iitk.ac.in/~tmk/courses/mth404/main.pdf`

[11] Lieb E.H., Loss M. : *Analysis*. Ed.AMS (Graduate Studies in Mathematics, V. 14), 1996,

[12] Marle R. : *Mesure et probabilités*. Hermann, 1974.

[13] Métivier M. : *Notions fondamentales de la théorie des probabilités*. Dunod, 1972.

[14] Métivier M. : *Théorie de la mesure et de l'intégration*. Ed. de l'Ecole Polytechnique, 1983.

[15] Neveu J. : *Bases mathématiques du calcul des probabilités*. Masson, 1970.

[16] Rudin W. : *Principles or Mathematical Analysis*. International Series in Pure and Applied Mathematics, McGraw-Hill, 1976.

[17] Rudin W. : *Analyse complexe et réelle*. Masson, 1995.

[18] Villani C. : `cedricvillani.org/wp-content/uploads/2013/03/IAF.pdf`

[19] Vo Khac K. : *Mesure, intégration, convolution, analyse de Fourier*. Ellipses, 1984.