

Théorie de la mesure et applications

Pierre-Alexandre Fournié

2018

Chapitre 1

Introduction à la théorie de la mesure

1.1 Cardinaux, sommes dénombrables, nouveau formalisme sur les ensembles

1.1.1 Cardinalité

On note $|E|$ le cardinal de E .

Deux ensembles E et F sont de même cardinal s'il existe une bijection entre eux.

Par ailleurs, on a $|E| \leq |F|$ lorsqu'il existe une injection de E dans F . C'est bien une relation d'ordre sur les cardinaux.

On définit de plus $|E| < |F|$ lorsque $|E| \neq |F|$ et $|E| \leq |F|$.

Un ensemble E est dit dénombrable lorsque $|E| = |\mathbb{N}|$.

On dispose du résultat fondamental suivant :

Proposition : Cardinalité de l'ensemble des parties de E

On a $|\mathcal{P}(E)| > |E|$.

Démonstration. C'est un argument de Cantor.

Il existe une injection de E dans $\mathcal{P}(E)$: l'injection canonique

$x \mapsto \{x\}$.

Supposons maintenant qu'il existe une injection de $\mathcal{P}(E)$ dans E .

Notons $f : \mathcal{P}(E) \rightarrow E$ cette injection.

Considérons $X = \{x \in E / x \notin f^{-1}(x)\}$. On peut ici parler de $f^{-1}(x)$ en tant que partie de E car f est injective.

Dans ce cas, que dire de $y = f(X)$? Si $y \in X$, on aboutit à une absurdité car $y \in f^{-1}(y)$ donc $y \notin X$. De même, si $y \notin X$, cela entraîne $y \in X$. f ne peut donc pas exister. \square

Les ensembles \mathbb{Z} , \mathbb{Q} sont dénombrables. De manière générale, pour tout entier p et pour tout ensemble E dénombrable, E^p est dénombrable. Il suffit pour cela d'utiliser un ordre de type alphabétique sur les p -uplets.

En utilisant la fonction bijective \tanh , on peut montrer que $] - 1; 1[$ est de même cardinal que \mathbb{R} .

De même, en utilisant les écritures binaires des nombres de $] - 1; 1[$, on peut montrer que $] - 1; 1[$ est de même cardinal que $\{0, 1\}^{\mathbb{N}}$.

Et enfin, toute partie de \mathbb{N} pouvant être désignée par un unique élément de $\{0, 1\}^{\mathbb{N}}$, on a la série d'égalités :

$$|\mathcal{P}(\mathbb{N})| = \left| \{0, 1\}^{\mathbb{N}} \right| =] - 1; 1[= |\mathbb{R}|$$

De même, si les $(E_n)_{n \in \mathbb{N}}$ sont tous dénombrables alors $\bigcup_{n \in \mathbb{N}} E_n$ l'est aussi. Il suffit pour cela de noter qu'un élément de $\bigcup_{n \in \mathbb{N}} E_n$ peut être désigné par un couple d'entiers.

1.1.2 Sommes de nombres positifs

Proposition : Somme bornée de nombres positifs

Si $\sum_{i \in I} x_i < +\infty$ et si, pour tout $i \in I$, $x_i \geq 0$ alors l'ensemble $K = \{i/x_i > 0\}$ est au plus dénombrable.

Démonstration. Notons que $K \cup \bigcup_{n \in \mathbb{N}^*} K_n$ avec $K_n = \left\{i/x_i > \frac{1}{n}\right\}$.

On sait d'autre part que

$$\sum_{i \in I} x_i \geq \sum_{i \in K_n} x_i \geq \frac{|K_n|}{n}$$

En particulier, pour tout n , $|K_n| \leq n \sum_{i \in I} x_i < +\infty$. Ainsi, les K_n sont tous de cardinaux finis. Cela permet de conclure ! □

Proposition : Permutations des termes d'une somme dénombrable de réels positifs

On considère la somme $\sum_{n \in \mathbb{N}} x_n$ où les x_n sont tous positifs ou nuls.

Soit une permutation σ de $\sigma(\mathbb{N})$.

Alors

$$\sum_{n \in \mathbb{N}} x_n = \sum_{n \in \mathbb{N}} x_{\sigma(n)}$$

Démonstration. On pose $S = \sum_{n \in \mathbb{N}} x_n$

On va commencer par supposer $S = +\infty$.

Soit $M \in \mathbb{R}$. Il existe N tel que $\sum_{0 \leq n \leq N} x_n > M$.

Il nous faut prouver qu'il existe aussi N' tel que $\sum_{0 \leq n \leq N'} x_{\sigma(n)} > M$.

Considérons pour cela l'ensemble $E_n = \{\sigma^{-1}(n) / 1 \leq n \leq N\}$. Cet ensemble admet un maximum noté N' . En particulier, $\llbracket 0; N' \rrbracket \supset E_n$ et donc $\sigma(\llbracket 0; N' \rrbracket) \supset \llbracket 0; N \rrbracket$. On a donc :

$$\sum_{0 \leq n \leq N'} x_{\sigma(n)} \geq \sum_{0 \leq n \leq N} x_n > M.$$

Ce qui prouve bien que $\sum_{n \in \mathbb{N}} x_{\sigma(n)} = +\infty$.

Le cas où $S < +\infty$ se traite de la même manière. □

Proposition : Réarrangement de termes dans une somme dénombrable

Soit $(I_i)_{i \in \mathbb{N}}$ une partition dénombrable de \mathbb{N} et une somme $\sum_{n \in \mathbb{N}} x_n$ où les x_n sont tous positifs ou nuls.

Alors,

$$\sum_{n \in \mathbb{N}} x_n = \sum_{i \in \mathbb{N}} \sum_{k \in I_i} x_k$$

Démonstration. Cela se prouve selon le même principe que précédemment. Pour changer, on va considérer que $\sum_{n \in \mathbb{N}} x_n = S < +\infty$.

Pour tout n et p , on a, $\sum_{i \in \llbracket 0; p \rrbracket} \sum_{k \in I_i \cap \llbracket 0; n \rrbracket} x_k \leq S$.

Ainsi, par passage à la limite sur n puis sur p , $\sum_{i \in \mathbb{N}} \sum_{k \in I_i} x_k \leq S$.

De plus, pour tout $\varepsilon > 0$, il existe un rang n tel que $S - \varepsilon \leq \sum_{k \in \llbracket 0; n \rrbracket} x_k$.

Mais dans ce cas, il existe un rang p tel que $\llbracket 0; n \rrbracket \subset \bigcup_{i \in \llbracket 0; p \rrbracket} I_i$. Et ainsi, pour tout $m \geq p$, $S - \varepsilon \leq$

$$\sum_{i \in \llbracket 0; m \rrbracket} \sum_{k \in I_i} x_k \leq S.$$

□

Proposition : Théorème de convergence absolue pour les séries

Soit une série $\sum_{n \in \mathbb{N}} x_n$ de nombres quelconques.

Si $\sum_{n \in \mathbb{N}} |x_n| < \infty$ alors cette série converge.

Démonstration. Cela se prouve en considérant $S_n = S_n^+ + S_n^-$ où les S_n^+ et les S_n^- sont les sous-séries de termes positifs (resp. négatifs). □

1.1.3 Limites d'ensembles, nouveau formalisme

Soit $(E_n)_{n \in \mathbb{N}}$ une suite d'ensembles tous inclus dans un ensemble E .

On pose $\sup E_n = \bigcup_{k \geq n} E_k$ et $\inf E_n = \bigcap_{k \geq n} E_k$.

La suite des $\sup E_n$ est une suite décroissante en ce sens que, pour tout n , $\sup E_n \supset \sup E_{n+1}$.

La suite des $\inf E_n$ est une suite croissante en ce sens que, pour tout n , $\inf E_n \subset \inf E_{n+1}$.

On remarque également que, pour tout n , $\inf E_n \subset E_n \subset \sup E_n$.

Enfin, on utilise la notation $\limsup E_n = \bigcap_{n \in \mathbb{N}} \sup E_n = \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} E_k$

et $\liminf E_n = \bigcup_{n \in \mathbb{N}} \inf E_n = \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} E_k$.

Dans le cas où $\limsup E_n = \liminf E_n$, on dit que la suite E_n converge.

Notons que $x \in \limsup E_n \iff \forall n, \exists p \geq n / x \in E_p$ et que

$x \in \liminf E_n \iff \exists n / \forall p \geq n, x \in E_p$.

En pratique, dire que les E_n convergent signifie que si x appartient à une infinité de E_p alors, à partir d'un certain rang x appartient à tous les E_p . C'est la traduction de $\limsup E_n \subset \liminf E_n$.

1.2 Tribu, mesure

1.2.1 Tribu

1.2.1.1 Définition

Une tribu \mathcal{T} sur E est un ensemble de parties de E :

- stable par union dénombrable ;
- stable par passage au complémentaire ;
- qui contient \emptyset .

À partir de ces trois hypothèses, on en déduit que \mathcal{T}

- est également stable par intersection dénombrable ;
- contient E .

Un ensemble E muni d'une tribu \mathcal{T} s'appelle un espace mesurable.

1.2.1.2 Exemples

$\mathcal{P}(E)$, $\{E; \emptyset\}$ sont des tribus appelées respectivement discrète et grossière.

Si \mathcal{T} est une tribu sur E et si $F \subset E$ alors l'ensemble des $\{T \cap F, T \in \mathcal{T}\}$ est une tribu sur F appelée tribu trace.

1.2.1.3 Tribu engendrée

Soit $\mathcal{S} \subset \mathcal{P}(E)$. Alors il existe une tribu \mathcal{T} qui contient \mathcal{S} et qui est la plus petite en ce sens que n'importe quelle tribu \mathcal{T}' qui contient \mathcal{S} contient également \mathcal{T} .

On construit cette tribu \mathcal{T} par l'extérieur, en posant

$$\mathcal{T} = \bigcap_{\substack{\mathcal{T}' \text{ est une tribu} \\ \mathcal{S} \subset \mathcal{T}'}} \mathcal{T}'$$

Un exemple classique s'appelle la tribu borélienne. C'est la tribu sur \mathbb{R} engendrée par les ouverts.

1.2.2 Mesure

1.2.2.1 Définition

On considère un espace mesurable $(E; \mathcal{T})$. Une mesure μ est une application de \mathcal{T} dans \mathbb{R}^+ qui vérifie :

- $\mu(\emptyset) = 0$.
- Pour tout les $(X_i)_{i \in I}$, famille dénombrable d'éléments de \mathcal{T} deux à deux disjoints,

$$\mu\left(\bigcup_{i \in I} X_i\right) = \sum_{i \in I} \mu(X_i)$$

Cette dernière propriété s'appelle la σ -additivité. Le triplet $(E; \mathcal{T}; \mu)$ ainsi construit s'appelle un espace mesuré.

Dans le cas où $\mu(E) < +\infty$, on dit que la mesure est finie.

Dans le cas où $\mu(E) = 1$, on parle de mesure de probabilité.

1.2.2.2 Propriétés de la mesure

Dans tout ce paragraphe, $(E; \mathcal{T}; \mu)$ désigne un espace mesuré.

Proposition : Soustraction

Soient L et M deux éléments de \mathcal{T} tels que $L \subset M$.

Alors

$$\mu(M - L) + \mu(L) = \mu(M)$$

Démonstration. Il suffit de constater que L et $M - L$ forment une partition de M . Par conséquent, $\mu(L) + \mu(M - L) = \mu(M)$. \square



On aimerait écrire $\mu(M - L) = \mu(M) - \mu(L)$ mais la soustraction par $\mu(L)$ n'a de sens que si L est de mesure finie. Cette hypothèse est très importante !

On en déduit la propriété très simple :

Proposition : Inclusion

Soient L et M deux éléments de \mathcal{T} tels que $L \subset M$. Alors

$$\mu(M) \geq \mu(L)$$

Proposition : Formule de l'union simple

Soient T_1 et T_2 deux éléments de \mathcal{T} tels que $\mu(T_1 \cap T_2) < +\infty$.

Alors

$$\mu(T_1 \cup T_2) + \mu(T_1 \cap T_2) = \mu(T_1) + \mu(T_2)$$

Démonstration. Il suffit de constater que $T_1 - (T_1 \cap T_2)$, $T_2 - (T_1 \cap T_2)$ et $T_1 \cap T_2$ sont trois ensembles disjoints qui vérifient :

- $T_1 = (T_1 - (T_1 \cap T_2)) \cup (T_1 \cap T_2)$
- $T_2 = (T_2 - (T_1 \cap T_2)) \cup (T_1 \cap T_2)$
- $T_1 \cup T_2 = (T_2 - (T_1 \cap T_2)) \cup (T_2 - (T_1 \cap T_2)) \cup (T_1 \cap T_2)$

En suite, on applique les formules de réarrangement de sommes. □

Proposition : Formule de l'union généralisée

Soient $(T_i)_{1 \leq i \leq n}$ des éléments de \mathcal{T} tels que pour tout $1 \leq i < j \leq n$, $\mu(T_i \cap T_j) < +\infty$.

Alors :

$$\begin{aligned} \mu\left(\bigcup_{1 \leq i \leq n} T_i\right) = & \sum_{1 \leq i \leq n} \mu(T_i) - \sum_{1 \leq i_1 < i_2 \leq n} \mu(T_{i_1} \cap T_{i_2}) + \cdots + (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \mu\left(\bigcap_{1 \leq j \leq k} T_{i_j}\right) + \cdots \\ & + (-1)^{n+1} \mu\left(\bigcap_{1 \leq j \leq n} T_j\right) \end{aligned}$$

Démonstration. C'est évident pour $n = 1$ et $n = 2$. On va raisonner par récurrence sur n .

Au rang $n + 1$, notons que $\bigcup_{1 \leq i \leq n+1} T_i = \left(\bigcup_{1 \leq i \leq n} T_i\right) \cup T_{n+1}$ et utilisons la formule qui précède :

$$\mu\left(\bigcup_{1 \leq i \leq n+1} T_i\right) = \mu\left(\bigcup_{1 \leq i \leq n} T_i\right) + \mu(T_{n+1}) - \mu\left(\bigcup_{1 \leq i \leq n} T_i \cap T_{n+1}\right)$$

En utilisant l'hypothèse de récurrence, on peut conclure :

$$\begin{aligned} \mu\left(\bigcup_{1 \leq i \leq n+1} T_i\right) = & \sum_{1 \leq k \leq n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \mu\left(\bigcap_{1 \leq j \leq k} T_{i_j}\right) + \mu(T_{n+1}) \\ & - \sum_{1 \leq k \leq n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \mu\left(\bigcap_{1 \leq j \leq k} (T_{i_j} \cap T_{n+1})\right) \end{aligned}$$

En effet, cette somme correspond à la version éclatée de $\sum_{1 \leq k \leq n+1} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n+1} \mu\left(\bigcap_{1 \leq j \leq k} T_{i_j}\right)$ selon que $i_k = n + 1$ ou non. □

Enfin, nous pouvons établir quelques résultats concernant les limites d'ensembles.

Proposition : Convergence monotone, version simple

Soit $(E_n)_{n \in \mathbb{N}}$ une suite croissante d'éléments de \mathcal{T} .

Alors $\lim \uparrow \mu(E_n) = \mu(\lim \uparrow E_n)$

Soit $(F_n)_{n \in \mathbb{N}}$ une suite décroissante d'éléments de \mathcal{T} .

On suppose de plus qu'il existe un rang p tel que $\mu(F_p) < +\infty$.

Alors $\lim \downarrow \mu(F_n) = \mu(\lim \downarrow F_n)$

Démonstration. Pour tout n , on pose $I_n = E_{n+1} - E_n$, de sorte que $\lim \uparrow E_n = \bigcup_{n \in \mathbb{N}} I_n$.

Cette famille est ainsi une partition de $\lim \uparrow E_n$. Ainsi, $\mu(\lim \uparrow E_n) = \sum_{n \in \mathbb{N}} \mu(I_n)$.

Or, par construction, $\sum_{1 \leq k \leq n} \mu(I_k) = \mu(E_n)$ et ainsi $\mu(\lim \uparrow E_n) = \lim \uparrow \mu(E_n)$.

Intéressons-nous au cas décroissant.

Pour tout $n \geq p$, on considère la suite des $(F_p - F_n)_{n \geq p}$. C'est une suite croissante et on peut conclure à l'aide du théorème de convergence monotone classique. En effet,

D'une part, $\lim \uparrow \mu(F_p - F_n) = \mu(F_p) - \lim \downarrow (\mu(F_n))$ car $\mu(F_p) < +\infty$.

D'autre part, $\lim \uparrow \mu(F_p - F_n) = \mu(\lim \uparrow (F_p - F_n)) = \mu(F_p - \lim \downarrow F_n) = \mu(F_p) - \mu(\lim \downarrow F_n)$ en raison de la convergence monotone et de l'hypothèse $\mu(F_p) < +\infty$.

Finalement, on obtient bien $\lim \downarrow (\mu(F_n)) = \mu(\lim \downarrow F_n)$ □

De la cas décroissant l'hypothèse de mesure finie est importante.

Par exemple, pour la mesure de comptage, si on considère $F_n = \left[0; \frac{1}{n}\right]$, pour tout n ,

$\mu(F_n) = +\infty$ or $\mu(\lim \downarrow F_n) = 1$ car $\lim \downarrow F_n = 0$.

Proposition : Majoration de $\mu\left(\bigcup_{n \in \mathbb{N}} E_n\right)$

Pour toute famille dénombrable $(E_n)_{n \in \mathbb{N}}$ de \mathcal{T} , telle que, pour tout n , $\mu(E_n) < +\infty$, on a :

$$\mu\left(\bigcup_{n \in \mathbb{N}} E_n\right) \leq \sum_{n \in \mathbb{N}} \mu(E_n)$$

Démonstration. On va le montrer dans le cas discret par récurrence sur le cardinal.

On considère une famille finie $(E_k)_{1 \leq k \leq p}$. Si $p = 1$, le résultat est évident.

Notons maintenant que $\bigcup_{1 \leq k \leq p+1} E_k = \left(\bigcup_{1 \leq k \leq p} E_k\right) \cup E_{p+1}$.

Ainsi,

$$\mu\left(\bigcup_{1 \leq k \leq p+1} E_k\right) + \mu\left(\left(\bigcup_{1 \leq k \leq p} E_k\right) \cap E_{p+1}\right) = \mu\left(\bigcup_{1 \leq k \leq p} E_k\right) + \mu(E_{p+1})$$

Finalement, en utilisant l'hypothèse de récurrence :

$$\mu\left(\bigcup_{1 \leq k \leq p+1} E_k\right) \leq \mu\left(\bigcup_{1 \leq k \leq p} E_k\right) + \mu(E_{p+1}) \leq \sum_{1 \leq k \leq p} \mu(E_k) + \mu(E_{p+1})$$

Par passage à la limite, pour $p \rightarrow +\infty$, sachant que la famille $\left(\bigcup_{1 \leq k \leq p} E_k\right)_{p \in \mathbb{N}}$ est croissante, on peut utiliser le théorème de convergence monotone. □

Proposition :

On reprend les mêmes hypothèses.

Si $\sum_{n \in \mathbb{N}} \mu(E_n) < +\infty$ alors $\mu(\limsup E_n) = 0$

Démonstration. On suppose que $\sum_{n \in \mathbb{N}} \mu(E_n) < +\infty$.

On pose alors $R_n = \sum_{p \geq n} \mu(E_p)$. Par hypothèse, on a donc $R_n \rightarrow 0$.

Or, $\mu\left(\bigcup_{p \geq n} E_p\right) \leq R_n$ et ainsi $\lim \downarrow \mu\left(\bigcup_{p \geq n} E_p\right) = 0$.

En utilisant la version décroissante du théorème de convergence monotone on peut conclure. □

1.2.2.3 Exemples

La mesure de comptage, à tout ensemble associe son cardinal.

La mesure de Lebesgue qui à tout intervalle $[a; b]$ associe $b - a$ constitue également une mesure mais il n'est pas aisé de construire l'extension de la mesure des intervalles à la mesure de n'importe quel élément de la tribu borélienne.

Chapitre 2

L'intégrale de Lebesgue

2.1 Fonctions mesurables

2.1.1 Fonctions et ensembles : résultat préliminaire

Proposition : Image réciproque

Soit E et F deux ensembles et $f : E \rightarrow F$ une application.

Soient $(F_i)_{i \in I}$ une famille quelconque d'éléments de $\mathcal{P}(F)$ et soit $L \subset F$ un sous-ensemble de F .

Alors :

$$f^{-1} \left\langle \bigcup_{i \in I} F_i \right\rangle = \bigcup_{i \in I} f^{-1} \langle F_i \rangle$$

$$f^{-1} \left\langle \bigcap_{i \in I} F_i \right\rangle = \bigcap_{i \in I} f^{-1} \langle F_i \rangle$$

$$f^{-1} \langle F - L \rangle = E - f^{-1} \langle L \rangle$$

Démonstration. On raisonne par équivalences :

D'une part :

$$\begin{aligned}
 x \in f^{-1} \left\langle \bigcup_{i \in I} F_i \right\rangle &\iff f(x) \in \bigcup_{i \in I} F_i \\
 &\iff \exists i \in I / f(x) \in F_i \\
 &\iff \exists i \in I / x \in f^{-1} \langle F_i \rangle \\
 &\iff x \in \bigcup_{i \in I} f^{-1} \langle F_i \rangle
 \end{aligned}$$

D'autre part :

$$\begin{aligned}
 x \in f^{-1} \left\langle \bigcap_{i \in I} F_i \right\rangle &\iff f(x) \in \bigcap_{i \in I} F_i \\
 &\iff \forall i \in I / f(x) \in F_i \\
 &\iff \forall i \in I / x \in f^{-1} \langle F_i \rangle \\
 &\iff x \in \bigcap_{i \in I} f^{-1} \langle F_i \rangle
 \end{aligned}$$

Et enfin :

$$\begin{aligned}
 f^{-1} \langle F - L \rangle &\iff f(x) \in F - L \\
 &\iff f(x) \notin L \\
 &\iff x \notin f^{-1} \langle L \rangle \\
 &\iff x \in E - f^{-1} \langle L \rangle
 \end{aligned}$$

□

En particulier les unions, intersections et passage au complémentaire sont conservés par l'image réciproque. Mais cela ne fonctionne pas pour l'image directe.

On va pour cela considérer $\sin : \mathbb{R} \rightarrow \mathbb{R}$.

On pose $I_1 = \left[-\frac{\pi}{2}; \frac{\pi}{2} \right]$ et $I_2 = \left[\frac{\pi}{2}; \frac{3\pi}{2} \right]$.

On a $I_1 \cap I_2 = \left\{ \frac{\pi}{2} \right\}$ et $\sin(I_1) = [-1; 1] = \sin(I_2)$. Ainsi, on n'a pas $\sin(I_1) \cap \sin(I_2) = \sin(I_1 \cap I_2)$.

Le passage au complémentaire ne fonctionne pas non plus par image directe !

On en déduit le résultat suivant.

Proposition : Transport de tribu

Soit E un espace et $(F; \mathcal{T})$ un espace mesuré.

On suppose que $f : E \rightarrow F$ est une application.

Alors l'ensemble $\mathcal{S} = \{f^{-1} \langle T \rangle, T \in \mathcal{T}\}$ constitue une tribu de E .

De même, si on considère un espace mesuré $(E; \mathcal{S})$, un espace F et une application $g : E \rightarrow F$, on peut fabriquer une tribu sur F en considérant $\{T \subset F / g^{-1} \langle T \rangle \in \mathcal{S}\}$. C'est la tribu image par g .

Démonstration. Cela se prouve très facilement en considérant la proposition qui précède. □

2.1.2 Définition

Soient $(E; \mathcal{S})$ et $(E; \mathcal{T})$ deux espaces mesurables.

On dit que $g : E \rightarrow F$ est mesurable lorsque, pour tout $T \in \mathcal{T}$, $g^{-1} \langle T \rangle \in \mathcal{S}$.

Dit autrement, g est mesurable si et seulement si la tribu des $\{g^{-1} \langle T \rangle, T \in \mathcal{T}\}$ est incluse dans \mathcal{S} .

2.1.3 Compléments sur les tribus

2.1.3.1 Transport de tribu, tribu engendrée

On en déduit une caractérisation de la mesurabilité de g dans le cas où \mathcal{T} est générée par une famille :

Proposition : Famille génératrice et mesurabilité d'une fonction

Soient $(E; \mathcal{S})$ et $(E; \mathcal{T})$ deux espaces mesurables.

Soit $g : E \rightarrow F$ une fonction.

On suppose enfin que \mathcal{T} est générée par une famille \mathcal{C} .

Alors, g est mesurable si et seulement si pour tout $C \in \mathcal{C}$, $g^{-1} \langle C \rangle \in \mathcal{S}$.

Démonstration. Le sens direct est évident. Examinons le sens réciproque.

On suppose que pour tout $C \in \mathcal{C}$, $g^{-1} \langle C \rangle \in \mathcal{S}$.

\mathcal{S} contient $g^{-1} \langle C \rangle$, elle contient donc la tribu engendrée par $g^{-1} \langle C \rangle$.

On va maintenant prouver que la tribu engendrée par $g^{-1} \langle C \rangle$ est en fait $g^{-1} \langle \mathcal{T} \rangle$. □

Lemme : Image réciproque d'un tribu engendrée par une famille

On reprend les mêmes hypothèses.

Alors la tribu engendrée par $g^{-1} \langle \mathcal{C} \rangle$ est l'image réciproque de la tribu engendrée par \mathcal{C} .

Démonstration. On pose \mathcal{S} la tribu engendrée par $g^{-1} \langle \mathcal{C} \rangle$ et $\tilde{\mathcal{S}}$ la tribu image réciproque de la tribu engendrée par \mathcal{C} .

$\tilde{\mathcal{S}}$ contient $g^{-1} \langle \mathcal{C} \rangle$. On a donc $\mathcal{S} \subset \tilde{\mathcal{S}}$.

Pour prouver l'inclusion réciproque, on considère \mathcal{H} la tribu image de \mathcal{S} par g . Cette tribu contient \mathcal{C} . Et ainsi, la tribu engendrée par \mathcal{C} est incluse dans \mathcal{H} .

Par image réciproque, on en déduit $\tilde{\mathcal{S}} \subset \mathcal{S}$.

Finalement, on a bien $\tilde{\mathcal{S}} = \mathcal{S}$. □

2.1.3.2 Propriétés des boréliens

On rappelle à toutes fins utiles que l'ensemble des rationnels est dénombrable.

On munit \mathbb{R} de la tribu des Boréliens.

Notons que cette tribu peut être engendrée par plusieurs types d'intervalles comme l'exprime cette proposition

Proposition : Familles générant la tribu des boréliens

Toutes les familles suivantes génèrent la tribu des boréliens :

\mathcal{F}_1 les intervalles $]a; b[$ avec $a < b$ réels ;

\mathcal{F}_2 les intervalles $]a; b[$ avec $a < b$ rationnels ;

\mathcal{F}_3 les intervalles $]a; +\infty[$ avec a rationnel ;

\mathcal{F}_4 les intervalles $] - \infty; b[$ avec b rationnel ;

\mathcal{F}_5 les intervalles $] - \infty; b[$ avec $b < 0$ rationnel et les intervalles $]a; +\infty[$ avec $a > 0$ rationnel.

\mathcal{F}_6 à \mathcal{F}_{10} toutes les familles ci-dessus en remplaçant les intervalles ouverts par des intervalles fermés.

Démonstration. On note \mathcal{T}_n la tribu engendrée par la famille \mathcal{F}_n .

Le principe général de cette démonstration est de montrer les inclusions successives des familles \mathcal{F}_{n+1} dans les tribus \mathcal{T}_n , ce qui prouvera $\mathcal{T}_{n+1} \subset \mathcal{T}_n$ puis de terminer la boucle en prouvant $\mathcal{F}_5 \subset \mathcal{T}_1$.

$\mathcal{F}_2 \subset \mathcal{F}_1 \subset \mathcal{T}_1$.

D'autre part, pour tout a rationnel, il existe un entier $N > a$ et on ainsi

$]a; +\infty[= \bigcup_{n \geq N}]a; n[$

Cela prouve que $\mathcal{F}_3 \subset \mathcal{T}_2$.

Pour la troisième implication, on considère b rationnel. Pour tout $c < b$ rationnel, le complémentaire de $]c; +\infty[$ est $] - \infty; c]$ qui est dans \mathcal{T}_3 . En remarquant que $] - \infty; b[= \bigcup_{c < b}] - \infty; c]$, c'est gagné, on a prouvé que $\mathcal{F}_4 \subset \mathcal{T}_3$.

Considérons maintenant la quatrième implication.

Les intervalles $] - \infty; b[$ avec $b < 0$ sont dans $\mathcal{F}_4 \subset \mathcal{T}_4$. Par le jeu des complémentaires, les intervalles $[c; +\infty[$ aussi, avec $c > 0$. En notant que, pour tout $a > 0$,

$]a; +\infty[= \bigcup_{c > a} [c; +\infty[$, on a montré que $\mathcal{F}_5 \subset \mathcal{T}_4$.

Reste à prouver la dernière implication, c'est à dire $\mathcal{F}_1 \subset \mathcal{T}_5$.

Par le jeu des complémentaires et des intersections, on montre que \mathcal{T}_5 contient tous les intervalles de la forme $[d; c[, [c; b],]b; a]$ où $d < c < 0 < b < a$ sont quatre rationnels.

Considérons deux réels $x < y$. On doit distinguer trois cas.

Si $x < y \leq 0$, on note que $]x; y[= \bigcup_{\substack{x < a < b < y \\ (a; b) \in \mathbb{Q}}} [a; b[$.

Si $x < 0 < y$, on note que $]x; y[= \bigcup_{\substack{x < a < 0 < b < y \\ (a; b) \in \mathbb{Q}}} [a; b]$.

Si $0 \leq x < y$, on note que $]x; y[= \bigcup_{\substack{x < a < b < y \\ (a; b) \in \mathbb{Q}}}]a; b]$.

Finalement, on a donc $\mathcal{F}_1 \subset \mathcal{T}_5$. □

2.1.4 Fonctions mesurables à valeurs réelles

Proposition : Opérations sur les fonctions mesurables à valeurs réelles

Soit $(E; \mathcal{T})$ un ensemble mesurable.

f et g sont deux fonctions mesurables à valeurs réelles. λ est un réel. Alors

- λf est mesurable ;
- $f + g$ est mesurable ;
- fg est mesurable ;
- f^+ est mesurable ;
- f^- est mesurable ;
- $|f|$ est mesurable ;
- si f ne s'annule pas alors $\frac{1}{f}$ est mesurable.

Démonstration. On va le prouver en utilisant ce qui précède ainsi que la caractérisation des fonctions mesurables dans le cas où la tribu d'arrivée est générée par une famille.

Pour le premier point, si λ est nul c'est évident. Si λ est positif, pour tout a , $\lambda f(x) > a \iff f(x) > \frac{a}{\lambda}$.

On en déduit que l'image réciproque de $]a; +\infty[$ par λf est $f^{-1}\left(\left]\frac{a}{\lambda}; +\infty\right[\right)$, ce qui permet de conclure car f est mesurable.

Pour le second point, on va utiliser la caractérisation par les intervalles $]a; +\infty[$ avec a rationnel.

On considère donc x tel que $f(x) + g(x) > a$. Il existe des nombres $r > 0$ tel que $f(x) + g(x) > a + r$. On a alors $g(x) > g(x) - r$ et $f(x) > a - g(x) + r$, c'est à dire $f(x) > a - (g(x) - r)$. Réciproquement, si $f(x) > a - (g(x) - r)$ alors $f(x) + g(x) > a + r > a$. En prenant r tel que $g(x) - r$ est rationnel, on prouve que

$$f(x) + g(x) > a \iff \exists b \in \mathbb{Q} / g(x) > b \text{ et } f(x) > b - a$$

En particulier, $(f + g)^{-1}\langle]a; +\infty[\rangle = \bigcup_{b \in \mathbb{Q}} f^{-1}\langle]a - b; +\infty[\rangle \cap g^{-1}\langle]b; +\infty[\rangle$, ce dernier ensemble étant une union dénombrable d'intersections finies d'éléments de \mathcal{T} puisque f et g sont mesurables !

Pour le produit, on va considérer la caractérisation par la famille \mathcal{F}_5 . Prenons par exemple $b < 0$ et considérons x tel que $f(x)g(x) < b$. Alors il existe des nombres $\eta > 1$ tels que $f(x)g(x) < \eta b$.

Il nous faut alors distinguer deux cas.

Si $g(x) > 0$, alors $g(x) > \frac{g(x)}{\eta}$. Pour η bien choisi, ce dernier nombre est rationnel et on le note c . Dans ce

cas, $f(x) < \frac{\eta b}{g(x)} = \frac{b}{c}$.

Si $g(x) > 0$ alors $f(x) > 0$ et donc $f(x) > \frac{f(x)}{\eta}$. Pour η bien choisi, $\frac{f(x)}{\eta}$ est rationnel et on note c ce nombre.

Et là, on a $g(x) > \frac{\eta b}{g(x)} = \frac{b}{c}$.

Finalement :

$$(fg)^{-1} \langle] - \infty; b[\rangle = \bigcup_{c \in \mathbb{Q}_*^+} (g^{-1} \langle]c; +\infty[\rangle \cap f^{-1} \langle] - \infty; \frac{b}{c}[\rangle) \cup (f^{-1} \langle]c; +\infty[\rangle \cap g^{-1} \langle] - \infty; \frac{b}{c}[\rangle)$$

De la même manière, on prouve que, pour $a > 0$:

$$(fg)^{-1} \langle]a; +\infty[\rangle = \bigcup_{c \in \mathbb{Q}_*^+} (g^{-1} \langle]c; +\infty[\rangle \cap f^{-1} \langle]\frac{a}{c}; +\infty[\rangle) \cup (f^{-1} \langle] - \infty; -c[\rangle \cap g^{-1} \langle] - \infty; -\frac{a}{c}[\rangle)$$

Pour la partie positive, notons que $(f^+)^{-1} \langle]a; +\infty[\rangle = \begin{cases} f^{-1} \langle]a; +\infty[\rangle & \text{si } a \geq 0 \\ f^{-1} \langle]0; +\infty[\rangle & \text{si } a < 0 \end{cases}$, ce qui permet de conclure.

La partie négative se traite de manière similaire.

La valeur absolue s'obtient en notant que $|f| = f^+ + f^-$.

Pour finir, l'inverse se traite en utilisant la caractérisation par la famille \mathcal{F}_5 .

Pour $a > 0$, $\left(\frac{1}{f}\right)^{-1} \langle]a; +\infty[\rangle = f^{-1} \langle]0; \frac{1}{a}[\rangle$ et, pour $b < 0$,

$$\left(\frac{1}{f}\right)^{-1} \langle] - \infty; b[\rangle = f^{-1} \langle]\frac{1}{b}; 0[\rangle.$$

□

On suppose que l'on complète maintenant les réels avec $+\infty$ et $-\infty$, avec tout ce que cela suppose d'ambiguïté pour les opérations usuelles. On note $\overline{\mathbb{R}}$ ce nouvel ensemble sur lequel on peut définir la même tribu des boréliens.

Proposition : Limite d'une suite de fonctions mesurables monotones.

Soit f_n une suite croissante de fonctions mesurables de $(E; \mathcal{T})$ dans \mathbb{R} .

Alors $\lim \uparrow f_n$ est une fonction mesurable de E dans $\overline{\mathbb{R}}$.

Ce résultat s'étend aux suites décroissantes de fonctions mesurables.

Démonstration. Soit $f = \lim \uparrow f_n$. En raison du théorème de convergence monotone, f existe.

De plus, en raison du sens de variation de f_n , pour tout a , $f^{-1} \langle]a; +\infty[\rangle = \bigcup_{n \in \mathbb{N}} f_n^{-1} \langle]a; +\infty[\rangle$. Cela permet de conclure puisque les f_n sont mesurables.

Le cas des suites décroissantes se traitent en remarquant que f_n mesurable $\iff -f_n$ mesurable et f_n décroissante $\iff -f_n$ croissante. □

Proposition : Borne supérieure d'une suite de fonctions mesurables

On considère f_n une suite de fonctions mesurables. Alors, pour tout n , les fonctions $\sup f_n : x \mapsto \sup \{f_p(x), p \geq n\}$ et $\inf f_n : x \mapsto \inf \{f_p(x), p \geq n\}$ sont mesurables

Démonstration. Le cas de $\inf f_n$ se traite également par passage à l'opposé en remarquant que $\inf f_n = -\sup(-f_n)$.

Le cas du $\sup f_n$ n'est pas compliqué. En effet, prouvons que, pour tout a ,

$$(\sup f_n)^{-1} \langle]a; +\infty[\rangle = \bigcup_{p \geq n} f_p^{-1} \langle]a; +\infty[\rangle.$$

Si $(\sup f_n)(x) > a$, il existe $p \geq n$ tel que $(\sup f_n)(x) \geq f_p(x) > a$. Réciproquement, s'il existe $p \geq n$ tel que $f_p(x) > a$ alors $(\sup f_n)(x) \geq f_p(x) > a$. □

Théorème : Convergence simple de suites de fonctions

Soit f_n une suite de fonctions mesurables qui converge simplement vers une fonction f à valeurs dans $\overline{\mathbb{R}}$. Alors f est mesurable.

Démonstration. Cela se prouve en notant que $f = \lim_{n \rightarrow +\infty} \downarrow \sup f_n = \lim_{n \rightarrow +\infty} \uparrow \inf f_n$. □

2.2 Fonctions échelonnées positives**2.2.1 Définition**

On considère $(E; \mathcal{T})$ un espace mesurable.

On dit qu'une fonction $f : E \rightarrow \mathbb{C}$ est échelonnée lorsque il existe une partition $(T_k)_{1 \leq k \leq n}$ de E par des éléments de \mathcal{T} et des nombres positifs $(\alpha_k)_{1 \leq k \leq n}$ tels que

$$f = \sum_{1 \leq k \leq n} \alpha_k \mathbb{1}_{T_k}$$

On note $\mathcal{E}(E; \mathbb{R}^+)$, l'ensemble des fonctions échelonnées de E dans \mathbb{R}^+ .

Proposition : Propriétés des fonctions échelonnées

L'ensemble des fonctions échelonnées est stable par produit et combinaisons linéaires.

Démonstration. Il suffit de se référer aux propriétés des fonctions indicatrices plus bas. □

Proposition : Fonctions indicatrices

Soit L et M deux sous ensembles d'un ensemble E .

Alors

$$\begin{aligned} \mathbb{1}_L \times \mathbb{1}_M &= \mathbb{1}_{L \cap M} \\ \mathbb{1}_{M-L} &= \mathbb{1}_M - \mathbb{1}_L \quad \text{si } L \subset M \\ \mathbb{1}_L + \mathbb{1}_M &= 2 \times \mathbb{1}_{L \cap M} + \mathbb{1}_{L-(L \cap M)} + \mathbb{1}_{M-(L \cap M)} \quad \text{et ces trois ensembles sont disjoints.} \\ \mathbb{1}_{L \cup M} &= \mathbb{1}_L + \mathbb{1}_M - \mathbb{1}_{L \cap M} \\ \mathbb{1}_{\overline{L}} &= 1 - \mathbb{1}_L \end{aligned}$$

Démonstration. Cela se prouve très facilement par disjonction de cas. □

2.2.2 Intégrale de fonction échelonnée positive

On considère maintenant $(E; \mathcal{T}; \mu)$ un espace mesuré et f une fonction échelonnée de E vers \mathbb{R}^+ en utilisant les mêmes notations qu'au paragraphe précédent.

L'intégrale de f est le nombre

$$\int f(x) d\mu(x) = \sum_{1 \leq k \leq n} \alpha_k \mu(T_k)$$

Proposition : Propriétés de base des fonctions échelonnées

Soient f et g deux fonctions échelonnées positives. Soit $\alpha \geq 0$ un nombre.

On a

$$\int (f + \alpha g) = \int f + \alpha \int g$$

Si, de plus, pour tout x , $f(x) \leq g(x)$ alors

$$\int f \leq \int g$$

2.3 Intégrale d'une fonction positive

Dans tout ce paragraphe, on considère les espaces $(E; \mathcal{T}; \mu)$ et $(\mathbb{R}^+; \mathcal{B})$ où \mathcal{B} est la tribu trace des boréliens sur \mathbb{R}^+ .

2.3.1 Définition**Théorème : Définition de l'intégrale**

Soit une fonction f mesurable de E dans \mathbb{R}^+ .

Alors il existe une suite croissante f_n de fonctions échelonnées qui tendent vers f .

On définit alors l'intégrale de f par

$$\int f(x) d\mu(x) = \lim \uparrow \int f_n(x) d\mu(x)$$

De plus, ce nombre ne dépend pas de la suite f_n de fonctions échelonnées choisie.

Démonstration. On va construire cette suite f_n puis on montrera l'unicité.

Soit n un entier. Pour tout $1 \leq k \leq 2^{2n}$, on pose

$$T_k^{(n)} = f^{-1} \left[\left[\frac{k-1}{2^n}; \frac{k}{2^n} \right[\right]. \text{ C'est une partition de } E \text{ d'éléments de } \mathcal{T}.$$

Puis on définit

$$f_n = \sum_{1 \leq k \leq 2^{2n}} \frac{k-1}{2^n} \mathbb{1}_{T_k^{(n)}}$$

C'est une suite croissante par construction.

Montrons la convergence simple. Soit $x \in E$.

Il existe p tel que $2^p > f(x)$. Par construction, ensuite, pour tout $n \geq p$, $0 \leq f(x) - f_n(x) \leq 2^{-n}$. Il suffit pour cela d'écrire l'approximation par défaut de $f(x)$ en base 2 au n -ième digit près.

Pour finir, il nous faut montrer l'unicité. On considère donc deux suites croissantes g_n et f_n de fonctions échelonnées qui tendent simplement vers f .

Soit $\eta \in]0; 1[$ un nombre.

Pour tout entier n et pour tout x de E , il existe un rang n' tel que, pour tout $p \geq n'$,

$$\eta f_n(x) \leq g_p(x).$$

On fixe maintenant n et $\eta \in]0; 1[$ et on note $S_p = \{x / g_p(x) \geq \eta f_n(x)\}$.

D'après la remarque qui précède, S_p est une suite croissante d'éléments de \mathcal{T} qui tendent vers E . En particulier, on a

$$\int g_p(x) d\mu(x) \geq \int \mathbb{1}_{S_p} g_p(x) d\mu(x) \geq \eta \int \mathbb{1}_{S_p} f_n(x) d\mu(x)$$

D'après la propriété de convergence monotone pour la mesure, on en déduit, par passage à la limite quand p tend vers l'infini :

$$\lim_{p \rightarrow +\infty} \int g_p(x) d\mu(x) \geq \eta \int f_n(x) d\mu(x)$$

En faisant tendre η vers 1 puis n vers l'infini, on obtient que

$$\lim_{p \rightarrow +\infty} \uparrow \int g_p(x) d\mu(x) \geq \lim_{p \rightarrow +\infty} \uparrow \int f_p(x) d\mu(x).$$

Mais, comme f et g jouent des rôles symétriques, cela permet de conclure. \square

2.3.2 Propriétés de base

Proposition : Opérations sur les fonctions positives

Soient f et g deux fonctions mesurables de E dans \mathbb{R}^+ . Soit $\alpha \geq 0$ un réel.

Alors

$$\int (f + \alpha g) = \int f + \alpha \int g$$

Si, pour tout x , $f(x) \geq g(x)$ et si $\int g < +\infty$ alors

$$\int (f - g) = \int f - \int g$$

Enfin,

$$f \geq g \implies \int f \geq \int g$$

Démonstration. Soient f_n et g_n des suites croissantes de fonctions échelonnées qui tendent respectivement vers f et g .

$f_n + \alpha g_n$ tend vers $f + \alpha g$. Or, pour tout n , d'après les propriétés des fonctions échelonnées,

$$\int (f_n + \alpha g_n) = \int f_n + \alpha \int g_n$$

Par passage à la limite sur n , on obtient le premier résultat.

Pour le second résultat, on utilise le premier. En effet,

$$\int (f - g) + \int g = \int (f - g + g) = \int f$$

On peut ensuite soustraire l'égalité par $\int g$ puisque $\int g < +\infty$.

Pour montrer la dernière implication, on considère deux cas.

Si $\int f = +\infty$, l'inégalité est vraie.

Si $\int f < +\infty$, toute suite croissante g_n de fonctions échelonnées positive qui tend vers g vérifie $\int g_n \leq \int f < +\infty$. En particulier, on en déduit que $\int g < +\infty$ et on peut donc utiliser le fait que $\int f - \int g = \int (f - g) \geq 0$ \square

2.3.3 Les grands théorèmes de convergence, cas positif

Théorème : Convergence monotone, pour les fonctions positives

Soit f_n une suite croissante de fonctions positives mesurables.

Alors, $\lim \uparrow \int f_n = \int (\lim \uparrow f_n)$.

La formulation de ce théorème rend possible que $f = \lim \uparrow f_n$ puisse prendre des valeurs infinies.

On raisonne donc ici sur $\overline{\mathbb{R}}^+$ et on étend les fonctions échelonnées à $\overline{\mathbb{R}}^+$, avec pour convention :

$$\int +\infty \times \mathbb{1}_F(x) d\mu(x) = 0 \text{ si et seulement si } \mu(F) = 0.$$

Démonstration. Il est évident que $\lim \uparrow \int f_n \leq \int (\lim \uparrow f_n)$ en raison de la positivité de l'intégrale.

On note $f = \lim \uparrow f_n$.

Notons $\tilde{E} = \{x / f(x) = +\infty\}$ puis distinguons deux cas.

Si $\mu(\tilde{E}) > 0$, alors $\int f = +\infty$. Il s'agit donc de montrer que $\lim \uparrow \int f_n = +\infty$.

Pour tout $M > 0$, on pose $H_n = \{x / f_n(x) \geq M\}$. C'est une suite croissante d'éléments de \mathcal{T} . De plus, on vérifie que $\int f_n \geq M\mu(H_n)$. Par passage à la limite sur n , on en déduit

$$\lim \uparrow \int f_n \geq M\mu(\lim \uparrow H_n)$$

Or $\tilde{E} \subset \lim \uparrow H_n$ et ainsi $\lim \uparrow \int f_n \geq M\mu(\tilde{E})$

En faisant tendre M vers l'infini, on obtient le résultat escompté.

Considérons maintenant le second cas.

Si $\mu(\tilde{E}) = 0$, on a $\int f = \int \mathbb{1}_{E-\tilde{E}} f$.

On applique maintenant sur $E - \tilde{E}$ une technique déjà utilisée lorsque l'on a défini l'intégrale d'une fonction positive. On considère ainsi, pour tout $\eta \in]0; 1[$, l'ensemble

$$E_n = \{x \in E - \tilde{E} / f_n(x) \geq \eta f(x)\}.$$

On sait que $\lim \uparrow E_n = E - \tilde{E}$. De plus,

$$\int f_n \geq \int (\mathbb{1}_{E_n} f_n) \geq \eta \int (\mathbb{1}_{E_n} f)$$

Il faudrait maintenant pouvoir écrire $\lim \uparrow \int (\mathbb{1}_{E_n} f) = \int (\lim \uparrow \mathbb{1}_{E_n} f)$ pour conclure.

Pour cela, on va établir un lemme. □

Lemme : Limite croissante d'ensembles et intégrale

Soit $T \subset E$ un élément de \mathcal{T} et soit f une fonction mesurable de E dans \mathbb{R}^+ .

Soit de plus T_n une suite croissante d'éléments de \mathcal{T} telle que $\lim \uparrow T_n = T$.

Alors :

$$\lim \uparrow \int (\mathbb{1}_{T_n} f) = \int (\mathbb{1}_T f)$$

Démonstration. Soit φ_n une suite croissante de fonctions échelonnées positives qui tend vers f .

On a déjà prouvé précédemment que pour tout η de $]0; 1[$, il existe n , tel que

$$\int (\mathbb{1}_T \varphi_n) \geq \eta \int (\mathbb{1}_T f).$$

Pour ce même η , il existe un rang p , tel que $\int (\mathbb{1}_{T_p} \varphi_n) \geq \eta \int \mathbb{1}_T \varphi_n \geq \eta^2 \int (\mathbb{1}_T f)$. On en déduit ainsi que, pour tout η , il existe p tel que pour tout $k \geq p$

$$\int (\mathbb{1}_{T_p} f) \geq \int (\mathbb{1}_{T_p} \varphi_n) \geq \eta^2 \int (\mathbb{1}_T f)$$

Donc, pour tout η de $]0; 1[$, $\lim \uparrow \int (\mathbb{1}_{T_p} f) \geq \eta^2 \int f$.

En faisant tendre η vers 1, on obtient le résultat escompté. □

Proposition : Convergence monotone dans le cas de suite décroissante, pour les fonctions positives

Soit f_n une suite décroissante de fonctions mesurables positives.

On suppose qu'il existe n tel que $\int f_n < +\infty$.

Alors $\int f_n = \int (\lim \downarrow f_n)$

Démonstration. Pour tout $k \geq n$, on pose $\tilde{f}_k = f_n - f_k$.

\tilde{f}_k est une suite croissante de fonctions positives mesurables et on a donc $\lim_k \uparrow \int (f_n - f_k) = \int \left(\lim_k \uparrow (f_n - f_k) \right)$

et comme, pour tout $k \geq n$, $\int f_k < +\infty$, on peut utiliser la propriété sur les intégrales de différences de fonctions pour obtenir le résultat escompté. \square

L'hypothèse que l'intégrale soit finie est importante. Pour s'en convaincre, considérer la suite $f_n = \mathbb{1}_{\left]0; \frac{1}{n}\right[}$ avec la mesure de comptage.

Proposition : Intégrale nulle

Soit f une fonction mesurable de E dans \mathbb{R}^+ .

$$\int f = 0 \iff \mu\{x/f(x) \neq 0\} = 0$$

Démonstration. La réciproque est aisée. En effet, d'après les conventions énoncées plus haut. Si on note $\tilde{E} = \{x/f(x) \neq 0\}$, on a $\mu(\tilde{E}) = 0$ et donc pour toute fonction échelonnée φ , $\int (\mathbb{1}_{\tilde{E}} \times \varphi) = 0$.

Ainsi, $\int f = \int (\mathbb{1}_{\tilde{E}} \times f) + (\mathbb{1}_{E-\tilde{E}} \times f) = 0 + 0$

Pour le sens direct, pour tout n , on pose $M_n = \left\{x/f(x) \geq \frac{1}{n}\right\}$.

Ainsi, $\int f \geq \frac{1}{n} \mu(M_n)$ et ainsi, on a $\mu(M_n) = 0$.

M_n étant une suite croissante d'ensemble mesurables, on peut utiliser le théorème de convergence monotone pour les mesure. On en déduit :

$$\mu(\lim \uparrow M_n) = 0$$

Or $\lim \uparrow M_n = \{x/f(x) \neq 0\}$, ce qui permet de conclure. \square

On énonce maintenant un lemme utile pour démontrer le théorème de convergence dominée.

Lemme : Fatou

Soit f_n une suite de fonctions mesurables positives.

Alors, pour tout entier n :

$$\int \inf_{k \geq n} f_k \leq \inf_{k \geq n} \int f_k \leq \sup_{k \geq n} \int f_k \leq \int \sup_{k \geq n} f_k$$

Démonstration. Pour tout $p \geq n$ et pour tout x :

$$\inf_{k \geq n} f_k(x) \leq f_p(x)$$

et

$$f_p(x) \leq \sup_{k \geq n} f_k(x)$$

En intégrant, il vient :

$$\int \inf_{k \geq n} f_k(x) d\mu(x) \leq \int f_p(x) d\mu(x)$$

et

$$\int f_p(x) d\mu(x) \leq \int \sup_{k \geq n} f_k(x) d\mu(x)$$

Et comme cela est vrai pour tout $p \geq n$, on a :

$$\int \inf_{k \geq n} f_k(x) d\mu(x) \leq \inf_{p \geq n} \int f_p(x) d\mu(x)$$

et

$$\sup_{p \geq n} \int f_p(x) d\mu(x) \leq \int \sup_{k \geq n} f_k(x) d\mu(x)$$

□

Théorème : Convergence dominée, pour les fonctions positives

Soit f_n une suite de fonctions mesurables et positives.

On suppose que les f_n tendent vers une fonction f et qu'il existe une fonction positive g telle que

- pour tout n , $f_n \leq g$
- $\int g < +\infty$

Alors la fonction f possède une intégrale finie et $\lim \int f_n = \int f$.

Démonstration. Pour tout n , $\sup f_n \leq g$ et par conséquent $\int (\sup f_n) \leq \int g < +\infty$.

On peut donc ici utiliser le théorème de convergence monotone pour les fonctions croissantes et décroissantes et ainsi

$$\lim \downarrow \int \sup f_n = \int (\lim \downarrow \sup f_n) = \int f \text{ car les } f_n \text{ convergent vers } f.$$

De même :

$$\lim \uparrow \int \inf f_n = \int (\lim \uparrow \inf f_n) = \int f \text{ pour la même raison.}$$

En appliquant le lemme de Fatou, on prouve la convergence des $\int f_n$ vers $\int f$.

□

2.4 Intégrale d'une fonction à valeurs réelles ou complexe

Dans toute la suite $(E; \mathcal{T}; \mu)$ désigne un espace mesuré.

2.4.1 Notion de propriété μ -presque partout, tribu complétée

Définition : Ensemble μ -négligeable

Un sous-ensemble $F \subset E$ est dit μ -négligeable lorsqu'il existe un élément T de \mathcal{T} tel que $F \subset T$ et $\mu(T) = 0$.

Soit $P(x)$ une proposition portant sur les éléments x de E .

On dit que P est vraie μ -presque partout si et seulement si l'ensemble $\{x / P(x) \text{ est fausse}\}$ est μ -négligeable.

Par exemple, si f et g sont deux fonctions mesurables, dire $f(x) = g(x)$ μ -presque partout signifie que $\mu\{x / f(x) \neq g(x)\} = 0$.

On rappelle les propriétés suivantes des ensembles de mesure nulle.

Proposition : Union dénombrable d'ensembles μ -négligeables

Une union dénombrable d'ensembles μ -négligeables est μ -négligeable.

Démonstration. C'est évident d'après les propriétés sur les sommes de nombres positifs.

□

Définition : Tribu complétée

L'ensemble $\tilde{\mathcal{T}} = \{T \cup N, T \in \mathcal{T} \text{ et } N \text{ négligeable}\}$ est une tribu appelée tribu complétée de \mathcal{T} . De plus, on peut sur cette tribu étendre la mesure μ en posant

$$\mu(T \cup N) = \mu(T)$$

Démonstration. $\tilde{\mathcal{T}}$ contient E et est stable par union dénombrable.

Reste à prouver la stabilité par passage au complémentaire. Soit N un ensemble négligeable, inclus dans S élément de \mathcal{T} de mesure nulle.

Soit enfin T un élément de \mathcal{T} . On veut montrer que le complémentaire de $T \cup N$ est dans $\tilde{\mathcal{T}}$.

Mais on a $\overline{T \cup N} = \overline{T \cup S} \cup (S \cap \overline{T \cup N})$.

Or $S \cap \overline{T \cup N}$ est négligeable car inclus dans S .

$\tilde{\mathcal{T}}$ est donc bien stable par passage au complémentaire. \square

2.4.2 Intégrabilité

On considérera dans ce paragraphe les fonctions mesurables de $(E; \mathcal{T}; \mu)$ dans \mathbb{R} ou \mathbb{C} .

Définition : Intégrabilité sur \mathbb{R}

Soit f une fonction mesurable de E dans \mathbb{R} .

On dit que f est intégrable lorsque $\int |f(x)| < +\infty$.

Dans ce cas, f^+ et f^- ont des intégrales finies et on pose

$$\int f = \int f^+ - \int f^-$$

Démonstration. C'est évident car f^+ et f^- sont des fonctions positives mesurables et majorées par $|f|$. \square

Définition : Mesurabilité et intégrabilité sur \mathbb{C}

Sur \mathbb{C} on définit la tribu produit des boréliens; celle qui rend mesurable les applications partie réelle et partie imaginaire.

En raison de la caractérisation par les cylindres, une fonction f à valeur complexe est mesurable si et seulement si sa partie réelle et sa partie imaginaire sont mesurables.

De plus, dans ce cas, on dira que f est intégrable si et seulement si $\int |f| < +\infty$.

En particulier, on pourra définir sans équivoque :

$$\int f = \int \mathcal{R}(f) + i \int \mathcal{I}(f)$$

Démonstration. Encore une fois, $|f| \geq |\mathcal{R}(f)|$ et $|f| \geq |\mathcal{I}(f)|$.

Cela permet de conclure. \square

2.4.3 Propriétés de l'intégrale des fonctions à valeurs réelles ou complexes

Pour alléger les notations, on définit :

Définition : Intégration sur une partie de E

Soit $T \in \mathcal{T}$. Lorsqu'elle existe, on définit

$$\int_T f = \int (\mathbb{1}_T f)$$

Proposition : Linéarité

Soient deux fonctions f et g mesurables de E dans \mathbb{R} ou \mathbb{C} . Soit λ un scalaire.

Si f et g sont intégrables alors $f + g$ est intégrable et

$$\int (f + g) = \int f + \int g$$

De plus λg est également intégrable et

$$\int (\lambda g) = \lambda \int g$$

Démonstration. On va considérer que f et g sont à valeurs réelles puis on généralisera aux complexes.

On va commencer par prouver que λg est intégrable. On sait que $\int |\lambda g| \leq |\lambda| \int |g| < +\infty$. Bien sûr, cela fonctionne aussi si λ est complexe.

Si $\lambda < 0$, $(\lambda g)^+ = -\lambda g^-$ et $(\lambda g)^- = -\lambda g^+$.

Ainsi,

$$\int (\lambda g) = \int (-\lambda g^-) - \int (-\lambda g^+) = -\lambda \int g^- + \lambda \int g^+ = \lambda \int g$$

D'autre part $f + g$ est intégrable (sur \mathbb{R} ou \mathbb{C}) en raison de l'inégalité triangulaire. Et là encore l'intégrabilité se généralise aux complexes puisque l'inégalité triangulaire s'applique également au module.

Reste maintenant à analyser les parties positives et négatives de $f + g$.

On va noter $E^+ = \{x / f(x) + g(x) \geq 0\} = (f + g)^{-1} \{[0; +\infty[\} \in \mathcal{T}$ car $f + g$ est mesurable.

Puis $E^- = E - E^+$.

Ainsi, $(f + g)^+ = \mathbb{1}_{E^+}(f + g)$ et $(f + g)^- = \mathbb{1}_{E^-}(f + g)$.

Ce que l'on peut réécrire, $(f + g)^+ = \mathbb{1}_{E^+}[(f^+ + g^+) - (f^- + g^-)]$. Les fonctions $(f^+ + g^+)$ et $(f^- + g^-)$ sont positives et vérifient, sur E^+ , $(f^+ + g^+) \geq (f^- + g^-)$. On peut donc utiliser la propriété sur l'intégration d'une différence de fonctions mesurables positives. Finalement :

$\int (f + g)^+ = \int_{E^+} (f^+ + g^+) - \int_{E^+} (f^- + g^-)$. En réarrangeant et en utilisant la linéarité des intégrales de fonctions positives, on obtient :

$$\int (f + g)^+ = \int_{E^+} f + \int_{E^+} g$$

De même, en utilisant ce qui précède sur la multiplication par un réel (ici -1)

$\int (f + g)^- = - \int_{E^-} [(f^+ + g^+) - (f^- + g^-)] = \int_{E^-} [(f^- + g^-) - (f^+ + g^+)]$. Or, sur E^- , $(f^- + g^-) \geq (f^+ + g^+)$.

En réarrangeant, on obtient le résultat escompté :

$$\int (f + g)^- = - \int_{E^-} f - \int_{E^-} g$$

Finalement, en calculant $\int (f + g)^+ - \int (f + g)^-$ on obtient le résultat attendu.

Pour l'extension de cette propriété aux complexes, on utilise la linéarité de la partie réelle et imaginaire.

Ainsi, on montre que si λ est un réel, on a :

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

Reste à examiner le cas de λg quand λ est complexe. Pour cela, on écrit $\lambda = \alpha + i\beta$ avec α et β réels. On note également $g_r = \mathcal{R}(g)$ et $g_i = \mathcal{I}(g)$.

Ainsi, $\lambda g = (\alpha g_r - \beta g_i) + i(\alpha g_i + \beta g_r)$. En utilisant ce qui précède, et en réarrangeant, on obtient bien :

$$\int (\lambda g) = \lambda \int g$$

□

Proposition : Positivité de l'intégrale

Soient f et g deux fonctions intégrables à valeurs réelles.

Si $f \geq g$ μ -presque partout alors $\int f \geq \int g$.

Démonstration. On raisonne sur la tribu complétée et on pose $\tilde{E} = \{x / f(x) \geq g(x)\}$.

On sait que $\int_{\tilde{E}} f = \int f$ et $\int_{\tilde{E}} g = \int g$ car le complémentaire de \tilde{E} est de mesure nulle.

De plus, sur \tilde{E} , $f - g \geq 0$. On peut donc conclure en utilisant ce qui précède sur la linéarité de l'intégrale. \square

Proposition : Égalité de fonctions

Soient f et g deux fonctions intégrables à valeurs réelles.

Si $f = g$ μ -presque partout alors $\int f = \int g$.

Démonstration. Très facile à partir de ce qui précède. \square

Proposition : Valeur absolue nulle

Soient f et g deux fonctions intégrables à valeurs réelles.

$\int |f - g| = 0 \iff f = g$ μ -presque partout.

Démonstration. On se réfère à la propriété équivalente portant sur les fonctions mesurables positives. On en déduit que $|f - g|(x) = 0$ pour μ -presque tout x . \square

2.4.4 Les grands théorèmes de convergence, cas général**Définition : Généralisation de la définition de l'intégrale réelle**

Soit f une fonction réelle positive.

Si $\int f^+ < +\infty$ ou $\int f^- < +\infty$, on pose

$$\int f = \int f^+ - \int f^-$$

On utilise alors les conventions habituelles sur les calculs portant sur l'infini.

En revanche, il faut utiliser avec précaution les propriétés établies au paragraphe précédent.

Théorème : Convergence monotone, pour les fonctions à valeurs réelles

Soit f_n une suite croissante de fonctions mesurables à valeurs réelles.

S'il existe n tel que $\int f_n^- < +\infty$ alors

$$\lim \uparrow \int f_n = \int (\lim \uparrow f_n)$$

.

De même, soit g_n une suite décroissante de fonctions mesurables à valeurs réelles.

S'il existe n tel que $\int f_n^+ < +\infty$ alors

$$\lim \downarrow \int f_n = \int (\lim \downarrow f_n)$$

.

Démonstration. On va juste étudier la première situation.

f_n^+ est une suite croissante de fonctions mesurables positives.

f_n^- est une suite décroissante de fonctions mesurables positives.

Comme $\int f_n^- < +\infty$ à partir d'un certain rang, on peut utiliser le théorème de convergence monotone sur les suites de fonctions mesurables positives.

Ainsi, à partir d'un certain rang, $\int f_n$ existe et on a

$$\int f_n = \int f_n^+ - \int f_n^-$$

On peut passer à la limite et on obtient le résultat escompté. \square

Théorème : Convergence dominée, cas général

Soit f_n une suite de fonctions mesurables.

On suppose que les f_n tendent vers une fonction f μ -presque partout et qu'il existe une fonction positive g telle que

- pour tout n , $|f_n| \leq g$ μ -presque partout ;
- $\int g < +\infty$.

Alors la fonction f est intégrable et $\lim \int f_n = \int f$.

Démonstration. On va le prouver dans le cas réel dans un premier temps. On raisonne encore une fois sur la tribu complétée.

On considère l'ensemble $\tilde{E} = \{x / \forall n, |f_n(x)| \leq g(x) \text{ et } f_n(x) \rightarrow f(x)\}$.

Le complémentaire de cet ensemble est de mesure nulle car c'est une union dénombrable de négligeables.

On raisonne donc sur \tilde{E} .

Considérons la suite des f_n^+ . On a les hypothèses :

- $f_n^+ \rightarrow f^+$
- pour tout n , $f_n^+ \leq g$

On peut donc appliquer le théorème de convergence dominée, dans sa version positive, à la suite des f_n^+ .

Pour les mêmes raisons, on peut appliquer le théorème de convergence dominée, dans sa version positive, à la suite des f_n^- .

Finalement, on a bien $\lim \int f_n = \int f$.

On généralise au cas complexe en notant que

- $\mathcal{R}(f_n) \rightarrow \mathcal{R}(f)$ et, pour tout n , $|\mathcal{R}(f_n)| \leq g$
- $\mathcal{I}(f_n) \rightarrow \mathcal{I}(f)$ et, pour tout n , $|\mathcal{I}(f_n)| \leq g$

\square

Chapitre 3

La mesure de Lebesgue

3.1 Théorème de Carathéodory

3.1.1 Définitions et premiers lemmes

Définition : Algèbre de Boole ou clan

Soit E un ensemble et $\mathcal{C} \subset \mathcal{P}(E)$ un ensemble de parties de E .

On dit que \mathcal{C} est une algèbre de Boole ou un clan de E lorsque \mathcal{C} :

- est stable par union finie ;
- contient E ;
- est stable par passage au complémentaire.

Définition : Mesure extérieure

Soit E un ensemble et soit $\mu : \mathcal{P}(E) \rightarrow \mathbb{R}^+$ une application.

On dit que μ est une mesure extérieure lorsque

- pour tout $A \subset B \subset E$, $\mu(A) \leq \mu(B)$;
- pour tout $(A_n)_{n \in \mathbb{N}} \in \mathcal{P}(E)^{\mathbb{N}}$, $\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$.

Avant d'aller plus loin, on va établir un lemme.

Lemme :

Soit E un ensemble sur lequel il existe une mesure extérieure μ et telle que $\mu(E) < +\infty$.

Alors l'ensemble $\mathcal{T} = \{A \subset E / \forall X \subset E, \mu(A \cap X) + \mu(A^c \cap X) = \mu(X)\}$ est une tribu de E et μ est une mesure sur cette tribu.

Démonstration. Il est clair que \mathcal{T} est stable par complémentaire et contient E .

On va maintenant prouver que \mathcal{T} est stable par union finie d'éléments.

Soient A et B deux éléments de \mathcal{T} .

Pour tout $X \subset E$, on a $\mu((A \cup B) \cap X) + \mu((A^c \cap B^c) \cap X) \geq \mu(X)$ en raison de la propriété de σ -sous-additivité de la mesure extérieure.

Mais d'autre part, on sait que $A \cup B = (A \cap B^c) \cup (A^c \cap B) \cup (A \cap B)$ donc on en déduit :

$\mu((A \cup B) \cap X) + \mu((A^c \cap B^c) \cap X) \leq \mu(A \cap B^c \cap X) + \mu(B \cap A^c \cap X) + \mu(B \cap A \cap X) + \mu(A^c \cap B^c \cap X)$ toujours en raison de la σ -sous-additivité.

En remarquant que $\mu(A^c \cap B^c \cap X) + \mu(A \cap B^c \cap X) = \mu(B^c \cap X)$ et $\mu(B \cap A^c \cap X) + \mu(B \cap A \cap X) = \mu(B \cap X)$ on a ainsi :

$\mu((A \cup B) \cap X) + \mu((A^c \cap B^c) \cap X) \leq \mu(B^c \cap X) + \mu(B \cap X) = \mu(X)$.

Finalement, on a bien, $\mu(X) \leq \mu((A \cup B) \cap X) + \mu((A^c \cap B^c) \cap X) \leq \mu(X)$ et ainsi $A \cup B \in \mathcal{T}$.

Ainsi, comme \mathcal{T} est stable par union finie, et complémentaire, il est stable par différence.

Reste maintenant à prouver qu'il est stable par union dénombrable.

Sans nuire à la généralité, on peut considérer une union croissante d'éléments.

Soit (A_n) une telle suite d'éléments de \mathcal{T} et examinons, pour tout $X \subset E$, l'expression

$\mu\left(\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap X\right) + \mu\left(\left(\bigcap_{n \in \mathbb{N}} A_n^c\right) \cap X\right)$. Encore une fois, en raison de la sous σ -additivité, on a

$\mu(X) \leq \mu\left(\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap X\right) + \mu\left(\left(\bigcap_{n \in \mathbb{N}} A_n^c\right) \cap X\right)$. Il nous faut donc prouver également :

$\mu\left(\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap X\right) + \mu\left(\left(\bigcap_{n \in \mathbb{N}} A_n^c\right) \cap X\right) \leq \mu(X)$.

Notons également, si $A \in \mathcal{T}$ et $B \subset E$ sont disjoints, on a :

$$\mu(A \cup B) = \mu((A \cup B) \cap A) + \mu((A \cup B) \cap A^c) = \mu(A) + \mu(B)$$

Cette dernière remarque nous prouve que μ est additive pour tout couple de $\mathcal{T} \times \mathcal{P}(E)$.

On va maintenant généraliser en montrant que μ vérifie la propriété de σ -additivité sur \mathcal{T} . Si (B_n) est une suite d'éléments de \mathcal{T} disjoints deux à deux, on a, pour tout p l'encadrement :

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

Cet encadrement, par passage à la limite sur p , nous prouve que $\sum_{n \in \mathbb{N}} \mu(B_n) = \mu\left(\bigcup_{n \in \mathbb{N}} B_n\right)$, c'est à dire que μ est bien σ -additive sur \mathcal{T} .

La σ -additivité entraîne la propriété de convergence monotone sur les suites croissantes d'éléments de \mathcal{T} .

Revenons maintenant à l'inégalité que l'on cherche à prouver. Nous savons, pour notre suite croissante A_n que :

$\lim \uparrow \mu(A_n) = \mu(\lim \uparrow A_n) < +\infty$ car $\mu(E) < +\infty$.

Pour tout $\varepsilon > 0$, il existe p tel que $\mu\left(\bigcup_{n > p} A_n\right) < \varepsilon$. On a donc, d'après ce qui précède :

$\mu\left(\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap X\right) + \mu\left(\left(\bigcap_{n \in \mathbb{N}} A_n^c\right) \cap X\right) \leq \mu(A_p \cap X) + \mu(A_p^c \cap X) + \varepsilon = \mu(X) + \varepsilon$

Et comme cela est vrai pour tout ε , on en déduit que la seconde inégalité est vraie.

Ainsi, \mathcal{T} est stable par limite croissante, par union finie, par passage au complémentaire.

C'est donc bien une tribu pour laquelle on a déjà prouvé la σ -additivité. \square

L'hypothèse de continuité à droite est importante et elle peut se formuler de manière équivalente de différentes façons.

Proposition : Continuité à droite et mesure extérieure

Soit \mathcal{C} un clan de E .

On suppose que \mathcal{C} est muni d'une mesure μ . On admet également que l'on a fabriqué une mesure extérieure μ^* à partir des éléments de μ .

Enfin, on suppose que E est σ -fini pour la mesure μ .

Alors les trois propositions suivantes sont équivalentes :

- μ et μ^* coïncident sur tous les éléments de \mathcal{C} de mesure finie ;
- μ est continue à droite ;
- pour tout élément C de \mathcal{C} de mesure finie et pour toute suite croissante C_n de \mathcal{C} telle que $\lim \uparrow C_n = C$, on a

$$\lim \uparrow \mu(C_n) = \mu(C)$$

Démonstration. On suppose que μ et μ^* coïncident sur tous les éléments de \mathcal{C} de mesure finie et que F_n est une suite décroissante d'éléments de \mathcal{C} tels que $\mu(F_0) < +\infty$ et $\lim \downarrow F_n = \emptyset$.

On réduit alors l'espace à F_0 et on se trouve alors dans les hypothèses du lemme 3.1.1. Ainsi, μ^* vérifie la convergence monotone et en particulier $\lim \downarrow \mu^*(F_n) = 0$. Mais comme μ et μ^* coïncident, on peut conclure quant à la continuité à droite.

Supposons maintenant la continuité à droite et considérons une suite C_n qui vérifie les hypothèses du troisième point. Comme \mathcal{C} est un clan et que μ est additive, on a

$$\mu(C) = \mu(C \cap C_n) + \mu(C \cap C_n^c).$$

Or $\lim \downarrow C \cap C_n^c = \emptyset$ et $\mu(C \cap C_0) \leq \mu(C) < +\infty$. On utilise donc la continuité à droite pour conclure sur la limite de $\mu(C \cap C_n) = \mu(C_n)$.

Supposons enfin le troisième point validé et montrons qu'alors les deux mesures coïncident sur tous les éléments de mesure finie de \mathcal{C} .

Notons tout d'abord que tout élément C de \mathcal{C} se recouvre lui-même. On a donc, $\mu^*(C) \leq \mu(C)$.

D'autre part, pour tout $\varepsilon > 0$, il existe un recouvrement de C par une famille $(D_i)_{i \in I}$ au plus dénombrable telle que

$$\sum_{i \in I} \mu(D_i) \leq \mu^*(C) + \varepsilon < +\infty.$$

Pour simplifier ici, on suppose que la famille est dénombrable et que $I = \mathbb{N}$, le cas où la famille est finie étant plus simple encore.

On pose alors, pour tout n , $C_n = C \cap \left(\bigcup_{k \leq n} D_k \right)$ de telle sorte que $\lim \uparrow C_n = C$.

Et comme μ est additive sur le clan, on en déduit, par croissance que, pour tout n , $\mu(C_n) \leq \sum_{i \in I} \mu(D_i) \leq \mu^*(C) + \varepsilon < +\infty$.

Par passage à la limite, on a donc $\mu(C) \leq \mu^*(C) + \varepsilon$; ce qui permet de conclure. \square

3.1.2 Le théorème

Théorème : Carathéodory

Soit E un ensemble.

On suppose que

- il existe une algèbre de Boole \mathcal{C} de E ;
- il existe une mesure $\mu : \mathcal{C} \rightarrow \mathbb{R}^+$ additive ;
- il existe une suite croissante E_n d'éléments de \mathcal{C} tels que
 - pour tout n , $\mu(E_n) < +\infty$;
 - $\lim \uparrow E_n = E$;
- pour toute suite décroissante F_n d'éléments de \mathcal{C} tels que $\lim \downarrow F_n = \emptyset$ et $\mu(F_0) < +\infty$; on a $\lim \mu(F_n) = 0$.

Alors il existe une tribu de E sur laquelle μ peut être étendue. De plus, cette extension est unique.

Pour prouver ce théorème, on va commencer par construire une mesure extérieure sur E .

Proposition : Mesure extérieure sur E

Soient E un ensemble, \mathcal{C} un clan sur E et μ une mesure additive sur \mathcal{C} .

Pour tout $X \subset E$, on pose

$$\mu^*(X) = \inf \left\{ \sum_{k \in K} \mu(C_k) / \left(\bigcup_{k \in K} C_k \right) \supset X \text{ et les } C_k \in \mathcal{C} \text{ sont au plus dénombrables} \right\}$$

Alors μ^* est une mesure extérieure sur E .

Démonstration. On va commencer par prouver la croissance. Soient $X \subset Y \subset E$.

Tout recouvrement $\bigcup_{k \in K} C_k$ de Y recouvre aussi X et on vérifie ainsi $\mu^*(X) \leq \sum_{k \in K} \mu(C_k)$.

Par passage à la borne inférieure sur les recouvrements de Y , on obtient bien $\mu^*(X) \leq \mu^*(Y)$.

Montrons maintenant la σ -sous-additivité.

Soient $(A_i)_{i \in \mathbb{N}}$ un ensemble dénombrable de parties de E disjointes deux à deux.

S'il existe $i \in \mathbb{N}$ tel que $\mu^*(A_i) = +\infty$, la σ -sous-additivité est évidente.

On suppose donc, pour tout i , $\mu^*(A_i) < +\infty$. Soit $\varepsilon > 0$ quelconque.

Pour tout $i \in \mathbb{N}$, il existe un ensemble dénombrable d'éléments $(C_k)_{k \in K_i}$ de \mathcal{C} qui recouvre A_i et tel que

$$\sum_{k \in K_i} \mu(C_k) \leq \mu^*(A_i) + \frac{\varepsilon}{2^{i+1}}.$$

Mais alors, l'ensemble dénombrable $(C_k)_{k \in \bigcup_i K_i}$ d'éléments de \mathcal{C} recouvre $\bigcup_{i \in \mathbb{N}} A_i$.

On en déduit

$$\mu^*\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \sum_{k \in K_i} \mu(C_k) \leq \sum_{i \in \mathbb{N}} \left(\mu^*(A_i) + \frac{\varepsilon}{2^{i+1}}\right) = \sum_{i \in \mathbb{N}} \mu^*(A_i) + \varepsilon$$

Et comme cela est vrai pour tout $\varepsilon > 0$, la σ -sous-additivité est bien prouvée. \square

On peut maintenant s'attaquer à la démonstration du théorème de Carathéodory.

Démonstration. L'unicité de la mesure découlera directement du théorème des classes monotones (l'étude des λ -systèmes engendrés par des π -systèmes réalisée dans un chapitre précédent).

En effet, un clan est un π -système et on a l'hypothèse de σ -finitude de E (E est union croissante d'éléments du clan de mesures finies).

Il faut donc s'attaquer à l'existence de la mesure. Pour tout entier p , on pose $\mathcal{C}_p = \{C \cap E_p, C \in \mathcal{C}\}$ et

$$\mathcal{T}_p = \{A \subset E_p / \forall X \subset E_p, \mu^*(A \cap X) + \mu^*((E_p - A) \cap X) = \mu^*(X)\}.$$

D'après le lemme, \mathcal{T}_p est une tribu sur laquelle μ^* est une mesure. On va maintenant montrer que \mathcal{T}_p contient en fait \mathcal{C}_p .

Considérons pour cela $X \subset E_p$ et $C \in \mathcal{C}_p$. On cherche à prouver que

$$\mu^*(C \cap X) + \mu^*(C^c \cap X) \leq \mu^*(X).$$

Par hypothèse $\mu^*(X) < +\infty$. Pour tout $\varepsilon > 0$, il existe un recouvrement $\bigcup_{n \in \mathbb{N}} D_n$ de X constitué d'éléments de

$$\mathcal{C}_p \text{ tel que : } \sum_{n \in \mathbb{N}} \mu(D_n) \leq \mu^*(X) + \varepsilon.$$

Or, $\bigcup_{n \in \mathbb{N}} D_n \cap C$ est un recouvrement de $X \cap C$ et $\bigcup_{n \in \mathbb{N}} D_n \cap C^c$ recouvre $X \cap C^c$.

D'autre part, pour tout n , on a $\mu(D_n) = \mu(D_n \cap C) + \mu(D_n \cap C^c)$ car μ est additive sur le clan \mathcal{C}_p .

Finalement,

$$\mu^*(X) + \varepsilon \geq \sum_{n \in \mathbb{N}} \mu(D_n \cap C) + \sum_{n \in \mathbb{N}} \mu(D_n \cap C^c) \geq \mu^*(X \cap C) + \mu^*(X \cap C^c)$$

On conclut que $C \in \mathcal{T}_p$; ce qui signifie que \mathcal{T}_p contient la tribu engendrée par \mathcal{C}_p .

Reste maintenant à passer à la limite sur p ... On considère donc \mathcal{T} la tribu engendrée par \mathcal{C} .

Pour tout p , on considère $\mathcal{T}_p = \{T \cap E_p, T \in \mathcal{T}\}$. Cette tribu est en fait engendrée par les \mathcal{C}_p (voir le lemme plus bas).

On en déduit que, pour tout p et pour toute famille dénombrable $(T_i)_{i \in \mathbb{N}}$ de \mathcal{T} ,

$$\mu^*\left(\bigcup_{i \in \mathbb{N}} (T_i \cap E_p)\right) = \sum_{i \in \mathbb{N}} \mu^*(T_i \cap E_p)$$

En utilisant le dernier lemme, on peut passer à la limite sur p et on obtient

$$\mu^*\left(\bigcup_{i \in \mathbb{N}} T_i\right) = \sum_{i \in \mathbb{N}} \mu^*(T_i)$$

\square

Pour conclure dans la démonstration précédente on a utilisé un premier résultat.

Lemme : Tribu trace et tribu engendrée

Soit \mathcal{C} un π -système qui engendre une tribu \mathcal{T} d'un ensemble E .

Soit $F \in \mathcal{C}$. On pose $\mathcal{D} = \{C \cap F, C \in \mathcal{C}\}$.

Alors la tribu trace de \mathcal{T} sur F est engendrée par \mathcal{D} .

Démonstration. On reprend ici des arguments portant sur les λ -systèmes engendrés par des π -systèmes.

On pose \mathcal{G} la tribu de F engendrée par \mathcal{D} et \mathcal{H} la tribu trace de \mathcal{T} sur F .

Il est clair que $\mathcal{H} \supset \mathcal{D}$ et, par suite, $\mathcal{H} \supset \mathcal{G}$.

Réciproquement, posons $\tilde{\mathcal{T}} = \{T \in \mathcal{T} / T \cap F \in \mathcal{G}\}$.

Il est clair que $\tilde{\mathcal{T}}$ est une sous-tribu de \mathcal{T} qui contient \mathcal{C} . On a donc en fait $\tilde{\mathcal{T}} = \mathcal{T}$.

Il en découle $\mathcal{H} \subset \mathcal{G}$. □

Reste maintenant à prouver le dernier lemme portant sur la convergence monotone.

Lemme : Famille σ -finie et convergence monotone

Soit E un ensemble. Soit \mathcal{C} un clan de E muni d'une mesure additive μ .

On suppose que :

- *on a construit une mesure extérieure μ^* à partir de μ ;*
- *il existe une suite croissante $(E_p)_{p \in \mathbb{N}}$ d'éléments de \mathcal{C} tous de mesure finies et telle que $\lim \uparrow E_p = E$;*
- *μ est continue à droite, c'est à dire que pour toute suite décroissante d'éléments F_p de \mathcal{C} telle que $\lim \downarrow F_p = \emptyset$ et $\mu(F_0) < +\infty$, on a $\lim \downarrow \mu(F_p) = 0$.*

Alors, pour tout $X \subset E$ tel que $\mu^(X) < +\infty$, on a $\lim \uparrow \mu^*(X \cap E_p) = \mu^*(X)$.*

La preuve de ce lemme constitue la fin de la démonstration du théorème de Carathéodory.

Démonstration. Pour tout p , on a $\mu^*(A_p \cap X) \leq \mu^*(X)$ donc $\lim \uparrow \mu^*(X \cap E_p) \leq \mu^*(X)$

Pour prouver l'inégalité réciproque, remarquons que, pour tout p ,

$$\mu^*(X) - \mu^*(X \cap E_p^c) \leq \mu^*(E_p \cap X)$$

Il suffit donc de prouver que $\lim \downarrow \mu^*(X \cap E_p^c) = 0$ pour conclure.

Mais nous savons qu'il existe une famille dénombrable $(C_k)_{k \in \mathbb{N}}$ d'éléments de \mathcal{C} qui recouvrent X et telle que $\mu^*(X) \leq \sum_{k \in \mathbb{N}} \mu(C_k) < +\infty$.

On en déduit, pour tout p que,

$$\mu^*(X \cap E_p^c) \leq \sum_{k \in \mathbb{N}} \mu(C_k \cap E_p^c) < +\infty$$

C'est là que l'on va utiliser l'argument de continuité à droite. Pour tout k , $\lim \downarrow \mu(C_k \cap E_p^c) = 0$. On peut donc utiliser le théorème de convergence monotone, dans sa version décroissante pour conclure... □

3.2 Applications

3.2.1 Théorème de Stieljes

Les fonctions continues à droites et croissantes occupent un rôle central dans cette mesure.

Proposition : Propriétés des fonction continue à droite et croissante

Soit F une fonction continue à droite et croissante.

Alors, l'ensemble des sauts de F est au plus dénombrable.

Démonstration. Considérons un intervalle $]n-1; n]$. Notons S_n l'ensemble des sauts de continuité de F sur cet intervalle.

Pour chaque réel a où F présente un saut de continuité, on pose $\phi(a) = F(a) - \lim_{t \rightarrow a^-} F(t) > 0$.

Et, pour tout entier $p > 0$, on pose $S_{n,p} = \left\{ a \in S_n / \phi(a) \geq \frac{1}{p} \right\}$.

$S_{n,p}$ est conçu de telle sorte que :

$$F(n) - F(n-1) \geq \sum_{a \in S_{n,p}} \phi(a) \geq \frac{1}{p} \sum_{a \in S_{n,p}} 1 = \frac{\#S_{n,p}}{n}$$

On en déduit que le cardinal de $S_{n,p}$ est fini. Or $S_n = \lim_{p \rightarrow +\infty} S_{n,p}$. Ainsi, S_n est dénombrable.

Or, l'ensemble des sauts s'obtient par union dénombrable des S_n . Ainsi, l'ensemble des sauts est dénombrable. \square

Proposition : Choix d'un clan

L'ensemble des unions finies d'intervalles de la forme $]a; b]$, $] - \infty; a]$ ou $]b; +\infty[$ avec $a < b$ forme un clan sur \mathbb{R} .

Démonstration. Assez simple. \square

Théorème : Stieljes

Soit une fonction F définie sur \mathbb{R} , croissante et continue à droite. Soient $a < b$ deux nombres.

Soient $a \leq b$ deux nombres.

En posant

$$\begin{aligned} s(]a; b]) &= F(b) - F(a) \\ s(]a; +\infty[) &= \lim_{t \rightarrow +\infty} F(t) - F(a) \\ s(]-\infty; a]) &= F(a) - \lim_{t \rightarrow -\infty} F(t) \end{aligned}$$

En étendant naturellement s sur les unions finies d'éléments deux à deux disjoints, on définit une mesure sur les éléments du clan de la proposition 3.2.1.

De plus, cette mesure peut être prolongée de manière unique à la tribu Borélienne.

On construit une première mesure extérieure en utilisant les intervalles ouverts *qui ne correspondent pas au clan*. On prouvera ensuite que cette mesure extérieure est égale à la mesure extérieure définie à partir d'éléments du clan.

Définition : Mesure extérieure de Stieljes à partir d'ouverts

Soient $a \leq b$ deux nombres.

On pose :

$$\begin{aligned} s_o(]a; b[) &= F(b) - F(a) \\ s_o(]a; +\infty[) &= \lim_{t \rightarrow +\infty} F(t) - F(a) \\ s_o(]-\infty; a[) &= F(a) - \lim_{t \rightarrow -\infty} F(t) \end{aligned}$$

Puis, on fabrique une mesure extérieure à partir de s_o notée s_o^ en utilisant des recouvrements dénombrables par des intervalles ouverts. On pose ainsi, pour tout $X \subset \mathbb{R}$:*

$$s_o^*(X) = \inf \left(\left\{ \sum_{k \in K} s_o(I_k) / (I_k)_{k \in K} \text{ est un recouvrement dénombrable de } X \text{ par des intervalles ouverts.} \right\} \right)$$

Démonstration. Il faut prouver que s_o^* est bien une mesure extérieure.

Soient X et Y deux sous-ensembles de \mathbb{R} tels que $X \subset Y$.

Tout recouvrement de Y recouvre aussi X . On obtient ainsi la croissance.

Traitons maintenant la σ -sous-additivité.

Considérons maintenant une famille $(X_i)_{i \in \mathbb{N}}$ d'éléments de $\mathcal{P}(\mathbb{R})$ disjoints deux à deux.

S'il existe $i \in \mathbb{N}$ tel que $s_o^*(X_i) = +\infty$, la propriété est évidente. On se place donc dans le cas contraire.

Pour tout $\varepsilon > 0$ et pour tout $i \in \mathbb{N}$, il existe un recouvrement dénombrable par des intervalles ouverts $(I_k)_{k \in K_i}$ tel que :

$$s_o^*(X_i) + \frac{\varepsilon}{2^{i+1}} \geq \sum_{k \in K_i} s_o(I_k)$$

Or, $(I_k)_{k \in \bigcup_{i \in \mathbb{N}} K_i}$ recouvre $\bigcup_{i \in \mathbb{N}} X_i$ et est dénombrable.

On obtient donc :

$$s_o^* \left(\bigcup_{i \in \mathbb{N}} X_i \right) \leq \sum_{\substack{k \in \bigcup \\ i \in \mathbb{N}} K_i} s_o(I_k) = \sum_{i \in \mathbb{N}} \sum_{k \in K_i} s_o(I_k) \leq \sum_{i \in \mathbb{N}} \left(s_o^*(X_i) + \frac{\varepsilon}{2^{i+1}} \right)$$

Cette dernière inégalité conduit à :

$$s_o^* \left(\bigcup_{i \in \mathbb{N}} X_i \right) \leq \varepsilon + \sum_{i \in \mathbb{N}} s_o^*(X_i)$$

Ce qui prouve la σ -sous-additivité. □

Proposition : Propriétés de la mesure extérieure de Stieljes à partir d'ouverts

Soient $a \leq b$ deux nombres. s_o^* désigne la mesure extérieure de Stieljes construite à partir d'une fonction F vérifiant les hypothèses adaptées.

Alors :

- $s_o^* (\{a\}) = F(a) - \lim_{t \rightarrow a^-} F(t)$;
- $s_o^* ([a; b]) = F(b) - \lim_{t \rightarrow a^-} F(t)$;
- $s_o^* (]a; b]) = s_o^* (]a; b])$;
- $s_o^* (]a; +\infty[) = s_o^* (]a; +\infty[)$;
- $s_o^* (]a - \infty; a]) = s_o^* (]a - \infty; a])$;
- s_o^* et s coïncident sur tous les éléments de clan de la proposition 3.2.1.

Démonstration. Pour tout entier naturel n , $]a - \frac{1}{n}; a + \frac{1}{n}[$ recouvre $\{a\}$ donc, par passage à la limite sur n , $s_o^* (\{a\}) \leq F(a) - F(a^-)$ car F est continue à droite.

Soit $\varepsilon > 0$ et un recouvrement de $\{a\}$ par des éléments du clan dont la somme des mesures est inférieure à $s_o^* (\{a\}) + \varepsilon$.

Bien sûr, il existe un intervalle unique $]c; d[$ qui recouvre a et dont la mesure demeure inférieure à $s_o^* (\{a\}) + \varepsilon$.

Dans ce cas, on a nécessairement $d > a > c$ et ainsi

$s_o^* (\{a\}) + \varepsilon \geq s_o^* (]c; d]) = F(d) - F(c) \geq F(a) - F(a^-)$ en raison de la croissance de F . Cela achève de prouver le premier point.

Attaquons maintenant le second point. Pour tout entier naturel n , $]a - \frac{1}{n}; b + \frac{1}{n}[$ recouvre $[a; b]$. Ainsi, par passage à la limite, en utilisant la continuité à droite de F :

$$s_o^* ([a; b]) \leq F(b) - F(a^-).$$

Réciproquement, pour tout $\varepsilon > 0$, il existe un recouvrement par un ensemble au plus dénombrable d'intervalles ouverts $(I_i)_{i \in I}$ tel que

$$\sum_{i \in I} s_o(I_i) \leq s_o^* ([a; b]) + \varepsilon.$$

Et comme $[a; b]$ est compact, on en déduit qu'on peut extraire un sous-recouvrement fini $(I_i)_{i \in J}$. De plus, on peut supposer sans risque que ce recouvrement est connexe.

Et on a ainsi :

$$\sum_{i \in J} s_o(I_i) \leq \sum_{i \in I} s_o(I_i) \leq s_o^* ([a; b]) + \varepsilon$$

Or, on peut montrer, en utilisant la croissance de F et la connexité du recouvrement, que $\sum_{i \in J} s_o(I_i) \geq F(\tilde{b}) - F(\tilde{a})$

avec $\tilde{b} = \max_{i \in J} b_i$ et $\tilde{a} = \min_{i \in J} a_i$. En particulier $\tilde{b} > b$ et $\tilde{a} < a$

Et comme $F(\tilde{b}) - F(\tilde{a}) \geq F(b) - F(a^-)$, on peut conclure concernant le second point.

Le troisième point se prouve en décomposant $[a; b]$ en $]a; b]$ et $\{a\}$. D'une part, il est clair que, pour tout entier naturel n ,

$$s_o^* (]a; b]) \leq s_o \left(\left] a; b + \frac{1}{n} \right[\right) \text{ car } \left] a; b + \frac{1}{n} \right[\text{ recouvre }]a; b]. \text{ Par passage à la limite sur } n, \text{ on obtient donc :}$$

$$s_o^* (]a; b]) \leq F(b) - F(a)$$

D'autre part, en raison de la σ -sous-additivité de la mesure extérieure :

$$s_o^*([a; b]) \leq s_o^*([a; b]) + s_o^*({a})$$

$$\text{Ainsi, } s^*([a; b]) \geq s^*([a; b]) - s^*({a}) = F(b) - F(a).$$

On va maintenant s'attaquer au quatrième point. En remarquant que $]a; +\infty[$ se couvre lui-même, il vient $s_o^*([a; +\infty]) \leq s([a; +\infty])$

Or, d'après le troisième point, pour tout entier naturel $n \geq a$, $s_o^*([a; n]) = s([a; n]) \leq s_o^*([a; +\infty])$.

Par passage à la limite sur n , on obtient $\lim_{n \rightarrow +\infty} F(n) - F(a) = s([a; +\infty]) \leq s_o^*([a; +\infty])$, ce qui permet de conclure quant à l'égalité des deux mesures pour cet intervalle.

La preuve du cinquième point est similaire à celle du quatrième point et nous ne la traiterons pas ici.

Pour prouver le sixième point, on considère un élément du clan. Il est composé d'une union finie d'intervalles $(A_k)_{k \in K}$ fermés à droite et ouverts à gauche ou de la forme $]-\infty; a]$ ou $]a; +\infty[$. De plus, on peut supposer que ces intervalles sont disjoints deux à deux et « espacés » entre eux.

Pour tout entier naturel n et pour tout $k \in K$, on pose

$$\widetilde{A}_k^{(n)} = \begin{cases} A_k & \text{si } \sup A_k = +\infty \\ \left] \inf A_k; \sup A_k + \frac{1}{n} \right[& \text{sinon} \end{cases}$$

Les $(\widetilde{A}_k^{(n)})_{k \in K}$ sont des ouverts qui couvrent $\bigcup_{k \in K} A_k$ et, par construction, on a :

$$s_o^*\left(\bigcup_{k \in K} A_k\right) \leq \sum_{k \in K} s_o\left(\widetilde{A}_k^{(n)}\right) = s\left(\bigcup_{k \in K} A_k\right) + \frac{\#A_k}{n}$$

Par passage à la limite sur n , on obtient :

$$s_o^*\left(\bigcup_{k \in K} A_k\right) \leq s\left(\bigcup_{k \in K} A_k\right).$$

Pour montrer l'inégalité réciproque, on suppose que $s_o^*\left(\bigcup_{k \in K} A_k\right) < +\infty$, le cas contraire étant trivial.

On considère ensuite $\varepsilon > 0$ et un ensemble dénombrable d'intervalles ouverts, notés $(I_i)_{i \in J}$ qui recouvre $\bigcup_{k \in K} A_k$ tels que

$$\sum_{i \in J} s(I_i) \leq s_o^*\left(\bigcup_{k \in K} A_k\right) + \varepsilon$$

Quitte à « rogner » un peu sur les I_i , on peut partitionner J en ensembles $(J_k)_{k \in K}$ tels que les familles $(I_i)_{i \in J_k}$ recouvrent les A_k .

On obtient ainsi, $\sum_{i \in J} s(I_i) = \sum_{k \in K} \sum_{i \in J_k} s(I_i) \geq \sum_{k \in K} s_o^*(A_k) = \sum_{k \in K} s(A_k)$, en raison de ce qui précède.

Finalement, on a :

$$s\left(\bigcup_{k \in K} A_k\right) = \sum_{k \in K} s(A_k) \leq \sum_{i \in J} s(I_i) \leq s_o^*\left(\bigcup_{k \in K} A_k\right) + \varepsilon, \text{ ce qui permet de conclure.} \quad \square$$

Enfin, on prouve que les mesures extérieures coïncident tout à fait.

Proposition : Les mesures extérieures coïncident.

Soit $X \subset \mathbb{R}$. Alors

$$s_o^*(X) = s^*(X)$$

Démonstration. Soit $X \subset \mathbb{R}$.

Avant de commencer, notons que, pour tout $a \leq b$, par construction, on a

$$s([a; b]) = s_o([a; b]).$$

Remarquons également que si $(]a_n; b_n])_{n \in \mathbb{N}}$ couvre X alors $(]a_n; b_n])_{n \in \mathbb{N}}$ couvre également X .

Les deux remarques précédentes prouvent que $s_o^*(X) \geq s^*(X)$.

Reste à prouver que $s^*(X) \geq s_o^*(X)$.

Si $s^*(X) = +\infty$, les deux mesures coïncident en raison de l'inégalité précédente.

On sait suppose donc $s^*(X) < +\infty$.

Soit $\varepsilon > 0$. Considérons alors un recouvrement dénombrable par des éléments $(A_n)_{n \in \mathbb{N}}$ du clan tels que $s^*(X) + \varepsilon \geq \sum_{n \in \mathbb{N}} s(A_n)$

Or, d'après ce qui précède, on sait que $s(A_n) = s_o^*(A_n)$.

En particulier, pour tout entier n , il existe un recouvrement dénombrable de A_n par des intervalles ouverts $(I_k)_{k \in K_n}$ tel que

$$s(A_n) \geq -\frac{\varepsilon}{2^{n+1}} + \sum_{k \in K_n} s(I_k)$$

Mais, $(I_k)_{k \in \bigcup_{n \in \mathbb{N}} K_n}$ est un ensemble dénombrable d'intervalles ouverts qui recouvrent X .

On en déduit, en sommant l'inégalité précédente sur n :

$$s^*(X) + \varepsilon \geq -\varepsilon + s_o^*(X)$$

Ce qui permet de conclure. □

On peut maintenant s'attaquer à la preuve du théorème de Stieljes qui devient trivial.

Démonstration. On dispose d'un clan défini dans la proposition 3.2.1 et d'une mesure s sur ce clan.

D'autre part \mathbb{R} est σ -fini puisque $\mathbb{R} = \bigcup_{n \in \mathbb{N}}]-n; n]$.

Enfin, la mesure s vérifie la propriété de continuité à droite puisque, d'après ce qui précède, pour tous les éléments C du clan, on a :

$$s^*(C) = s_o^*(C) = s(C)$$

Ainsi, la proposition 3.1.1 nous assure qu'on a bien la continuité à droite.

On est donc dans les hypothèses d'application du théorème de Carathéodory, ce qui nous permet de conclure. □

3.2.2 Mesure de Lebesgue

Proposition : Existence et unicité de la mesure de Lebesgue

Il existe une unique mesure λ sur les boréliens telle que $\lambda([a; b]) = b - a$ pour tous nombres $a < b$.

Démonstration. C'est une application du théorème de Stieljes avec $F(x) = x$. □

3.2.3 Quelques résultats de densité dans L^1

Pour prouver l'existence de la mesure de Stieljes, on a défini une mesure extérieure un peu particulière à l'aide d'intervalles ouverts et on a montré que cette mesure extérieure coïncide avec la mesure des éléments du clan.

En interprétant cette démarche en terme d'indicatrices, on dispose donc d'un résultat de densité dans L^1 .

Proposition : Densité des indicatrices d'intervalles dans L^p

On suppose que l'on a muni \mathbb{R} de la tribu des Boréliens et d'une mesure de Stieljes sur cette tribu.

On dispose des résultats de densité suivants, pour la norme L^1

- *l'ensemble des combinaisons linéaires finies d'indicatrices d'intervalles ouverts est dense dans l'ensemble des indicatrices d'éléments du clan ;*
- *l'ensemble des combinaisons linéaires finies d'indicatrices d'éléments du clan est dense dans l'ensemble des indicatrices de borélien ;*
- *l'ensemble des combinaisons linéaires finies d'indicatrices de boréliens est dense dans L^p .*

Démonstration. On reprend les hypothèses et notations du paragraphe concernant le théorème de Stieljes. Soit C un élément du clan de mesure finie. On sait que $s(C) = s_o^*(C)$.

Donc, pour $\varepsilon > 0$, il existe un recouvrement dénombrable d'intervalles ouverts $(I_i)_{i \in I}$ tels que

$$s(C) \leq \sum_{i \in I} s(I_i) \leq s(C) + \varepsilon$$

Et, en particulier, il existe un sous-ensemble fini $J \subset I$ tel que :

$$\sum_{i \in I} s(I_i) - \varepsilon \leq \sum_{i \in J} s(I_i) \leq \sum_{i \in I} s(I_i)$$

On en déduit finalement :

$$s(C) - \varepsilon \leq \sum_{i \in I} s(I_i) - \varepsilon \leq \sum_{i \in J} s(I_i) \leq \sum_{i \in I} s(I_i) \leq s(C) + \varepsilon$$

On démontre ainsi le premier point.

$$\left| \sum_{i \in J} s(I_i) - s(C) \right| \leq \varepsilon$$

Le second point se montre de manière tout à fait équivalente en utilisant le théorème de Carathéodory et la mesure extérieure induite par les éléments du clan.

Enfin, le troisième point est une conséquence directe de la définition de l'intégrale. \square

Proposition : Densité des fonctions continues à support compact dans L^p

L'ensemble des fonctions continues à support compact est dense dans L^p

3.2.4 Deux exemples : ensemble de Cantor, mesure non continue à droite

3.2.4.1 Importance de la continuité à droite

Considérons la fonction $F : x \mapsto \begin{cases} x & \text{si } x \leq 0 \\ x + 1 & \text{sinon} \end{cases}$.

La mesure de Stieljes de l'intervalle $]0; 1]$ donne $s(]0; 1]) = 2$.

Considérons maintenant la suite d'intervalles $\left] \frac{1}{n}; 1 \right]$. On vérifie aisément que

$$\lim_{n \rightarrow +\infty} s\left(\left] \frac{1}{n}; 1 \right]\right) = 1 \neq s(]0; 1])$$

Ainsi, on n'est pas dans les conditions de la proposition 3.1.1. Une telle mesure ne satisfait pas les attentes essentielles en terme de convergence...

3.2.4.2 Ensemble de Cantor

On pose $T_0 =]0; 1]$.

On suppose que l'on a construit T_n et qu'il est formé d'une union de K_n intervalles $(I_k^{(n)})_{k \in K_n}$ « espacés ».

On construit T_{n+1} en supprimant de chacun des $I_k^{(n)}$ un morceau central $\left[\frac{2 \inf I_k^{(n)} + \sup I_k^{(n)}}{3}; \frac{2 \sup I_k^{(n)} + \inf I_k^{(n)}}{3} \right]$.

Ainsi, T_{n+1} est composé de $2 \times K_n$ intervalles séparés et on a $\lambda(T_{n+1}) = \frac{2}{3} \lambda(T_n)$.

Par une récurrence immédiate, la suite de Boréliens ainsi construite est décroissante et, par le théorème de convergence monotone, $\lambda(\lim \downarrow T_n) = 0$.

Mais pour autant $\lim \downarrow T_n$ est non dénombrable.

En effet, on peut montrer que tout élément de $]0; 1]$ dont l'écriture en base 3 ne comprend pas de 1 appartient à cet ensemble. En particulier, il existe une injection de $\{0; 2\}^{\mathbb{N}}$ vers cet ensemble et comme $\{0; 2\}^{\mathbb{N}}$ n'est pas dénombrable on en déduit que $\lim \downarrow T_n$ ne l'est pas non plus.

L'ensemble de Cantor montre donc qu'il existe des Boréliens de mesure nulle et non dénombrable.

3.3 Le théorème de Riesz

Le théorème de Riesz établit que, si l'intégrale d'une fonction est une forme linéaire positive, alors réciproquement, sous certaines conditions, on peut définir une mesure à partir d'une forme linéaire positive définie sur un espace de fonctions.

Théorème : Théorème de Riesz

Soit $(X; d)$ un espace métrique localement compact et séparable.

Soit ϕ une forme linéaire positive sur l'ensemble $\mathcal{C}_K(X; \mathbb{R})$ des fonctions continues à support compact de X vers \mathbb{R} .

Alors il existe une unique mesure μ définie sur la tribu des boréliens de X et associée à ϕ , en ce sens que

$$\phi: f \mapsto \int f d\mu$$

Plus précisément μ vérifie :

- Pour tout ouvert ω , $\mu(\omega) = \sup \{ \phi(f) / f \in \mathcal{C}_K \text{ et } f \leq \mathbb{1}_\omega \}$;
- Pour tout compact K , $\mu(K) = \inf \{ \phi(f) / f \in \mathcal{C}_K \text{ et } f \geq \mathbb{1}_K \}$.

3.3.1 Le lemme d'Urysohn

Ce lemme technique permet de produire les fonctions continues à support compact qui vérifient des propriétés cruciales, en rapport avec les indicatrices mentionnées dans le théorème de Riesz.

Lemme : Urysohn

Soit $(X; d)$ un espace métrique localement compact. Alors :

- Pour tous fermés F et G disjoints, il existe deux ouverts U et V disjoints tels que $F \subset U$ et $G \subset V$. De plus, si F est compact on peut aussi avoir \overline{U} et \overline{V} disjoints et \overline{U} compact.
- Pour tout ouvert U non vide et $K \subset U$ compact, il existe une fonction φ continue à support compact telle que $\mathbb{1}_K \leq \varphi \leq \mathbb{1}_U$ et telle que le support de φ est inclus dans U .
- Pour tout compact K et ensemble fini d'ouverts $(U_i)_{i \in I}$ couvrant K , il existe une famille de fonctions continues à support compact $(\varphi_i)_{i \in I}$ telle que, pour tout $i \in I$, $\varphi_i \leq \mathbb{1}_{U_i}$, et le support de φ_i est inclus dans U_i et qui vérifie enfin :

$$\mathbb{1} \geq \sum_{i \in I} \varphi_i \geq \mathbb{1}_K$$

(Remarque 0) On rappelle que si F et G sont deux sous-ensembles de X , on note :

$$d(F; G) = \inf_{(x; y) \in F \times G} d(x; y)$$

(Remarque 1) On rappelle aussi que pour tout élément x , écrire $d(x; F) = 0$ signifie que tout voisinage de x rencontre F , ce qui revient à affirmer que $x \in \overline{F}$.

Démonstration. Montrons le premier point.

Pour tout $x \in F$. On a $d(x; G) > 0$ car $x \notin \overline{G} = G$. On pose donc :

$$U = \bigcup_{x \in F} B\left(x; \frac{d(x; G)}{3}\right)$$

De même, on peut poser :

$$V = \bigcup_{y \in G} B\left(y; \frac{d(y; F)}{3}\right)$$

U et V ainsi construits vérifient les hypothèses du premier point.

Si de plus F est compact, on peut trouver pour chaque $x \in F$ un voisinage compact V_x qui ne rencontre pas G . En particulier, il existe une union finie de ces voisinages compacts qui couvre F sans rencontrer G , ce qui permet de construire un compact L qui ne rencontre pas G et dont l'intérieur contient F .

Prouvons aussi que $d(L; G) > 0$. En effet, il existe deux suites (x_n) et (y_n) d'éléments de L et G tels que $\lim d(x_n; y_n) = d(L; G)$. Et comme L est compact, on peut extraire une sous-suite de (x_n) qui converge vers une limite $x \in L$, ce qui confirme que $d(L; G) = d(x; G) > 0$.

On termine en construisant V de telle sorte que $L \cap \overline{V} = \emptyset$.

$$V = \bigcup_{y \in G} B\left(y; \frac{d(L; G)}{3}\right)$$

Montrons maintenant le second point en supposant que $U \neq X$ (qui constitue un cas trivial). Il y a alors plusieurs manières de construire la fonction φ . On peut par exemple poser pour tout x :

$$\varphi(x) = \left[1 - 2 \frac{d(x; K)}{d(K; U^c)} \right]^+$$

φ est continue, majorée par 1 et son support (compact) est inclus dans U .

Enfin, pour tout $x \in K$, $\varphi(x) = 1$.

Une autre manière de construire φ consiste à se donner un compact L tel que $K \subset \overset{\circ}{L}$, et $L \subset U$, et de poser pour tout x :

$$\varphi(x) = \frac{d(x; L^c)}{d(x; L^c) + d(x; K)}$$

L'existence de L est garanti par le premier point du lemme.

Montrons enfin le dernier point en exploitant la compacité locale de X . D'après le premier point, pour tout $x \in K$, on peut trouver un voisinage compact V_x qui est inclus dans l'un des U_i . Et réciproquement pour chacun des U_i , on peut trouver un voisinage compact V_x d'un élément $x \in K$ inclus dans U_i .

Par compacité on peut avoir un recouvrement fini de K par des compacts $(L_j)_{j \in J}$ tel que :

- pour tout $i \in I$, il existe $j \in J$ vérifiant $L_j \subset U_i$;
- pour tout $j \in J$, il existe $i \in I$ vérifiant $L_j \subset U_i$.

Pour tout $i \in I$, on pose $M_i = \bigcup_{\substack{j \in J \\ L_j \subset U_i}} L_j$. Les $(M_i)_{i \in I}$ sont des compacts qui recouvrent K et on peut aussi

trouver, d'après le premier point, des compacts $(N_i)_{i \in I}$ tels que, pour tout i , $M_i \subset \overset{\circ}{N_i} \subset N_i \subset U_i$.

On construit maintenant les (φ_i) en s'inspirant du second point et en posant, pour tout i et pour tout $x \in X$:

$$\varphi_i(x) = \frac{d(x; M_i^c)}{d(x; K) + \sum_{k \in I} d(x; M_k^c)}$$

Par construction, pour tout $x \notin M_i \subset U_i$, $\varphi_i(x) = 0$ et on vérifie aussi $\mathbb{1} \geq \sum_{i \in I} \varphi_i \geq \mathbb{1}_K$. □

3.3.2 Le théorème de Riesz

Chapitre 4

L'intégrale de Riemann

Dans toute la suite E désigne un espace de Banach et $[a; b]$ est un intervalle fermé avec $a < b$.

4.1 Définition

4.1.1 Fonctions en escalier

Définition : Subdivision d'un intervalle fermé bornée, pas

On dit que $\sigma = (a_i)_{0 \leq i \leq n}$ est une subdivision de $[a; b]$ lorsque l'on a $a = a_0 < a_1 < a_2 < \dots < a_n = b$.
Le nombre $\max \{a_{i+1} - a_i, 0 \leq i \leq n-1\}$ s'appelle le pas de la subdivision.

Définition : Fonction en escalier, subdivision adaptée, intégrale de fonction en escalier

Soit $f : [a; b] \rightarrow E$ une fonction.

On dit que f est en escalier lorsqu'il existe une subdivision $\sigma = (a_i)_{0 \leq i \leq n}$ de $[a; b]$ telle que, pour tout $0 \leq i \leq n-1$, f est constante sur chacun des $]a_i; a_{i+1}[$; c'est à dire lorsqu'il existe des éléments de E notés $(y_i)_{0 \leq i \leq n-1}$ tels que $f|_{]a_i; a_{i+1}[} = y_i \times \mathbb{1}_{]a_i; a_{i+1}[}$.

Dans ce cas, on définit l'intégrale de Riemann f sur $[a; b]$ par

$$\int_a^b = \sum_{0 \leq i \leq n-1} y_i (a_{i+1} - a_i)$$

Et on dit que σ est une subdivision adaptée à f en ce sens que f est constante sur chacun des $]a_i; a_{i+1}[$.
En particulier la valeur de l'intégrale ne dépend pas de la subdivision adaptée choisie.

Pour aborder correctement cette notion, il faut donc détailler les propriétés des subdivisions.

Définition : Finesse de subdivision

Soient σ et σ' deux subdivisions de $[a; b]$. On dit que σ' est plus fine que σ lorsque $\sigma \subset \sigma'$. La finesse définit une relation d'ordre partielle sur l'ensemble des subdivisions de $[a; b]$.

Dans ce cas, le pas de σ' est inférieur au pas de σ .

Proposition : Subdivision la plus grossière pour une fonction en escalier

Soit f une fonction en escalier sur $[a; b]$. Il existe une subdivision adaptée la plus grossière σ avec le sens suivant :

Toute subdivision adaptée à f est plus fine que σ .

Plus généralement, si σ' est une subdivision adaptée à f alors toute subdivision σ'' plus fine que σ est également adaptée.

Démonstration. Soit $\sigma = (a_i)_{0 \leq i \leq n}$ une subdivision adaptée à f .

Pour tous les $i \in \llbracket 1; n-1 \rrbracket$ tels que $f(]a_{i-1}; a_i]) = f(a_i) = f(]a_i; a_{i+1}[)$, on supprime a_i de σ . On obtient ainsi une subdivision $\tilde{\sigma} = (\alpha_i)_{0 \leq i \leq r}$ plus grossière que σ .

On peut affirmer que toute subdivision $\sigma' = (\beta_i)_{0 \leq i \leq p}$ adaptée à f est plus fine que $\tilde{\sigma}$. Pourquoi?

On sait que $\alpha_0 \in \sigma'$ et $\alpha_r \in \sigma'$. Supposons maintenant par l'absurde qu'il existe $i \in \llbracket 1; r-1 \rrbracket$ tel que $\alpha_i \notin \sigma'$. Dans ce cas, il existe $j \in \llbracket 0; p-1 \rrbracket$ tel que $\alpha_i \in]\beta_j; \beta_{j+1}[$. Mais dans ce cas, on aurait $f(]a_{i-1}; \alpha_i]) = f(\alpha_i) = f(]a_i; a_{i+1}[)$ car f est constante sur $] \beta_j; \beta_{j+1}[$.

On aboutirait donc, par construction, à une situation absurde.

La dernière partie de la proposition est évidente. \square

Cette petite réflexion permet de revenir sur l'unicité de l'intégrale de Riemann d'une fonction en escalier.

Démonstration. Soit $\sigma = (a_i)_{0 \leq i \leq n}$ une subdivision adaptée à f avec $(y_i)_{0 \leq i \leq n-1}$ les valeurs de f correspondantes et $\tilde{\sigma} = (\alpha_i)_{0 \leq i \leq r}$ la subdivision adaptée la plus grossière avec $(\tilde{y}_i)_{0 \leq i \leq r-1}$ les valeurs de f correspondantes.

Alors,

$$\sum_{0 \leq i \leq n-1} y_i (a_{i+1} - a_i) = \sum_{0 \leq i \leq r-1} \tilde{y}_i (\alpha_{i+1} - \alpha_i)$$

En effet, pour tout $i \in \llbracket 0; r-1 \rrbracket$, on pose $(a_{j_k})_{1 \leq k \leq n_i}$ les éléments de σ appartenant à $] \alpha_i; \alpha_{i+1}[$. On vérifie alors, par construction, que :

$$\sum_{1 \leq k \leq n_i} y_{j_k-1} (a_{j_k} - a_{j_k-1}) = y_i (\alpha_{i+1} - \alpha_i)$$

En réarrangeant ainsi les termes de la première somme « par paquets », on obtient le résultat escompté. \square

Proposition : Propriétés des fonctions en escalier, linéarité de l'intégrale

Soient f et g deux fonctions en escalier sur $[a; b]$ et à valeurs dans E un espace de Banach.

Soit λ un scalaire.

Alors $f + \lambda g$ est une fonction en escalier.

En particulier l'ensemble des fonctions en escalier est un sous-espace vectoriel des fonctions de $[a; b]$ dans E et, sur cet espace vectoriel, l'intégrale de Riemann est une application linéaire.

Enfin, une fonction en escalier est bornée.

Démonstration. Si σ_f et σ_g sont des subdivisions adaptées respectivement à f et à g alors $\sigma_f \cup \sigma_g$ est une subdivision adaptée aux deux fonctions à la fois.

Partant de cette remarque, sur $\sigma_f \cup \sigma_g = (a_i)_{0 \leq i \leq n}$, il est clair que $f + \lambda g$ est en escalier car, pour tout $i \in \llbracket 0; n-1 \rrbracket$, $(f + \lambda g)_{|]a_i; a_{i+1}[} = f_{|]a_i; a_{i+1}[} + \lambda g_{|]a_i; a_{i+1}[}$ et cette restriction est donc constante.

L'application de la formule de l'intégrale à $f + \lambda g$ permet de conclure quant à la linéarité de l'intégrale.

Une fonction en escalier est bien sûr bornée car elle ne prend qu'un nombre fini de valeurs. \square

Proposition : Majoration de la norme de l'intégrale d'une fonction en escalier

Soit f une fonction en escalier à valeurs dans E et $\sigma = (a_i)_{0 \leq i \leq n}$ une subdivision adaptée à f et $(y_i)_{0 \leq i \leq n-1}$ les valeurs de f associées.

Alors

$$\left\| \int_a^b f \right\| \leq \int_a^b \|f\|$$

et, en particulier :

$$\left\| \int_a^b f \right\| \leq |b-a| \|f\|_\infty$$

Avec $\|f\|_\infty = \max_{x \in [a; b]} \|f(x)\|$.

Démonstration. Ce sont des conséquences assez directes de l'inégalité triangulaire. \square

4.1.2 Intégrale de fonctions

Proposition : Fonctions en escalier à valeurs réelles

Si f et g sont deux fonctions en escalier à valeurs réelles alors

- fg est en escalier ;
- $\min(f, g)$ est en escalier ;
- $\max(f, g)$ est en escalier.

Démonstration. Cela se prouve très facilement en considérant une subdivision adaptée à f et g . \square

Définition : Intégrale de fonction

Soit f une fonction de $[a; b]$ dans E . On dit que f admet une intégrale lorsqu'il existe une suite de fonctions en escaliers φ_n à valeurs dans E et une suite de fonctions en escaliers ψ_n à valeurs dans \mathbb{R}^+ telles que :

- pour tout n et pour tout x de $[a; b]$, $\|\varphi_n(x) - f(x)\| \leq \psi_n(x)$;
- $\lim_{n \rightarrow +\infty} \int_a^b \psi_n = 0$.

Dans ce cas, $\lim_{n \rightarrow +\infty} \int_a^b \varphi_n$ existe et on pose :

$$\int_a^b f = \lim_{n \rightarrow +\infty} \int_a^b \varphi_n$$

En particulier cette définition ne dépend pas des suites φ_n et ψ_n choisies.

Démonstration. On va commencer par prouver la convergence puis l'unicité.

C'est ici qu'intervient la complétude de E .

Pour tout n et p ,

$$\left\| \int_a^b \varphi_{n+p} - \int_a^b \varphi_n \right\| \leq \int_a^b \|\varphi_{n+p} - \varphi_n\|$$

Or, pour tout x , $\|\varphi_{n+p} - \varphi_n\| \leq \|\varphi_{n+p}(x) - f(x)\| + \|f(x) - \varphi_n(x)\| \leq \psi_{n+p}(x) + \psi_n(x)$. Finalement,

$$\left\| \int_a^b \varphi_{n+p} - \int_a^b \varphi_n \right\| \leq \int_a^b \psi_n + \int_a^b \psi_{n+p}$$

Cette inégalité permet de prouver que la suite $\int_a^b \varphi_n$ est de Cauchy puisque la suite $\int_a^b \psi_n$ tend vers 0.

Reste à prouver maintenant l'unicité.

On considère donc $\tilde{\varphi}_n$ et $\tilde{\psi}_n$ un couple de suites de fonctions vérifiant les deux hypothèses de la définition.

Par des techniques semblables, on peut prouver, pour tout n , que

$$\left\| \int_a^b \varphi_n - \int_a^b \tilde{\varphi}_n \right\| \leq \int_a^b \psi_n + \int_a^b \tilde{\psi}_n$$

Et là encore, on peut conclure quant à l'unicité de la limite ! \square

Proposition : Propriétés des fonctions qui admettent une intégrale de Riemann

Ces fonctions forment un sous-espace vectoriel des fonctions bornées de $[a; b]$ dans E .

De plus, l'intégrale est une application linéaire sur cet espace.

Démonstration. Cela se prouve très facilement avec quelques inégalités triangulaires.

Ainsi, si f_1 et f_2 admettent des intégrales, que l'on associe à f_1 les suites de fonctions en $\varphi_n^{(1)}$ escalier et $\psi_n^{(1)}$, que l'on associe à f_2 les suites de fonctions en $\varphi_n^{(2)}$ escalier et $\psi_n^{(2)}$, on peut vérifier que les suites $\varphi_n^{(1)} + \lambda\varphi_n^{(2)}$ et $\psi_n^{(1)} + \|\lambda\| \psi_n^{(2)}$ vérifient pour tout x de $[a; b]$

$$\left\| (f_1 + \lambda f_2)(x) - \left(\varphi_n^{(1)} + \lambda \varphi_n^{(2)} \right) (x) \right\| \leq \left(\psi_n^{(1)} + \|\lambda\| \psi_n^{(2)} \right) (x)$$

et qu'on a bien $\lim_{n \rightarrow +\infty} \int_a^b \left(\psi_n^{(1)} + \|\lambda\| \psi_n^{(2)} \right) = 0$.

Par passage à la limite et grâce à la linéarité de l'intégrale sur les fonctions en escalier, on peut conclure.

Pour finir, il est clair que, pour tout x ,

$\|f(x)\| \leq \|\varphi_0(x)\| + \psi_0(x)$, ce qui prouve qu'une fonction qui admet une intégrale est minorée en norme par une fonction bornée de \mathbb{R}^+ . Elle est donc bornée. \square

Un lemme qui sera utile pour la suite.

Lemme : Construction d'une suite « optimale » de fonctions en escaliers

Soit f une fonction intégrable au sens de Riemann sur un intervalle $[a; b]$.

Alors, il est possible de construire les suites φ_n et ψ_n de fonctions en escalier telle sorte que

- pour tout n , et pour tout $t \in [a; b]$, $\|f(t) - \varphi_n(t)\| \leq \psi_n(t)$;
- ψ_n est la plus petite fonction en escalier compte-tenu de φ_n ;
- ψ_n décroît vers une fonction d'intégrale nulle.

En particulier, les suites ψ_n et φ_n sont bornées.

Démonstration. On considère une suite φ_n et ψ_n vérifiant les hypothèses de la définition de l'intégrale.

On sait que f est bornée.

Au rang 0 : on considère σ_0 la subdivision la plus grossière adaptée à φ_0 .

Sur chacun des sous-intervalles ouverts de σ_0 , on détermine la valeur de $\tilde{\psi}_0$ comme étant la borne supérieure de $\|f - \varphi_0\|$. Enfin, on pose $\tilde{\varphi}_0 = \varphi_0$. Sur les points de la subdivision, on définit $\tilde{\phi}_0 = f$.

On suppose construite ainsi la suite jusqu'à un certain rang n .

Au rang $n + 1$, on considère la subdivision σ_{n+1} la plus grossière adaptée à φ_{n+1} et $\tilde{\phi}_n$.

Sur chacun des sous-intervalles de cette subdivision, on choisit la valeur de $\tilde{\varphi}_{n+1}$ entre φ_{n+1} et $\tilde{\phi}_n$ de telle sorte que la borne supérieure de $\|f - \tilde{\phi}_{n+1}\|$ soit minimale et on pose $\tilde{\psi}_{n+1}$ cette borne supérieure.

Par construction, on dispose donc d'une suite $\tilde{\psi}_n$ inférieure à ψ_n et d'une suite $\tilde{\varphi}_n$ qui vérifient les hypothèses de la définition de l'intégrale et qui satisfont les conditions du lemme ! \square

Théorème : Mesure de Lebesgue et intégrale de Riemann

Soit f une fonction intégrable au sens de Riemann et à valeurs dans E .

Alors, f est presque partout limite simple d'une suite de fonctions en escaliers.

En particulier, si $E = \mathbb{R}$ ou \mathbb{C} , f est alors mesurable au sens de Lebesgue.

Démonstration. On reprend les notations et hypothèses de l'intégrabilité au sens de Riemann.

Notons qu'une fonction en escalier à valeurs réelles est mesurable, comme combinaisons linéaires de fonctions mesurables. D'autre part, il est clair que l'intégrale de Riemann d'une telle fonction coïncide avec l'intégrale de Lebesgue.

En particulier, on peut utiliser le théorème de convergence monotone et on obtient :

$$\lim \downarrow \int_a^b \psi_n = \lim \downarrow \int_{[a; b]} \psi_n(x) d\lambda(x) = \int_{[a; b]} \lim \downarrow \psi_n(x) d\lambda(x)$$

Par conséquent, la fonction $\lim \downarrow \psi_n$ est presque partout nulle.

Mais comme, pour tout x , $\|\varphi_n(x) - f(x)\| \leq \psi_n(x)$, on en déduit la convergence simple presque partout.

Dans le cas réel, cette remarque nous montre que f est mesurable. \square

Proposition : Produit de fonctions intégrables

Soit f et θ deux fonctions définies sur $[a; b]$ à valeurs respectivement dans E et \mathbb{R} .
Si f et θ sont intégrables au sens de Riemann alors θf l'est aussi.

Démonstration. Soient (φ_n, ψ_n) et (ρ_n, χ_n) les suites de fonctions en escaliers associées aux intégrales de f et θ . Pour tout n et pour tout $t \in [a; b]$:

$$\|\theta(t)f(t) - \rho_n(t)\varphi_n(t)\| \leq \|\theta(t)f(t) - \rho_n(t)f(t)\| + \|\rho_n(t)f(t) - \rho_n(t)\varphi_n(t)\| \leq \|f(t)\| \times \chi_n(t) + \|\rho_n(t)\| \times \psi_n(t)$$

On peut conclure car on sait que f et ρ_n sont bornées, indépendamment de n et t . \square

4.2 Applications

4.2.1 Comparaison séries-intégrales

Lemme : Règle de Riemann sur les séries à termes positifs

Soit $\sum_{n \in \mathbb{N}} a_n$ une série à termes positifs.

S'il existe $\alpha > 1$ tel que $n^\alpha a_n \xrightarrow{n \rightarrow +\infty} 0$ alors la série converge.

S'il existe $0 \leq \alpha < 1$ tel que $n^\alpha a_n \xrightarrow{n \rightarrow +\infty} 0$ alors la série diverge.

On rappelle que l'on connaît la nature des séries de Riemann $\sum \frac{1}{n^\alpha}$ grâce à l'étude de la série $\sum \frac{1}{n}$ et grâce aux comparaisons séries-intégrales.

Démonstration. Le cas $\alpha > 1$ est facile car on a alors $a_n = o\left(\frac{1}{n^\alpha}\right)$.

On va s'intéresser au cas $0 \leq \alpha < 1$ et supposer $n^\alpha a_n \xrightarrow{n \rightarrow +\infty} 0$. En particulier, il existe $\eta > 0$ et une extractrice φ telle que $\varphi(n)^\alpha \times a_{\varphi(n)} \geq \eta$, ce qui donne $a_{\varphi(n)} \geq \frac{\eta}{\varphi(n)^\alpha}$. Or la série des $\sum \frac{1}{n^\alpha}$ tend vers $+\infty$, ce qui permet de conclure ! \square

4.2.2 Intégrale de Riemann et suites

Théorème : Convergence

Soit une suite de fonctions f_n définies sur un intervalle $[a; b]$ et à valeurs dans un espace de Banach E .

On suppose que :

- les f_n sont Riemann intégrables ;
- la suite f_n converge uniformément vers une fonction f ;
- f est Riemann intégrable.

Alors

$$\lim_{n \rightarrow +\infty} \int_a^b f_n = \int_a^b f$$

Démonstration. Soit $\varepsilon > 0$. Il existe N tel que, pour tout $n \geq N$, $\|f_n - f\|_\infty < \frac{\varepsilon}{b-a}$, ce qui entraîne :

$$\left\| \int_a^b f_n - \int_a^b f \right\| \leq (b-a) \|f_n - f\|_\infty < \varepsilon$$

\square

Une conséquence de ce théorème concerne la dérivation sous le signe intégrale.

Théorème : Dérivation et intégrale de Riemann

Soit I un intervalle ouvert et f une fonction définie sur $[a; b] \times I$ à valeurs dans \mathbb{R} .

On suppose que :

- pour tout y de I , la fonction $x \mapsto f(x; y)$ est Riemann intégrable ;
- pour tout x de $[a; b]$, la fonction $y \mapsto f(x; y)$ est dérivable ;
- pour tout y de I , la fonction $y \mapsto \frac{\partial f(x; y)}{\partial y}$ est Riemann intégrable ;
- la fonction $(x; y) \mapsto \frac{\partial f(x; y)}{\partial y}$ est uniformément continue sur $[a; b] \times I$.

Alors, la fonction $y \mapsto \int_a^b f(x; y) dx$ est dérivable et sa dérivée est $y \mapsto \int_a^b \frac{\partial f(x; y)}{\partial y} dx$.

Démonstration. On pose $\varphi : y \mapsto \int_a^b f(x; y) dx$.

Pour tout y de I , on considère une suite h_n de réels non nuls tels que, pour tout n , $h_n + y \in I$ et $\lim_{n \rightarrow +\infty} h_n = 0$.

Alors, pour tout n , $\frac{\varphi(y + h_n) - \varphi(y)}{h_n} - \int_a^b \frac{\partial f(x; y)}{\partial y} dx = \int_a^b \left(\frac{f(x; y + h_n) - f(x; y)}{h_n} - \frac{\partial f(x; y)}{\partial y} \right) dx$ en raison de la linéarité de l'intégrale.

On conclut ensuite grâce à l'uniforme continuité de la dérivée partielle. \square

4.2.3 Le cas des dimensions finies

On rappelle qu'en dimension finie, toutes les normes sont équivalentes.

4.2.3.1 Le cas réel

On étudie ici le cas où l'espace image est \mathbb{R} . On dispose alors d'une formulation équivalente beaucoup plus commode.

Lemme : Premier critère d'intégration au sens de Riemann pour les fonctions à valeur réelle

Une fonction $f : [a; b] \rightarrow \mathbb{R}$ est intégrable au sens de Riemann si et seulement s'il existe deux suites de fonctions en escalier $(\varphi_n^b)_{n \in \mathbb{N}}$ et $(\varphi_n^h)_{n \in \mathbb{N}}$ telles que :

- pour tout x de $[a; b]$, et pour tout n , $\varphi_n^b(x) \leq f(x) \leq \varphi_n^h(x)$;
- la suite φ_n^b est croissante et la suite φ_n^h est décroissante ;
- $\lim_{n \rightarrow +\infty} \int_a^b (\varphi_n^h - \varphi_n^b) = 0$

Démonstration. On reprend les conditions d'existence d'intégrale avec les mêmes notations.

On a alors pour tout n et tout x , $\varphi_n(x) - \psi_n(x) \leq f(x) \leq \varphi_n(x) + \psi_n(x)$.

On pose ainsi $\varphi_0^h(x) = \varphi_0(x) + \psi_0(x)$ et $\varphi_0^b(x) = \varphi_0(x) - \psi_0(x)$.

On suppose que l'on a construit jusqu'à un certain rang n la suite des φ_n^h et φ_n^b . Au rang $n + 1$, il suffit de poser

$\varphi_{n+1}^h = \min(\varphi_n^h, \varphi_{n+1} + \psi_{n+1})$ et $\varphi_{n+1}^b = \max(\varphi_n^b, \varphi_{n+1} - \psi_{n+1})$ qui sont des fonctions en escalier.

On fabrique ainsi par récurrence les suites φ_n^h et φ_n^b et ces suites valident par construction le sens direct du lemme.

Le sens réciproque est évident. Il suffit pour cela de poser $\psi_n = \varphi_n^h - \varphi_n^b$. \square

De plus, dans le cas réel, l'intégrale de Riemann est positive.

Proposition : Positivité de l'intégrale

Soient f et g deux fonctions Riemann intégrables sur $[a; b]$ et à valeurs réelles.

Si, pour tout $x \in [a; b]$, $f(x) \leq g(x)$ alors $\int_a^b f \leq \int_a^b g$.

Démonstration. Cela se prouve en intégrant $g - f$ qui est positive et en utilisant la linéarité. \square

Définition : Sommes de Darboux de f associées à une subdivision

Soit f une fonction de $[a; b]$ dans \mathbb{R} bornée.

Soit $\sigma = (a_i)_{0 \leq i \leq n}$ une subdivision de $[a; b]$.

Pour tout $i \in \llbracket 0; n-1 \rrbracket$, on pose

$$\begin{aligned} M_i &= \sup_{]a_i; a_{i+1}[} f \\ \overline{M}_i &= \sup_{[a_i; a_{i+1}[} f \\ m_i &= \inf_{]a_i; a_{i+1}[} f \\ \overline{m}_i &= \inf_{[a_i; a_{i+1}[} f \end{aligned}$$

On définit alors

$$\begin{aligned} S_\sigma^h &= \sum_{i \in \llbracket 0; n-1 \rrbracket} M_i \times \mathbb{1}_{]a_i; a_{i+1}[} + \sum_{i \in \llbracket 0; n \rrbracket} f(a_i) \times \mathbb{1}_{a_i} \\ S_\sigma^d &= \sum_{i \in \llbracket 0; n-1 \rrbracket} m_i \times \mathbb{1}_{]a_i; a_{i+1}[} + \sum_{i \in \llbracket 0; n \rrbracket} f(a_i) \times \mathbb{1}_{a_i} \\ \overline{S}_\sigma^h &= \sum_{i \in \llbracket 0; n-1 \rrbracket} \overline{M}_i \times \mathbb{1}_{[a_i; a_{i+1}[} + \sum_{i \in \llbracket 1; n \rrbracket} f(a_i) \times \mathbb{1}_{a_i} \\ \overline{S}_\sigma^d &= \sum_{i \in \llbracket 0; n-1 \rrbracket} \overline{m}_i \times \mathbb{1}_{[a_i; a_{i+1}[} + \sum_{i \in \llbracket 1; n \rrbracket} f(a_i) \times \mathbb{1}_{a_i} \end{aligned}$$

Démonstration. L'existence de ces sommes est garantie par le fait que f soit bornée. \square

Cette construction justifie la proposition suivante

Proposition : Propriété fondamentale des sommes de Darboux

On reprend les mêmes notations et hypothèses que dans la définition.

Pour toutes fonctions en escaliers φ^h et φ^b associées à σ et telles que $\varphi^b \leq f \leq \varphi^h$, on a

$$\varphi^b \leq S_\sigma^b \leq f \leq S_\sigma^h \leq \varphi^h$$

Démonstration. Pour tout $i \in \llbracket 0; n-1 \rrbracket$, et pour tout $x \in]a_i; a_{i+1}[$, on a $y_i^b \leq f(x)$ et $f(x) \leq y_i^h$ où y_i^b et y_i^h représentent les valeurs (constantes) de φ^b et φ^h sur l'intervalle considéré. Par passage à la borne inférieure sur $f(x)$ dans la première inégalité et à la borne supérieure dans la seconde, on obtient le résultat escompté. \square

Par construction les sommes de Darboux réalisent un encadrement optimal de f par des fonctions en escalier pour une subdivision donnée. On en déduit ainsi le corollaire.

Corollaire : Sommes de Darboux et encadrement d'intégrale

On reprend les mêmes notations. Pour toute subdivision σ et pour toute fonction f intégrable :

$$\int_a^b S_\sigma^b \leq \int_a^b f \leq \int_a^b S_\sigma^h$$

En particulier :

$$\left(\inf_{[a; b]} f \right) \times (b - a) \leq \int_a^b f \leq \left(\sup_{[a; b]} f \right) \times (b - a)$$

Proposition : Sommes de Darboux et subdivisions plus fines

On reprend les mêmes hypothèses.

Soit $\sigma \subset \sigma'$ deux subdivisions.

Alors :

$$S_\sigma^b \leq S_{\sigma'}^b \leq f \leq S_{\sigma'}^h \leq S_\sigma^h$$

et

$$\overline{S_\sigma^b} \leq \overline{S_{\sigma'}^b} \leq f \leq \overline{S_{\sigma'}^h} \leq \overline{S_\sigma^h}$$

Démonstration. C'est évident par construction des sommes de Darboux. □

Théorème : Convergence des sommes de Darboux vers l'intégrale

Soit f une fonction bornée de $[a; b]$ dans \mathbb{R} .

f est intégrable au sens de Riemann si et seulement si, pour tout $\varepsilon > 0$, il existe un nombre $\eta > 0$ tel que, pour toute subdivision σ de pas inférieur à η ,

$$\int_a^b (S_\sigma^h - S_\sigma^b) < \varepsilon$$

Démonstration. Le sens réciproque est évident.

Montrons le sens direct. Ainsi, on suppose f intégrable au sens de Riemann.

On va considérer deux suites de fonctions en escaliers φ_n^h et φ_n^b telles qu'exposées dans le lemme du début de ce paragraphe.

Soit $\varepsilon > 0$ un nombre.

On sait qu'il existe n tel que $\int_a^b (\varphi_n^h - \varphi_n^b) < \frac{\varepsilon}{2}$

Soit $\sigma = (a_i)_{0 \leq i \leq p}$ la subdivision adaptée à ces deux fonctions en escalier.

On pose enfin M un majorant de $|\varphi_n^b|$ et $|\varphi_n^h|$ sur $[a; b]$.

Pour tout intervalle, $]c; d[\subset]a; b[$, on sait que

$$\int_c^d \left(\sup_{]c; d[} f - \inf_{]c; d[} f \right) \leq 2M(d - c).$$

Fort de cette remarque, et du fait que φ_n^h et φ_n^b contiennent au plus $p + 1$ sauts, on va poser $\eta = \frac{\varepsilon}{4M(p + 1)}$.

Pour toute subdivision $\sigma' = (\alpha_i)_{0 \leq i \leq q}$ de pas plus fin que η , la majoration de $\int_a^b (S_{\sigma'}^h - S_{\sigma'}^b)$ se fait en distinguant deux cas sur les $i \in \llbracket 0; q - 1 \rrbracket$:

- s'il existe $j \in \llbracket 0; p-1 \rrbracket$ tel que $] \alpha_i; \alpha_{i+1}[\subset] a_j; a_{j+1}[,$ alors

$$\int_{\alpha_i}^{\alpha_{i+1}} (S_{\sigma'}^h - S_{\sigma'}^b) \leq \int_{\alpha_i}^{\alpha_{i+1}} (\varphi_n^h - \varphi_n^b)$$
- s'il existe $j \in \llbracket 1; p-1 \rrbracket$ tel que $a_j \in] \alpha_i; \alpha_{i+1}[,$ alors

$$\int_{\alpha_i}^{\alpha_{i+1}} (S_{\sigma'}^h - S_{\sigma'}^b) \leq 2M(\alpha_{i+1} - \alpha_i) \leq 2M\eta = \frac{\varepsilon}{2(p+1)}$$

Mais on rappelle qu'il y a au plus $p+1$ discontinuités, correspondant au cas n° 2.

En sommant sur tous les $i \in \llbracket 0; q-1 \rrbracket$, on obtient la majoration recherchée :

$$\int_a^b (S_{\sigma'}^h - S_{\sigma'}^b) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

□

On peut prouver que ce théorème s'étend aussi aux sommes de Darboux « étendues » \overline{S}_{σ}^h et \overline{S}_{σ}^b .

Une des conséquences de ce théorème concerne les sommes de Riemann, sujet sur lequel nous ne nous étendrons pas.

Définition : Sommes de Riemann, subdivisions pointées

Une subdivision pointée d'un intervalle $[a; b]$ est la donnée d'une subdivision $\sigma = (a_i)_{0 \leq i \leq n}$ de $[a; b]$ et de nombres $t = (t_i)_{0 \leq i \leq n-1}$ tels que pour tout $i \in \llbracket 0; n-1 \rrbracket$, $t_i \in [a_i; a_{i+1}]$.

Pour toute subdivision pointée $(\sigma; t)$ de $[a; b]$, on construit la somme de Riemann associée à la fonction f de la manière suivante :

$$R_{\sigma, t}(f) = \sum_{i \in \llbracket 0; n-1 \rrbracket} (a_{i+1} - a_i) f(t_i)$$

On a alors le théorème suivant

Théorème : Sommes de Riemann et intégration

Une fonction f est intégrable au sens de Riemann sur $[a; b]$ si et seulement si il existe un nombre $\int_a^b f$ tel que, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, pour toute subdivision pointée $(\sigma; t)$ de $[a; b]$ de pas inférieur à η , on a

$$\left| R_{\sigma, t}(f) - \int_a^b f \right| < \varepsilon$$

Démonstration. C'est une conséquence assez immédiate des sommes de Darboux.

□

4.2.3.2 Extension aux espaces de dimensions finies

On suppose cette fois que f est à valeurs dans un espace E de dimension finie, notée n .

On peut alors étendre ce qui précède concernant les sommes de Riemann.

Corollaire : Sommes de Riemann et intégration pour un espace de dimension finie

Une fonction f est intégrable au sens de Riemann sur $[a; b]$ si et seulement si il existe un nombre $\int_a^b f$ tel que, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, pour toute subdivision pointée $(\sigma; t)$ de $[a; b]$ de pas inférieur à η , on a

$$\left\| R_{\sigma, t}(f) - \int_a^b f \right\| < \varepsilon$$

Démonstration. Soit $(e_i)_{1 \leq i \leq n}$ une base de E et $f^{(i)}$ les coordonnées de f sur chacun des e_i .

Nous savons que toutes les normes de E sont équivalentes. Nous choisissons la norme infinie associée à cette base.

Si f est Riemann intégrable, il existe deux suites de fonctions en escaliers φ_n et ψ_n telles que

- pour tout $t \in [a; b]$, $\|\varphi_n(t) - f(t)\|_\infty \leq \psi_n(t)$;
- $\int_a^b \psi_n \xrightarrow{n \rightarrow +\infty} 0$.

En considérant les coordonnées de φ_n sur chacun des e_i , fonctions notées $\varphi_n^{(i)}$, on est ramené au cas réel pour chacune des coordonnées.

Ainsi, pour tout $\varepsilon > 0$ et pour tout i , il existe un nombre $\int_a^b f^{(i)}$ et nombre $\eta_i > 0$ tel que pour toute subdivision pointée $(\sigma_i; t_i)$ de pas inférieur à η_i , on a

$$\left| R_{\sigma_i, t_i}(f^{(i)}) - \int_a^b f^{(i)} \right| < \varepsilon$$

En considérant, $\eta = \min_{1 \leq i \leq n} (\eta_i)$, on obtient le résultat escompté.

La réciproque se traite de la même manière, en raisonnant sur les coordonnées. □

4.2.3.3 Formules de changement de variable

Dans toute la suite E désigne un \mathbb{R} -espace vectoriel de dimension n .

Définition : Chemin, équivalence de chemins

Un chemin Γ est une application \mathcal{C}^1 d'un intervalle non vide $[a; b]$ vers E .

On dit que deux chemins Γ et $\tilde{\Gamma}$ définis respectivement de $[a; b]$ et $[\tilde{a}; \tilde{b}]$ vers \mathbb{R}^n sont équivalents lorsqu'il existe un \mathcal{C}^1 -difféomorphisme $\phi : [a; b] \rightarrow [\tilde{a}; \tilde{b}]$ tel que

$$\Gamma \circ \phi = \tilde{\Gamma}$$

Démonstration. La relation d'équivalence est assez facilement prouvée : la réflexivité est évidente, la transitivité et la symétrie ne sont pas très compliquées à prouver. □

Théorème : Intégrale d'une fonction sur un chemin

Soit $f : E \rightarrow E$ une fonction. Soit Γ un chemin défini sur un intervalle $[a; b]$.

Dans le cas où la fonction $f \circ \Gamma$ est Riemann-intégrable, on posera :

$$\int_{\Gamma} f = \int_a^b \Gamma'(t) f \circ \Gamma(t) dt$$

De plus, si Γ et $\tilde{\Gamma}$ sont équivalents, alors $f \circ \tilde{\Gamma}$ est également Riemann-intégrable et on vérifie :

$$\int_{\Gamma} f = \int_{\tilde{\Gamma}} f$$

Démonstration. On sait que le produit de deux fonctions Riemann-intégrable l'est aussi, ce qui nous prouve l'existence de $\int_{\Gamma} f$.

Reste à prouver l'égalité concernant les deux chemins équivalents. On va pour cela utiliser la caractérisation à l'aide subdivisions pointées.

On suppose donc que $\tilde{\Gamma} : [\tilde{a}; \tilde{b}] \xrightarrow{\phi} [a; b] \xrightarrow{\Gamma} \mathbb{C}$ où ϕ est un \mathcal{C}^1 -difféomorphisme.

On considère alors une subdivision pointée $(\sigma; t)$ de $[\tilde{a}; \tilde{b}]$ et la somme de Riemann associée :

$$R_{\sigma, t}(f) = \sum_{0 \leq i \leq n-1} \tilde{\Gamma}'(\tilde{t}_i) \times f \circ \tilde{\Gamma}(\tilde{t}_i) \times (\tilde{a}_{i+1} - \tilde{a}_i)$$

On pose, pour la suite $a_i = \phi(\tilde{a}_i)$ et $t_i = \phi(\tilde{t}_i)$. En utilisant définition de $\tilde{\Gamma}$, ainsi que la dérivation composée on obtient :

$$R_{\sigma, t}(f) = \sum_{0 \leq i \leq n-1} \phi'(\tilde{t}_i) \times \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (\tilde{a}_{i+1} - \tilde{a}_i)$$

Mais, d'après l'égalité des accroissements finis, on sait que, pour tout i , il existe $\tilde{c}_i \in]\tilde{a}_i; \tilde{a}_{i+1}[$ tel que $(\phi(\tilde{a}_{i+1}) - \phi(\tilde{a}_i)) = \phi'(\tilde{c}_i) \times (\tilde{a}_{i+1} - \tilde{a}_i)$. Or ϕ' ne s'annule pas. On en déduit :

$$R_{\sigma, t}(f) = \sum_{0 \leq i \leq n-1} \frac{\phi'(\tilde{t}_i)}{\phi'(\tilde{c}_i)} \times \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (a_{i+1} - a_i)$$

On va maintenant utiliser deux arguments : d'une part l'uniforme continuité de ϕ' sur $[\tilde{a}; \tilde{b}]$ et d'autre part, le fait que $|\phi'|$ possède un minimum.

On pose ainsi $m = \min_{[\tilde{a}; \tilde{b}]} |\phi'| > 0$ et M un majorant de $\Gamma' \times f \circ \Gamma$ sur $[a; b]$.

Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que si la subdivision pointée $(\sigma; t)$ est de pas inférieur à η , on a, pour tout i ,

$$|\phi'(\tilde{t}_i) - \phi'(\tilde{c}_i)| \leq \frac{m\varepsilon}{2 \times M \times (b - a)}$$

Ce qui donne, par construction :

$$\left| \frac{\phi'(\tilde{t}_i)}{\phi'(\tilde{c}_i)} - 1 \right| \leq \frac{\varepsilon}{2 \times M \times (b - a)}$$

En raison de la définition de $\int_{\Gamma} f$, il existe un nombre $\eta' > 0$ éventuellement plus petit que η tel que, si le pas de la subdivision pointée $(\sigma; t)$ est plus petit que η' alors le pas de l'image de cette subdivision par ϕ sera suffisamment petit pour que l'on ait :

$$\left\| \sum_{0 \leq i \leq n-1} \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (a_{i+1} - a_i) - \int_{\Gamma} f \right\| \leq \frac{\varepsilon}{2}$$

En effet, il suffit de constater que le pas de la subdivision image est contrôlé par le maximum de ϕ' en raison de l'inégalité des accroissements finis.

On obtient dans ce cas :

$$\begin{aligned} & \left\| \sum_{0 \leq i \leq n-1} \tilde{\Gamma}'(\tilde{t}_i) \times f \circ \tilde{\Gamma}(\tilde{t}_i) \times (\tilde{a}_{i+1} - \tilde{a}_i) - \int_{\Gamma} f \right\| \\ & \leq \left\| \sum_{0 \leq i \leq n-1} \tilde{\Gamma}'(\tilde{t}_i) \times f \circ \tilde{\Gamma}(\tilde{t}_i) \times (\tilde{a}_{i+1} - \tilde{a}_i) - \sum_{0 \leq i \leq n-1} \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (a_{i+1} - a_i) \right\| \\ & \quad + \left\| \sum_{0 \leq i \leq n-1} \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (a_{i+1} - a_i) - \int_{\Gamma} f \right\| \\ & \leq \left\| \sum_{0 \leq i \leq n-1} \tilde{\Gamma}'(\tilde{t}_i) \times f \circ \tilde{\Gamma}(\tilde{t}_i) \times (\tilde{a}_{i+1} - \tilde{a}_i) - \sum_{0 \leq i \leq n-1} \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (a_{i+1} - a_i) \right\| + \frac{\varepsilon}{2} \end{aligned}$$

Or, par construction, en utilisant l'inégalité triangulaire :

$$\begin{aligned} & \left\| \sum_{0 \leq i \leq n-1} \tilde{\Gamma}'(\tilde{t}_i) \times f \circ \tilde{\Gamma}(\tilde{t}_i) \times (\tilde{a}_{i+1} - \tilde{a}_i) - \sum_{0 \leq i \leq n-1} \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (a_{i+1} - a_i) \right\| = \\ & \left\| \sum_{0 \leq i \leq n-1} \left(\frac{\phi'(\tilde{t}_i)}{\phi'(\tilde{c}_i)} - 1 \right) \times \Gamma'(t_i) \times f \circ \Gamma(t_i) \times (a_{i+1} - a_i) \right\| \\ & \leq \frac{\varepsilon}{2 \times M \times (b - a)} \times M \times \sum_{0 \leq i \leq n-1} |a_{i+1} - a_i| = \frac{\varepsilon}{2} \end{aligned}$$

Finalement, on a bien

$$\left\| \sum_{0 \leq i \leq n-1} \tilde{\Gamma}'(\tilde{t}_i) \times f \circ \tilde{\Gamma}(\tilde{t}_i) \times (\tilde{a}_{i+1} - \tilde{a}_i) - \int_{\Gamma} f \right\| \leq \varepsilon$$

□

Corollaire : Changement de variable

Soit une fonction f définie sur un intervalle $[a; b]$ et à valeurs dans E un espace vectoriel de dimension finie.

On suppose que f est Riemann-intégrable.

Soit ϕ un \mathcal{C}^1 -difféomorphisme de $[a; b]$ dans $[\tilde{a}; \tilde{b}]$.

Alors $\frac{1}{\phi'} \times f \circ \phi^{-1}$ est Riemann intégrable sur $[\tilde{a}; \tilde{b}]$ et

$$\int_{\tilde{a}}^{\tilde{b}} \frac{1}{\phi'} \times f \circ \phi^{-1} = \int_a^b f$$

Chapitre 5

Méthodes de quadratures

5.1 Autour de la formule d'Euler Mac Laurin

5.1.1 Première version

Dans ce paragraphe nous allons établir et donner des premières applications d'une célèbre formule.

Proposition : Formule de Euler-Mac Laurin, polynômes et nombres de Bernoulli

Soit f une fonction de classe \mathcal{C}^n , avec $n \geq 2$, définie sur $[0; 1]$.

Alors :

$$\int_0^1 f = \frac{1}{2} (f(0) + f(1)) + \sum_{k=1}^{n-1} \frac{(-1)^k b_{k+1}}{(k+1)!} (f^{(k)}(1) - f^{(k)}(0)) + \frac{(-1)^n}{n!} \int_0^1 B_n f^{(n)} \quad (5.1)$$

avec (B_n) les polynômes de Bernoulli définis par :

- $B_0 = 1$;
- $\forall n, B'_{n+1} = (n+1)B_n$ et $\int_0^1 B_{n+1} = 0$.

Démonstration. On le montre par récurrence.

Notons que $B_1(X) = X - \frac{1}{2}$ est une primitive de 1 et que $\frac{1}{2}B_2$ est une primitive de B_1 . Supposons que f est de classe \mathcal{C}^2 . On réalise une double intégration par parties :

$$\begin{aligned} \int_0^1 f &= [B_1 f]_0^1 - \int_0^1 B_1 f' = \frac{1}{2} (f(0) + f(1)) - \frac{1}{2} [B_2 f']_0^1 + \frac{1}{2} \int_0^1 B_2 f'' \\ &= \frac{1}{2} (f(0) + f(1)) - \frac{b_2}{2} (f'(1) - f'(0)) + \frac{1}{2} \int_0^1 B_2 f'' \text{ car } B_2(1) = B_2(0) = b_2 \end{aligned}$$

La formule est donc initialisée.

Si on suppose maintenant que f est de classe \mathcal{C}^{n+1} avec $n \geq 2$, on peut réaliser une intégration par parties du reste intégral.

$$\begin{aligned} \frac{(-1)^n}{n!} \int_0^1 B_n f^{(n)} &= \frac{(-1)^n}{(n+1)!} [B_{n+1} f^{(n)}]_0^1 + \frac{(-1)^{n+1}}{(n+1)!} \int_0^1 B_{n+1} f^{(n+1)} \\ &= \frac{(-1)^n b_{n+1}}{(n+1)!} (f^{(n)}(1) - f^{(n)}(0)) + \frac{(-1)^{n+1}}{(n+1)!} \int_0^1 B_{n+1} f^{(n+1)} \end{aligned}$$

En regroupant avec l'hypothèse de récurrence, on obtient bien :

$$\begin{aligned} \int_0^1 f &= \frac{1}{2} (f(0) + f(1)) + \sum_{k=1}^{n-1} \frac{(-1)^k b_{k+1}}{(k+1)!} (f^{(k)}(1) - f^{(k)}(0)) + \frac{(-1)^n}{n!} \int_0^1 B_n f^{(n)} \\ &= \frac{1}{2} (f(0) + f(1)) + \sum_{k=1}^{n-1} \frac{(-1)^k b_{k+1}}{(k+1)!} (f^{(k)}(1) - f^{(k)}(0)) + \frac{(-1)^n b_{n+1}}{(n+1)!} (f^{(n)}(1) - f^{(n)}(0)) + \frac{(-1)^{n+1}}{(n+1)!} \int_0^1 B_{n+1} f^{(n+1)} \\ &= \frac{1}{2} (f(0) + f(1)) + \sum_{k=1}^n \frac{(-1)^k b_{k+1}}{(k+1)!} (f^{(k)}(1) - f^{(k)}(0)) + \frac{(-1)^{n+1}}{(n+1)!} \int_0^1 B_{n+1} f^{(n+1)} \end{aligned}$$

□

5.1.2 Étude des polynômes de Bernoulli

Définition : Polynômes et nombres de Bernoulli

On définit une suite de polynômes (B_n) par $B_0 = 1$ et, pour tout $n \geq 1$:

- $(B_n)' = nB_{n-1}$;
- $\int_0^1 B_n = 0$;

Enfin, on pose pour tout n , $b_n = B_n(0)$.

(Remarque 0) Un petit calcul donne :

n	$B_n(X)$
0	1
1	$X - \frac{1}{2}$
2	$X^2 - X + \frac{1}{6}$
3	$X^3 - \frac{3}{2}X^2 + \frac{1}{2}X$
4	$X^4 - 2X^3 + X^2 - \frac{1}{30}$

(Remarque 1) On montre facilement que, pour tout n , B_n est unitaire, de degré n et à coefficients rationnels.

Proposition : Propriétés des polynômes de Bernoulli

Pour tout $n \geq 2$, $B_n(1) = B_n(0)$.

Pour tout n , les coefficients des polynômes de Bernoulli s'expriment en fonction des B_n :

$$B_n(X) = \sum_{k=0}^n \binom{n}{k} b_k X^{n-k}$$

On a aussi, pour tout $n \geq 2$, une formule qui permet d'obtenir les (b_n) par récurrence :

$$b_n = \sum_{k=0}^n \binom{n}{k} b_k$$

On a également :

$$B_n(1-X) = (-1)^n B_n(X)$$

Et ainsi, pour tout entier impair $n \geq 3$, on vérifie $b_n = 0$.

Pour finir, on prouve que, pour tout n , $\frac{1}{n+1} (B_{n+1}(X+1) - B_{n+1}(X)) = X^n$, ce qui entraîne pour tout $(n; p) \in (\mathbb{N}^*)^2$:

$$\sum_{k=1}^n k^p = \frac{1}{p+1} (B_{p+1}(n+1) - B_{p+1}(0))$$

Démonstration. Pour tout $n \geq 2$, $B_n(1) - B_n(0) = \int_0^1 B_n' = n \int_0^1 B_{n-1} = 0$ donc $B_n(1) = B_n(0)$.

La seconde formule se montre par récurrence. L'initialisation est vraie. Supposons-la vraie au rang n . Notons qu'au rang $n+1$, le coefficient constant vaut $B_{n+1}(0) = b_{n+1}$ et tous les autres coefficients s'obtiennent par

intégration à partir de la relation $B'_{n+1} = (n+1)B_n$. Ainsi :

$$\begin{aligned} B_{n+1}(X) &= (n+1) \sum_{k=0}^n \frac{1}{n+1-k} b_k \binom{n}{k} X^{n+1-k} + b_{n+1} \\ &= \sum_{k=0}^n b_k \binom{n+1}{k} X^{n+1-k} + b_{n+1} = \sum_{k=0}^{n+1} b_k \binom{n+1}{k} X^{n+1-k} \text{ car } \frac{n+1}{n+1-k} \binom{n}{k} = \binom{n+1}{k} \end{aligned}$$

La troisième formule s'obtient en écrivant, pour tout $n \geq 2$, $b_n = B_n(0) = B_n(1)$ et en exploitant la seconde.

La troisième formule se montre par récurrence. Elle est vraie aux rangs 0 et 1.

Supposons-la vraie jusqu'à un certain rang n .

Au rang suivant, on sait que $(B_{n+1}(1-X))' = -B'_{n+1}(1-X) = -(n+1)B_n(1-X) = (n+1)(-1)^{n+1}B_n(X)$. Et d'autre part $((-1)^{n+1}B_{n+1}(X))' = (n+1)(-1)^{n+1}B'_n(X)$. On en déduit que $B_{n+1}(1-X)$ et $(-1)^{n+1}B_{n+1}(X)$

sont égaux à une constante près. Or, on a aussi $\int_0^1 B_{n+1}(1-t) dt = \int_0^1 (-1)^{n+1} B_{n+1}(t) dt = 0$.

Ce qui permet de conclure : ces deux polynômes sont égaux.

La dernière formule est vraie au rang 0. En effet, $\frac{1}{1}(B_1(X+1) - B_1(X)) = 1 = X^0$.

Supposons-la vraie jusqu'à un certain rang n . Au rang $n+1$, on vérifie que $\frac{1}{n+2}(B_{n+2}(X+1) - B_{n+2}(X))' = \frac{n+2}{n+2}(B_{n+1}(X+1) - B_{n+1}(X)) = (n+1)X^n$, par hypothèse de récurrence. Et on a aussi $(X^{n+1})' = (n+1)X^n$. On en déduit que les polynômes $\frac{1}{n+2}(B_{n+2}(X+1) - B_{n+2}(X))$ et X^{n+1} sont égaux à une constante près. Et pour $X=0$, on a $B_{n+2}(1) - B_{n+2}(0) = 0$ et $0^{n+1} = 0$ donc ces deux polynômes sont en fait égaux.

La toute dernière formule s'obtient grâce à une somme télescopique.

$$\begin{aligned} \sum_{k=1}^n k^p &= \sum_{k=0}^n k^p = \frac{1}{p+1} \sum_{k=0}^n (B_{p+1}(k+1) - B_{p+1}(k)) \\ &= \frac{1}{p+1} (B_{p+1}(n+1) - B_{p+1}(0)) \end{aligned}$$

□

Proposition : Coefficients de Fourier des polynômes de Bernoulli

Pour tout $(k; n) \in \mathbb{Z} \times \mathbb{N}^*$, on note $c_k^{(n)} = \sum_0^1 B_n(x) e^{-2ik\pi x} dx$.

Alors :

- pour tout $n \geq 1$, $c_0^{(n)} = 0$;
- pour tout $k \in \mathbb{Z}^*$, $c_k^{(n+1)} = \frac{n+1}{2ik\pi} c_k^{(n)}$;
- les coefficients de Fourier de B_1 notés $(c_k^{(1)})_{k \in \mathbb{N}}$ valent :

$$c_k^{(1)} = \frac{-1}{2ik\pi} \text{ pour } k \in \mathbb{Z}^*$$

À partir des deux dernières relations, on obtient :

- pour tout $p \in \mathbb{N}^*$ et pour tout $k \in \mathbb{Z}^*$, $c_k^{(2p)} = \frac{(-1)^{p+1} (2p)!}{(2\pi k)^{2p}}$, ce qui donne, pour tout x :

$$\tilde{B}_{2p}(x) = 2(-1)^{p+1} (2p)! \sum_{k \in \mathbb{N}^*} \frac{1}{(2\pi k)^{2p}} \cos(2\pi kx)$$

- pour tout $p \in \mathbb{N}^*$ et pour tout $k \in \mathbb{Z}^*$, $c_k^{(2p+1)} = \frac{(-1)^{p+1} (2p+1)!}{i(2\pi k)^{2p+1}}$, ce qui donne, pour tout x :

$$\tilde{B}_{2p+1}(x) = 2(-1)^{p+1} (2p+1)! \sum_{k \in \mathbb{N}^*} \frac{1}{(2\pi k)^{2p+1}} \sin(2\pi kx)$$

On « périodise » les polynômes de Bernoulli, en posant $\tilde{B}_n : x \mapsto B_n(x - [x])$, la version périodique des Bernoulli qui est de classe C^1 à partir de $n=2$ et de classe C^1 par morceaux pour $n=1$.

Démonstration. La première relation provient de $c_0^{(n)} = \int_0^1 B_n = 0$

La seconde relation se déduit de l'égalité valable pour tout n , $B'_{n+1} = (n+1)B_n$.

Reste maintenant à calculer les coefficients de Fourier de B_1 . On réalise un changement de variable $t = x - \frac{1}{2}$ puis une intégration par parties :

$$\begin{aligned} c_k^{(1)} &= \int_0^1 \left(x - \frac{1}{2}\right) e^{-2ik\pi x} dx = \int_{-1/2}^{1/2} t e^{-2ik\pi(t+1/2)} dt \\ &= e^{-ik\pi} \int_{-1/2}^{1/2} t e^{-2ik\pi t} dt \\ &= (-1)^k \frac{(-1)}{2ik\pi} \left([t e^{-2ik\pi t}]_{-1/2}^{1/2} - \int_{-1/2}^{1/2} e^{-2ik\pi t} dt \right) \\ &= (-1)^k \frac{(-1)}{2ik\pi} \left(\frac{1}{2} (e^{ik\pi} + e^{-ik\pi}) - 0 \right) \\ &= (-1)^k \frac{(-1)}{2ik\pi} (-1)^k = \frac{-1}{2ik\pi} \end{aligned}$$

Montrons par récurrence sur n que, pour tout $k \in \mathbb{Z}^*$, $c_k^{(n)} = \frac{-n!}{(2i\pi k)^n}$. Cette relation est vraie pour $n = 1$. Et au rang suivant :

$$c_k^{(n+1)} = \left(\frac{n+1}{2ik\pi} \right) c_k^{(n)} = \left(\frac{n+1}{2ik\pi} \right) \frac{-n!}{(2i\pi k)^n} = \frac{-(n+1)!}{(2i\pi k)^{n+1}}$$

Dans le cas où $n = 2p$ avec $p \geq 1$, cela donne $c_k^{(2p)} = \frac{-(2p)!}{(2i\pi k)^{2p}} = \frac{(-1)^{p+1} (2p)!}{(2\pi k)^{2p}}$. Et dans le cas où $n = 2p+1$, on obtient $c_k^{(2p+1)} = \frac{(-1)^{p+1} (2p+1)!}{i(2\pi k)^{2p+1}}$.

Enfin les toutes dernières formules s'obtiennent en remarquant que, pour tout $k \in \mathbb{N}^*$, et pour tout x réel :

$$c_k^{(n)} e^{2i\pi kx} + c_{-k}^{(n)} e^{-2i\pi kx} = 2\Re \left(c_k^{(n)} e^{2i\pi kx} \right)$$

□

(Remarque 2) Au moyen des sommes de Fourier, on retrouve $b_{2p+1} = 0$ et on obtient aussi, dans le cas pair, une expression des coefficients de Bernoulli au moyen de la fonction zêta de Riemann.

Corollaire : Expression des coefficients de Bernoulli

Pour tout p non nul :

$$b_{2p} = \frac{2(-1)^{p+1} (2p)!}{(2\pi)^{2p}} \zeta(2p) \quad (5.2)$$

On rappelle que, pour tout $s > 1$, $\zeta(s) = \sum_{k \in \mathbb{N}^*} \frac{1}{k^s}$.

En considérant les propriétés de la fonction ζ , on obtient :

$$b_{2p} \underset{p \rightarrow +\infty}{\sim} \frac{2(-1)^{p+1} (2p)!}{(2\pi)^{2p}} \quad (5.3)$$

Enfin, on dispose de la majoration pour le polynôme périodicisé :

$$|\tilde{B}_{2p}| \leq b_{2p} \quad (5.4)$$

5.1.3 Version améliorée de la formule d'Euler-Mac-Laurin



Dans toute la suite, les notations (B_n) désigneront les fonctions polynômiales de Bernoulli périodiciées $x \mapsto B_n(x - \lfloor x \rfloor)$.

Fort de l'étude des polynômes de Bernoulli, nous pouvons réécrire une version améliorée de la formule d'Euler-Mac-Laurin que nous allons exploiter pour deux applications.

Théorème : Formule d'Euler-Mac-Maurin, version définitive

Soient $a < b$ deux entiers et f une fonction de classe \mathcal{C}^{2n} sur $[a; b]$. Alors :

$$\sum_{k=a}^b f(k) = \frac{1}{2} (f(a) + f(b)) + \int_a^b f + \sum_{j=1}^n \frac{b_{2j}}{(2j)!} \left(f^{(2j-1)}(b) - f^{(2j-1)}(a) \right) - \frac{1}{(2n)!} \int_a^b B_{2n} f^{(2n)} \quad (5.5)$$

Dans le cas où f est de classe \mathcal{C}^{2n+1} , on améliore l'ordre du reste :

$$\sum_{k=a}^b f(k) = \frac{1}{2} (f(a) + f(b)) + \int_a^b f + \sum_{j=1}^n \frac{b_{2j}}{(2j)!} \left(f^{(2j-1)}(b) - f^{(2j-1)}(a) \right) + \frac{1}{(2n+1)!} \int_a^b B_{2n+1} f^{(2n+1)} \quad (5.6)$$

Démonstration. On reprend la formule 5.1 d'origine en tirant parti de la nullité des $(b_{2i+1})_{i \geq 1}$. Pour tout $k \in \llbracket a; b-1 \rrbracket$:

$$\int_k^{k+1} f = \frac{1}{2} (f(k) + f(k+1)) - \sum_{j=1}^n \frac{b_{2j}}{(2j)!} \left(f^{(2j-1)}(k+1) - f^{(2j-1)}(k) \right) + \frac{1}{(2n)!} \int_k^{k+1} B_{2n} f^{(2n)}$$

En sommant pour tous les k , on obtient :

$$\int_a^b f = \sum_{k=a}^b f(k) - \frac{1}{2} (f(a) + f(b)) - \sum_{j=1}^n \frac{b_{2j}}{(2j)!} \left(f^{(2j-1)}(b) - f^{(2j-1)}(a) \right) + \frac{1}{(2n)!} \int_a^b B_{2n} f^{(2n)}$$

En réarrangeant, on obtient l'égalité 5.5.

Et toujours parce que $b_{2n+1} = 0$, on déduit facilement la formule améliorée. □

5.1.4 Approximation asymptotique de sommes

5.1.4.1 Résultat général

Théorème : Approximation asymptotique de sommes

On suppose que :

- $a < b$ sont deux entiers f une fonction de classe \mathcal{C}^∞ sur $[a; +\infty[$;
- il existe $n_0 \in \mathbb{N}$ et $\alpha \in \mathbb{R}$ tel que pour tout $k \geq n_0$, $f^{(k)}$ ne change pas de signe sur $[\alpha; +\infty[$ et $\lim_{+\infty} f^{(k)} = 0$.

On pose enfin $T_b(f) = \sum_{k=a}^b f(k)$.

Alors, pour tout entier $b \geq \alpha$ et pour tout $n > \frac{n_0}{2}$, on a :

$$T_b(f) = \int_a^b f + \frac{1}{2} f(b) + \sum_{j=1}^{n-1} \frac{b_{2j}}{(2j)!} f^{(2j-1)}(b) + C_n + R_{n,b} \quad (5.7)$$

Avec :

$$C_n = \frac{1}{2} f(a) - \sum_{j=1}^n \frac{b_{2j}}{(2j)!} f^{(2j-1)}(a) - \frac{1}{(2n)!} \int_a^{+\infty} B_{2n} f^{(2n)} \\ R_{n,b} = \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b) + \frac{1}{(2n)!} \int_b^{+\infty} B_{2n} f^{(2n)}$$

De plus, on vérifie que :

- C_n ne dépend que de a ;
- il existe $\theta \in [0; 1]$ tel que $R_{n,b} = \theta \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b)$.

Démonstration. Commençons par prouver la convergence des intégrales. D'après l'inégalité 5.4, il est clair que, pour tout $N \geq b$, $\int_b^N |B_{2n} f^{(2n)}| \leq |b_{2n}| \int_b^N |f^{(2n)}| = \left| b_{2n} \int_b^N f^{(2n)} \right|$ car $f^{(2n)}$ ne change pas de signe. On obtient donc :

$$\int_b^N |B_{2n} f^{(2n)}| \leq |b_{2n}| |f^{(2n-1)}(N) - f^{(2n-1)}(b)|$$

Et par passage à la limite, cela donne $\int_b^N |B_{2n} f^{(2n)}| \leq |b_{2n} f^{(2n-1)}(b)| < +\infty$.

Nous allons maintenant prouver que C_n est constant. Pour ce faire, on écrit la formule au rang $n+1$. On obtient :

$$T_b(f) = \int_a^b f + \frac{1}{2} f(b) + \sum_{j=1}^{n-1} \frac{b_{2j}}{(2j)!} f^{(2j-1)}(b) + \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b) + C_{n+1} + R_{n+1,b}$$

Et ainsi, on extrait $\frac{b_{2n}}{(2n)!} f^{(2n-1)}(b) + C_{n+1} + R_{n+1,b} = C_n + R_{n,b}$, soit $C_{n+1} - C_n = R_{n,b} - R_{n+1,b} - \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b)$. L'inégalité sur l'intégrale établie plus haut ainsi que les hypothèses sur les dérivées successives de f permet de prouver que $\lim_{b \rightarrow +\infty} R_{n,b} - R_{n+1,b} - \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b) = 0$ ce qui montre que $C_{n+1} = C_n$.

Pour finir, on va montrer que $R_{n,b} = \theta \theta \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b)$ avec θ un nombre de $[0; 1]$. Sachant que f^{2n} ne change pas de signe et que $\int_b^{+\infty} f^{(2n)} = -f^{(2n-1)}(b)$, on peut invoquer l'égalité de la moyenne généralisée pour écrire :

$$R_{n,b} = \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b) + \frac{1}{(2n)!} \int_b^{+\infty} B_{2n} f^{(2n)} = \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b) \left(1 - \frac{\overline{B_{2n}}}{b_{2n}} \right)$$

Le nombre $\overline{B_{2n}}$ désigne la valeur moyenne de B_{2n} sur $[b; +\infty[$, pondéré par $f^{(2n)}$. En particulier, à partir de la majoration 5.4, on sait que $\theta = 1 - \frac{\overline{B_{2n}}}{b_{2n}} \in [0; 2]$.

Pour affiner le contrôle de θ , procédons à une analyse de signe en reprenant la relation entre $R_{n,b}$ et $R_{n+1,b}$ établie plus haut.

On a montré que $R_{n,b} = R_{n+1,b} + \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b)$. On vient de prouver que $R_{n,b}$ est de même signe que $\frac{b_{2n}}{(2n)!} f^{(2n-1)}(b)$, ce qui entraîne que $R_{n+1,b}$ est du signe de $\frac{b_{2n+2}}{(2n+2)!} f^{(2n+1)}(b)$. Mais on sait que $f^{(2n+1)}(b)$ est du même signe que $f^{(2n-1)}(b)$ et que b_{2n+2} est du signe opposé de b_{2n} en raison de la formule 5.2, sur les coefficients de Bernoulli.

Finalement, $R_{n+1,b}$ est du signe opposé à $\frac{b_{2n}}{(2n)!} f^{(2n-1)}(b)$, ce qui prouve que $\theta \in [0; 1]$ et achève cette démonstration. \square

5.1.4.2 Formule de Stirling

Nous allons appliquer la formule qui précède à la fonction $f = \ln$ avec $a = 1$ et $b > 1$ quelconque.

Soit ainsi $n \in \mathbb{N}^*$, la formule donne

$$T_b(f) = \ln(b!) = \frac{1}{2} \ln(b) + \int_1^b \ln(t) dt + \sum_{j=1}^{n-1} \frac{b_{2j}}{(2j)!} f^{(2j-1)}(b) + C + R_{n,b}$$

On vérifie assez facilement que, pour tout entier $k \geq 1$, $f^{(k)}(b) = \frac{(-1)^{k-1}(k-1)!}{b^k}$ ce qui simplifie le terme général de la somme $\frac{b_{2j}}{(2j)!} f^{(2j-1)}(b) = \frac{(-1)^{2j-2} (2j-2)! b_{2j}}{(2j)! b^{2j-1}} = \frac{b_{2j}}{2j(2j-1)b^{2j-1}}$.

En outre, on a $R_{n,b} = \theta \frac{b_{2n}}{(2n)!} f^{(2n-1)}(b) = \theta \frac{b_{2n}}{2n(2n-1)b^{2n-1}}$.

Le calcul de l'intégrale donne :

$$\int_1^b \ln(t) dt = b \ln(b) - b = \ln(b^b) - \ln(e^b) = \ln\left(\left(\frac{b}{e}\right)^b\right)$$

On obtient ainsi, par passage à l'exponentielle, la formule de Stirling :

$$b! = e^C \sqrt{b} \left(\frac{b}{e}\right)^b \exp\left(\sum_{j=1}^{n-1} \frac{b_{2j}}{2j(2j-1)b^{2j-1}} + \theta \frac{b_{2n}}{2n(2n-1)b^{2n-1}}\right) \quad (5.8)$$

Pour $n = 4$, cela donne :

$$b! = e^C \sqrt{b} \left(\frac{b}{e}\right)^b \exp\left(\frac{1}{12b} - \frac{1}{360b^3} + \frac{1}{1260b^5} - \theta \frac{1}{1680b^7}\right)$$

5.1.4.3 Évaluation de la constante de Stirling : intégrales de Wallis

Définition : Intégrales de Wallis

Pour tout n , on pose $I_n = \int_0^{\pi/2} \sin^n(t) dt$.

On définit ainsi la suite des intégrales de Wallis

(Remarque 3) Il est clair que $\lim_{n \rightarrow \infty} I_n = 0$.

(Remarque 4) On calcule $I_0 = \frac{\pi}{2}$, $I_1 = 1$, $I_2 = \int_0^{\pi/2} \sin^2(t) dt = \int_0^{\pi/2} \frac{1 - \cos(2t)}{2} dt = \frac{\pi}{4}$.

Proposition : Relation de récurrence, formule explicite

Pour tout n , $(n+2)I_{n+2} = (n+1)I_n$. On en déduit :

$$I_{2n} = \frac{\prod_{1 \leq k \leq n} (2k-1)}{\prod_{1 \leq k \leq n} 2k} I_0 \qquad I_{2n+1} = \frac{\prod_{1 \leq k \leq n} 2k}{\prod_{1 \leq k \leq n} (2k+1)} I_1$$

Démonstration. On réalise une intégration par parties

$$\begin{aligned} I_{n+2} &= \int_0^{\pi/2} \sin^{n+2}(t) dt \\ &= \left[-\cos(t) \sin^{n+1}(t) \right]_0^{\pi/2} + (n+1) \int_0^{\pi/2} \cos^2(t) \sin^n(t) dt \\ &= (n+1)I_n - (n+1)I_{n+2} \end{aligned}$$

Cela donne le résultat attendu $(n+2)I_{n+2} = (n+1)I_n$.

Le reste se montre facilement par récurrence à partir de cette formule. □

Proposition : Autres propriété des intégrales de Wallis

On a, pour tout p , $I_{2p}I_{2p+1} = \frac{\pi}{2(2p+1)}$.

On vérifie également les comportements asymptotiques suivants :

$$I_{2p-1} \sim I_{2p} \sim I_{2p+1} \qquad (I_{2p})^2 \sim \frac{\pi}{4p}$$

Démonstration. On calcule $I_{2p}I_{2p+1} = \frac{\prod_{1 \leq k \leq n} (2k-1)}{\prod_{1 \leq k \leq n} 2k} \times \frac{\prod_{1 \leq k \leq n} 2k}{\prod_{1 \leq k \leq n} (2k+1)} I_0 I_1 = \frac{(2p)!}{(2p+1)!} I_0 I_1 = \frac{\pi}{2(2p+1)}$.

D'après la relation de récurrence des intégrales de Wallis et leurs sens de variation, on a $\frac{n+1}{n+2} = \frac{I_{n+2}}{I_n} \leq \frac{I_{n+1}}{I_n} \leq 1$. On en déduit que : $I_{n+1} \sim I_n$.

En combinant les deux résultats précédents, on obtient :

$$I_{2p}I_{2p+1} = \frac{\pi}{2(2p+1)} \sim \frac{\pi}{4p} \quad \text{et} \quad I_{2p}I_{2p+1} \sim (I_{2p})^2$$

Ce qui prouve $(I_{2p})^2 \sim \frac{\pi}{4p}$ □

Proposition : Version définitive de la formule de Stirling

On a :

$$b! \sim \sqrt{2\pi b} \left(\frac{b}{e} \right)^b$$

Cette formule s'obtient à partir de la formule 5.8 et de ce que l'on vient de prouver sur les intégrales de Wallis.

Démonstration. Reprenons la formule 5.8, que l'on peut réécrire, sachant que le terme en exponentielle tend vers 1 lorsque b tend vers l'infini.

$$b! \sim K\sqrt{b}\left(\frac{b}{e}\right)^b \text{ en notant } K = e^C$$

$$\text{Maintenant nous pouvons travailler sur } I_{2p} = \frac{\prod_{1 \leq k \leq n} (2k-1)}{\prod_{1 \leq k \leq n} 2k} I_0 = \frac{(2k)!}{\left(\prod_{1 \leq k \leq n} 2k\right)^2} I_0 = \frac{(2k)!}{2^{2p}(p!)^2} I_0.$$

On a donc :

$$I_{2p} \sim \frac{K\sqrt{2p}\left(\frac{2p}{e}\right)^{2p}}{2^{2p}K^2p\left(\frac{p}{e}\right)^{2p}} \frac{\pi}{2} = \frac{1}{K} \sqrt{\frac{2}{p}} \frac{\pi}{2}$$

Cela donne $I_{2p}^2 \sim \frac{1}{K^2} \frac{\pi^2}{2p}$. Compte-tenu de l'équivalence $I_{2p}^2 \sim \frac{\pi}{4p}$, on en déduit :

$$K = \sqrt{2\pi}$$

□

5.2 Premiers résultats sur les quadratures

5.2.1 Résultats et notations préliminaires

Proposition : Égalité de la moyenne généralisée

Soit μ une mesure non triviale définie sur un intervalle $[a; b]$ et f une fonction continue sur ce même intervalle. Alors, il existe $\gamma \in [a; b]$ tel que :

$$f(\gamma) = \frac{1}{\mu([a; b])} \int_{[a; b]} f \, d\mu$$

On dit que $\frac{1}{\mu([a; b])} \int_{[a; b]} f \, d\mu$ est la moyenne de f sur $[a; b]$.

Démonstration. C'est une application du théorème des valeurs intermédiaires et de l'encadrement :

$$m\mu([a; b]) \leq \int_{[a; b]} f \, d\mu \leq M\mu([a; b])$$

Avec m et M les extrema de f sur $[a; b]$.

□

(Remarque 5) En particulier, si μ admet une densité notée g , la moyenne vaut :

$$\frac{1}{\int_a^b g(t) \, dt} \int_a^b f(t)g(t) \, dt$$

(Remarque 6) Ce résultat demeure vrai si $a > b$.

Proposition : Formule de Taylor avec reste intégrale, ou Taylor-Laplace

Soit f une fonction de classe $C^{n+1}(I)$ où I est un intervalle et n un entier naturel. Alors, pour tout $(a; x) \in I^2$:

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) \, dt$$

Dans la suite on notera $T_n : x \mapsto \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k$ et $R_n : x \mapsto \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) \, dt$. Ce sont les polynômes de Taylor et le reste associés à la formule de Taylor en a .

Démonstration. Facile à prouver par récurrence sur n et en réalisant une intégration par parties. \square

(Remarque 7) Sous cette hypothèse un peu plus contraignante, on peut ainsi retrouver les formules de Taylor-Young et de Taylor-Lagrange.

En notant K_{n+1} le maximum local de $|f^{(n+1)}|$, on montre $|R_n(x)| \leq \frac{K_{n+1}}{(n+1)!} |x-a|^{n+1}$ et donc $R_n(x) = o((x-a)^n)$ (Taylor-Young).

Mais, grâce à l'égalité de la moyenne généralisée, on a aussi :

$$\exists \gamma \in [a; x] / R_n(x) = \frac{f^{(n+1)}(\gamma)}{(n+1)!} (x-a)^{n+1} \quad (\text{Taylor-Lagrange})$$

Définition : Approximation d'intégrales par quadrature

Soit f une fonction définie sur un intervalle $[a; b]$ et μ une mesure définie sur les boréliens.

On choisit $n+1$ points $a_0 = a < a_1 < \dots < a_n = b$ avec $a_0 = a$ et $a_n = b$ et pour tout $i \in \llbracket 0; n-1 \rrbracket$, on choisit l_i points $(\xi_{ij}) \in [a_i; a_{i+1}]^{l_i}$ et l_i poids $(\omega_{ij}) \in (\mathbb{R}_*^+)^{l_i}$ tels que $\sum_{j=1}^{l_i} \omega_{ij} = 1$. Enfin, on pose :

$$\Phi(f) = \sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=1}^{l_i} \omega_{ij} f(\xi_{ij})$$

En notant $h = \max_{0 \leq i \leq n-1} a_{i+1} - a_i$, on a, sous certaines conditions :

$$\Phi(f) \xrightarrow{h \rightarrow 0} \int_{[a; b]} f d\mu$$

On dit que $\Phi(f)$ est une approximation de $\int_{[a; b]} f d\mu$ par quadrature.

(Remarque 8) Dans la plupart des cas, μ sera la mesure de Lebesgue ou une mesure à densité et la fonction considérée sera intégrable au sens de Riemann.

(Remarque 9) On obtient une bonne quadrature en choisissant avec soin les points et les poids.

Proposition : Cas où f est continue, pour la mesure de Lebesgue

Dans le cas où la fonction est continue et où μ est la mesure de Lebesgue, on a convergence pour h tendant vers 0.

Démonstration. Soit $\varepsilon > 0$. Par l'uniforme continuité de f il existe $\eta > 0$ tel que, pour tout $|x-y| < \eta$, $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$.

Considérons maintenant $h < \eta$. Pour tout $i \in \llbracket 0; n-1 \rrbracket$, il existe $\gamma_i \in [a_i; a_{i+1}]$ tel que $\sum_{j=1}^{l_i} \omega_{ij} f(\xi_{ij}) = f(\gamma_i)$.

Cela s'obtient par application du théorème de valeurs intermédiaires compte-tenu du fait que $\sum_{j=1}^{l_i} \omega_{ij} f(\xi_{ij})$ est une valeur moyenne de f , donc située entre son maximum et son minimum du sous-intervalle considéré. Mais alors, on obtient, par intégration :

$$\left| \int_{a_i}^{a_{i+1}} f - \sum_{j=1}^{l_i} \omega_{ij} f(\xi_{ij}) \right| < \frac{\varepsilon}{b-a} (a_{i+1} - a_i)$$

En sommant sur tous les i , cela donne :

$$\left| \int_a^b f - \Phi(f) \right| < \varepsilon$$

□

Définition : Ordre d'une méthode

On dit qu'une méthode d'approximation est d'ordre $n \in \mathbb{N}$ lorsque $\Phi(P) = \int_a^b P$ pour tout polynôme P de degré inférieur ou égal à n et $\Phi(Q) \neq \int_a^b Q$ pour un polynôme de degré $n+1$.

(Remarque 10) Toutes les méthodes sont d'ordre au moins 0.

(Remarque 11) On définira les méthodes (points et poids) en fonction de l'ordre souhaité.

5.2.2 Calcul de l'erreur

On conserve les notations établies juste avant.

Définition : Fonction erreur

On pose $E(f) = \int_a^b f - \Phi(f)$.

(Remarque 12) Φ et E sont des formes linéaires sur l'ensemble des fonctions intégrables sur $[a; b]$.

Proposition : Calcul de l'erreur, noyau de Peano

Soit f une fonction de classe \mathcal{C}^{n+1} sur $[a; b]$. On suppose que Φ est d'ordre n . Alors :

$$E(f) = \int_a^b \frac{f^{(n+1)}(t)}{n!} K(t) dt$$

Avec, pour tout t , $K(t) = E(x \mapsto ((x-t)^+)^n)$, fonction appelée noyau de Peano.

On utilise la notation, pour tout x et t réels, $(x-t)^+ = \begin{cases} x-t & \text{si } x-t \geq 0 \\ 0 & \text{sinon} \end{cases}$.

Démonstration. On exploite la formule de Taylor-Laplace en a et la linéarité de l'erreur :

$$\begin{aligned} E(f) &= E(T_n(f)) + E(R_n(f)) \\ &= E(R_n(f)) \text{ car la méthode est d'ordre } n \end{aligned}$$

Calculons maintenant plus spécifiquement l'erreur sur le reste.

$$E(R_n(f)) = \int_a^b \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt dx - \sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=1}^{l_i} \omega_{ij} \int_a^{\xi_{ij}} \frac{(\xi_{ij}-t)^n}{n!} f^{(n+1)}(t) dt$$

Avec la notation introduite plus haut, on peut exploiter le théorème de Fubini :

$$\begin{aligned} E(R_n(f)) &= \int_a^b \int_a^b \frac{((x-t)^+)^n}{n!} f^{(n+1)}(t) dt dx - \sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=1}^{l_i} \omega_{ij} \int_a^b \frac{((\xi_{ij}-t)^+)^n}{n!} f^{(n+1)}(t) dt \\ &= \int_a^b \int_a^b \frac{((x-t)^+)^n}{n!} f^{(n+1)}(t) dx dt - \int_a^b f^{(n+1)}(t) \sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=1}^{l_i} \omega_{ij} \frac{((\xi_{ij}-t)^+)^n}{n!} dt \\ &= \int_a^b \frac{f^{(n+1)}(t)}{n!} \left[\int_a^b ((x-t)^+)^n dx - \sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=1}^{l_i} \omega_{ij} ((\xi_{ij}-t)^+)^n \right] dt \end{aligned}$$

On reconnaît que le terme entre crochets représente l'erreur pour la fonction $x \mapsto ((x-t)^+)^n$ et on obtient la formule attendue. □

5.2.3 Polynômes orthogonaux

Définition : Polynômes orthogonaux

Soit μ une mesure finie sur l'ensemble des boréliens d'un intervalle $[a; b]$.

On suppose que, pour tout borélien B , si $\mu(B) = \mu([a; b])$ alors B est de cardinal infini.

Pour tous polynômes P et Q à coefficients réels, on pose :

$$\langle P|Q \rangle = \int_{[a; b]} PQ \, d\mu$$

On définit ainsi un produit scalaire sur les polynômes.

Démonstration. On montre assez facilement que cette forme est bilinéaire, définie et positive car $\int_{[a; b]} P^2 \, d\mu = 0$ lorsque P s'annule sur un ensemble de mesure $\mu([a; b])$, ce qui entraîne que P possède une infinité de racines donc est nul. \square

Proposition : Base orthogonale de polynômes

On peut fabriquer une famille $(T_n)_{n \in \mathbb{N}^*}$ de polynômes tels que :

- pour tout $n \in \mathbb{N}$, $\deg(T_n) = n$;
- pour tout $i \neq j$, $\langle T_i|T_j \rangle = 0$.

En particulier les $(T_k)_{0 \leq k \leq n}$ forment une base de $\mathbb{R}_n[X]$.

Démonstration. On peut construire une telle famille par le procédé d'orthogonalisation de Graham-Schmidt. On pose $T_0 = 1$ et, pour tout $n \in \mathbb{N}^*$:

$$T_n = X^n - \sum_{k=0}^{n-1} \frac{1}{\|T_k\|^2} \langle X^n|T_k \rangle T_k$$

\square

La propriété qui suit aura son importance pour la définition des méthodes de Gauss.

Proposition : Racines des polynômes orthogonaux

On reprend les mêmes hypothèses et notations. Alors, pour tout $n \in \mathbb{N}^*$, T_n possède n racines distinctes dans $[a; b]$.

Démonstration. Notons $(\alpha_i)_{1 \leq i \leq p}$ les racines distinctes de T_n situées dans $[a; b]$. Si on suppose $p < n$ alors on peut trouver un polynôme Q tel que :

- $\deg(Q) < n$;
- les racines de $T_n Q$ situées dans l'intervalle $[a; b]$ sont de multiplicités paires.

Il suffit pour cela de poser $Q = \prod_{i \in I} (x - \alpha_i)$ avec $I \subset \llbracket 1; p \rrbracket$ défini de telle sorte que, pour tout $i \in I$, la multiplicité de α_i dans le polynôme T_n est impaire.

En particulier, $Q \in \text{Vec}((T_k)_{0 \leq k \leq n-1})$ donc $\langle T_n|Q \rangle = \int_{[a; b]} T_n Q \, d\mu = 0$, ce qui est absurde puisque le polynôme $T_n Q$ ne change pas de signe sur $[a; b]$. \square

5.3 Description et propriétés des méthodes

Dans tout ce qui suit, on va s'intéresser uniquement à l'approximation de l'intégrale sur l'un des sous-intervalles $[a_i; a_{i+1}]$. On écrira donc plus simplement la formule de quadrature :

$$\Phi(f) = \sum_{i=0}^{n-1} \omega_i f(a_i) \text{ avec } a \leq a_0 < a_1 < \dots < a_{n-1} \leq b \text{ et } \forall i, \omega_i > 0$$

Et cette formule servira à approximer $\int_{[a; b]} f d\mu$.

Pour la mesure de Lebesgue, on peut se ramener à l'intervalle $[-1; 1]$ en considérant le changement de variable $t \mapsto u = \frac{2}{b-a} \left(t - \frac{a+b}{2} \right)$.

5.3.1 Newton-Cotes

En exploitant la remarque précédente, on se place sur $[-1; 1]$.

Définition : Méthode de Newton-Cotes

On choisit les (a_i) uniformément répartis sur $[-1; 1]$ avec $a_0 = -1$ et $a_{n-1} = 1$.

Dans ce cas, la méthode est d'ordre au moins $n-1$ si et seulement si les (ω_i) valent, pour tout i :

$$\omega_i = \int_{[a; b]} L_i$$

Avec les (L_i) qui sont les polynômes de Lagrange définis pour tout $i \leq n-1$ par :

$$L_i(X) = \frac{1}{\prod_{j \neq i} (a_i - a_j)} \prod_{j \neq i} (X - a_j)$$

Démonstration. Soit P un polynôme de degré inférieur ou égal à $n-1$. On a $P = \sum_{i \leq n-1} P(a_i) L_i$. Donc les (ω_i) ainsi définis donnent bien une méthode d'ordre au moins $n-1$. De plus, ils sont définis de manière unique. En effet, on a, pour tout i et de manière univoque, $\Phi(L_i) = \omega_i$. \square

(Remarque 13) Il est clair que, pour tout i , $a_{n-1-i} = -a_i$ et ceci entraîne que $\omega_i = \omega_{n-1-i}$ car les courbes de L_i et L_{n-1-i} sont symétriques par rapport à l'axe des ordonnées.

Proposition : Augmentation de l'ordre

Lorsque $n-1$ est pair, la méthode est d'ordre n .

Démonstration. Dans ce cas, le polynôme $P = X^n$ est impair donc $\Phi(P) = \int_{[-1; 1]} P = 0$ en raison de la remarque précédente sur la symétrie des poids ω_i .

Ainsi, par linéarité, la méthode est d'ordre n . \square

(Remarque 14) En pratique, on choisira un nombre n impair de points afin d'atteindre un ordre n .

5.3.2 Gauss

On reprend les hypothèses et notations du paragraphe sur les polynômes orthogonaux.

Proposition : Méthode de Gauss

Soit $n \in \mathbb{N}$. Il existe une unique manière de choisir $n + 1$ points $(a_i)_{0 \leq i \leq n}$ dans $[a; b]$ associés à des poids $(\omega_i)_{0 \leq i \leq n}$ telle que la quadrature suivante approxime $\int_{[a; b]} f d\mu$ avec un ordre $2n + 1$.

$$\Phi(f) = \sum_{i=0}^n \omega_i f(a_i)$$

En particulier :

- les (a_i) sont les racines du polynôme T_{n+1} orthogonal à $\mathbb{R}_n[X]$ et de degré $n + 1$;
- les poids valent $\omega_i = \int_{[a; b]} L_i d\mu$ avec L_i le polynôme de Lagrange associé au point a_i :

$$L_i(X) = \frac{1}{\prod_{j \neq i} (a_i - a_j)} \prod_{j \neq i} (X - a_j)$$

Démonstration. Supposons qu'une telle méthode existe. Posons $P(X) = \prod_i (X - a_i)$. On a $\deg(P) = n + 1$

donc pour tout Q de degré inférieur ou égal à n , on vérifie $\langle P|Q \rangle = \int_{[a; b]} PQ d\mu = \sum_{i=0}^n \omega_i P(a_i)Q(a_i) = 0$ car $\deg(PQ) \leq 2n + 1$.

Cela prouve que P est un polynôme orthogonal à l'espace $\mathbb{R}_n[X]$.

Réciproquement, on sait, d'après le paragraphe sur l'orthogonalité, qu'un tel polynôme P existe, est défini à un multiple scalaire non nul près, et possède $n + 1$ racines distinctes dans $[a; b]$. En notant (a_i) ces racines, il nous faut maintenant prouver que la quadrature fonctionne et calculer les poids (ω_i) . Soit un polynôme A de degré maximal $2n + 1$. Réalisons la division euclidienne de A par P . Il existe un polynôme D et un polynôme R tels que $A = PD + R$ avec $\deg(R) \leq n$. On a aussi $\deg(D) \leq n$ ce qui entraîne $\langle P|D \rangle = 0$ et par suite $\int_{[a; b]} A d\mu = \int_{[a; b]} R d\mu$. De plus, pour tout i , $A(a_i) = R(a_i)$ car $P(a_i) = 0$ et on a ainsi

$$R(X) = \sum_i R(a_i) L_i(X) = \sum_i A(a_i) L_i(X), \text{ ce qui entraîne } \int_{[a; b]} A d\mu = \int_{[a; b]} R d\mu = \sum_i A(a_i) \int_{[a; b]} L_i d\mu.$$

Les poids (ω_i) définis par $\omega_i = \int_{[a; b]} L_i d\mu$ conviennent et ce sont les seuls. Pour s'en convaincre, il suffit d'appliquer la formule de quadrature aux (L_i) .

Reste maintenant à prouver que l'ordre de la méthode ne dépasse pas $2n + 1$. Mais cela est facile par l'absurde en reprenant les arguments de début de démonstration. En effet, le polynôme $P(X) = \prod_i (X - a_i)$ de degré $n + 1$ serait orthogonal à $\mathbb{R}_{n+1}[X]$ donc orthogonal à lui-même, ce qui conduit à une contradiction. \square

5.3.3 Accélération de Romberg-Richardson

Voici le principe général de la méthode :

- Moyennant une transformation affine sur les abscisses, la formule d'Euler-Mac-Laurin nous offre un développement limité de la quadrature des trapèzes dont le terme constant est $\int f$.
- Avec une astuce de calcul, on élimine de manière récursive les termes polynomiaux non constants du développement limité, *sans recalculer de points intermédiaires dans notre quadrature*.

Dans tout ce paragraphe, on se donne $(n; p) \in (\mathbb{N}^*)^2$ et f une fonction de classe \mathcal{C}^{2p+1} sur un intervalle $[a; b]$.

Commençons par réaliser la transformation qui nous permettra d'exploiter la formule d'Euler-Mac-Laurin. Pour tout n , on pose $\varphi_n : s \mapsto f\left(a + s \frac{b-a}{n}\right)$, de sorte que :

- φ_n est de classe \mathcal{C}^{2p+1} sur $[0; n]$;
- $\int_a^b f = \frac{b-a}{n} \int_0^n \varphi_n$.

Dans toute la suite, on posera :

$$T_n = \frac{b-a}{n} \sum_{k=0}^{n-1} \frac{1}{2} \left(f\left(a + k \frac{b-a}{n}\right) + f\left(a + (k+1) \frac{b-a}{n}\right) \right) = \frac{b-a}{n} \sum_{k=0}^{n-1} \frac{1}{2} (\varphi_n(k) + \varphi_n(k+1))$$

C'est l'approximation par la méthode des trapèzes de $\int_a^n f$ en réalisant n subdivisions.

La formule d'Euler-Mac-Laurin appliquée à φ_n donne :

$$T_n = \frac{b-a}{n} \left[\int_0^n \varphi_n + \sum_{k=1}^p \frac{b_{2k}}{(2k)!} \left(\varphi_n^{(2k-1)}(n) - \varphi_n^{(2k-1)}(0) \right) - \frac{1}{(2p+1)!} \int_0^n B_{2p+1} \varphi_n^{(2p+1)} \right]$$

Or, pour tout $0 \leq k \leq 2p+1$ et pour tout $s \in [0; n]$, $\varphi_n^{(k)}(s) = \left(\frac{b-a}{n}\right)^k f^{(k)}\left(a + s \frac{b-a}{n}\right)$, ce qui permet de réécrire :

$$T_n = \int_a^b f + \sum_{k=1}^p \left(\frac{b_{2k}}{(2k)!} \right) (f^{2k-1}(b) - f^{2k-1}(a)) \frac{(b-a)^{2k}}{n^{2k}} + \frac{1}{(2p+1)!} \left(\int_0^n B_{2p+1}(s) f^{(2p+1)}\left(a + s \frac{b-a}{n}\right) ds \right) \frac{(b-a)^{2p+1}}{n^{2p+2}}$$

Remarquons que $\int_0^n B_{2p+1}(s) f^{(2p+1)}\left(a + s \frac{b-a}{n}\right) ds = O(n)$, ce qui prouve que le reste intégral est un $O\left(\frac{(b-a)^{2p+1}}{n^{2p+1}}\right)$.

Cette remarque permet de construire cette proposition.

Proposition : Développement limité de l'approximation des trapèzes

On reprend les mêmes hypothèses et notations. Alors :

$$T_n = c_0 + \sum_{k=1}^p c_k \frac{(b-a)^{2k}}{n^{2k}} + O\left(\frac{(b-a)^{2p+1}}{n^{2p+1}}\right)$$

Avec, pour tout $1 \leq k \leq p$, $c_k = \left(\frac{b_{2k}}{(2k)!}\right) (f^{2k-1}(b) - f^{2k-1}(a))$ et $c_0 = \int_a^b f$

L'objectif du paragraphe suivant est d'obtenir, en jouant sur les différentes valeurs de n , une approximation fine du terme constant $\int_a^b f$ sans devoir recalculer des points.

Remarquons que, en formant la quantité $2^2 T_{2n} - T_n$, le terme de degré 2 s'élimine. Sur ce principe d'élimination, nous allons construire l'algorithme de Romberg-Richardson qui élimine les termes de degré $2k$ par ordre croissant.

Posons pour tout $0 \leq j \leq p$, on pose $R_j = T_{2^j}$. On a ainsi :

$$\begin{aligned} R_0 &= c_0 + \sum_{k=1}^p c_k (b-a)^{2k} + O((b-a)^{2p+1}) \\ R_1 &= c_0 + \sum_{k=1}^p c_k \frac{(b-a)^{2k}}{2^{2k}} + O\left(\frac{(b-a)^{2p+1}}{2^{2p+1}}\right) \\ R_2 &= c_0 + \sum_{k=1}^p c_k \frac{(b-a)^{2k}}{2^{4k}} + O\left(\frac{(b-a)^{2p+1}}{2^{4p+2}}\right) \\ &\dots \\ R_j &= c_0 + \sum_{k=1}^p c_k \frac{(b-a)^{2k}}{2^{2jk}} + O\left(\frac{(b-a)^{2p+1}}{2^{j(2p+1)}}\right) \\ R_{j+1} &= c_0 + \sum_{k=1}^p c_k \frac{(b-a)^{2k}}{2^{2(j+1)k}} + O\left(\frac{(b-a)^{2p+1}}{2^{(j+1)(2p+1)}}\right) \\ &\dots \\ R_p &= c_0 + \sum_{k=1}^p c_k \frac{(b-a)^{2k}}{2^{2pk}} + O\left(\frac{(b-a)^{2p+1}}{2^{p(2p+1)}}\right) \end{aligned}$$

Pour tout $0 \leq j \leq p-1$, formons la quantité $\frac{4R_{j+1}-R_j}{3}$:

$$\frac{4R_{j+1}-R_j}{3} = c_0 + \sum_{k=2}^p c_k (b-a)^{2k} \frac{1}{3} \left(\frac{1}{2^{2(j+1)k-2}} - \frac{1}{2^{2jk}} \right) + O\left(\frac{(b-a)^{2p+1}}{2^{(j+1)(2p+1)-2}}\right) + O\left(\frac{(b-a)^{2p+1}}{2^{j(2p+1)}}\right)$$

Commençons par examiner les restes. On a $(j+1)(2p+1)-2 = j(2p+1) + 2p-1 > j(2p+1)$ ce qui montre que $O\left(\frac{(b-a)^{2p+1}}{2^{(j+1)(2p+1)-2}}\right) + O\left(\frac{(b-a)^{2p+1}}{2^{j(2p+1)}}\right) = O\left(\frac{(b-a)^{2p+1}}{2^{j(2p+1)}}\right)$.

Calculons maintenant le terme général :

$$\frac{1}{2^{2(j+1)k-2}} - \frac{1}{2^{2jk}} = \frac{1}{2^{2jk+2(k-1)}} - \frac{1}{2^{2jk}} = \frac{1}{2^{2jk}} \frac{(1-2^{2(k-1)})}{2^{2(k-1)}}$$

On en déduit :

$$\frac{4R_{j+1}-R_j}{3} = c_0 + \sum_{k=2}^p c_k (b-a)^{2k} \frac{(1-2^{2(k-1)})}{3 \times 2^{2(k-1)}} \frac{1}{2^{2jk}} + O\left(\frac{(b-a)^{2p+1}}{2^{j(2p+1)}}\right)$$

Pour tout $2 \leq k \leq p$, les nouveaux coefficients sont $c_k^{(1)} = c_k \frac{(1-2^{2(k-1)})}{3 \times 2^{2(k-1)}}$ et sont indépendants de j .

Nous voici munis de l'initialisation et de la première étape de la méthode de Romberg-Richardson. Reste à éliminer le terme de degré 4, puis le terme de degré 6 et ainsi de suite. Le théorème suivant nous fournit l'algorithme complet.

Théorème : Algorithme d'accélération de Romberg Richardson

On reprend les notations et hypothèses du paragraphe. On considère l'algorithme suivant :

Pour $0 \leq i \leq p$:

On assigne $R_i^{(0)} = T_{2^i}$

Fin pour

Pour $1 \leq k \leq p$:

Pour $0 \leq i \leq p-k$:

On assigne $R_i^{(k)} = \frac{1}{4^k-1} \left(4^k R_{i+1}^{(k-1)} - R_i^{(k-1)} \right)$

Fin pour

Fin pour

On renvoie $R_0^{(p)}$

Alors cet algorithme renvoie $\int_a^b f$ *avec un reste négligeable devant* $(b-a)^{2p}$.

Démonstration. Cela se fait par récurrence sur p en généralisant ce que l'on vient de faire au paragraphe précédent. □

Chapitre 6

Compléments sur l'intégration

6.1 Tribu, mesure produit, théorème de Fubini-Tonelli

6.1.1 Théorème d'unicité de mesure

Définition : π -système

Un ensemble \mathcal{C} de parties de E est appelé π -système lorsque $\forall C_1, C_2 \in \mathcal{C}, C_1 \cap C_2 \in \mathcal{C}$.

Définition : λ -système

Un ensemble \mathcal{L} de parties de E est appelé λ -système lorsque :

- $E \in \mathcal{L}$;
- \mathcal{L} est stable par union dénombrable croissante ; i.e. $\forall (L_i)_{i \in \mathbb{N}} / \forall i, L_{i+1} \supset L_i, \bigcup_{i \in \mathbb{N}} L_i \in \mathcal{L}$;
- \mathcal{L} est stable par différence, i.e. $\forall L, M \in \mathcal{L} / L \subset M, M - L \in \mathcal{L}$.

On peut de la même manière fabriquer des λ -système engendrés par des sous-ensembles de parties de E exactement comme avec les tribus. La notion de λ -système trouve son intérêt dans les preuves d'unicité de mesures ; comme exposé plus bas.

Proposition : λ -système engendré par un π -système

Soit \mathcal{C} un π -système de E contenant E . Soit d'autre part \mathcal{L} le λ -système engendré par \mathcal{C} .

Alors \mathcal{L} est la tribu engendrée par \mathcal{C} .

Démonstration. On va considérer \mathcal{T} la tribu engendrée par \mathcal{C} . Il est évident qu'une tribu est un λ -système donc $\mathcal{L} \subset \mathcal{T}$.

Pour prouver l'inclusion réciproque on va prouver que \mathcal{L} est une tribu. Il nous faut donc prouver pour cela qu'elle est stable par réunion dénombrable, la stabilité par complémentaire étant évidente.

Ainsi, considérant une famille $(L_i)_{i \in \mathbb{N}}$, on veut prouver que $\bigcup_{i \in \mathbb{N}} L_i \in \mathcal{L}$. Mais, si, pour tout entier i , on pose

$\hat{L}_i = \bigcup_{0 \leq n \leq i} L_n$, et que l'on montre que $\hat{L}_i \in \mathcal{L}$, c'est gagné car alors les \hat{L}_i constituent une famille dénombrable

et croissante d'éléments de \mathcal{L} . Il s'agit donc de montrer que \mathcal{L} est stable par réunion finie, ce qui est équivalent à prouver que \mathcal{L} est stable par intersection finie, en faisant la remarque que $L_1 \cup L_2 = \overline{\overline{L_1} \cap \overline{L_2}}$.

Pour ce faire, on procède en deux temps :

- on montre que \mathcal{L} est stable par intersection avec les éléments de \mathcal{C} ;
- puis on en déduit que \mathcal{L} est stable par intersection finie.

Considérons donc dans un premier temps l'ensemble $\mathcal{M} = \{L \in \mathcal{L} / \forall C \in \mathcal{C}, L \cap C \in \mathcal{L}\}$. On sait que $\mathcal{C} \subset \mathcal{M} \subset \mathcal{L}$. De plus, pour toute famille dénombrable et croissante de $(L_i)_{i \in \mathbb{N}}$ de \mathcal{M} , $\bigcup_{i \in \mathbb{N}} L_i \in \mathcal{M}$ car, pour tout C de \mathcal{C} ,

$\left(\bigcup_{n \in \mathbb{N}} L_i\right) \cap C = \bigcup_{n \in \mathbb{N}} (L_i \cap C)$ et les $L_i \cap C$ sont aussi une famille croissante et dénombrable de \mathcal{L} . Enfin, de la même manière, \mathcal{M} est stable par différence car pour tous les éléments M et L de \mathcal{M} tels que $L \subset M$, on a $(M - L) \cap C = (M \cap C) - (L \cap C)$. Ainsi, \mathcal{M} est un λ -système. En particulier, cela prouve que $\mathcal{L} \subset \mathcal{M}$.

On a donc prouvé que $\mathcal{L} = \mathcal{M}$ en raison de la double inclusion.

Dans un second temps, considérons $\mathcal{N} = \{L \in \mathcal{L} / \forall M \in \mathcal{L}, L \cap M \in \mathcal{L}\}$. D'après ce que l'on vient d'établir, $\mathcal{C} \subset \mathcal{N}$. De plus, on peut, en répétant le même raisonnement montrer que \mathcal{N} est un λ -système et ainsi en déduire $\mathcal{L} = \mathcal{N}$. Finalement, \mathcal{L} est bien stable par intersection finie, ce qui achève la démonstration. \square

Théorème : Unicité de la mesure

Soit un ensemble mesurable $(E; \mathcal{T})$ dotée de deux mesures μ_1 et μ_2 .

On suppose que :

- \mathcal{T} est engendrée par un π -système \mathcal{C} ;
- il existe une suite E_n croissante d'éléments de \mathcal{C} telle que $E = \lim \uparrow E_n$ et pour tout n , $\mu_1(E_n) < +\infty$ et $\mu_2(E_n) < +\infty$;
- pour tout $C \in \mathcal{C}$, $\mu_1(C) = \mu_2(C)$.

Alors les mesures μ_1 et μ_2 sont identiques.

Démonstration. Pour tout n , on pose $\mathcal{C}_n = \{C \cap E_n, C \in \mathcal{C}\}$, $\mathcal{L}_n = \{T \in \mathcal{T} / \mu_1(T) = \mu_2(T) \text{ et } T \subset E_n\}$ et enfin $\mathcal{T}_n = \{T \in \mathcal{T} / T \subset E_n\}$.

En raison des propriétés des mesures μ_1 et μ_2 , \mathcal{L}_n est stable par différence, par union croissante dénombrable et contient E_n . C'est donc un λ -système qui contient \mathcal{C}_n .

Donc, d'après le théorème précédent, \mathcal{L}_n contient \mathcal{T}_n .

Finalement, pour tout élément T de \mathcal{T} et pour tout n , $\mu_1(T \cap E_n) = \mu_2(T \cap E_n)$.

On conclut ensuite par passage à la limite sur n . \square

6.1.2 Définition de la tribu produit

On considère deux ensembles mesurés $(E_1; \mathcal{T}_1; \mu_1)$ et $(E_2; \mathcal{T}_2; \mu_2)$.

Alors, on peut construire la tribu produit de $E_1 \times E_2$ en considérant la tribu engendrée les cylindres $\{(T_1; E_2), (E_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\}$. On note $\mathcal{T}_1 \otimes \mathcal{T}_2$ cette tribu produit.

C'est la plus petite tribu qui rende les applications « coordonnées » mesurables. De plus, cette tribu est également engendrée par un π -système :

Proposition :

La tribu engendrée par $\{(T_1; E_2), (E_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\}$ est également la tribu engendrée par $\{(T_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\}$.

Ce dernier ensemble constitue d'ailleurs un π -système.

Démonstration. On note \mathcal{T} , la tribu engendrée par $\{(T_1; E_2), (E_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\}$ et \mathcal{T}' celle engendrée par $\{(T_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\}$.

Remarquons que $\{(T_1; E_2), (E_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\} \subset \{(T_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\}$ donc $\mathcal{T} \subset \mathcal{T}'$.

Pour montrer l'implication réciproque, il suffit de constater que $(T_1; E_2) \cap (E_1; T_2) = (T_1; T_2)$. Ainsi, tous les éléments de $\{(T_1; T_2), T_1 \in \mathcal{T}_1 \text{ et } T_2 \in \mathcal{T}_2\}$ sont dans \mathcal{T} et, par suite, $\mathcal{T}' \subset \mathcal{T}$. \square

6.1.3 Définition de la mesure produit

Pour définir une mesure sur cette tribu produit, il nous faut énoncer le théorème de Fubini-Tonelli que nous démontrerons en utilisant les λ -systèmes.

Théorème : Fubini-Tonelli, définition de la mesure produit

Soient $(E_1; \mathcal{T}_1; \mu_1)$ et $(E_2; \mathcal{T}_2; \mu_2)$ deux espaces mesurés et $\mathcal{T}_1 \otimes \mathcal{T}_2$ la tribu produit associée.

On suppose que $E_1 = \lim_{n \rightarrow +\infty} \uparrow E_1^{(n)}$ et que $E_2 = \lim_{n \rightarrow +\infty} \uparrow E_2^{(n)}$ avec pour tout n , $\mu_1(E_1^{(n)}) < +\infty$ et $\mu_2(E_2^{(n)}) < +\infty$.

Alors, pour tout $T \in \mathcal{T}_1 \otimes \mathcal{T}_2$:

- la fonction $x_2 \mapsto \int_{E_1} \mathbb{1}_T(x_1; x_2) d\mu_1(x_1)$ est mesurable ;
- la fonction $x_1 \mapsto \int_{E_2} \mathbb{1}_T(x_1; x_2) d\mu_2(x_2)$ est mesurable ;
- et on a l'égalité :

$$\int_{E_2} \left(\int_{E_1} \mathbb{1}_T(x_1; x_2) d\mu_1(x_1) \right) d\mu_2(x_2) = \int_{E_1} \left(\int_{E_2} \mathbb{1}_T(x_1; x_2) d\mu_2(x_2) \right) d\mu_1(x_1)$$

On note $\mu(T) = \int_{E_2} \left(\int_{E_1} \mathbb{1}_T(x_1; x_2) d\mu_1(x_1) \right) d\mu_2(x_2) = \int_{E_1} \left(\int_{E_2} \mathbb{1}_T(x_1; x_2) d\mu_2(x_2) \right) d\mu_1(x_1)$ et μ constitue la mesure produit sur $\mathcal{T}_1 \otimes \mathcal{T}_2$.

Démonstration. On considère l'ensemble \mathcal{C} des pavés de $E_1 \times E_2$; c'est à dire les éléments de la forme $(T_1; T_2) \in \mathcal{T}_1 \times \mathcal{T}_2$.

Pour tout $(T_1; T_2)$ de cet ensemble :

- $x_2 \mapsto \int_{E_1} \mathbb{1}_{(T_1; T_2)}(x_1; x_2) d\mu_1(x_1)$ est la fonction $\mathbb{1}_{T_2} \mu_1(T_1)$ et est donc mesurable.
- $x_1 \mapsto \int_{E_2} \mathbb{1}_{(T_1; T_2)}(x_1; x_2) d\mu_2(x_2)$ est la fonction $\mathbb{1}_{T_1} \mu_2(T_2)$ et est donc mesurable.
- et on vérifie :

$$\int_{E_2} \left(\int_{E_1} \mathbb{1}_{(T_1; T_2)}(x_1; x_2) d\mu_1(x_1) \right) d\mu_2(x_2) = \mu_1(T_1) \times \mu_2(T_2)$$

et

$$\int_{E_1} \left(\int_{E_2} \mathbb{1}_{(T_1; T_2)}(x_1; x_2) d\mu_2(x_2) \right) d\mu_1(x_1) = \mu_2(T_2) \times \mu_1(T_1)$$

Ce qui prouve que :

$$\int_{E_1} \left(\int_{E_2} \mathbb{1}_T(x_1; x_2) d\mu_2(x_2) \right) d\mu_1(x_1) = \int_{E_1} \left(\int_{E_2} \mathbb{1}_T(x_1; x_2) d\mu_2(x_2) \right) d\mu_1(x_1)$$

On définit d'autre part $P_n = (E_1^{(n)}, E_2^{(n)})$ et on pose, pour tout n ,

$$\mathcal{L}_1^{(n)} = \left\{ T \subset P_n / x_2 \mapsto \int_{E_1} \mathbb{1}_T(x_1; x_2) d\mu_1(x_1) \text{ est mesurable} \right\}$$

et

$$\mathcal{L}_2^{(n)} = \left\{ T \subset P_n / x_1 \mapsto \int_{E_2} \mathbb{1}_T(x_1; x_2) d\mu_2(x_2) \text{ est mesurable} \right\}$$

$\mathcal{L}_1^{(n)}$ et $\mathcal{L}_2^{(n)}$ sont stables par union croissante, par différence, et contiennent le π -système

$$\mathcal{C}^{(n)} = \{ C \cap P_n / C \in \mathcal{C} \}.$$

C'est donc, d'après les résultats sur les λ -systèmes, la tribu trace de $\mathcal{T}_1 \otimes \mathcal{T}_2$ sur P_n . Le théorème de convergence monotone et les résultats sur les limites de suites de fonctions permet ensuite d'étendre ce résultat, ce qui prouve la mesurabilité des intégrales des deux premiers points du théorème.

Consacrons-nous maintenant au troisième point du théorème.

Pour tout $T \in \mathcal{T}_1 \otimes \mathcal{T}_2$, on pose

$$\tilde{\mu}(T) = \int_{E_1} \left(\int_{E_2} \mathbb{1}_T(x; y) d\mu_1(x) \right) d\mu_2(y) \text{ et } \check{\mu}(T) = \int_{E_1} \left(\int_{E_2} \mathbb{1}_T(x; y) d\mu_1(x) \right) d\mu_2(y)$$

$\tilde{\mu}(T)$ et $\check{\mu}(T)$ sont des mesures en raison des propriétés de l'intégrale et des fonctions indicatrices.

Tous les éléments C de \mathcal{C} vérifient $\tilde{\mu}(C) = \check{\mu}(C)$. D'autre part, \mathcal{C} forme un π -système. Enfin, tous les P_n sont dans \mathcal{C} , de mesures finies et on a $\lim \uparrow P_n = E_1 \times E_2$.

Ainsi, d'après le théorème d'unicité de la mesure, on peut conclure que $\check{\mu} = \tilde{\mu}$ sur $\mathcal{T}_1 \otimes \mathcal{T}_2$. \square

6.1.4 Théorème de Fubini-Tonelli, cas général

Théorème : Fubini-Tonelli

Soient $(E_1; \mathcal{T}_1; \mu_1)$ et $(E_2; \mathcal{T}_2; \mu_2)$ deux espaces mesurés et $\mathcal{T}_1 \otimes \mathcal{T}_2$ la tribu produit associée.

On suppose que $E_1 = \lim_{n \rightarrow +\infty} \uparrow E_1^{(n)}$ et que $E_2 = \lim_{n \rightarrow +\infty} \uparrow E_2^{(n)}$ avec pour tout n , $\mu_1(E_1^{(n)}) < +\infty$ et $\mu_2(E_2^{(n)}) < +\infty$.

Dans ce cas, on peut définir la mesure produit μ sur $\mathcal{T}_1 \otimes \mathcal{T}_2$.

On considère enfin g une fonction mesurable de $E_1 \times E_2$ dans \mathbb{R} (ou \mathbb{C}).

On suppose que :

- Pour μ_1 -presque tout x_1 de E_1 , la fonction $g_2 : x_2 \mapsto g(x_1; x_2)$ est μ_2 -intégrable.
- Pour μ_2 -presque tout x_2 de E_2 , la fonction $g_1 : x_1 \mapsto g(x_1; x_2)$ est μ_1 -intégrable.

Alors, les fonctions $x_2 \mapsto \int_{E_1} g(x_1; x_2) d\mu_1(x_1)$ et $x_1 \mapsto \int_{E_2} g(x_1; x_2) d\mu_2(x_2)$ sont mesurables et on a l'équivalence entre les trois propositions suivantes :

- $x_2 \mapsto \int_{E_1} |g(x_1; x_2)| d\mu_1(x_1)$ est μ_2 -intégrable ;
- $x_1 \mapsto \int_{E_2} |g(x_1; x_2)| d\mu_2(x_2)$ est μ_1 -intégrable ;
- g est intégrable.

Et, dans ce cas, on a :

$$\iint g(x_1; x_2) d\mu(x_1; x_2) = \int_{E_2} d\mu_2(x_2) \int_{E_1} g(x_1; x_2) d\mu_1(x_1) = \int_{E_1} d\mu_1(x_1) \int_{E_2} g(x_1; x_2) d\mu_2(x_2)$$

Démonstration. Dans cette démonstration, on va adopter la démarche suivante :

1. prouver les résultats pour les fonctions positives en passant à la limite sur les suites croissantes de fonctions échelonnées ;
2. étendre les résultats à \mathbb{R} en utilisant les parties positives et négatives ;
3. étendre les résultats à \mathbb{C} en utilisant les parties réelles et imaginaires.

On va commencer par supposer que g est mesurable et à valeurs positives sans faire l'hypothèse de l'intégrabilité.

Dans ce cas g est limite croissante de fonctions échelonnées positives. Mais la première version du théorème de Fubini-Tonelli nous montre que les fonctions échelonnées vérifient les trois points du théorème (mesurabilités et égalités d'intégrales).

D'autre part, en raison du théorème de convergence monotone, ces trois points restent valides par passage à la limite croissante.

On en déduit que le théorème de Fubini-Tonelli peut s'appliquer aux fonctions mesurables positives, sans d'ailleurs condition particulière sur l'intégrabilité !

On va étendre le résultat aux réels en considérant les parties positives et négatives de la fonction. C'est ici que le critère d'intégrabilité intervient.

On commence par définir $M = \{(x_1; x_2) / g_2 \text{ est } \mu_2\text{-intégrable et } g_1 \text{ est } \mu_1\text{-intégrable}\}$. Le complémentaire de M est négligeable. On peut donc, à partir de maintenant, raisonner sur $M \subset E_1 \times E_2$.

Les fonctions $x_2 \mapsto \int_{E_1} g(x_1; x_2) d\mu_1(x_1)$ et $x_1 \mapsto \int_{E_2} g(x_1; x_2) d\mu_2(x_2)$ sont effectivement mesurables. Il suffit pour prouver cela d'écrire $g = g^+ - g^-$ et d'exploiter ce qui précède.

Reste à prouver l'équivalence des trois points suivants.

On suppose que $x_2 \mapsto \int_{E_1} |g(x_1; x_2)| d\mu_1(x_1)$ est μ_2 -intégrable.

Cela signifie que :

$$\int_{E_2} d\mu_2(x_2) \int_{E_1} |g(x_1; x_2)| d\mu_1(x_1) < +\infty.$$

Ce qui précède concernant les fonctions positives permet d'inverser les signes d'intégration. On en déduit ainsi le second point et le troisième point.

Les égalités finales découlent des intégrales des parties positives et négatives de g .

Enfin, le passage aux complexes se fait par l'étude des parties réelles et imaginaires. \square

6.2 Inégalités, espaces L^p , dérivation

6.2.1 Inégalités de Hölder et de Minkowski

On se réfère ici au petit document sur la convexité dans lequel on a prouvé ces deux inégalités dans le cas des sommes finies.

Proposition : Inégalité de Hölder et de Minkowski dans le cas des fonctions positives

Soient f et g deux fonctions mesurables de $(E; \mathcal{T}; \mu)$ à valeurs positives.

Soient p et q deux nombres positifs tels que $\frac{1}{p} + \frac{1}{q}$

Alors

$$\int (f \times g) \leq \left(\int f^p \right)^{1/p} \times \left(\int g^q \right)^{1/q} \quad (\text{Hölder})$$

et

$$\left(\int (f + g)^p \right)^{1/p} \leq \left(\int f^p \right)^{1/p} + \left(\int g^p \right)^{1/p} \quad (\text{Minkowski})$$

Démonstration. Si $\int f = 0$ ou $\int g = 0$ alors l'inégalité de Hölder devient $0 \leq 0$ et elle est vraie.

Si $\int f = +\infty$ ou $\int g = +\infty$ alors l'inégalité de Hölder est également vraie.

On se placera donc dans le cas où $0 < \int f < +\infty$ et $0 < \int g < +\infty$.

Le principe est le même que pour le cas des sommes discrètes.

On part de l'inégalité de Young et on « normalise ».

Ainsi, on pose $\tilde{f} = \frac{f^p}{\int f^p}$ et $\tilde{g} = \frac{g^q}{\int g^q}$.

L'inégalité de Young donne donc, pour tout x de E :

$$\widetilde{f(x)}^{1/p} \widetilde{g(x)}^{1/q} \leq \frac{1}{p} \widetilde{f(x)} + \frac{1}{q} \widetilde{g(x)}$$

On intègre cette inégalité sur E en remarquant que $(f^p)^{1/p} \times (g^q)^{1/q} = fg$ et que $\int \tilde{f} = \int \tilde{g} = 1$. On obtient ainsi :

$$\frac{\int (fg)}{\left(\int f^p \right)^{1/p} \left(\int g^q \right)^{1/q}} \leq \frac{1}{p} + \frac{1}{q} = 1$$

Maintenant, on peut prouver l'inégalité de Minkowski.

Pour tout $p \geq 1$, on a en effet, $(f + g)^p = (f + g)^{p-1}(f + g) = f(f + g)^{p-1} + g(f + g)^{p-1}$.

Il suffit alors de poser $q = \frac{p}{p-1}$ et d'appliquer l'inégalité de Hölder à $f(f + g)^{p-1}$ puis à $g(f + g)^{p-1}$, en remarquant que $q(p-1) = p$ pour obtenir l'inégalité de Minkowski. Plus précisément, Hölder donne

$$\int (f(f + g)^{p-1}) \leq \left(\int f^p \right)^{1/p} \left(\int (f + g)^{(p-1)q} \right)^{1/q}$$

ce que l'on peut réécrire

$$\int (f(f+g)^{p-1}) \leq \left(\int f^p \right)^{1/p} \left(\int (f+g)^p \right)^{1/q}$$

Et de la même manière :

$$\int (g(f+g)^{p-1}) \leq \left(\int g^p \right)^{1/p} \left(\int (f+g)^p \right)^{1/q}$$

En sommant ces deux inégalités et en divisant par $\left(\int (f+g)^p \right)^{1/q}$ (avec toutes les mesures de prudence qui s'imposent), il vient

$$\left(\int (f+g)^p \right)^{1-1/q} \leq \left(\int f^p \right)^{1/p} + \left(\int g^p \right)^{1/p}$$

En remarquant que $1 - \frac{1}{q} = \frac{1}{p}$, on retrouve le résultats recherché. \square

6.2.2 Notion d'espaces L^p

Définition : Espaces L^p

Soit f une fonction mesurable de $(E; \mathcal{T}; \mu)$ à valeurs dans \mathbb{R} (ou \mathbb{C}).

On dit que f appartient à l'ensemble \mathcal{L}^p lorsque f^p est intégrable.

Par ailleurs, sur l'ensemble \mathcal{L}^p , on définit une relation d'équivalence

$$f \sim g \iff \int |f - g| = 0$$

Les classes d'équivalences de \mathcal{L}^p forment un espace vectoriel sur lequel on définit une norme.

$$\|f\|_p = \left(\int |f|^p \right)^{1/p}$$

Démonstration. L'inégalité de Minkowski nous offre une preuve de la stabilité par addition et de la définition de la norme. \square

Examinons maintenant les convergences dans les espaces L^p pour lesquels on dispose de théorèmes équivalents.

Proposition : Convergence dominée dans les espaces L^p

Soit $p \geq 1$. On considère une suite f_n de fonctions telles que :

- les f_n sont L^p ;
- il existe une fonction g de classe L^p qui domine chacun des f_n μ -presque partout ;
- les f_n converge simplement μ -presque partout vers une fonction f .

Alors f est de classe L^p et $\lim \int f_n = \int \lim f_n$

Démonstration. Soit la suite de fonctions $\delta_n = |f_n - g|$.

On veut montrer que f est de classe L^p et que δ_n tend vers 0 pour la norme L^p .

Le fait que f soit de classe L^p provient de l'application du théorème de convergence dominée aux fonctions intégrables f_n^p dominées par la fonction intégrable g^p et qui tendent simplement vers f^p μ -presque partout.

On peut donc examiner la quantité $\|\delta_n\|_p^p$. Un tout petit peu de calcul montre que :

$$|\delta_n|^p \leq 2^p g^p$$

On en déduit que $(\delta_n)^p$ est dominée par $2^p g^p$ μ -presque partout. Or δ_n converge simplement vers 0 μ -presque partout.

Le théorème de convergence dominée entraîne ainsi que $\|\delta_n\|_p \rightarrow 0$, ce qui achève la démonstration. \square

Théorème : Les espaces L^p sont complets*Soit $p \geq 1$.**Alors L^p est complet.*

Avant de démontrer ce théorème, il faut établir un lemme.

Lemme : Caractérisation de la complétude grâce aux séries*Soit un espace vectoriel normé E .**Alors E est complet si et seulement si pour toute série $\sum_{n \in \mathbb{N}} u_n$, la série converge si elle est absolument convergente.*

Démonstration. Si E est complet et que la série est absolument convergente alors la suite $S_n = \sum_{k=0}^n u_k$ est de Cauchy. En effet :

$$|S_{n+p} - S_n| \leq \sum_{n+1 \leq k \leq n+p} |u_k|$$

Ce qui entraîne la convergence de la suite S_n .

Réciproquement, si on a la propriété de convergence absolue, on va considérer une suite u_n de Cauchy. On peut donc, en utilisant une extractrice φ fabriquer une série télescopique $u_{\varphi(n+1)} - u_{\varphi(n)}$ telle que, pour tout n :

$$|u_{\varphi(n+1)} - u_{\varphi(n)}| \leq \frac{1}{2^n}$$

En particulier cette série est absolument convergente donc convergente, ce qui prouve que la suite $u_{\varphi(n)}$ converge simplement. Et donc, u_n également. \square

Prouvons maintenant le théorème précédent.

Démonstration. On va utiliser le lemme. On va donc considérer une série $\sum_{n \in \mathbb{N}} f_n$ absolument convergente dans L^p .

On pose, pour tout n , $G_n = \sum_{k \leq n} |f_k|$.

Pour tout n , $\|G_n\|_p \leq \sum_{k \in \mathbb{N}} \|f_k\|_p < +\infty$.

Par le théorème de convergence monotone, on en déduit que $\|G_\infty\|_p < +\infty$.

Ainsi, pour presque tout x , $G_\infty(x) < +\infty$, ce qui prouve que la série $\sum_{k \in \mathbb{N}} f_k(x)$ est absolument convergente donc convergente.

Enfin, le théorème de convergence dominée s'applique ici. En effet, pour tout n , $\left| \sum_{k \leq n} f_k \right| \leq G_\infty$ qui est L^p .

On en déduit que la série $\sum_{k \leq n} f_k$ converge dans L^p . \square

6.2.3 Dérivation et intégrale

Théorème : Dérivation sous le signe intégrale

On considère $(E; \mathcal{T}; \mu)$ un espace mesuré et $I \subset \mathbb{R}$ un intervalle ouvert.

Soit $f : E \times I \rightarrow \mathbb{R}$ une fonction.

On suppose que :

- pour tout y de I , $x \mapsto f(x; y)$ est mesurable et μ -intégrable ;
- pour μ -presque tout x de E , $y \mapsto f(x; y)$ est dérivable sur I ;
- il existe une fonction $g : E \rightarrow \mathbb{R}^+$ mesurable et intégrable telle que

$$\forall y \in I, \left| \frac{\partial f(x; y)}{\partial y} \right| \leq g(x)$$

Alors :

- pour tout y de I , $x \mapsto \frac{\partial f(x; y)}{\partial y}$ est mesurable et μ -intégrable ;
- la fonction $y \mapsto \int_E f(x; y) d\mu(x)$ est dérivable sur I ;
- on a l'égalité :

$$\frac{\partial}{\partial y} \int_E f(x; y) d\mu(x) = \int_E \frac{\partial f(x; y)}{\partial y} d\mu(x)$$

Démonstration. On considère un élément y quelconque de I , et une suite α_n d'éléments de I qui vérifie $\alpha_n \rightarrow y$ et $\forall n, \alpha_n \neq y$.

On pose alors $\varphi_n : x \mapsto \frac{f(x; \alpha_n) - f(x; y)}{\alpha_n - y}$.

La suite des φ_n est une suite de fonctions mesurables de E dans \mathbb{R} .

Par ailleurs, cette suite converge pour μ -presque tout x vers $x \mapsto \frac{\partial f(x; y)}{\partial y}$.

Enfin, en raison de l'égalité des accroissements finis, les valeurs absolues des termes de cette suite sont toutes dominées par une fonction g intégrable.

On est donc dans les hypothèses du théorème de convergence dominée.

Ainsi, on en déduit que $x \mapsto \frac{\partial f(x; y)}{\partial y}$ est mesurable et μ -intégrable et, par linéarité de l'intégrale, que

$$\lim_{n \rightarrow +\infty} \frac{\int_E f(x; \alpha_n) d\mu(x) - \int_E f(x; y) d\mu(x)}{\alpha_n - y} = \int_E \frac{\partial f(x; y)}{\partial y} d\mu(x)$$

Comme la suite des α_n est a priori quelconque, cela prouve la dérivabilité de $y \mapsto \int_E f(x; y) d\mu(x)$ ainsi que la dernière égalité. \square

Chapitre 7

Convolution et transformée de Fourier

7.1 Convolution

Sauf indication contraire, on se place dans \mathbb{R}^d muni de la mesure de Lebesgue.

7.1.1 Définition, premières propriétés

Définition : Convolution

Soient f et g deux fonctions mesurables et $x \in \mathbb{R}^d$. On suppose que $t \mapsto f(t)g(x-t)$ est intégrable.

Alors, on pose $f * g(x) = \int f(t)g(x-t) d\lambda(t)$.

(Remarque 0) Il existe plein de conditions différentes garantissant l'existence de ce produit de convolution. La plus large étant celle de la définition.

Proposition : Premières propriétés

(Mêmes hypothèses et notations)

Si $f * g(x)$ existe alors $g * f(x)$ existe aussi et $g * f(x) = f * g(x)$.

Si h est une autre fonction mesurable telle que $f * h(x)$ existe alors, pour tout $(\lambda; \mu) \in \mathbb{C}^2$, $f * (\lambda g + \mu h)(x)$ existe aussi et on a :

$$f * (\lambda g + \mu h)(x) = \lambda f * g(x) + \mu f * h(x)$$

Démonstration. La commutativité se prouve à l'aide d'un changement de variable et la linéarité provient de la linéarité de l'intégrale. \square

Nous allons maintenant voir un cas particulier : les fonctions L^1 .

Proposition : Convolution L^1 et L^1

Soient f et g deux fonctions de L^1 .

Alors, pour presque tout x , $f * g(x)$ existe.

De plus, $f * g$ est de classe L^1 et on a :

$$\|f * g\| \leq \|f\|_1 \times \|g\|_1$$

Démonstration. On utilise le théorème de Fubini-Tonelli.

Ainsi, pour tout x , $\int |f(t)g(x-t)| d\lambda(t)$ est λ -mesurable et positive.

De plus :

$$\int d\lambda(x) \int d\lambda(t) |f(t)g(x-t)| = \int d\lambda(t) \int d\lambda(x) |f(t)g(x-t)| = \|g\|_1 \|f\|_1$$

On en déduit que $(t; x) \mapsto f(t)g(x-t)$ est intégrable sur \mathbb{R}^2 muni de la mesure produit.

En particulier, $f * g(x)$ existe pour presque tout x et est de classe L^1 et on a :

$$\|f * g\| \leq \|f\|_1 \times \|g\|_1$$

□

Une propriété importante de la convolution L^1 : l'associativité.

Proposition : Associativité

Soient f, g et h trois fonctions L^1 . Alors $(f * g) * h = f * (g * h)$.

Démonstration. On sait que $f * g$ est L^1 donc, pour presque tout x , $(f * g) * h(x)$ existe.

En particulier, on a :

$$\begin{aligned} (f * g) * h(x) &= \int d\lambda(t) h(x-t) \int d\lambda(u) f(u) g(t-u) \\ &= \int \int h(x-t) f(u) g(t-u) d\lambda(t) d\lambda(u) \end{aligned}$$

De même, pour presque tout x , $f * (g * h)(x)$ existe et on a :

$$\begin{aligned} f * (g * h)(x) &= \int d\lambda(t) f(t) \int d\lambda(u) g(u) h(x-t-u) \\ &= \int \int f(t) g(u) h(x-t-u) d\lambda(t) d\lambda(u) \\ &= \int \int f(u) g(t) h(x-t-u) d\lambda(t) d\lambda(u) \text{ on pose } v = t + u \\ &= \int \int f(u) g(v-u) h(x-v) d\lambda(v) d\lambda(u) \text{ on pose } v = t \\ &= \int \int f(u) g(t-u) h(x-t) d\lambda(t) d\lambda(u) = (f * g) * h(x) \end{aligned}$$

□

7.1.2 Régularisation

On raisonne ici sur des fonctions de \mathbb{R} dans \mathbb{R} muni de la tribu des boréliens.

7.1.2.1 Une fonction C^∞ à support compact

Lemme : Construction d'une fonction C^∞ à support compact

Soit la fonction $\phi : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-1/x} & \text{si } x > 0 \end{cases}$

Alors la fonction ϕ est C^∞ .

Et ainsi, la fonction :

$$g : x \mapsto \phi(1-x) \times \phi(1+x)$$

est de classe C^∞ à support compact.

Démonstration. On peut prouver par récurrence que, pour tout n , il existe un polynôme P_n tel que, pour tout $x > 0$:

$$\phi^{(n)}(x) = P_n\left(\frac{1}{x}\right)e^{-1/x}$$

En particulier, cela montre que $\lim_{\substack{x \rightarrow 0 \\ x > 0}} \phi^{(n)}(x) = \phi^{(n)}(0)$. On en déduit que ϕ est bien \mathcal{C}^∞ .

Et, par construction, g est également \mathcal{C}^∞ et à support sur $[-1; 1]$. \square

7.1.2.2 Approximations de l'unité

Définition : Suite approximante de l'unité

Soit $(g_n)_{n \in \mathbb{N}}$ une suite de fonctions L^1 .

On dit que cette suite est une approximation de l'unité lorsque :

- $\sup_{n \in \mathbb{N}} \|g_n\|_1 < +\infty$;
- pour tout n , $\int g_n = 1$;
- pour tout $\eta > 0$, $\lim_{n \rightarrow +\infty} \int |g_n(t)| \times \mathbb{1}_{t > \eta} d\lambda(t) = 0$

On peut aisément fabriquer une suite approximante de l'unité à partir d'une fonction positive L^1 .

Proposition : Création d'une suite approximante de l'unité

Soit f une fonction positive de classe L^1 non identiquement nulle.

On pose $\tilde{f} = \frac{f}{\|f\|_1}$ et pour tout $n \in \mathbb{N}^*$, $f_n : t \mapsto n\tilde{f}(nt)$.

Alors f_n est une suite approximante de l'unité.

Démonstration. Par construction, pour tout n , $\|f_n\| = \int f_n = 1$ (cela s'obtient avec un changement de variable $t \mapsto \frac{t}{n}$).

De plus, pour tout $\eta > 0$, par un changement de variable et compte-tenu de ce qui précède :

$$\int_{t > \eta} f_n(t) d\lambda(t) = 1 - \int_{t \leq \eta} f_n(t) d\lambda(t) = 1 - \int_{u \leq n\eta} \tilde{f}(u) d\lambda(u)$$

Or, par le théorème de convergence monotone, $\lim_{n \rightarrow +\infty} \int_{u \leq n\eta} \tilde{f}(u) d\lambda(u) = 1$, ce qui permet de conclure. \square

7.1.2.3 Le cas des fonctions à support compact

Proposition : Fonctions continues à support compact ou à limite nulle en l'infini

On définit les ensembles \mathcal{C}_k et \mathcal{C}_0 des fonctions définies sur \mathbb{R}^d .

$$f \in \mathcal{C}_k \iff f \text{ est continue et } \exists K \subset \mathbb{R}^d / K \text{ est compact et } \forall x \notin K, f(x) = 0$$

$$f \in \mathcal{C}_0 \iff f \text{ est continue et } \lim_{|x| \rightarrow +\infty} f(x) = 0$$

On peut alors munir ces deux ensembles de la norme infinie et ils constituent dans ce cas des \mathbb{R} -espaces vectoriels.

De plus, toute fonction appartenant à l'un de ces ensembles est uniformément continue.

Enfin, l'espace \mathcal{C}_0 est complet.

Démonstration. Il est assez évident que $\mathcal{C}_k \subset \mathcal{C}_0$. On va donc montrer qu'une fonction f de \mathcal{C}_0 est bornée, uniformément continue.

Considérons une telle fonction f . On sait qu'il existe un disque $\overline{D(0; R)}$ en dehors duquel la fonction $|f|$ est majorée par 1 (en raison de sa limite en l'infini). Sur ce compact, $|f|$ est également majorée. Ainsi, f est bornée.

Soit $\varepsilon > 0$ quelconque et soit $\rho > 0$ tel que pour tout $x \notin \overline{D(0; \rho)}$, $|f(x)| \leq \frac{\varepsilon}{2}$. Sur $\overline{D(0; \rho+1)}$, f est uniformément continue en raison du théorème de Heine.

Donc $\exists \eta \geq \eta > 0$ tel que, $\forall (x; y) \in \overline{D(0; \rho+1)}$, $|f(x) - f(y)| \leq \varepsilon$.

En particulier, pour tout $(x; y) / |x - y| \leq \eta$, on distingue des cas :

- Si $x \notin \overline{D(0; \rho)}$ et si $y \notin \overline{D(0; \rho)}$ en raison de l'inégalité triangulaire et on a ainsi $|f(x) - f(y)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.
- Si $x \notin \overline{D(0; \rho)}$ et si $y \in \overline{D(0; \rho)}$ alors on sait que $x \in \overline{D(0; \rho+1)}$ en raison de l'inégalité triangulaire. On a donc également $|f(x) - f(y)| \leq \varepsilon$ car les deux éléments sont dans $\overline{D(0; \rho+1)}$.
- Enfin, si $x \in \overline{D(0; \rho)}$ alors $y \in \overline{D(0; \rho+1)}$ en raison de l'inégalité triangulaire et dans ce cas $|f(x) - f(y)| \leq \varepsilon$ car les deux éléments sont dans $\overline{D(0; \rho+1)}$.

Finalement, dans tous les cas, $|f(x) - f(y)| \leq \varepsilon$, ce qui prouve que f est uniformément continue.

La stabilité par addition et multiplication par un réel sont faciles à montrer.

Reste donc à prouver que, si $(f_n)_{n \in \mathbb{N}}$ est une suite de Cauchy de \mathcal{C}_0 alors cette suite converge dans \mathcal{C}_0 .

Il est évident que f_n converge simplement. Soit f sa limite.

Pour tout $\varepsilon > 0$, il existe N tel que pour tout $p \geq N$, pour tout $n \geq N$, pour tout x , $|f_p(x) - f_n(x)| \leq \varepsilon$. Ce qui donne par passage à la limite sur p :

$|f(x) - f_n(x)| \leq \varepsilon$, ce qui prouve que f_n converge uniformément vers f et donc que f est continue et bornée.

Reste à prouver que f a pour limite 0 en $+\infty$. Mais cela est relativement simple.

En effet, pour tout $\varepsilon > 0$, il existe N tel que $\|f_N - f\|_\infty \leq \frac{\varepsilon}{2}$.

Or $f_N \in \mathcal{C}_0$ donc il existe $A > 0$ tel que $|x| > A \implies |f_N(x)| \leq \frac{\varepsilon}{2}$.

En particulier, $|f(x)| \leq |f_N(x) - f(x)| + |f_N(x)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. □

Définition : Convolution par des fonctions à support compact

Soit $p \geq 1$ un nombre et n un entier.

Soit g une fonction \mathcal{C}^n à support compact et f une fonction L^p .

Alors la fonction suivante est \mathcal{C}^n :

$$f * g : x \mapsto \int f(t)g(x-t)d\lambda(t)$$

En particulier, pour tout $1 \leq k \leq n$:

$$(f * g)^{(k)}(x) = \int f(t)g^{(k)}(x-t)d\lambda(t)$$

Démonstration. La mesurabilité et la définition de $f * g$ sont des conséquences du théorème de Fubini-Tonelli. Supposons que K est un support compact de g . C'est à dire que K est un compact tel, pour tout $x \notin K$, $g(x) = 0$.

Pour tout $k \leq k \leq n$, on note $M_k = \max |g^{(k)}|$.

Notons ainsi que, pour tout $q \geq 1$, g est L^q car $|g^q| \leq M^q \mathbb{1}_K$. l'inégalité de Hölder nous garantit donc l'existence de $f * g$.

Posons $\theta : (x; t) \mapsto f(t)g(x-t)$. Cette application est n fois dérivable par rapport à x . De plus, pour tout x , et pour tout $1 \leq k \leq n$ $\left| \frac{\partial^k}{\partial x^k} \theta(x; t) \right| \leq M_k \mathbb{1}_K |f(t)|$ qui est de classe L^1 , toujours d'après Hölder.

Le théorème de convergence dominée nous garantit donc que $f * g$ est dérivable k fois et que $(f * g)^{(k)}(x) = \int f(t)g^{(k)}(x-t)d\lambda(t)$.

Reste à prouver que $(f * g)^{(n)}$ est continue pour conclure. Pour cela, on va juste supposer que g est continue à support compact et prouver que $f * g$ est continue, cela suffira.

Pour tout x et y , d'après l'inégalité triangulaire :

$$|f * g(y) - f * g(x)| \leq \int |f(t)| |g(y-t) - g(x-t)| d\lambda(t)$$

On pose maintenant q tel que $\frac{1}{p} + \frac{1}{q} = 1$

Soit $g_y : t \mapsto g(y-t)$ et $g_x : t \mapsto g(x-t)$. L'inégalité de Hölder donne :

$$|f * g(y) - f * g(x)| \leq \|f\|_p \times \|g_y - g_x\|_q$$

On pose $\tilde{K} = \{x+h, |h| \leq 1 \text{ et } x \in K\}$. \tilde{K} est également un compact.

Comme g est continue à support compact, elle est donc uniformément continue. Par conséquent, pour tout $\varepsilon > 0$, il existe $\eta > 0$ éventuellement plus petit que 1, tel que, pour tout u et v tels que $|v-u| \leq \eta$, on a $\|f\|_p^q \times |g(v) - g(u)|^q \times \lambda(\tilde{K}) \leq \varepsilon^q$.

Ainsi, par construction, pour tout x et y tels que $|y-x| \leq \eta \leq 1$, la fonction $g_x - g_y$ a pour support un ensemble de mesure au plus égal à \tilde{K} (en raison de l'invariance de la mesure de Lebesgue par translation et symétrie) et on a donc :

$$\|f\|_p \times \|g_x - g_y\|_q \leq \varepsilon$$

Ainsi, on vérifie que $f * g$ est continue, ce qui achève la démonstration. \square

7.1.2.4 Densité des fonctions \mathcal{C}^∞ à support compact dans L^p

On utilise ici un résultat établi page 33.

Proposition : Approximation d'une indicatrice par convolution

On reprend ici la fonction g définie au paragraphe 7.1.2.1. On pose $\tilde{g} = \frac{g}{\|g\|_1}$

On considère $\chi = \mathbb{1}_{[a; b]}$ et, pour tout $n \in \mathbb{N}^*$, $g_n : t \mapsto n\tilde{g}(nt)$.

Alors, pour tout $p \in [1; +\infty]$, $g_n * \chi \xrightarrow{L^p} \chi$.

Démonstration. Nous allons exploiter le théorème de convergence dominée L^p .

Remarquons que, pour tout n , le support de g_n est $[\frac{-1}{n}; \frac{1}{n}]$. Ainsi, pour $x > b + \frac{1}{n}$ ou $x < a - \frac{1}{n}$, il est impossible d'avoir simultanément $|t| \leq \frac{1}{n}$ et $(x-t) \in [a; b]$, ce qui entraîne que le support de $\chi * g_n$ est inclus dans $[a-1; b+1]$.

Remarquons aussi que, pour tout x , $0 \leq \chi * g_n(x) \leq 1$. Nous sommes donc dans les conditions d'application du théorème de convergence dominée L^p . Il suffit maintenant de montrer que $\chi * g_n$ converge simplement presque partout vers χ .

Or, pour $x > a$ ou $x < b$, il existe N tel que, pour tout $n \geq N$, $x \notin [a - \frac{1}{n}; b + \frac{1}{n}]$, ce qui entraîne $\chi * g_n(x) = 0$. Et, pour $x \in]a; b[$, il existe N tel que, pour tout $n \geq N$, $[x - \frac{1}{n}; x + \frac{1}{n}] \subset [a; b]$, ce qui entraîne $\chi * g_n(x) = 1$. \square

Corollaire : Densité des fonctions \mathcal{C}^∞

Pour $p \in [1; +\infty[$, \mathcal{C}_k^∞ , les fonctions de classe \mathcal{C}^∞ à support compact, sont denses dans L^p

Démonstration. Les combinaisons linéaires d'indicatrices d'intervalles sont denses dans L^p . Or, d'après ce que l'on vient de prouver par l'approximation de l'unité, on sait que les fonctions \mathcal{C}_k^∞ sont denses dans les combinaisons linéaires d'indicatrices d'intervalles. Cela permet de conclure. \square

7.1.2.5 Étude de l'opérateur de translation

Dans ce paragraphe, on peut éventuellement se placer dans \mathbb{R}^d sans que cela nuise au raisonnement qui va suivre.

Définition : Opérateur de translation

Soit f une fonction et $x \in \mathbb{R}^d$. On pose $\tau_x[f] : t \mapsto f(t - x)$.

Quelques propriétés élémentaires.

Proposition : Propriétés de la translation

Soit $p \in [1; +\infty]$ et $x \in \mathbb{R}^d$.

Alors, τ_x est linéaire. De plus, pour tout $f \in L^p$, on a $\tau_x[f] \in L^p$ et $\|\tau_x(f)\|_p = \|f\|_p$.

Enfin, pour $p < +\infty$, on vérifie $\|\tau_x[f] - f\|_p \xrightarrow{x \rightarrow 0} 0$.

Démonstration. Seul le dernier point est délicat. Nous allons exploiter la densité des fonction \mathcal{C}_k dans L^p .

Considérons ainsi une fonction f à support compact K . Nous savons que f est uniformément continue. Considérons l'ensemble $\tilde{K} = \{t/t - x \in K \text{ avec } \|x\| \leq 1\}$.

Remarquons que \tilde{K} est de mesure finie.

Sortons maintenant $\varepsilon > 0$. Il existe $0 < \eta \leq 1$ tel que, pour tout $\|x\| < \eta$, et pour tout t , $\lambda(\tilde{K}) \|f(t - x) - f(t)\|^p < \varepsilon^p$. On en déduit :

$$\|\tau_x(f) - f\|_p = \left(\int |f(x - t) - f(t)|^p d\lambda(t) \right)^{1/p} \leq \varepsilon$$

Ce qui permet de conclure pour ce cas particulier.

Considérons maintenant $f \in L^p$ quelconque. On sait qu'il existe une suite de fonctions (f_n) continues à support compact qui convergent dans L^p vers f .

En particulier, pour tout $\varepsilon > 0$, il existe n tel que $\|f_n - f\|_p \leq \frac{\varepsilon}{3}$. Mais alors, pour tout x :

$$\|\tau_x(f) - f\|_p \leq \|\tau_x(f) - \tau_x(f_n)\|_p + \|\tau_x(f_n) - f_n\|_p + \|f_n - f\|_p \leq \frac{2\varepsilon}{3} + \|\tau_x f_n - f_n\|_p$$

Et comme f_n est continue à support compact, on peut exploiter ce qui vient d'être fait ! □

7.1.2.6 Résultats généraux sur les approximations de l'unité

Proposition : Approximation de l'unité, cas général

Soit f une fonction de classe L^1 . Soit (g_n) une suite approximante de l'unité.

Alors $g_n * f$ converge vers f dans L^1 .

Si f est uniformément continue alors $g_n * f$ converge uniformément vers f .

Démonstration. Pour tout n , on a :

$$\|g_n * f - f\|_1 = \int \left| \int g_n(x) f(t - x) d\lambda(x) - f(t) \right| d\lambda(t) = \int \left| \int g_n(x) (f(t - x) - f(t)) d\lambda(x) \right| d\lambda(t)$$

Ce qui donne, après application du théorème de Fubini :

$$\|g_n * f - f\|_1 \leq \int |g_n(x)| \int |f(t - x) - f(t)| d\lambda(t) d\lambda(x) = \int |g_n(x)| \|\tau_x(f) - f\|_1 d\lambda(x)$$

Notons $M > 0$ un majorant des $(\|g_n\|_1)$.

Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que, pour tout $\|x\| < \eta$, $\|\tau_x(f) - f\|_1 < \frac{\varepsilon}{2M}$.

On en déduit :

$$\|g_n * f - f\|_1 \leq \int \mathbb{1}_{\|x\| < \eta} |g_n(x)| \frac{\varepsilon}{2M} d\lambda(x) + \int 2 \times \mathbb{1}_{\|x\| \geq \eta} |g_n(x)| \|f\|_1 d\lambda(x)$$

D'après notre hypothèse sur les suites approximantes de l'unité, il existe N tel que, pour tout $n \geq N$, $\int \mathbb{1}_{\|x\| \geq \eta} |g_n(x)| d\lambda(x) < \frac{\varepsilon}{4\|f\|_1}$. Ce qui donne, par construction :

$$\|g_n * f - f\|_1 < \varepsilon$$

Considérons maintenant le cas où f est uniformément continue. On reprend ce qui précède, pour tout t :

$$|g_n * f(t) - f(t)| = \left| \int g_n(x)(f(t-x) - f(t)) d\lambda(x) \right| \leq \int |g_n(x)| |f(t-x) - f(t)| d\lambda(x)$$

Comme f est uniformément continue, il existe $\eta > 0$ tel que, pour tout $\|x\| < \eta$, et pour tout t , $|f(t-x) - f(t)| < \frac{\varepsilon}{2M}$. On achève la majoration de la même manière en séparant l'espace selon les cas $\|x\| < \eta$ et $\|x\| \geq \eta$ et on obtient ainsi la convergence uniforme. \square

7.1.3 Des cas plus généraux

Définition : Convolution L^p et L^q

Soient p et q deux éléments de $[1; +\infty]$ tels que $\frac{1}{p} + \frac{1}{q} = 1$.

Soient deux fonctions f, g appartenant respectivement aux espaces L^p et L^q .

Alors, pour tout x , la fonction $f * g : x \mapsto \int f(t)g(x-t)d\lambda(t)$ appartient à \mathcal{C}_0 .

Démonstration. Notons que $f * g$ est bornée pour tout x en raison de l'inégalité de Hölder.

On exploite maintenant la densité des fonctions continues à support compact dans L^p .

Soient f_n et g_n des fonctions continues à support compact qui tendent respectivement vers f et g dans L^p et L^q .

D'après ce qui précède, pour tout n , $f_n * g_n$ est continue et à support compact. En particulier, $f_n * g_n \in \mathcal{C}_0$.

De plus, en utilisant l'inégalité triangulaire et l'inégalité de Hölder, on obtient, pour tout x :

$$|f_n * g_n(x) - f * g(x)| \leq \|f_n\|_p \|g_n - g\|_q + \|g\|_q \|f_n - f\|_p$$

Cette inégalité nous prouve que $f_n * g_n$ converge uniformément vers $f * g$.

Par conséquent $f * g \in \mathcal{C}_0$. \square

Théorème : Convolution $L^p \times L^q$, cas général

Soient $(p; q) \in [1; +\infty]^2$ tels que $\frac{1}{p} + \frac{1}{q} > 1$.

Soient f et g deux fonctions respectivement L^p et L^q .

On pose r le nombre tel que $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$.

Alors :

- $r \geq \max(p; q) \geq 1$;
- $f * g$ existe pour presque tout x ;
- $f * g$ est de classe L^r .

De plus :

$$\|f * g\|_r \leq \|f\|_p \times \|g\|_q$$

Démonstration. Notons que $\frac{1}{q} \leq 1$ et $\frac{1}{p} \leq 1$. Ainsi :

$0 < \frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1 \leq \frac{1}{p}$ et $\frac{1}{r} \leq \frac{1}{q}$, ce qui permet d'en déduire la première inégalité.

On va supposer que f et g sont à valeurs positives, sans nuire à la généralité du problème, en raison du théorème de Fubini et de l'inégalité triangulaire.

Remarquons que $\frac{1}{p} - \frac{1}{r} + \frac{1}{q} - \frac{1}{r} + \frac{1}{r} = 1$ et comme $\frac{1}{p} - \frac{1}{r} \geq 0$ et $\frac{1}{q} - \frac{1}{r} \geq 0$, cela nous place dans les conditions d'application de l'inégalité de Hölder généralisée.

On pose ainsi, $\frac{1}{\alpha} = \frac{1}{p} - \frac{1}{r} \iff \alpha = \frac{pr}{r-p}$, $\frac{1}{\beta} = \frac{1}{q} - \frac{1}{r} \iff \beta = \frac{qr}{r-q}$, de sorte que $\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{r} = 1$.

Ainsi, pour tout x ,

$$\begin{aligned} f * g(x) &= \int f(t)g(x-t) d\lambda(t) \\ &= \int f(t)^{1-p/r} g(x-t)^{1-q/r} f(t)^{p/r} g(x-t)^{q/r} d\lambda(t) \\ &= \int f(t)^{1-p/r} g(x-t)^{1-q/r} f(t)^{p/r} g(x-t)^{q/r} d\lambda(t) \\ &\leq \left(\int f(t)^{\alpha \times (r-p)/r} d\lambda(t) \right)^{1/\alpha} \times \left(\int g(x-t)^{\beta \times (r-q)/r} d\lambda(t) \right)^{1/\beta} \times \left(\int f(t)^{r \times p/r} g(x-t)^{r \times q/r} d\lambda(t) \right)^{1/r} \\ &\leq \|f\|_p^{p/\alpha} \times \|g\|_q^{q/\beta} \times (f^p * g^q(x))^{1/r} \end{aligned}$$

Or, d'après ce qui précède, nous savons que $f^p * g^q$ existe et est L^1 car f^p et g^q le sont. Cela nous prouve que $f * g(x)$ existe pour presque tout x .

D'autre part, on en déduit :

$$(f * g(x))^r \leq \|f\|_p^{pr/\alpha} \times \|g\|_q^{qr/\beta} f^p * g^q(x)$$

Mais on sait que $\frac{pr}{\alpha} = r-p$ et $\frac{qr}{\beta} = r-q$

Ainsi,

$$\int (f * g(x))^r d\lambda(x) \leq \|f\|_p^{r-p} \times \|g\|_q^{r-q} \times \int f^p * g^q(x) d\lambda(x) \leq \|f\|_p^r \times \|f\|_q^r$$

Finalement, on a bien :

$$\|f * g\|_r \leq \|f\|_p \times \|f\|_q$$

□

7.2 Transformée de Fourier : définitions et premières propriétés

Dans toute la suite, on munit \mathbb{R} de la tribu des Boréliens et de la mesure de Lebesgue.

7.2.1 Définition générale

Définition : Transformée de Fourier d'une mesure, d'une fonction

Soit f une fonction de \mathbb{R} dans \mathbb{R} de classe L^1 et soit μ une mesure finie sur \mathbb{R} .

La transformée de Fourier de f est la fonction :

$$\widehat{f} : \omega \mapsto \int e^{i\omega t} f(t) d\lambda(t)$$

La transformée de Fourier de μ est la fonction :

$$\widehat{\mu} : \omega \mapsto \int e^{i\omega t} d\mu(t)$$

Démonstration. Ces deux fonctions sont bien définies. En effet, pour tout t et pour tout ω , $|e^{i\omega t} f(t)| \leq |f(t)|$ qui est λ -intégrable et d'autre part $|e^{i\omega t}| \leq 1$ qui est μ -intégrable. \square

7.2.2 Transformée de Fourier d'une fonction : propriétés

Proposition : Premières propriétés de la transformée de Fourier

On peut définir une application

$$\begin{aligned} \phi: L^1(\mathbb{R}, \mathbb{R}) &\rightarrow \mathcal{C}_b(\mathbb{R}, \mathbb{C}) \\ f &\mapsto \widehat{f} \end{aligned}$$

De plus, ϕ est linéaire et continue.

Démonstration. Il faut commencer par prouver que \widehat{f} et $\widehat{\mu}$ sont continues. Or, pour tout ω_1 et ω_2 ,

$$\begin{aligned} |\widehat{f}(\omega_1) - \widehat{f}(\omega_2)| &= \left| \int e^{i\omega_1 t} f(t) d\lambda(t) - \int e^{i\omega_2 t} f(t) d\lambda(t) \right| \\ &\leq \int |e^{i\omega_1 t} - e^{i\omega_2 t}| |f(t)| d\lambda(t) \end{aligned}$$

Soit $A > 0$ tel que $\int_{\mathbb{R}[-A; A]} |f| \leq \frac{\varepsilon}{3}$.

En raison de la convergence croissante de la suite $\mathbb{1}_{[-n; n]} |f|$ vers $|f|$, un tel nombre A existe.

En particulier, on en déduit :

$$|\widehat{f}(\omega_1) - \widehat{f}(\omega_2)| \leq \frac{2\varepsilon}{3} + \int_{[-A; A]} |e^{i\omega_1 t} - e^{i\omega_2 t}| |f(t)| d\lambda(t)$$

Or, pour tout t :

$$|e^{i\omega_1 t} - e^{i\omega_2 t}| \leq |\omega_1 t - \omega_2 t|$$

En particulier, pour tout $t \in [-A; A]$:

$$|e^{i\omega_1 t} - e^{i\omega_2 t}| \leq A |\omega_1 - \omega_2|$$

Ainsi :

$$|\widehat{f}(\omega_1) - \widehat{f}(\omega_2)| \leq \frac{2\varepsilon}{3} + A |\omega_1 - \omega_2| \|f\|_1$$

On en déduit, pour $\|f\|_1 \neq 0$ pour $|\omega_1 - \omega_2| < \frac{\varepsilon}{3A \|f\|_1}$, on obtient :

$$|\widehat{f}(\omega_1) - \widehat{f}(\omega_2)| < \varepsilon$$

Et si $\|f\|_1 = 0$, la majoration est triviale puisque dans ce cas $\widehat{f} = 0$.

Remarquons que, pour tout ω , $|\widehat{f}(\omega)| \leq \|f\|_1$. On en déduit que \widehat{f} est bien bornée. L'inégalité précédente montre en particulier que

$$\|\widehat{f}\|_\infty \leq \|f\|_1$$

La linéarité de la transformée de Fourier est simple à prouver.

Reste à prouver que ϕ est bien continue.

Mais nous venons de noter que, pour tout fonction f , $\|\widehat{f}\|_\infty \leq \|f\|_1$, ce qui prouve que ϕ est bornée.

Comme elle est linéaire, on en déduit qu'elle est continue (et même 1 lipschitzienne pour être précis). \square

Lemme : \widehat{f} appartient à \mathcal{C}_0 dans le cas où f est une indicatrice d'intervalle.

Soit $[a; b]$ un intervalle fermé borné. Alors $\widehat{\mathbb{1}_{[a; b]}}$ appartient à \mathcal{C}_0 .

Démonstration. On sait déjà que $\widehat{\mathbb{1}_{[a; b]}}$ est continue. Il suffit de prouver par exemple que la partie réelle de la transformée de Fourier tend vers 0 en $+\infty$. La partie imaginaire et le lieu $-\infty$ se traitant de manière similaire.

$$\lim_{\lambda \rightarrow +\infty} \int_{[a; b]} \cos(\lambda t) dt = 0$$

Mais cela est facile.

En effet, la fonction $t \mapsto \cos(\lambda t)$ est périodique de période $\frac{2\pi}{\lambda}$ et son intégrale est nulle sur une période.

Sur l'intervalle $[a; b]$, il y a $\left\lfloor \frac{\lambda(b-a)}{2\pi} \right\rfloor$ périodes.

Notons b_λ le nombre $a + \left\lfloor \frac{\lambda(b-a)}{2\pi} \right\rfloor \times \frac{2\pi}{\lambda} \leq b$.

On a donc :

$$\left| \int_{[a; b]} \cos(\lambda t) dt \right| = \left| \int_{[b_\lambda; b]} \cos(\lambda t) dt \right| \leq (b - b_\lambda)$$

Or, $b - b_\lambda = (b-a) - \left\lfloor \frac{\lambda(b-a)}{2\pi} \right\rfloor \times \frac{2\pi}{\lambda}$ et on montre facilement par encadrement que $\lim_{\lambda \rightarrow +\infty} \left\lfloor \frac{\lambda(b-a)}{2\pi} \right\rfloor \times \frac{2\pi}{\lambda} = (b-a)$.

On obtient donc la limite escomptée. \square

Théorème : \widehat{f} appartient à \mathcal{C}_0

Si f est L^1 alors \widehat{f} appartient à \mathcal{C}_0 .

Démonstration. On exploite le lemme qui précède. Ainsi les transformées de Fourier de combinaisons linéaires d'indicatrices de compact sont de classe \mathcal{C}_0 . Or ces fonctions sont denses dans L^1 .

Soit ainsi f une fonction de L^1 et f_n une suite de combinaison linéaires d'indicatrices de compact qui converge vers f dans L^1 .

D'après la continuité de la transformée de Fourier, on en déduit que $\lim \widehat{f_n} = \widehat{f} \in \mathcal{C}_0$. \square

Proposition : Transformée de Fourier et produit de convolution

Soient deux fonctions f et g appartenant à L^1 .

Alors $\widehat{f * g}$ existe et vérifie

$$\widehat{f * g} = \widehat{f} \times \widehat{g}$$

Démonstration. $f * g$ est dans L^1 puisque f et g sont dans L^1 . La transformée de Fourier de $f * g$ existe donc.

D'autre part, les transformées de Fourier de f et g existent puisque ces deux fonctions sont également dans L^1 .

Calculons la transformée de Fourier de $f * g$

$$\begin{aligned} \widehat{f * g}(\omega) &= \int e^{i\omega x} d\lambda(x) \int d\lambda(t) f(t)g(x-t) \\ &= \int \int e^{i\omega x} f(t)g(x-t) d\lambda(t)d\lambda(x) \\ &= \int \int e^{i\omega(x-t+t)} f(t)g(x-t) d\lambda(t)d\lambda(x) \\ &= \int \int e^{i\omega(x-t)} e^{i\omega t} f(t)g(x-t) d\lambda(t)d\lambda(x) \end{aligned}$$

Faisons maintenant un changement de variable dans \mathbb{R}^2 :

$(t; x) \mapsto (t; u)$ avec $u = x - t$.

On obtient ainsi, par application du théorème de Fubini-Tonelli :

$$\begin{aligned}\widehat{f * g}(\omega) &= \int \int e^{i\omega u} e^{i\omega t} f(t) g(u) d\lambda(t) d\lambda(u) \\ &= \left(\int e^{i\omega u} g(u) d\lambda(u) \right) \left(\int e^{i\omega t} f(t) d\lambda(t) \right) \\ &= \widehat{f}(\omega) \widehat{g}(\omega)\end{aligned}$$

□

Proposition : Transformée de Fourier et dérivation

Soit une fonction $f \in L^1$ et n fois dérivable et dont les dérivées appartiennent à L^1 .

Alors, pour tout $0 \leq k \leq n$ on vérifie

$$\widehat{f^{(k)}} : \omega \mapsto (-i\omega)^k \widehat{f}(\omega)$$

Démonstration. On fait une récurrence sur n et k assez directe...

On calcule

$$\widehat{f'}(\omega) = \int e^{i\omega x} f'(x) d\lambda(x)$$

Ici, on va utiliser des techniques d'intégration par partie. On pose :

$$\begin{cases} u' = f' \\ v = e^{i\omega x} \end{cases}$$

On obtient

$$\begin{cases} u = f \\ v' = i\omega e^{i\omega x} \end{cases}$$

Et ainsi :

$$\widehat{f'}(\omega) = \lim_{x \rightarrow +\infty} [e^{i\omega x} f(x) - e^{-i\omega x} f(-x)] - i\omega \int e^{i\omega x} f(x) d\lambda(x)$$

Or $\lim_{x \rightarrow +\infty} [e^{i\omega x} f(x) - e^{-i\omega x} f(-x)] = 0$ car f' est L^1 .

On obtient donc bien :

$$\widehat{f'}(\omega) = -i\omega \widehat{f}(\omega)$$

□

Proposition : Dérivation de la transformée de Fourier

Soit une fonction f de classe L^1 . On suppose que, pour tout $1 \leq k \leq n$, $x \mapsto x^k f(x)$ est de classe L^1 .

Alors la transformée de Fourier de f est n fois dérivable et on a

$$\frac{\partial^k \widehat{f}(\omega)}{\partial \omega^k} = \int (ix)^k f(x) e^{i\omega x} d\lambda(x)$$

Démonstration. C'est une application directe du théorème de convergence dominée.

□

Une conséquence de ce dernier résultat concerne les transformées de Fourier de noyaux gaussiens.

Proposition : Transformée de Fourier d'un noyau Gaussien

Soit la fonction $g : x \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Alors, \widehat{g} est infiniment dérivable et vérifie l'équation différentielle :

$$\widehat{g}'(\omega) + \omega \widehat{g}(\omega) = 0$$

avec la condition initiale :

$$\widehat{g}(0) = 1$$

Et ainsi, on a :

$$\widehat{g}(\omega) = e^{-\omega^2/2}$$

Démonstration. En raison des croissances comparées, pour tout $k \in \mathbb{N}$, $x \mapsto x^k g(x)$ est continue et tend vers 0 en $+\infty$ et $-\infty$.

La règle de comparaison de Riemann nous permet d'en déduire que ces fonctions sont intégrables.

En particulier, d'après ce qui précède, $\widehat{g} \in \infty_0$.

D'autre part :

$$\begin{aligned} \widehat{g}'(\omega) &= \frac{1}{\sqrt{2\pi}} \int i x e^{i\omega x - x^2/2} \\ &= \frac{i}{\sqrt{2\pi}} \left[\int (x - i\omega) e^{i\omega x - x^2/2} d\lambda(x) + \int i\omega e^{i\omega x - x^2/2} d\lambda(x) \right] \\ &= \frac{i}{\sqrt{2\pi}} \left[\left[-e^{i\omega x - x^2/2} \right]_{-\infty}^{+\infty} + i\omega \int e^{i\omega x - x^2/2} d\lambda(x) \right] \\ &= 0 - \frac{\omega}{\sqrt{2\pi}} \int e^{i\omega x - x^2/2} d\lambda(x) \\ &= -\omega \widehat{g}(\omega) \end{aligned}$$

Et on a également $\widehat{g}(0) = \int g = 1$.

La seule solution de cette équation différentielle est la fonction

$$\widehat{g} : \omega \mapsto e^{-\omega^2/2}$$

□

Chapitre 8

Espace de Schwartz, transformée de Fourier et mesures

8.1 Généralités sur l'espace de Schwartz

8.1.1 Notations

On se place dans \mathbb{R}^d que l'on munit de la mesure $\tilde{\lambda}_d = \frac{1}{(2\pi)^{n/2}} \lambda_d$ où λ_d est la mesure de Lebesgue. Grâce à cette manipulation, le noyau gaussien noté $g : x \mapsto e^{-\|x\|^2/2}$ vérifie $\|g\|_1 = 1$.

Pour tout $t \in \mathbb{R}^d$, on pose $e_t : x \mapsto e^{ix \cdot t}$ où « \cdot » désigne le produit scalaire « classique » sur \mathbb{R}^d .

Pour tout $x = (x_1; x_2; \dots; x_d) \in \mathbb{R}^d$ et pour tout multi-indice $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_d) \in \mathbb{N}^d$, on note $x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}$.

Concernant la dérivation, pour tout multi-indice $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_d) \in \mathbb{N}^d$, on note $D^\alpha = \frac{\partial^{\alpha_1} \partial^{\alpha_2} \dots \partial^{\alpha_d}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_d^{\alpha_d}}$.

Et on notera aussi $D_\alpha = \frac{1}{(i)^{|\alpha|}} D^\alpha$ où $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_d$.

Ces notations seront utiles lorsque nous établirons un lien entre polynômes, dérivation et transformée de Fourier.

Enfin, si P est un polynôme de $\mathbb{C}[X_1, X_2, \dots, X_d]$ dont les coefficients sont $(c_\alpha)_{\alpha \in \mathbb{N}^d}$, on notera $P(D) = \sum_{\alpha \in \mathbb{N}^d} c_\alpha D_\alpha$.

8.1.2 Définition : l'espace de Schwartz

Définition : Espace de Schwartz

On dit qu'une fonction $f : \mathbb{R}^d \rightarrow \mathbb{C}$ appartient à l'espace de Schwartz, noté \mathcal{S} , lorsque :

- $f \in \mathcal{C}^\infty$;
- $\forall (\alpha; p) \in \mathbb{N}^d \times \mathbb{N}, \sup_{x \in \mathbb{R}^d} (1 + \|x\|^2)^p \|D^\alpha f(x)\| < +\infty$.

Dans ce cas, pour tout $(\alpha; p) \in \mathbb{N}^d \times \mathbb{N}$, on notera $N_{\alpha; p}(f) = \sup_{x \in \mathbb{R}^d} (1 + \|x\|^2)^p \|D^\alpha f(x)\|$.

Notons que la famille $N_{\alpha; p}$ est une famille dénombrable et séparatrice de semi-normes puisque, pour $\alpha = 0$, il s'agit en fait d'une norme. On travaillera donc avec la topologie induite par cette famille, telle qu'elle a été définie dans le document *normes*.



On rappelle qu'une base d'ouverts de cette topologie est formée des intersections finies de boules $V_{N_{\alpha; p}; q_{\alpha; p}} = \left\{ f \in \mathcal{S} / N_{\alpha; p}(f) < \frac{1}{q_{\alpha; p}} \right\}$, avec $(\alpha; p) \in \mathbb{N}^d \times \mathbb{N}$ et $q_{\alpha; p} \in \mathbb{N}^*$.

(Remarque 0) Notons que $\mathcal{S} \subset \mathcal{C}_0$, l'ensemble des fonctions continues qui tendent vers 0 en $+\infty$. En effet, toute fonction $f \in \mathcal{S}$ est dominée par $x \mapsto \frac{1}{1 + \|x\|^2}$.

On rappelle au passage que $\mathcal{C}_0 \subset \mathcal{C}_b$, les fonctions continues et bornées.

(Remarque 1) Remarquons aussi que, pour deux entiers $p \geq q$, et pour tout $x \in \mathbb{R}^d$, on a $(1 + \|x\|^2)^p \geq (1 + \|x\|^2)^q$, ce qui entraîne, pour tout multi-indice α , $N_\alpha; p \geq N_\alpha; q$.

(Remarque 2) On appelle aussi *fonctions à décroissance rapide* cet ensemble.

Lemme : Majoration grossière pour un polynôme

Pour tout $x \in \mathbb{R}^d$ et pour tout $\alpha \in \mathbb{N}^d$:

$$|x^\alpha| = \prod_{i=1}^d |x_i|^{\alpha_i} \leq (1 + \|x\|^2)^{|\alpha|}$$

On en déduit que, si P est un polynôme de degré q , il existe un réel positif M_P , tel que, pour tout $x \in \mathbb{R}^d$:

$$|P(x)| \leq M_P(1 + \|x\|^2)^q$$

Démonstration. La première majoration n'est pas difficile à condition de distinguer deux cas.

- Si $\max_i |x_i| < 1$, dans ce cas $|x^\alpha| < 1$ et donc la majoration est triviale.
- Si $\max_i |x_i| \geq 1$, dans ce cas, $\|x\| \geq 1$ et, pour tout i , $|x_i| \leq \|x\| \leq (1 + \|x\|^2)^{1/2}$. On en déduit que $|x^\alpha| \leq (1 + \|x\|^2)^{|\alpha|}$

Pour la seconde majoration, remarquons que pour tout multi-indice β tel que $|\beta| \leq q$, $|x^\beta| \leq (1 + \|x\|^2)^{|\beta|} \leq (1 + \|x\|^2)^q$. Ainsi, en posant $K = \max_\beta |c_\beta|$ avec c_β les coefficients de P , on a la majoration grossière suivante :

$$|P(x)| \leq \sum_{\beta \in \mathbb{N}^d} |c_\beta| |x^\beta| \leq (1 + q)^d K (1 + \|x\|)^q$$

En effet, $(1 + q)^d$ est une majoration grossière du nombre de coefficients non nuls. □

Cela nous permet d'en déduire facilement la propriété suivante.

Proposition : Caractérisation équivalente

Pour toute fonction f , on a :

$f \in \mathcal{S}$ si et seulement si f est de classe \mathcal{C}^∞ et, pour tout $\alpha \in \mathbb{N}^d$ et pour tout polynôme P , $PD^\alpha f$ est borné.

Un second lemme qui nous sera utile pour la suite.

Lemme : Formule de Liebnitz en dimension finie

Soit $\alpha \in \mathbb{N}^d$ et soient f et h deux fonctions définies sur \mathbb{R}^d et à valeurs dans \mathbb{C} .

On suppose que $D^\alpha f$ et $D^\alpha h$ existent.

Alors $D^\alpha(fh)$ existe. Plus précisément il existe des coefficients entiers naturels $(c_\beta, \gamma)_{\beta+\gamma=\alpha}$ tels que :

$$D^\alpha(fh) = \sum_{\beta+\gamma=\alpha} c_\beta, \gamma D^\beta f D^\gamma h$$

Démonstration. Ce lemme se montre sans trop de difficulté par récurrence sur d , la dimension de l'espace de départ. Je ne le fais pas car les notations multi-indices sont lourdingues. L'idée est simple : appliquer $\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}$ à $\frac{\partial^{\alpha_2} \dots \partial^{\alpha_d}}{\partial x_2^{\alpha_2} \dots \partial x_d^{\alpha_d}}(fh)$ en appliquant l'hypothèse de récurrence à cette dernière expression. \square

Proposition : Propriétés des fonctions de Schwartz

\mathcal{S} forme une \mathbb{C} -algèbre. Muni de la topologie évoquée plus haut, \mathcal{S} est un espace complet.

En outre, pour $h \in \mathcal{S}$, pour P un polynôme et pour $\alpha \in \mathbb{N}^d$, les applications linéaires suivantes sont des endomorphismes continus :

$$f \mapsto hf \qquad f \mapsto Pf \qquad f \mapsto D^\alpha f$$

Démonstration. On sait que \mathcal{C}^∞ est une algèbre. Il nous faut donc montrer les propriétés associées aux semi-normes.

Il est clair que si $(f; h) \in \mathcal{S}^2$ et si λ et μ sont des complexes alors, pour tout $(\alpha; p) \in \mathbb{N}^d \times \mathbb{N}$, $N_{\alpha; p}(\lambda f + \mu h) \leq |\lambda| N_{\alpha; p}(f) + |\mu| N_{\alpha; p}(h) < +\infty$ donc \mathcal{S} est bien un espace vectoriel.

De plus, sous les mêmes hypothèses, en reprenant les notations de la formule de Liebnitz énoncée plus haut, on a :

$$N_{\alpha; p}(fh) \leq \sum_{\beta+\gamma=\alpha} c_{\beta, \gamma} N_{\beta; p}(f) N_{\gamma; p}(h) < +\infty$$

Cela confirme que \mathcal{S} est une algèbre.

Montrons maintenant que \mathcal{S} est complet. Considérons une suite de Cauchy (f_n) d'éléments de \mathcal{S} . Pour tout multi-indice α , $D^\alpha f_n$ est de Cauchy pour la norme infinie (celle de la convergence uniforme).

En particulier, on en déduit que (f_n) converge uniformément vers une fonction f et que $(D^\alpha f_n)$ converge également uniformément vers $D^\alpha f$. Et cela prouve au passage que cette limite f est \mathcal{C}^∞ . Il ne nous reste plus qu'à prouver :

- $f \in \mathcal{S}$;
- (f_n) converge vers f pour la topologie de \mathcal{S} .

Soit ainsi $p \in \mathbb{N}$ quelconque et soit $\varepsilon > 0$. Il existe un rang N tel que, pour tout $q \in \mathbb{N}$ et pour tout $n \geq N$, on a $N_{\alpha; p}(f_{n+q} - f_n) < \varepsilon$. En particulier, pour $q \rightarrow +\infty$, on obtient $N_{\alpha; p}(f - f_n) \leq \varepsilon$, ce qui achève la démonstration. En effet, cela prouve d'une part que $N_{\alpha; p}(f) < +\infty$ et d'autre part la convergence pour la topologie de \mathcal{S} , quitte à éventuellement réitérer la manipulation pour plusieurs valeurs de α et p et à considérer le maximum des rangs obtenus.

D'après les deux lemmes plus haut, il est clair que les trois applications considérées sont des endomorphismes.

La continuité concernant la dérivation est triviale. En effet, pour tout multi-indices α et β et, pour tout $p \in \mathbb{N}$, on a $N_{\beta; p}(D^\alpha f) = N_{\beta+\alpha; p}(f)$. La convergence des (f_n) vers (f) dans \mathcal{S} entraîne donc la convergence des $(D^\alpha f_n)$ vers $D^\alpha f$.

Prouvons maintenant la continuité des deux autres applications considérées. Soit (f_n) une suite de fonctions qui converge vers f dans \mathcal{S} . Pour tout $(\alpha; p) \in \mathbb{N}^d \times \mathbb{N}$, on a, en reprenant les notations de la formule de Liebnitz :

$$N_{\alpha; p}(h(f_n - f)) \leq \sum_{\beta+\gamma=\alpha} |c_{\beta, \gamma}| N_{\beta; p}(f_n - f) N_{\gamma; 0}(h)$$

On en déduit en particulier que $\lim_n N_{\alpha; p}(h(f_n - f)) = 0$, ce qui prouve que $f \mapsto hf$ est continue.

La continuité de $f \mapsto Pf$ est un peu plus pénible. On suppose que P est de degré q . Pour tout $x \in \mathbb{R}^d$:

$$|(1 + \|x\|^2)^p D^\alpha [Pf](x)| \leq (1 + \|x\|^2)^p \sum_{\beta+\gamma=\alpha} |c_{\beta, \gamma}| |D^\beta [f](x)| |D^\gamma [P](x)|$$

Or $D^\gamma[P]$ est un polynôme de degré inférieur ou égal à q , ce qui permet, d'après le lemme de majoration grossière, l'existence d'un réel positif M_γ tel que, pour tout x , $|D^\gamma[P](x)| \leq M_\gamma(1 + \|x\|^2)^q$. On obtient donc :

$$|(1 + \|x\|^2)^p D^\alpha[Pf](x)| \leq (1 + \|x\|^2)^{p+q} \sum_{\beta+\gamma=\alpha} M_\gamma |c_{\beta, \gamma}| |D^\beta[f](x)|$$

De cette inégalité, on déduit :

$$N_{\alpha; p}(Pf) \leq \sum_{\beta+\gamma=\alpha} M_\gamma |c_{\beta, \gamma}| N_{\beta; p+q}(f)$$

Ce qui permet de conclure. Pour toute suite (f_n) qui converge vers f dans \mathcal{S} , alors (Pf_n) converge vers Pf dans \mathcal{S} . \square

Énonçons quelques propriétés simples sur l'intégrabilité de fonctions de \mathcal{S} .

Proposition : \mathcal{S} et l'intégration

Soit $f \in \mathcal{S}$. Alors, pour tout $p \in [1; +\infty]$, et pour tout multi-indice α , $D^\alpha f \in L^p$.

En outre, \mathcal{S} est dense dans L^p pour $p \in [1; +\infty[$.

Démonstration. La première affirmation est facile car $D^\alpha f$ est continue (donc mesurable) et dominée par $\frac{1}{1+\|x\|^2}$ qui est L^p .

Remarquons pour la seconde affirmation que les fonctions \mathcal{C}^∞ à support compact sont incluses dans \mathcal{S} . Or ces premières fonctions sont denses dans L^p . \square

8.2 Transformée de Fourier dans \mathcal{S}

8.2.1 Définition et premières propriétés

On s'appuie ici sur les notations établies en introduction.

Définition : Transformée de Fourier

Soit $f \in \mathcal{S}$.

On pose $\widehat{f}: t \mapsto \int_{\mathbb{R}^d} e^{-it \cdot x} f(x) d\tilde{\lambda}_d(x) = \int_{\mathbb{R}^d} e_t(x) f(x) d\tilde{\lambda}_d(x)$.

\widehat{f} est une application de \mathbb{R}^d vers \mathbb{C} .

On sait que $\mathcal{S} \subset L^1$ donc cette application est bien définie. Quelques rappels concernant la transformée de Fourier dans L^1 .

Proposition : Transformée de Fourier de fonction L^1

Pour tout $f \in L^1$, $\widehat{f} \in \mathcal{C}_0$. En outre, l'application suivante est un morphisme continu :

$$\begin{array}{ccc} \psi: & L^1 & \rightarrow \mathcal{C}_0 \\ & f & \mapsto \widehat{f} \end{array}$$

Démonstration. Nous l'avions prouvé dans le document *mesure*. \square

1. Nous l'avions montré dans le document *mesure*

Proposition : Transformée de Fourier du noyau Gaussien

On rappelle que $g : t \mapsto e^{-\|t\|^2/2}$ est le noyau Gaussien. Alors :

$$\widehat{g} = g$$

Démonstration. Fait dans le document *mesure*, de deux manières différentes. □

Définition : Translation et dilatation

Pour tout $x \in \mathbb{R}^d$ et pour tout fonction f définie sur \mathbb{R}^d , on pose $\tau_x[f]$ la fonction :

$$\tau_x[f] : t \mapsto f(t - x)$$

Pour tout $\alpha > 0$, on pose $\delta_\alpha[f]$ la fonction :

$$\delta_\alpha[f] : t \mapsto f\left(\frac{t}{\alpha}\right)$$

Proposition : Premières propriétés de la transformée de Fourier

Soit P un polynôme, $x \in \mathbb{R}^d$ et f et h deux fonctions de \mathcal{S} .

Alors :

$$\begin{aligned} \widehat{\tau_x[f]} &= e_{-x} \widehat{f} & \widehat{\delta_\alpha[f]} &= \alpha^n \delta_{\frac{1}{\alpha}} [\widehat{f}] & \widehat{P(D)[f]} &= P \widehat{f} \\ \widehat{e_x f} &= \tau_x [\widehat{f}] & \widehat{f * h} &= \widehat{f} \widehat{h} & \widehat{P f} &= P(-D) [\widehat{f}] \end{aligned}$$

Enfin, l'application $\psi : f \mapsto \widehat{f}$ est un endomorphisme continu sur \mathcal{S} .

Démonstration. Dans toute cette démonstration $t \in \mathbb{R}^d$ est quelconque.

Petit changement de variable pour la première.

$$\begin{aligned} \widehat{\tau_x[f]}(t) &= \int e^{-it \cdot y} f(y - x) d\tilde{\lambda}_d(y) = \int e^{-it \cdot (z+x)} f(z) d\tilde{\lambda}_d(z) \\ &= e^{-it \cdot x} \int e^{-it \cdot z} f(z) d\tilde{\lambda}_d(z) = e_{-x}(t) \widehat{f}(t) \end{aligned}$$

Inutile pour montrer la seconde. Cela s'obtient directement :

$$\widehat{e_x f}(t) = \int e^{-it \cdot y} e^{ix \cdot y} f(y) d\tilde{\lambda}_d(y) = \int e^{-i(t-x) \cdot y} f(y) d\tilde{\lambda}_d(y) = \tau_x [\widehat{f}](t)$$

Encore un changement de variable pour la troisième :

$$\widehat{\delta_\alpha[f]}(t) = \int e^{-it \cdot y} f\left(\frac{y}{\alpha}\right) d\tilde{\lambda}_d(y) = \alpha^n \int e^{-i\alpha t \cdot z} f(z) d\tilde{\lambda}_d(y) = \alpha^n \delta_{\frac{1}{\alpha}} [\widehat{f}](t)$$

Pour la propriété portant sur le produit de convolution, il faut réaliser un changement de variable puis appliquer le théorème de Fubini-Tonelli :

$$\begin{aligned} \widehat{f * h}(t) &= \int e^{-it \cdot y} \int f(z) h(y - z) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) = \int \int e^{-it \cdot y} e^{-it \cdot z} e^{it \cdot z} f(z) h(y - z) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) \\ &= \int \int e^{-it \cdot (y-z)} f(z) h(y - z) e^{-it \cdot z} d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) = \int \int e^{-it \cdot u} h(u) e^{-it \cdot z} f(z) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(u) \\ &= \left(\int e^{-it \cdot u} h(u) d\tilde{\lambda}_d(u) \right) \left(\int e^{-it \cdot z} f(z) d\tilde{\lambda}_d(z) \right) = \widehat{h}(t) \widehat{f}(t) \end{aligned}$$

Maintenant, concernant la propriété portant sur la dérivation, on va réaliser la transformée de Fourier de $\frac{\partial}{\partial x_1} f$, la généralisation étant très simple pour $D^\alpha f$ et enfin, par combinaison linéaire, pour $P(D)[f]$.

Ici, il s'agit de réaliser une intégration par parties, en identifiant l'intégrale de Riemann avec l'intégrale de Lebesgue (cas d'une fonction continue et intégrable). En pratique, on écrira, pour tout $y = (y_1, y_2, \dots, y_d) \in \mathbb{R}^d$, $y = (y_1, \tilde{y})$ avec $\tilde{y} \in \mathbb{R}^{d-1}$, et on utilisera la même convention pour t , de sorte que :

$$\begin{aligned} \widehat{\frac{\partial f}{\partial x_1}}(t) &= \int e^{-it \cdot y} \frac{\partial f}{\partial x_1}(y) d\tilde{\lambda}_d(y) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^{d-1}} e^{-i\tilde{t} \cdot \tilde{y}} \left(\lim_{N \rightarrow +\infty} \int_{-N}^N e^{-it_1 y_1} \frac{\partial f}{\partial x_1}(y_1, \tilde{y}) dy_1 \right) d\mu_{d-1}(\tilde{y}) \end{aligned}$$

On traite à part, avec une intégration par parties, l'intégrale portant sur y_1 :

$$\int_{-N}^N e^{-iy_1 t_1} \frac{\partial f}{\partial x_1}(y_1, \tilde{y}) dy_1 = [e^{-it_1 y_1} f(y_1, \tilde{y})]_{-N}^N + it_1 \int_{-N}^N e^{-iy_1 t_1} f(y_1, \tilde{y}) dy_1$$

On obtient ainsi, compte-tenu de la propriété de décroissance rapide de f :

$$\lim_{N \rightarrow +\infty} \int_{-N}^N e^{-it_1 y_1} \frac{\partial f}{\partial x_1}(y_1, \tilde{y}) dy_1 = it_1 \int_{\mathbb{R}} e^{-iy_1 t_1} f(y_1, \tilde{y}) dy_1$$

En réinjectant, cela donne :

$$\widehat{\frac{\partial f}{\partial x_1}}(t) = (it_1) \widehat{f}(t)$$

Compte-tenu de la définition de D_α , en réitérant ce que l'on fait pour la première variable à des variables quelconques et autant de fois que nécessaire, on obtient bien $\widehat{D_\alpha f}(t) = t^\alpha \widehat{f}(t)$ et cela se généralise avec $P(D)$, par linéarité !

La dernière égalité se montre en appliquant le théorème de convergence dominée de Lebesgue dont on vérifie aisément que les conditions d'application sont réunies.

En effet, en reprenant les mêmes notations et supposant que le polynôme considéré est $P = X_1$

$$\widehat{Pf}(t) = \int y_1 e^{-it \cdot y} f(y) d\tilde{\lambda}_d(y) = \int \frac{-1}{i} \frac{\partial}{\partial t_1} e^{-it \cdot y} f(y) d\tilde{\lambda}_d(y) = \frac{-1}{i} \frac{\partial}{\partial t_1} \int e^{-it \cdot y} f(y) d\tilde{\lambda}_d(y)$$

En réitérant l'opération sur autant de variables et autant de fois que nécessaire et en exploitant la linéarité, on retrouve le résultat escompté.

Reste maintenant à prouver que la transformée de Fourier est un endomorphisme continu ! D'après la dernière égalité prouvée, il est clair que, pour $f \in \mathcal{S}$, \widehat{f} est \mathcal{C}_b^∞ , c'est à dire l'ensemble des fonctions indéfiniment dérivables et dont toutes les dérivées sont bornées.

Montrons qu'elle aussi à décroissance rapide. Soit $\alpha \in \mathbb{N}^d$ et $p \in \mathbb{N}$. Posons $Q(x) = (1 + \|x\|^2)^p$ et $P = X^\alpha$.

D'après ce que l'on vient de montrer, $Q(D)Pf \in \mathcal{S}$ et $\widehat{Q(D)Pf} \in \mathcal{S} = (-1)^{|\alpha|} Q(D_\alpha \widehat{f})$. Or, nous savons que $\widehat{Q(D)Pf} \in \mathcal{S}$ est bornée, ce qui prouve que \widehat{f} est à décroissance rapide.

Reste enfin à prouver la continuité. Soit ainsi une suite de fonctions (f_n) qui tend vers f dans \mathcal{S} . Pour tout t , et pour tout n , on a :

$$|QD_\alpha \widehat{f_n}(t) - QD_\alpha \widehat{f}(t)| = |Q(D) \widehat{P(f_n - f)}(t)|$$

Posons $g_n = Q(D)Pf_n$ et $g = Q(D)Pf$. D'après ce qui précède, on sait que les opération de multiplication par un polynôme et de dérivation sont des endomorphismes continus de \mathcal{S} , on en déduit ainsi que g_n tend vers g dans \mathcal{S} . Il suffit donc de prouver que, pour g_n tendant vers g avec la topologie de \mathcal{S} , $\widehat{g_n - g}$ converge uniformément vers 0. Mais cela est facile, en effet :

$$|\widehat{g_n - g}(t)| \leq \int \frac{1}{1 + \|y\|^2} (1 + \|y\|^2) |g_n - g|(y) d\tilde{\lambda}_d(y) \leq KN_{0,1}(g_n - g)$$

Avec $K = \int \frac{1}{1 + \|y\|^2} d\tilde{\lambda}_d(y)$.

Finalement, la transformée de Fourier est bien un endomorphisme continu de \mathcal{S} . □

8.2.2 Théorème d'inversion

8.2.2.1 Approximation de l'unité

On rappelle la définition du noyau Gaussien, $g : t \mapsto e^{-\|t\|^2/2}$. On rappelle plus bas quelques résultats minimaux dont on aura besoin pour la suite.

Proposition : Approximation de l'unité

Pour tout n , on note $g_n : t \mapsto n^d g(nt)$.

Alors, pour toute fonction f de classe L^1 , $g_n * f$ tend vers f dans L^1 .

Démonstration. Fait dans le document *mesure*. □

(Remarque 3) Il y a beaucoup d'autres choses à dire. D'abord, $g_n * f$ est de classe \mathcal{C}^∞ . D'autre part, dans le cas où f est continue, on a aussi convergence uniforme de $g_n * f$ vers f .

8.2.2.2 Théorème d'inversion

Définition : Un nouvel opérateur

Soit f une fonction de classe L^1 . Alors, on définit :

$$\check{f} : t \mapsto \widehat{f}(-t)$$

(Remarque 4) Par composition, cet opérateur est continu et linéaire de L^1 vers \mathcal{C}_0 , munis de leurs topologies usuelles.

(Remarque 5) Cet opérateur est également un endomorphisme continu et linéaire de \mathcal{S} .

Théorème : Inversion de la transformée de Fourier

Soit f une fonction de classe L^1 telle que \widehat{f} est aussi de classe L^1 . Alors :

$$\check{\check{f}} = f$$

En particulier, la transformée de Fourier est un isomorphisme continu de \mathcal{S} dont g est un élément neutre.

Démonstration. Pour tout x et pour tout n , calculons, $\widehat{g_n * \check{f}}(x)$, sachant que cette quantité existe puisque

$\widehat{g_n * f}$ est de classe L^1 . On peut aussi exploiter le théorème de Fubini-Tonnelli :

$$\begin{aligned}
\widehat{g_n * f}(x) &= \int e^{ix \cdot y} \int e^{-iy \cdot t} \int g_n(t-z) f(z) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(t) d\tilde{\lambda}_d(y) \\
&= \int \int \int e^{-iy \cdot (t-x)} g_n(t-z) f(z) d\tilde{\lambda}_d(t) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) \\
&= \int \int \int e^{-iy \cdot (z+u-x)} g_n(u) f(z) d\tilde{\lambda}_d(u) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) \\
&= \int \int e^{-iy \cdot (z-x)} f(z) \int e^{-iy \cdot u} g_n(u) d\tilde{\lambda}_d(u) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) \text{ on a posé } u = t - z \\
&= \int \int e^{-iy \cdot (z-x)} f(z) \widehat{g_n}(y) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) \text{ on sait que } \widehat{g_n}(y) = g\left(\frac{y}{n}\right) \\
&= \int \int e^{-i(z-x)} f(z) g\left(\frac{y}{n}\right) d\tilde{\lambda}_d(z) d\tilde{\lambda}_d(y) \\
&= \int \int e^{-iy \cdot v} f(v+x) g\left(\frac{y}{n}\right) d\tilde{\lambda}_d(v) d\tilde{\lambda}_d(y) \text{ on a posé } v = z - x \\
&= \int f(v+x) \int e^{-iy \cdot v} g\left(\frac{y}{n}\right) d\tilde{\lambda}_d(y) d\tilde{\lambda}_d(v) \\
&= \int f(v+x) \widehat{\delta_n[g]}(v) d\tilde{\lambda}_d(v) \text{ on rappelle que } \delta \text{ est la dilatation} \\
&= \int f(v+x) n^d g(nv) d\tilde{\lambda}_d(v) \\
&= \int f(x-w) g_n(w) d\tilde{\lambda}_d(w) \text{ on a posé } w = -v \text{ sachant que } g \text{ est paire} \\
&= f * g_n(x)
\end{aligned}$$

Ainsi, $\widehat{g_n * f}$ converge vers \widehat{f} dans L^1 .

Il nous faut maintenant prouver que $\widehat{g_n * f}$ tend aussi vers \widehat{f} dans L^1 .

On sait que $\widehat{g_n * f} = \widehat{g_n} \widehat{f} = \delta_n[g] \widehat{f} : t \mapsto \widehat{t} g\left(\frac{t}{n}\right)$. Cette fonction est dominée par $|\widehat{f}|$ qui est de classe L^1 et on a convergence simple de $\delta_n[g] \widehat{f}$ vers \widehat{f} .

Par le théorème de convergence dominée, on obtient la convergence de $\widehat{g_n * f}$ vers \widehat{f} dans L^1 .

Exploitions maintenant la remarque faite sur la continuité de l'opérateur $f \mapsto \check{f}$ pour écrire, dans L^1 :

$$\lim_{n \rightarrow +\infty} \widehat{g_n * f} = \check{\check{f}} = f$$

Considérons maintenant $f \in \mathcal{S} \subset L^1$. Il est clair que $\widehat{f} \in L^1$ car $\widehat{f} \in \mathcal{S}$. Nous sommes donc dans les conditions d'application du résultat précédent. On sait ainsi que, pour presque tout x , $\check{\check{f}}(x) = f(x)$. Or ces deux fonctions sont dans \mathcal{S} . On en déduit que l'égalité est valable pour tout x , ce qui prouve que la transformée de Fourier est un automorphisme continu de \mathcal{S} dont la réciproque est $f \mapsto \check{f}$. \square

8.3 Étude des suites de mesures finies

Ce premier paragraphe nous donne des outils pour aborder les suites de mesures finies sous un angle fréquentiel.

8.3.1 Transformée de Fourier de mesures finies

Définition : Transformée de Fourier d'une mesure

Soit μ une mesure finie définie sur $(\mathbb{R}^d; \mathcal{B})$ avec \mathcal{B} la tribu borélienne. On pose :

$$\begin{aligned}
\widehat{\mu} : \mathbb{R} &\rightarrow \mathbb{C} \\
x &\mapsto \int e^{-ix \cdot t} d\mu(t)
\end{aligned}$$

(Remarque 6) La fonction $t \mapsto e^{it \cdot x}$ est μ -intégrable car μ est finie ;

(Remarque 7) On a, pour tout x , $|\widehat{\mu}(x)| \leq \widehat{\mu}(0) = \mu(\mathbb{R}^d)$.

Théorème : Cette transformée est injective

Soit μ et ν deux mesures finies.

On a $\mu = \nu$ si et seulement si $\widehat{\mu} = \widehat{\nu}$.

De plus, si f et g sont deux fonctions L^1 et positives, on a aussi $f = g$, λ -presque partout, si et seulement si $\widehat{f} = \widehat{g}$.

Démonstration. La seconde partie se montre facilement à partir de la première. En effet, il suffit de considérer μ et ν les mesures dont f et g sont les fonctions de densité.

Pour montrer la première partie de cette proposition, on va exploiter encore une fois les noyaux gaussiens et la densité des fonctions continues à support compact \mathcal{C}_K dans L^1 . \square

Commençons par établir un lemme.

Lemme : Convolution de mesures, et convergence

On rappelle que g est le noyau gaussien et que, pour tout n , $g_n : t \mapsto n^d g(nt)$ constitue une suite approximante de l'unité.

Pour toute mesure finie μ , on pose :

$$g_n * \mu : x \mapsto \int g_n(x-t) d\mu(t)$$

Alors, pour toute fonction f continue à support compact, on a :

$$\lim_{n \rightarrow +\infty} \int f(x) g_n * \mu(x) d\tilde{\lambda}_d(x) = \int f(x) d\mu(x)$$

Démonstration. Il est clair que ces deux intégrales existent et satisfont les conditions d'application du théorème de Fubini-Tonelli. En effet, $g_n * \mu$ est une fonction bornée car g_n est bornée et μ est finie.

Calculons, pour tout n :

$$\begin{aligned} \int f(x) g_n * \mu(x) d\tilde{\lambda}_d(x) - \int f(x) d\mu(x) &= \int \int f(x) g_n(x-t) d\mu(t) d\tilde{\lambda}_d(x) - \int f(x) \int g_n(t) d\tilde{\lambda}_d(t) d\mu(x) \\ &= \underbrace{\int \int f(u+t) g_n(u) d\tilde{\lambda}_d(u) d\mu(t)}_{\text{on pose } u=x-t} - \int f(x) \int g_n(t) d\tilde{\lambda}_d(t) d\mu(x) \\ &= \underbrace{\int \int f(t+x) g_n(t) d\tilde{\lambda}_d(t) d\mu(x)}_{\text{on pose } x=t \text{ puis } t=u} - \int f(x) \int g_n(t) d\tilde{\lambda}_d(t) d\mu(x) \\ &= \int \int (f(t+x) - f(x)) g_n(t) d\tilde{\lambda}_d(t) d\mu(x) \end{aligned}$$

On va maintenant exploiter l'uniforme continuité de f . Notons $M = \max(\mu(\mathbb{R}^d); 1)$. Soit $\varepsilon > 0$. Il existe η tel que pour tout $|t| < \eta$, $|f(t+x) - f(x)| < \frac{\varepsilon}{2M}$. À l'aide de l'inégalité triangulaire, on obtient ainsi :

$$\begin{aligned} \left| \int f(x) g_n * \mu(x) d\tilde{\lambda}_d(x) - \int f(x) d\mu(x) \right| &\leq \frac{\varepsilon}{2M} \int \int \mathbb{1}_{|t| < \eta} g_n(t) d\tilde{\lambda}_d(t) d\mu(x) + 2 \|f\|_\infty \int \int \mathbb{1}_{|t| \geq \eta} g_n(t) d\tilde{\lambda}_d(t) d\mu(x) \\ &\leq \frac{\varepsilon}{2} + 2 \|f\|_\infty \int \int \mathbb{1}_{|t| \geq \eta} g_n(t) d\tilde{\lambda}_d(t) d\mu(x) \end{aligned}$$

On sait que $\lim_{n \rightarrow +\infty} \int \mathbb{1}_{|t| \geq \eta} g_n(t) d\tilde{\lambda}_d(t) = 0$ donc il existe un rang N tel que, pour tout $n \geq N$,
 $\int \int \mathbb{1}_{|t| \geq \eta} g_n(t) d\tilde{\lambda}_d(t) d\mu(x) < \frac{\varepsilon}{2 \max(2 \|f\|_\infty \mu(\mathbb{R}^d); 1)}$. Ce qui donne :

$$\left| \int f(x) g_n * \mu(x) d\tilde{\lambda}_d(x) - \int f(x) d\mu(x) \right| \leq \varepsilon$$

□

Ce second lemme nous permettra de démontrer le théorème sur l'injectivité.

Lemme :

On reprend les notations précédentes. Alors, pour tout n :

$$\widehat{g_n * \mu} = g_n * \widehat{\mu}$$

En définissant l'opérateur suivant, pour tout h de L^1 par $\check{h} : x \mapsto \int e^{ix \cdot t} h(t) d\tilde{\lambda}_d(t)$.

(Remarque 8) On montre assez facilement que $\widehat{g_n * \mu} = \widehat{g_n} \widehat{\mu}$. Mais comme $\widehat{\mu}$ est bornée et $\widehat{g_n} \in \mathcal{S}$, cela nous prouve que $\widehat{g_n * \mu}$ est intégrable.

Montrons maintenant ce second lemme.

Démonstration. Pour tout y , avec les précautions d'usage pour l'application du théorème de Fubini-Tonelli

$$\begin{aligned} \widehat{g_n * \mu}(y) &= \int e^{iy \cdot x} \int e^{-ix \cdot t} \int g_n(t - u) d\mu(u) d\tilde{\lambda}_d(t) d\tilde{\lambda}_d(x) \\ &= \int e^{iy \cdot x} \int e^{-ix \cdot t} \int g_n(t - u) d\mu(u) d\tilde{\lambda}_d(t) d\tilde{\lambda}_d(x) \\ &= \int \int \int e^{ix \cdot (y - t)} g_n(t - u) d\tilde{\lambda}_d(t) d\tilde{\lambda}_d(x) d\mu(u) \end{aligned}$$

Calculons séparément l'intégrale centrale :

$$\begin{aligned} \int e^{ix \cdot (y - t)} g_n(t - u) d\tilde{\lambda}_d(t) &= \int e^{ix \cdot (y - u - v)} g_n(v) d\tilde{\lambda}_d(v) \\ &= e^{ix \cdot (y - u)} \int e^{-ix \cdot v} g_n(v) d\tilde{\lambda}_d(v) \\ &= e^{ix \cdot (y - u)} \widehat{g_n}(x) = e^{ix \cdot (y - u)} g\left(\frac{x}{n}\right) \end{aligned}$$

Ainsi :

$$\begin{aligned} \widehat{g_n * \mu}(y) &= \int \int e^{ix \cdot (y - u)} g\left(\frac{x}{n}\right) d\tilde{\lambda}_d(x) d\mu(u) \\ &= \int \int e^{-ix \cdot (y - u)} g\left(\frac{x}{n}\right) d\tilde{\lambda}_d(x) d\mu(u) \text{ car } g \text{ est paire} \\ &= \int \widehat{\delta_n[g]}(y - u) d\mu(u) \text{ avec } \delta \text{ l'opérateur de dilatation} \\ &= \int n^d g(n(y - u)) d\mu(u) = g_n * \mu(y) \end{aligned}$$

□

Voici maintenant la preuve du théorème de l'injectivité de la transformée de Fourier.

Démonstration. On suppose ainsi que deux mesures finies μ et ν vérifient $\widehat{\mu} = \widehat{\nu}$. Dans ce cas, on a, pour tout n :

$$\widehat{g_n \mu} = \widehat{g_n \nu} \text{ donc } \widehat{g_n * \mu} = \widehat{g_n * \nu}$$

Ainsi, d'après le second lemme, $g_n * \mu = g_n * \nu$. On en déduit, à l'aide du premier lemme et par passage à la limite que, pour toute fonction f continue à support compact :

$$\int f d\mu = \int f d\nu$$

On peut, pour tout compact K fabriquer une suite de fonctions (f_n) continues à support compact, bornée, et qui tendent vers $\mathbb{1}_K$, ce qui prouve, par passage à la limite et application du théorème de convergence dominée, $\mu(K) = \nu(K)$.

On conclut enfin avec le théorème des classes monotones pour en déduire $\mu = \nu$. \square

8.3.2 Différents types de convergence

Dans toute la suite d désigne un entier naturel. On munit \mathbb{R}^d de la tribu borélienne produit.

Définition : Convergence d'une suite de mesures bornées

Soit $(\mu_n)_{n \in \mathbb{N}}$ une suite de mesures bornées sur \mathbb{R}^d .

Soit μ une mesure sur \mathbb{R}^d .

On dit que la suite de mesures (μ_n) converge étroitement (resp. faiblement, vaguement) vers μ lorsque, pour toute fonction f de $\mathcal{C}_b(\mathbb{R}^d, \mathbb{R})$ (resp. $\mathcal{C}_0(\mathbb{R}^d, \mathbb{R})$, $\mathcal{C}_K(\mathbb{R}^d, \mathbb{R})$) :

$$\lim_{n \rightarrow +\infty} \int f(x) d\mu_n(x) = \int f(x) d\mu(x)$$

(Remarque 9) Notons que, si μ n'est pas finie, on ne peut pas avoir de convergence étroite. Pour s'en convaincre, considérer la fonction bornée $\mathbb{1}$.

(Remarque 10) Dans tout ce qui suit, nous allons établir des liens entre ces différents types de convergence.

Proposition : Lien entre les différents types de convergence

De la série d'inclusions $\mathcal{C}_b(\mathbb{R}^d, \mathbb{R}) \supset \mathcal{C}_0(\mathbb{R}^d, \mathbb{R}) \supset \mathcal{C}_K(\mathbb{R}^d, \mathbb{R})$, nous déduisons que la convergence étroite entraîne la convergence faible qui elle-même entraîne la convergence vague.

Proposition : La limite vague d'une suite de mesures est unique

Tout est dans le titre !

Démonstration. Supposons qu'il existe deux limites (pour la convergence vague) μ et ν . Soit $P \subset \mathbb{R}^d$ un pavé compact.

Posons, pour tout $p \in \mathbb{N}^*$, $\varphi_p : x \mapsto (1 - d(x; P))^+$ où d désigne la distance.

Le support de φ_p est un compact. De plus, $\lim_{p \in \mathbb{N}} \downarrow \varphi_p = \mathbb{1}_P$.

Par passage à la limite sur n , on a, pour tout p , $\int \varphi_p d\mu = \int \varphi_p d\nu$.

Par passage à la limite sur p , on en déduit $\mu(P) = \nu(P)$. On peut ensuite conclure à l'aide du théorème des classes monotones : $\mu = \nu$. \square

Les résultats suivants précisent un peu le lien entre les différents types de convergence.

Théorème : Lien entre convergence faible et étroite

Réciproquement, si (μ_n) converge faiblement vers μ et si $\lim_{n \rightarrow +\infty} \mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d)$, alors μ_n converge étroitement vers μ .

Cette dernière condition s'appelle la convergence des masses totales.



En pratique, lorsqu'on examine une suite de probabilités, convergence étroite et convergence faible sont équivalentes. Un second résultat nous montrera que la convergence vague sera elle-aussi équivalente.

Démonstration. L'inclusion de $\mathcal{C}_0(\mathbb{R}^d, \mathbb{R})$ dans $\mathcal{C}_b(\mathbb{R}^d, \mathbb{R})$ rend évident la première implication.

Étudions la réciproque.

Considérons une fonction f continue et bornée. Sans réduire les hypothèses, on peut montrer le résultat en supposant f positive.

En effet, f est bornée et continue si et seulement si f^+ et f^- le sont.

Et, par définition de l'intégrale, la convergence des intégrales de f^+ et f^- entraînera bien la convergence de l'intégrale de f .

Pour tout $q \in \mathbb{N}$, on définit sur \mathbb{R} la fonction $\varphi_q : x \mapsto \begin{cases} 1 & \text{si } x \leq p \\ p+1-x & \text{si } x \in]p; p+1] \\ 0 & \text{sinon} \end{cases}$.

Cette fonction est continue. On définit également sur \mathbb{R}^d la fonction $\Pi_p : x \mapsto \varphi_q(\|x\|)$, peu importe le choix de la norme.

Cette fonction est continue, comme composée de fonctions continues et elle est à support le disque fermé de centre 0 et de rayon $p+1$ qui est compact.

En particulier, pour tout q et pour tout n , toutes les intégrales de l'égalité suivante sont bien définies et bornées :

$$\int f(x) d\mu_n(x) - \int f(x) d\mu(x) = \int \Pi_q(x) f(x) d\mu_n(x) - \int \Pi_q(x) f(x) d\mu(x) + \int (1 - \Pi_q(x)) f(x) d\mu_n(x) - \int (1 - \Pi_q(x)) f(x) d\mu(x)$$

En notant $M > 0$ un majorant de f , on a donc la majoration :

$$\left| \int f(x) d\mu_n(x) - \int f(x) d\mu(x) \right| \leq \left| \int \Pi_q(x) f(x) d\mu_n(x) - \int \Pi_q(x) f(x) d\mu(x) \right| + M \left(\int (1 - \Pi_q(x)) d\mu_n(x) + \int (1 - \Pi_q(x)) d\mu(x) \right)$$

On va commencer par majorer $M \left(\int (1 - \Pi_q(x)) d\mu_n(x) + \int (1 - \Pi_q(x)) d\mu(x) \right)$.

Notons que $\lim \uparrow \Pi_q = \mathbb{1}$ et comme μ est bornée, pour tout $\varepsilon > 0$, il existe q tel que

$M \int (1 - \Pi_q(x)) d\mu(x) \leq \frac{\varepsilon}{4}$. On fixe maintenant cette valeur de q .

Et comme Π_q est continue à support compact, en raison de la convergence étroite et de la convergence des masses totales, il existe N tel que pour tout $n \geq N$,

$M \left| \int (1 - \Pi_q(x)) d\mu_n(x) - \int (1 - \Pi_q(x)) d\mu(x) \right| \leq \frac{\varepsilon}{4}$.

On en déduit, en utilisant la majoration de $M \int (1 - \Pi_q(x)) d\mu(x)$ que

$M \int (1 - \Pi_q(x)) d\mu_n(x) \leq \frac{\varepsilon}{2}$.

Finalement, il nous reste à majorer $\left| \int \Pi_q(x) f(x) d\mu_n(x) - \int \Pi_q(x) f(x) d\mu(x) \right|$. Mais comme la fonction

$\Pi_q \times f$ est continue et à support compact, c'est très simple ! Il existe un rang \tilde{N} éventuellement plus grand que N , tel que pour tout $n \geq \tilde{N}$,

$\left| \int \Pi_q(x) f(x) d\mu_n(x) - \int \Pi_q(x) f(x) d\mu(x) \right| \leq \frac{\varepsilon}{4}$.

On obtient donc bien :

$$\left| \int f(x) d\mu_n(x) - \int f(x) d\mu(x) \right| \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon$$

Ce qui prouve la convergence étroite de μ_n vers μ . □



Pour toute fonction f , μ -intégrable, on pose $\mu(f) = \int f d\mu$. Notons que c'est une forme linéaire positive et que le théorème de Riesz établit que réciproquement, sous certaines conditions, on peut définir une mesure à partir d'une forme linéaire positive sur un espace de fonctions continue à support compact.

Un premier résultat de convergence étroite.

Proposition : Convergence d'une suite de noyaux gaussiens

En utilisant la mesure de Lebesgue normalisée, pour toute suite de réels strictement positifs (ρ_n) tendant vers 0, la suite de mesures (μ_n) de densités $\frac{1}{\rho_n^d} g_{a, \rho_n} : t \mapsto \frac{1}{\rho_n^d} e^{-\frac{\|t-a\|^2}{2\rho_n^2}}$ converge étroitement vers δ_a la mesure de Dirac en a .

Démonstration. Notons que (μ_n) est une suite de probabilités.

Fixons $\eta > 0$. Par un changement de variable, on a :

$$\int \mathbb{1}_{\|x-a\| \leq \eta} g_{a, \rho_n}(x) d\tilde{\lambda}(x) = \int \mathbb{1}_{\|x\| \leq \frac{1}{\rho_n} \eta} g(x) d\tilde{\lambda}(x) \xrightarrow{n \rightarrow +\infty} 1$$

Pour toute fonction f bornée et continue, et pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, pour tout $x \in [a - \eta; a + \eta]$, $f(x) \in [f(a) - \varepsilon; f(a) + \varepsilon]$, ce qui donne, par découpage de l'intégrale et passage à la limite :

$$f(a) - \varepsilon \leq \liminf \mu_n(f) \leq \limsup \mu_n(f) \leq f(a) + \varepsilon$$

ε étant arbitraire on obtient $\lim \mu_n(f) = f(a) = \delta_a(f)$. □



En fait la suite $\frac{1}{\rho_n^d} g_{a, \rho_n}$ constitue, à une translation près, une suite approximante de l'unité. D'ailleurs, on aurait pu traiter ce lemme en exploitant les convolutions.

On va maintenant établir une petite proposition qui possède de nombreuses conséquences.

Proposition : Convergence faible et famille totale

Soit \mathcal{F} une famille totale dans \mathcal{C}_0 . Soit (μ_n) une suite de mesure.

Alors (μ_n) converge faiblement si et seulement si :

- (μ_n) est bornée ;
- pour tout $f \in \mathcal{F}$, $\mu_n(f) = \int f d\mu_n$ converge.

(Exemple 0) Les gaussiennes forment une famille totale de \mathcal{C}_0 . Voir le document sur le théorème de Stone-Weierstrass pour s'en convaincre.

(Exemple 1) On peut préciser l'exemple précédent. On dispose ainsi d'une famille totale dénombrable dans \mathcal{C}_0 . Il suffit de considérer par exemple les fonctions $g_{a,b} : x \mapsto e^{-a\|x\|^2 + b \cdot x}$ avec $a \in \mathbb{Q}_*^+$ et $b \in \mathbb{Q}^d$.

(Exemple 2) Les fonctions continues à support compact \mathcal{C}_K forment une famille totale de \mathcal{C}_0 .

Démonstration. Nous allons montrer uniquement le sens réciproque, en nous appuyant sur le théorème de Riesz. Posons $M > 0$ un majorant de $(\mu_n(\mathbb{R}^d))$.

Soit $g \in \mathcal{C}_0$ et $\varepsilon > 0$. Il existe une sous-famille finie $(f_i)_{i \in I} \in \mathcal{F}^I$ et des scalaires $(\lambda_i)_{i \in I}$ tous non nuls tels que la combinaison linéaire $\tilde{f} = \sum_{i \in I} \lambda_i f_i$ vérifie :

$$\|\tilde{f} - g\|_\infty < \frac{\varepsilon}{3M}$$

Et, pour chaque $i \in I$, il existe n_i tel que, pour tout p et q supérieurs à n_i , $|\mu_p(f_i) - \mu_q(f_i)| < \frac{\varepsilon}{3|\lambda_i|\#I}$ où $\#I$ désigne le cardinal de I .

En particulier, par construction, en exploitant la linéarité de l'intégrale et l'inégalité triangulaire, on obtient, pour tout p et q supérieurs à $\max_i n_i$, $|\mu_p(\tilde{f}) - \mu_q(\tilde{f})| < \frac{\varepsilon}{3}$. Ce qui permet de conclure :

$$|\mu_p(g) - \mu_q(g)| \leq |\mu_p(g) - \mu_p(\tilde{f})| + |\mu_p(\tilde{f}) - \mu_q(\tilde{f})| + |\mu_q(\tilde{f}) - \mu_q(g)| < \varepsilon$$

La suite $(\mu_n(g))$ est de Cauchy donc converge. On définit ainsi, pour tout $g \in \mathcal{C}_0$, la forme linéaire $\mu(g) = \lim_n \mu_n(g)$. Par passage à la limite, on vérifie qu'il s'agit d'une forme linéaire positive.

Sachant que $\mathcal{C}_K \subset \mathcal{C}_0$, on peut appliquer le théorème de Riesz et associer à la forme linéaire μ une unique mesure μ telle que $\mu(g) = \int g d\mu$. \square

(Remarque 11) La réciproque devient fausse si on enlève l'hypothèse « suite bornée. »

Considérer par exemple la suite de mesure, définie pour tout $n > 0$, par $\mu_n = n\delta_n$ où δ_n est la mesure de Dirac qui vaut 1 en n et 0 partout ailleurs.

En effet, pour toute fonction Gaussienne g , $\lim \mu_n(g) = 0$. Or, pour la fonction $f : x \mapsto \begin{cases} 1 & \text{si } x \in [-1; 1] \\ \frac{1}{|x|} & \text{sinon} \end{cases}$,

on a $\mu_n(f) = 1$ pour tout n , ce qui est contradictoire car les Gaussiennes forment une famille totale.

Cette proposition possède un corollaire bien commode.

Corollaire : Compacité d'une suite bornée de mesure

De toute suite (μ_n) de mesures bornées on peut extraire une sous-suite qui converge étroitement.

Démonstration. On exploite la proposition précédente. Soit $(g_n)_{n \in \mathbb{N}}$ une famille totale dénombrable dans \mathcal{C}_0 .

On construit une suite d'extractrices (ψ_n) par l'algorithme suivant.

Pour $n = 0$, il existe une extractrice φ_0 telle que $(\mu_{\varphi_0(n)}(g_0))_{n \in \mathbb{N}}$ converge.

On pose $\psi_0 = \varphi_0$.

Pour $p \in \mathbb{N}$ quelconque, on suppose que l'on a construit une extractrice ψ_p telle que, pour tout $k \leq p$, la suite $(\mu_{\psi_p(n)}(g_k))_{n \in \mathbb{N}}$ converge.

Mais on sait qu'il existe une extractrice φ_{p+1} telle que la suite $(\mu_{\psi_p \circ \varphi_{p+1}(n)}(g_{p+1}))_{n \in \mathbb{N}}$ converge. On pose alors $\psi_{p+1} = \psi_p \circ \varphi_{p+1}$ et ψ_{p+1} vérifie que, pour tout $k \leq p+1$, la suite $(\mu_{\psi_{p+1}(n)}(g_k))_{n \in \mathbb{N}}$ converge.

On pose ensuite, pour tout n , $\tilde{\psi}(n) = \psi_n(n)$. Par construction, pour tout $k \in \mathbb{N}$ et pour tout $n \geq k$, $\tilde{\psi}(n) = \psi_k(p)$ avec $p \geq n$. En particulier, la suite $(\mu_{\tilde{\psi}(n)}(g_k))_{n \in \mathbb{N}}$ converge. Nous sommes donc dans les conditions d'application de la proposition précédente et pouvons conclure :

La suite $(\mu_{\tilde{\psi}(n)})_{n \in \mathbb{N}}$ converge étroitement. \square

8.3.3 Convergence de mesures et transformée de Fourier

Nous allons ici établir un résultat mettant en correspondance la convergence simple de la transformée de Fourier et la convergence étroite d'une suite de mesures.

Voici un premier résultat.

Proposition : Premier résultat de convergence avec Fourier

Soit (μ_n) une suite bornée de mesures. Soit μ une mesure finie.

Alors (μ_n) converge étroitement vers μ si et seulement si la suite des transformées de Fourier $(\hat{\mu}_n)$ converge vers la transformée de Fourier $\hat{\mu}$ de μ .

Pour réaliser cette démonstration, on pose :

- pour tout $\rho > 0$, l'opérateur de dilatation qui, à une fonction f , associe $\delta_\rho[f] : t \mapsto f\left(\frac{t}{\rho}\right)$;
- pour tout $a \in \mathbb{R}^d$, l'opérateur de translation qui, à une fonction f , associe $\tau_a[f] : t \mapsto f(t - a)$;
- pour tout $a \in \mathbb{R}^d$, la fonction $e_a : t \mapsto e^{ia \cdot t}$.

On rappelle que pour tout f de classe L^1 :

$$\widehat{e_a f} = \tau_a[\widehat{f}] \text{ et } \widehat{\delta_\rho[f]} = \rho^d \delta_{1/\rho}[\widehat{f}]$$

Enfin, on sait que, pourvu que l'on normalise la mesure de Lebesgue, le noyau Gaussien g vérifie $\widehat{g} = g$.

Démonstration. Montrons le sens réciproque de cette proposition. On va ici exploiter la densité des Gaussiennes et reprendre les mêmes hypothèses sur ρ et a . Posons :

$$g_{a,\rho} : t \mapsto e^{-\frac{\|t-a\|^2}{2\rho^2}}$$

Compte-tenu des remarques précédentes, $g_{a,\rho} = \rho^d e_a \widehat{\delta_{1/\rho}[g]}$, avec $g : t \mapsto e^{-\frac{\|t\|^2}{2}}$. Pour alléger, posons $\psi_{a,\rho} = \rho^d e_a \delta_{1/\rho}[g]$. C'est une fonction de l'espace de Schwartz.

Pour tout n , on a, en exploitant le théorème de Fubini-Tonelli

$$\begin{aligned} \int g_{a,\rho} d\mu_n &= \int \widehat{\psi_{a,\rho}} d\mu_n = \int \int d\mu_n(\omega) d\lambda(t) e^{-i\omega \cdot t} \psi_{a,\rho}(t) \\ &= \int d\lambda(t) \psi_{a,\rho}(t) \int d\mu_n(\omega) e^{-i\omega \cdot t} \\ &= \int d\lambda(t) \psi_{a,\rho}(t) \widehat{\mu_n}(t) \end{aligned}$$

Nous sommes dans les conditions d'application du théorème de convergence monotone. Pour n tendant vers l'infini, cette quantité tend vers :

$$\int d\lambda(t) \psi_{a,\rho}(t) \widehat{\mu}(t) = \int g_{a,\rho} d\mu$$

Ce qui prouve la convergence étroite de la suite de mesures (μ_n) .

Pour montrer le sens direct, on exploite le fait, pour tout ω , la fonction $t \mapsto e^{it\omega \cdot t}$ est continue et bornée, ce qui donne, en raison de la convergence étroite :

$$\lim_n \int e^{it\omega \cdot t} d\mu_n(t) = \int e^{it\omega \cdot t} d\mu(t) \text{ c'est à dire } \lim_n \widehat{\mu_n}(\omega) = \widehat{\mu}(\omega)$$

□

Chapitre 9

Probabilites

9.1 Quelques résultats préliminaires

9.1.1 Premières définitions, propriétés et notations

Définition : Probabilité, évènement

Un espace mesuré $(\Omega; \mathcal{T}; P)$ est dit probabilisé lorsque $P(E) = 1$.

Un élément de la tribu \mathcal{T} d'un espace probabilisé s'appelle un évènement.

Proposition : Convergence monotone

En particulier, un espace probabilisé vérifie le théorème de convergence monotone dans sa version ascendante et descendante.

Définition : Variable aléatoire, fonction de répartition, espérance

Soit $(\Omega; \mathcal{T}; P)$ un espace probabilisé. On munit R de la tribu des Boréliens.

On dit que $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire réelle si et seulement si X est mesurable.

En particulier, une telle application permet de transporter une mesure de probabilité sur R que l'on notera P_X .

La fonction de répartition de X est la fonction

$F_X : t \mapsto P(X \leq t)$.

Enfin, on dit que la variable X est L^1 lorsque $\int_{\Omega} |X(\omega)| dP(\omega) < +\infty$ et dans ce cas on note

$$E(X) = \int X(\omega) dP(\omega)$$

Définition : Notations

Soit (X_n) une suite de variables aléatoires. Soient X et Y deux variables aléatoires. On suppose que toutes ces variables partagent le même espace probabilisé $(\Omega; \mathcal{T}; P)$. Soit enfin a un réel.

On utilise alors les notations suivantes :

$$\begin{aligned} \{X \geq Y\} &= \{\omega \in \mathcal{T} / X(\omega) \geq Y(\omega)\} & \{X_n \xrightarrow[n \rightarrow +\infty]{} X\} &= \left\{ \omega \in \mathcal{T} / X_n(\omega) \xrightarrow[n \rightarrow +\infty]{} X(\omega) \right\} \\ \{\sup X_n = +\infty\} &= \{\omega \in \mathcal{T} / \sup X_n(\omega) = +\infty\} & \{\sup X_n \leq a\} &= \{\omega \in \mathcal{T} / \sup X_n(\omega) \leq a\} \end{aligned}$$

On rappelle les définitions suivantes :

$$\limsup X_n = \lim_{n \rightarrow +\infty} \downarrow \sup_{k \geq n} X_k \qquad \liminf X_n = \lim_{n \rightarrow +\infty} \uparrow \inf_{k \geq n} X_k$$

On peut aussi exploiter les notations ensemblistes pour caractériser les évènements.

$$\begin{aligned} \left\{ \lim_{n \rightarrow +\infty} X_n = X \right\} &= \bigcap_{\varepsilon \in \mathbb{Q}_+^*} \bigcup_n \bigcap_{k \geq n} \{|X_n - X| < \varepsilon\} & \left\{ \sup_{n \rightarrow +\infty} X_n = +\infty \right\} &= \bigcap_{M \in \mathbb{Q}_+^*} \bigcup_n \{X_n \geq M\} \\ \{X_n \text{ est bornée}\} &= \bigcup_{M \in \mathbb{Q}_+^*} \bigcap_n \{|X_n| \leq M\} & \{\limsup X_n < a\} &= \bigcup_n \bigcap_{k \geq n} \{X_n < a\} \end{aligned}$$

Définition : Fonction de densité

Soit X une variable aléatoire réelle définie sur Ω .

On dit que X admet une densité lorsqu'il existe une fonction mesurable f telle que, pour tout borélien B :

$$P(X \in B) = \int_B f(x) d\lambda(x)$$

λ désigne la mesure de Lebesgue.

En particulier lorsqu'elle existe la fonction de densité est L^1 et est positive λ -presque partout.

Proposition : La densité est définie de manière unique λ -presque partout

Soit X une variable aléatoire admettant deux densités f et g .

Alors $f = g$, λ -presque partout.

Démonstration. Pour tout rationnel a strictement positif, on pose $T_a = \{x, /f(x) \geq g(x) + a\}$. On a

$$P(X \in T) = \int_T f(x) d\lambda(x) = \int_T g(x) d\lambda(x)$$

Or $\int_T f(x) d\lambda(x) \geq \int_T f(x) d\lambda(x) + a\lambda(T_a)$.

On en déduit $\lambda(T_a) = 0$. Comme cela est vrai pour tout a , on en déduit par convergence monotone que l'ensemble $T = \{x, /f(x) > g(x)\}$ est λ -négligeable.

Donc, par symétrie de f et g , $f = g$ λ -presque partout. □

Définition : Probabilité transportée

Soit X une variable aléatoire. La probabilité transportée P_X est une mesure définie sur les boréliens par :

$$P_X(B) = P(X^{-1} < B >) \quad \forall B \text{ borélien}$$

Un dernier résultat très important : le lemme de Borel-Cantelli.

Lemme : Borel-Cantelli

Soit (A_n) une suite d'évènements. On suppose que :

$$\sum_n P(A_k) < +\infty$$

Alors :

$$P\left(\bigcap_n \bigcup_{k \geq n} A_k\right) = 0$$

(Remarque 0) On note aussi $\limsup A_n = \bigcap_n \bigcup_{k \geq n} A_k$: c'est la limite supérieure ensembliste déjà abordée auparavant.

(Remarque 1) Dans le cas où les (A_n) sont indépendants, la loi du tout ou rien, page 107, renforce cet énoncé en une équivalence.

Démonstration. Soit n . Notons $R_n = \sum_{k \geq n} P(A_k)$. Pour tout p , on a :

$$P\left(\bigcup_{n+p \leq k \leq n} A_k\right) \leq \sum_{n+p \leq k \leq n} P(A_k)$$

Par passage à la limite sur p , cela donne $P\left(\bigcup_{k \geq n} A_k\right) \leq R_n$.

Et, par passage à la limite sur n , sachant que $\lim R_n = 0$, on obtient le résultat escompté. \square

9.1.2 Indépendance, construction de suites d'expériences indépendantes et identiquement distribuées

Définition : Indépendance de deux évènements

Soient A et B deux évènements. On dit que A et B sont indépendants lorsque

$$P(A \cap B) = P(A) \times P(B)$$

Proposition : Indépendance et contraire

Deux évènements sont indépendants si et seulement si leurs contraires le sont.

Démonstration.

$$\begin{aligned} P(A \cap B) = P(A) \times P(B) &\iff P(B) - P(A \cap B) = P(B) - P(A) \times P(B) \\ &\iff P(A^c \cap B) = P(B)(1 - P(A)) = P(B) \times P(A^c) \end{aligned}$$

\square

Définition : Indépendance d'une famille d'évènement

Soit $(A_i)_{i \in I}$ une famille d'évènements.

Dans le cas où I est finie, on dit que cette famille est indépendante lorsque

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

Dans le cas où I est dénombrable, on dit que cette famille est indépendante lorsque toute sous-famille finie est indépendante.

Proposition : Indépendance et contraire

Une famille est indépendante si et seulement si la famille formée des contraires l'est aussi.

Démonstration. Évident par récurrence. □

Définition : Indépendance d'une famille de tribus

Soit $(\mathcal{T}_i)_{i \in I}$ une famille de sous-tribus de \mathcal{T} au plus dénombrable.

On dit que cette famille est indépendante lorsque toute famille d'évènements $(A_i)_{i \in I} \in \prod_{i \in I} \mathcal{T}_i$ est indépendante.

Proposition : Caractérisation de l'indépendance à l'aide de classes monotones

On reprend les hypothèses de la définition précédente.

On suppose que les \mathcal{T}_i sont engendrées par des π -systèmes \mathcal{C}_i .

Les \mathcal{T}_i sont indépendantes si et seulement si toute famille d'éléments $(C_i)_{i \in I} \in \prod_{i \in I} \mathcal{C}_i$ est indépendante.

Démonstration. Le sens direct est évident puisque les \mathcal{C}_i sont des sous-familles des \mathcal{T}_i .

On se consacre donc au sens réciproque. Il suffit de le prouver dans le cas d'une famille finie.

Soient ainsi les tribus $(\mathcal{T}_i)_{1 \leq i \leq n}$ engendrée par les π -systèmes indépendants $(\mathcal{C}_i)_{1 \leq i \leq n}$. On va utiliser l'argumentaire classique des classes monotones.

On pose $\mathcal{L}_1 = \left\{ T_1 \in \mathcal{T}_1 / \forall (C_i)_{2 \leq i \leq n} \in \prod_{2 \leq i \leq n} \mathcal{C}_i, P(T_1 \cap C_2 \cap \dots \cap C_n) = P(T_1) \times P(C_2) \times \dots \times P(C_n) \right\}$.

Cet ensemble contient \mathcal{C}_1 . De plus, c'est un λ -système car il est stable par différence et par union croissante dénombrable et il contient Ω .

Ainsi, $\mathcal{L}_1 = \mathcal{T}_1$.

Considérons maintenant

$\mathcal{L}_2 = \left\{ T_2 \in \mathcal{T}_2 / \forall T_1 \in \mathcal{T}_1, \forall (C_i)_{2 \leq i \leq n} \in \prod_{2 \leq i \leq n} \mathcal{C}_i, P(T_1 \cap T_2 \cap C_3 \cap \dots \cap C_n) = P(T_1) \times P(T_2) \times P(C_3) \times \dots \times P(C_n) \right\}$.

\mathcal{L}_2 est également un λ -système qui contient \mathcal{C}_2 . On a donc $\mathcal{L}_2 = \mathcal{T}_2$.

Par une récurrence immédiate, on arrive au résultat. □

Définition : Variables aléatoires indépendantes

Soit $(X_i)_{i \in I}$ une famille au plus dénombrable de variables aléatoires définies sur le même espace probabilisé.

On dit que cette famille de variables aléatoires est indépendante lorsque les tribus réciproques de la tribu borélienne $(X_i^{-1}(<\mathcal{B}>))_{i \in I}$ sont indépendantes.

Proposition : Caractérisation par des classes de réels de l'indépendance de variables aléatoires

Soit \mathcal{C} un π -système engendrant la tribu Borélienne.

Alors les $(X_i)_{i \in I}$ sont indépendantes si et seulement si les $(X_i^{-1}(<\mathcal{C}>))_{i \in I}$ sont indépendants.

Démonstration. On a prouvé précédemment dans le cadre de l'étude de fonctions mesurables que les $X_i^{-1}(<\mathcal{B}>)$ étaient engendrées par les $X_i^{-1}(<\mathcal{C}>)$; c'est à dire que l'image réciproque d'une tribu engendrée par une classe était engendrée par l'image réciproque de cette même classe.

Ainsi, on peut donc se ramener au résultat de la proposition précédente. \square

Corollaire : Caractérisation de l'indépendance de variables aléatoires par leurs fonctions de répartition, par leurs espérances

Soit $(X_i)_{i \in I}$ une famille au plus dénombrable de variables aléatoires

Cette famille est indépendante si et seulement si, pour toute sous-famille finie $(X_i)_{i \in J}$, l'une des propositions suivantes est vérifiée :

- pour tout élément $(a_i)_{i \in J}$ de \mathbb{R}^d , avec $d = |J|$:

$$P\left(\bigcap_{i \in J} X_i \leq a_i\right) = \prod_{i \in J} F_{X_i}(a_i)$$

- pour tout pavé borélien $\Pi = (B_i)_{i \in J}$ de \mathbb{R}^d , avec $d = |J|$:

$$P((X_i)_{i \in J} \in \Pi) = \prod_{i \in J} P(B_i)$$

- la mesure transportée par $(X_i)_{i \in J}$ sur \mathbb{R}^d , avec $d = |J|$, est la mesure produit $\bigotimes_{i \in J} P_{X_i}$;
- pour toute famille $(f_i)_{i \in J}$ de fonctions positives mesurables :

$$E\left(\prod_{i \in J} f_i(X_i)\right) = \prod_{i \in J} E(f_i(X_i))$$

- pour toute famille $(f_i)_{i \in J}$ de fonctions continues et bornées :

$$E\left(\prod_{i \in J} f_i(X_i)\right) = \prod_{i \in J} E(f_i(X_i))$$

On pourrait étendre ce corollaire aux fonctions continues à support compact, aux fonctions \mathcal{C}^∞ à support compact.

Démonstration. C'est une conséquence directe de la proposition précédente. Le premier point s'obtient en remarquant que la famille des pavés $] - \infty; a_i]$ est un π -système qui engendre la tribu produit de \mathbb{R}^d .

Pour le second point, c'est d'autant plus vrai que les pavés boréliens engendrent eux aussi la tribu produit de \mathbb{R}^d .

Puisque la mesure coïncide avec la mesure produit sur les pavés boréliens ; en raison du théorème des classes monotones, elle coïncide sur la tribu produit de \mathbb{R}^d .

Le quatrième point s'obtient par le théorème de Fubini-Tonelli.

Le cinquième point s'obtient en raison de la densité dans L^1 des fonctions continues et bornées. \square

Proposition : Probabilité de somme de v.a.r. indépendantes

Soient X et Y deux variables aléatoires définies sur le même espace probabilisé. Alors :

$$P_{X+Y} = P_X * P_Y$$

La probabilité transportée de $X + Y$ est la convolution des deux probabilités.

Démonstration. Soit B un borélien. On a, en raison de l'indépendance des événements :

$$P_{X+Y}(B) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_B(y+x) dP_X(x) dP_Y(y)$$

Avec le changement de variable $z = x + y$ dans la seconde intégrale et l'application du théorème de Fubini-Tonelli, on obtient bien :

$$P_{X+Y}(B) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_B(z) dP_X(x) dP_Y(z-x) = \int_{\mathbb{R}} \mathbb{1}_B(z) dP_X * P_Y(z)$$

□

Proposition : Indépendance et densité

Soient une famille de variables $(X_i)_{i \in I}$ au plus dénombrables admettant des densités respectives $(f_i)_{i \in I}$. Alors ces variables sont indépendante si et seulement si, toute sous famille finie $(X_i)_{i \in J}$ admettant sur \mathbb{R}^d , avec $d = |J|$, une fonction de densité

$$(t_i)_{i \in J} \mapsto \prod_{i \in J} f_i(t_i)$$

Démonstration. Évident en utilisant la proposition 9.1.1 et le corollaire 9.1.2. □

9.1.3 Loi du tout ou rien, tribu asymptotique

Abordons maintenant la loi dite du tout ou rien que l'on doit à Kolmogorov. Avant cela, un résultat préliminaire très simple.

Lemme : Évènement indépendant de lui-même

Soit A un événement.

A est indépendant de lui-même si et seulement si $P(A) = 0$ ou $P(A) = 1$.

Démonstration. La condition d'indépendance s'écrit $P(A)^2 = P(A) \iff P(A) = 0 \text{ ou } P(A) = 1$. □

Définition : Tribu asymptotique

Soit $(\mathcal{T}_n)_{n \in \mathbb{N}}$ une suite de tribus. On définit la tribu asymptotique par :

$$\mathcal{A}_\infty = \bigcap_{n \in \mathbb{N}} \sigma \left(\bigcup_{k \geq n} \mathcal{T}_k \right)$$

Cette définition est assez abstraite et on dispose d'une caractérisation plus commode.

Proposition : Caractérisation

On note $\mathcal{B}_\infty = \sigma \left(\bigcup_{k \geq 0} \mathcal{T}_k \right)$. Alors un événement de \mathcal{B}_∞ n'appartient pas à \mathcal{A}_∞ si et seulement s'il s'obtient à partir d'événements d'un sous-ensemble fini de tribus (\mathcal{T}_k) .

- (Remarque 2) Cela signifie que la tribu asymptotique ne contient que les événements de \mathcal{B}_∞ qui ne s'obtiennent qu'à partir d'événements choisis dans un nombre infini de (\mathcal{T}_k) .
- (Remarque 3) En pratique, les événements qui sont construits un nombre infini de tribus (\mathcal{T}_k) sont donc dans \mathcal{A}_∞ .
- (Remarque 4) Dans le cas où la suite (\mathcal{T}_n) est constituée des réciproques de la tribu borélienne par une suite de variables aléatoires (X_n) , des événements portant sur $\sum_{n \in \mathbb{N}} X_n$, ou sur $\limsup X_n$ appartiendront à \mathcal{A}_∞ .

Théorème : Loi du tout ou rien

On considère $(\mathcal{T}_n)_{n \in \mathbb{N}}$ une suite de sous-tribus indépendantes d'un espace probabilisé $(\Omega; \mathcal{T}; P)$. Alors la tribu asymptotique est presque sûrement grossière.

Démonstration. On va utiliser la caractérisation par les classes monotones pour prouver que cette tribu est indépendante d'elle-même.

Posons $\mathcal{T}_\infty = \bigcap_{n \in \mathbb{N}} \sigma \left(\bigcup_{k \geq n} \mathcal{T}_k \right)$ et $\mathcal{S}_\infty = \sigma \left(\bigcup_{n \in \mathbb{N}} \mathcal{T}_n \right)$.

Il est clair que $\mathcal{T}_\infty \subset \mathcal{S}_\infty$.

Pour tout entier naturel n , on va noter $\mathcal{U}_n = \sigma \left(\bigcup_{k < n} \mathcal{T}_k \right)$ et $\mathcal{V}_n = \sigma \left(\bigcup_{k \geq n} \mathcal{T}_k \right)$.

On considère d'autre part les ensembles $\mathcal{C}_n = \left\{ \bigcap_{k < n} T_k, T_k \in \mathcal{T}_k \right\}$ et $\mathcal{D}_n = \left\{ \bigcap_{n+m \geq k \geq n} T_k, m \in \mathbb{N}, T_k \in \mathcal{T}_k \right\}$.

Les \mathcal{C}_n et \mathcal{D}_n sont des π -systèmes respectivement inclus dans \mathcal{U}_n et \mathcal{V}_n et qui contiennent respectivement $\bigcup_{k < n} \mathcal{T}_k$ et $\bigcup_{k \geq n} \mathcal{T}_k$.

Ainsi, ces π -systèmes engendrent respectivement \mathcal{U}_n et \mathcal{V}_n .

Or, on vérifie aisément que pour tout $(C_n; D_n) \in \mathcal{C}_n \times \mathcal{D}_n$, C_n et D_n sont indépendants.

En exploitant la proposition 9.1.2 sur l'indépendance et les classes monotones, on a ainsi prouvé que les éléments de \mathcal{V}_n sont indépendants des éléments de \mathcal{C}_n .

Considérons maintenant un événement $A \in \mathcal{T}_\infty$.

A est, pour tout n , dans \mathcal{V}_n donc est indépendant des éléments de \mathcal{C}_n pour tout n .

Ainsi, A est indépendant des éléments de $\bigcup_{n \in \mathbb{N}^*} \mathcal{C}_n$.

Or ce dernier ensemble est un π -système qui contient $\bigcup_{n \in \mathbb{N}^*} \mathcal{T}_n$ et qui appartient à \mathcal{S}_∞ . Il engendre donc \mathcal{S}_∞ .

Toujours en raison du résultat sur les classes monotones, on en déduit que A est indépendant de \mathcal{S}_∞ donc de lui-même. \square

L'une des conséquences de ce théorème est le théorème probabiliste de Borel-Cantelli.

Théorème : Borel-Cantelli

Soit $(\Omega; \mathcal{T}; P)$ un espace probabilisé.

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements indépendants de \mathcal{T} . Alors :

$$\begin{aligned} \sum_{n \in \mathbb{N}} P(A_n) < +\infty &\iff P \left(\bigcap_{n \in \mathbb{N}} \bigcup_{p \geq n} A_p \right) = 0 \\ \sum_{n \in \mathbb{N}} P(A_n) = +\infty &\iff P \left(\bigcap_{n \in \mathbb{N}} \bigcup_{p \geq n} A_p \right) = 1 \end{aligned}$$

Démonstration. On considère la suite de tribus $\mathcal{T}_n = \{A_n; A_n^c; \Omega; \emptyset\}$.

Par construction cette suite est indépendante. Or l'événement $\limsup_{n \rightarrow +\infty} A_n = \bigcap_{n \in \mathbb{N}} \bigcup_{p \geq n} A_p$ appartient à \mathcal{T}_∞ en reprenant les notations précédentes.

On a donc $P\left(\limsup_{n \rightarrow +\infty} A_n\right) = 0$ ou bien $P\left(\limsup_{n \rightarrow +\infty} A_n\right) = 1$.

Dans le cas où $\sum_{n \in \mathbb{N}} P(A_n) < +\infty$, le lemme classique de Borel-Cantelli nous dit que $P\left(\limsup_{n \rightarrow +\infty} A_n\right) = 0$.

Dans le cas contraire, si $\sum_{n \in \mathbb{N}} P(A_n) = +\infty$, on va montrer que $P\left(\liminf_{n \rightarrow +\infty} A_n^c\right) = 0$.

Or, pour tout $n \in \mathbb{N}$, $P\left(\bigcap_{k \geq n} A_k^c\right) = \prod_{k \geq n} (1 - P(A_k))$.

Par passage au logarithme sur cette dernière expression :

$$\begin{aligned} \ln\left(\prod_{k \geq n} (1 - P(A_k))\right) &= \sum_{k \geq n} \ln(1 - P(A_k)) \\ &\leq -\sum_{k \geq n} P(A_k) = -\infty \end{aligned}$$

Ainsi, $P\left(\bigcap_{k \geq n} A_k^c\right) = 0$. Par convergence monotone, on obtient le résultat escompté. \square

9.1.4 Variables aléatoires L^p

Dans toute la suite $(\Omega; \mathcal{T}; P)$ désigne un univers probabilisé.

Définition : Moments d'ordre p

Soit $p \in [1; +\infty[$. Soit X une variable aléatoire définie sur Ω . On dit que X admet un moment d'ordre p lorsque :

$$E(|X|^p) = \int_{\Omega} |X(\omega)|^p dP(\omega) < +\infty$$

On note L^p l'ensemble des variables aléatoires définies sur Ω et qui admettent un moment d'ordre p .

En fait, L^p désigne un ensemble des classes de variables aléatoires pour la relation d'équivalence d'égalité presque partout.

Proposition : Propriétés des espaces L^p

Soient X et Y deux variables aléatoires définies sur Ω et de classe L^p . Soient λ un réel. Alors :

- λX est de classe L^p .
- $\max(|X|; |Y|)$ est de classe L^p .
- $X + Y$ est de classe L^p .

Démonstration. Le premier point est facile. Considérons le second point.

Pour tout $\omega \in \Omega$, $\max(|X|; |Y|)^p = \mathbb{1}_{|X| \geq |Y|} |X(\omega)|^p + \mathbb{1}_{|X| < |Y|} |Y(\omega)|^p$.

On en déduit :

$$E(\max(|X|; |Y|)^p) \leq E(|X|^p) + E(|Y|^p) < +\infty$$

Le dernier point est une conséquence du second.

En effet, pour tout $\omega \in \Omega$, $|X(\omega) + Y(\omega)|^p \leq (\max(|X|; |Y|) + \max(|X|; |Y|))^p = 2^p \max(|X|; |Y|)^p$, ce qui permet d'obtenir :

$$E(|X + Y|^p) \leq 2^p E(\max(|X|; |Y|)^p) < +\infty$$

\square

Passons maintenant aux deux inégalités classiques : Hölder et Minkowski.

Proposition : Inégalités de Hölder et Minkowski

Soient p et q deux réels positifs tels que $\frac{1}{p} + \frac{1}{q} = 1$. Soient X , Y et Z trois variables aléatoires de classes respectives L^p , L^q et L^p . Alors :

- XY est de classe L^1 et $E(|XY|) \leq E(|X|^p)^{1/p} E(|Y|^q)^{1/q}$;
- $X + Z$ est de classe L^p et $E(|X + Z|^p)^{1/p} \leq E(|X|^p)^{1/p} + E(|Z|^p)^{1/p}$.

En particulier, cette dernière inégalité montre que L^p est un espace vectoriel normé, muni de la norme $\|X\|_p = E(|X|^p)^{1/p}$.

Démonstration. Si X ou Y est nul presque partout les deux inégalités sont triviales. On se place donc dans le cas contraire. Posons $\tilde{X} = \frac{|X|^p}{E(|X|^p)}$ et $\tilde{Y} = \frac{|Y|^q}{E(|Y|^q)}$. D'après l'inégalité de Young, pour tout ω , on a :

$$\frac{1}{p} \tilde{X}(\omega) + \frac{1}{q} \tilde{Y}(\omega) \geq \tilde{X}(\omega)^{1/p} \tilde{Y}(\omega)^{1/q} \iff \frac{1}{p} \tilde{X}(\omega) + \frac{1}{q} \tilde{Y}(\omega) \geq \frac{|XY|}{E(|X|^p)^{1/p} E(|Y|^q)^{1/q}}$$

En intégrant cette dernière inégalité, sachant que \tilde{X} et \tilde{Y} sont normalisés, on obtient :

$$\frac{1}{p} + \frac{1}{q} \geq \frac{1}{E(|X|^p)^{1/p} E(|Y|^q)^{1/q}} \int |XY|(\omega) dP(\omega) \text{ soit } \int |XY|(\omega) dP(\omega) \leq E(|X|^p)^{1/p} E(|Y|^q)^{1/q} < +\infty$$

Montrons maintenant l'inégalité de Minkowski. Remarquons que $q = \frac{p}{p-1}$ vérifie $\frac{1}{p} + \frac{1}{q} = 1$. D'autre part, si X et Z sont L^p alors $|X + Z|^{p-1}$ est L^q car $(|X + Z|^{p-1})^q = |X + Z|^p$ est intégrable. Fort de ces deux remarques, on va exploiter l'inégalité de Hölder juste établie :

$$\begin{aligned} E(|X + Z|^p) &\leq E(|X| |X + Z|^{p-1}) + E(|Z| |X + Z|^{p-1}) \\ &\leq E(|X|^p)^{1/p} E\left((|X + Z|^{p-1})^q\right)^{1/q} + E(|Z|^p)^{1/p} E\left((|X + Z|^{p-1})^q\right)^{1/q} < +\infty \end{aligned}$$

Or, $E\left((|X + Z|^{p-1})^q\right)^{1/q} = E(|X + Z|^p)^{(p-1)/p}$. On obtient ainsi :

$$E(|X + Z|^p) \leq E(|X + Z|^p)^{1-1/p} \left(E(|X|^p)^{1/p} + E(|Z|^p)^{1/p} \right) \text{ soit } E(|X + Z|^p)^{1/p} \leq E(|X|^p)^{1/p} + E(|Z|^p)^{1/p}$$

□

Proposition : Autres propriétés des espaces L^p

Les espaces L^p munis de la norme définie plus haut sont complets. En outre, pour tout $p' > p$, on a $L^{p'} \subset L^p$.

Démonstration. Il est facile de constater que X est $L^{p'}$ si et seulement si $\max(1; |X|)$ l'est également.

En particulier, on a alors $\max(1; |X|)^{p'} \geq \max(1; |X|)^p$ ce qui permet de conclure.

Pour montrer que ces espaces sont complets, on peut par exemple exploiter la caractérisation de la complétude par des séries comme on l'avait fait précédemment. □



Cette inclusion des espaces intégrables se produit pour toute mesure finie. Dans le cas général, cela n'est plus valable. Par exemple la fonction $t \mapsto \mathbb{1}_{[1; +\infty[}(t) \frac{1}{t}$ est L^2 mais pas L^1 pour la mesure de Lebesgue.

9.2 Convergences presque sûrement, convergence \mathcal{L}^p , convergence en probabilité

9.2.1 Convergence \mathcal{L}^p , convergence presque sûre

Définition : Convergence \mathcal{L}^p , convergence presque sûre

Soit (X_n) une suite de variables aléatoires réelles.

Soit X une variable réelle.

On dit que la suite (X_n) converge presque sûrement vers X lorsque

$$P(\{\omega / X_n(\omega) \rightarrow X(\omega)\}) = 1$$

Si de plus les (X_n) et X sont \mathcal{L}^p , on dit que la suite X_n converge vers X dans \mathcal{L}^p lorsque

$$\lim_{n \rightarrow +\infty} E(|X_n - X|^p) = 0$$

Proposition : Inégalités de Markov et de Bienaymé-Tchebychev

Soit X une variable \mathcal{L}^1 . Alors, pour tout $a > 0$

$$P(|X| > a) \leq \frac{E(|X|)}{a}$$

Si de plus, X est \mathcal{L}^p avec p un entier strictement positif, on a :

$$P(|X - E(X)| > a) \leq \frac{E(|X - E(X)|^p)}{a^p}$$

Démonstration. Pour tout $a > 0$, on pose $S_a = \{\omega / |X(\omega)| > a\}$.

On a ainsi, $|X| \geq a \mathbb{1}_{S_a}$. En intégrant, on obtient

$aP(S_a) \leq E(|X|)$, ce qui donne l'inégalité de Markov.

Plaçons-nous dans les hypothèses de l'inégalité de Bienaymé-Tchebychev et considérons

$T_a = \{\omega / |X(\omega) - E(X)| > a\}$.

On a, de même, $|X - E(X)|^p \geq a^p \mathbb{1}_{S_a}$. En intégrant, on obtient la seconde inégalité. □

Voici une petite propriété qui permet de caractériser une convergence presque sûre.

Proposition : Convergence presque sûre

Soit une suite (X_n) de variables aléatoires et X une autre variable aléatoire. On suppose que, pour tout $\varepsilon > 0$, on a :

$$\sum_n P(|X_n - X| \geq \varepsilon) < +\infty$$

Alors la suite (X_n) converge presque sûrement vers X .

Si de plus les (X_n) sont indépendantes, cette condition devient nécessaire à la convergence presque sûre.

Démonstration. Plaçons-nous dans les hypothèses de la proposition. Pour tout $\varepsilon > 0$, le lemme de Borel-Cantelli entraîne que :

$$P\left(\bigcap_n \bigcup_{k \geq n} \{|X_k - X| \geq \varepsilon\}\right) = 0$$

Par passage au complémentaire, cela donne :

$$P\left(\bigcup_n \bigcap_{k \geq n} \{|X_n - X| < \varepsilon\}\right) = 1$$

Ainsi la suite (X_n) converge presque sûrement vers X .

Supposons maintenant que les (X_n) sont indépendantes et que $\sum_n P(|X_n - X| \geq \varepsilon) = +\infty$.

Par le théorème de Borel-Cantelli, pour tout $\varepsilon > 0$, $P\left(\bigcap_n \bigcup_{k \geq n} \{|X_n - X| \geq \varepsilon\}\right) = 1$, ce qui entraîne que, presque sûrement, les (X_n) ne tendent pas vers X . \square

9.2.2 Convergence en probabilité

Définition : Convergence en probabilité

Soit (X_n) une suite de variables aléatoires et X une variable aléatoire.

On dit que (X_n) converge vers X en probabilité lorsque, pour tout $a > 0$,

$$P(|X_n - X| > a) \xrightarrow{n \rightarrow +\infty} 0$$

Proposition : Convergence en norme \mathcal{L}^p , convergence presque-partout et convergence en probabilité

Si X_n converge presque partout vers X alors X converge en probabilité vers X .

Si les X_n convergent en norme \mathcal{L}^p vers X alors X_n converge en probabilité vers X .

Démonstration. Supposons que X_n converge presque sûrement vers X . Alors, pour tout $a > 0$, $A_n = \{\omega / |X_n(\omega) - X(\omega)| > a\}$ est inclus dans $\tilde{A}_n = \left\{\omega / \sup_{p \geq n} |X_p(\omega) - X(\omega)| > a\right\}$.

Or la suite des \tilde{A}_n décroît vers un ensemble de probabilité nulle, ce qui permet de conclure. Supposons que X_n converge en norme \mathcal{L}^p vers X . On applique l'inégalité de Markov :

$$P(|X_n - X| > a) \leq \frac{E(|X_n - X|^p)}{a^p} \xrightarrow{n \rightarrow +\infty} 0$$

\square

Proposition : Convergence en probabilité et suite de Cauchy en probabilité

Une suite X_n converge en probabilité si et seulement si elle est de Cauchy en probabilité, c'est à dire

$$\forall a > 0, \forall p \in \mathbb{N}, \lim_{n \rightarrow +\infty} P(|X_{n+p} - X_n| > a) = 0$$

De plus, dans ce cas, on peut extraire une sous-suite qui converge presque partout.

Démonstration. Si la suite converge en probabilité, notons que pour tout n et p entiers,

$$\{|X_{n+p} - X_n| \leq a\} \supset \left(\left\{|X_n - X| \leq \frac{a}{2}\right\} \cup \left\{|X_{n+p} - X| \leq \frac{a}{2}\right\} \right).$$

En considérant le contraire, on obtient :

$$P(\{|X_{n+p} - X_n| > a\}) \leq P\left(\left\{|X_n - X| > \frac{a}{2}\right\}\right) + P\left(\left\{|X_{n+p} - X| > \frac{a}{2}\right\}\right).$$

Par passage à la limite sur n , on prouve que X_n est de Cauchy en probabilité.

Supposons maintenant que X_n est de Cauchy en probabilité. On va utiliser le lemme de Borel-Cantelli.

Soit la suite extractrice α_n telle que, pour tout n , $P(\{|X_{\alpha_{n+1}} - X_{\alpha_n}| > 2^{-n}\}) < 2^{-n}$

Notons A_n l'évènement $|X_{\alpha_{n+1}} - X_{\alpha_n}| > 2^{-n}$. Par construction, $\sum_{n \rightarrow +\infty} P(A_n) < +\infty$.

Ainsi, $P(\limsup A_n) = 0$.

Considérons maintenant le contraire de $\limsup A_n$. Sa probabilité vaut :

$$P(\{\omega / \exists n, \forall p \geq n, |X_{\alpha_{n+1}}(\omega) - X_{\alpha_n}(\omega)| \leq 2^{-n}\}) = 1$$

Or pour tout ω de cet ensemble, la suite $X_{\alpha_n}(\omega)$ est de Cauchy donc converge vers $X(\omega)$. On définit presque partout la variable aléatoire X comme étant cette limite.

Supposons maintenant que la suite X_n est de Cauchy en probabilité et montrons que l'on peut en extraire une sous-suite convergente presque partout.

Soit $\varepsilon > 0$. Comme la suite est de Cauchy, il existe un certain rang N tel que pour tout $n \geq N$ et pour tout entier p , $P(|X_n - X_{\alpha_{n+p}}| > \frac{\varepsilon}{2}) \leq \frac{\varepsilon}{2}$.

Et de même, il existe Q tel que pour $p \geq Q$, $P(|X_{\alpha_{n+p}} - X| > \frac{\varepsilon}{2}) \leq \frac{\varepsilon}{2}$ car la suite X_{α_n} converge presque partout donc en probabilité vers X .

Or :

$$\{\omega / |X_n(\omega) - X(\omega)| > a\} \subset \left\{ \omega / |X_{\alpha_{n+p}}(\omega) - X_n(\omega)| > \frac{a}{2} \right\} \cup \left\{ \omega / |X_{\alpha_{n+p}}(\omega) - X(\omega)| > \frac{a}{2} \right\}$$

On obtient donc bien que, pour tout $n \geq N$, $P(\{\omega / |X_n(\omega) - X(\omega)| > a\}) < \varepsilon$ □

Proposition : Fonction uniformément continue et convergence en probabilité

Soit X_n une suite de variables aléatoires qui convergent en probabilité vers X . Soit f une fonction uniformément continue.

Alors $f(X_n)$ converge en probabilité vers $f(X)$.

Démonstration. Si f est uniformément continue, notons que pour tout $a > 0$, il existe $\eta > 0$ tel que $\{\omega / |f(X_n(\omega)) - f(X_n(\omega))| > a\} \subset \{\omega / |f(X_n(\omega)) - f(X_n(\omega))| > \eta\}$.

En prenant les probabilités, par majoration, cela prouve que les $f(X_n)$ convergent en probabilité vers $f(X)$. □

Lemme : Suite de variables L^2 et convergence en probabilité

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables indépendantes, identiquement distribuées et de carré intégrable.

Alors $\frac{1}{n} \sum_{1 \leq k \leq n} X_k$ converge en probabilité vers $E(X_1)$.

Démonstration. Notons que

$$E\left(\frac{1}{n} \sum_{1 \leq k \leq n} X_k\right) = E(X_1)$$

Donc, d'après l'inégalité de Bienaymé-Tchbychev, pour tout $a > 0$:

$$P\left(\left|\sum_{1 \leq k \leq n} \frac{X_k}{n} - E(X_1)\right| > a\right) \leq \frac{V\left(\sum_{1 \leq k \leq n} \frac{X_k}{n}\right)}{a^2}$$

Or $V\left(\sum_{1 \leq k \leq n} \frac{X_k}{n}\right) = \frac{V(X_1)}{n} \xrightarrow{n \rightarrow +\infty} 0$; ce qui permet de conclure. □

Théorème : Loi faible des grands nombres

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables indépendantes, identiquement distribuées et intégrable. Alors $\frac{1}{n} \sum_{1 \leq k \leq n} X_k$ converge en probabilité vers $E(X_1)$.

Démonstration. Soit $(X_n)_{n \in \mathbb{N}}$ une telle suite de variables indépendantes.

L'indépendance des X_n est équivalente à l'indépendance des $\tilde{X}_n = X_n - E(X_0)$. On peut utiliser une caractérisation par la fonction de répartition, par exemple, pour prouver cela.

Et dire que $\overline{X_n}$ converge en loi vers $E(X_1)$ est équivalent à dire que $\widetilde{\overline{X_n}}$ converge en loi vers 0.

On peut donc, supposer sans nuire à la généralité du problème que les X_n sont centrés.

Pour tout nombre $M > 0$, on pose alors

$$\begin{aligned} Y_n &= X_n \mathbb{1}_{|X_n| > M} - E(X_n \mathbb{1}_{|X_n| > M}) \\ Z_n &= X_n \mathbb{1}_{|X_n| \leq M} - E(X_n \mathbb{1}_{|X_n| \leq M}) \end{aligned}$$

On a alors, par construction, $X_n = Y_n + Z_n$ et les Y_n et Z_n sont centrées.

D'autre part, $|Z_n| \leq 2M$ donc Z_n est L^∞ .

On va maintenant utiliser une technique déjà éprouvée lors de raisonnements sur la critère de Cauchy en probabilité.

Soit $\delta > 0$. L'inclusion de $\{|\overline{X_n}| > \delta\}$ dans $\{|\overline{Y_n}| > \frac{\delta}{2}\} \cup \{|\overline{Z_n}| > \frac{\delta}{2}\}$ donne l'inégalité :

$$P(|\overline{X_n}| > \delta) \leq P(|\overline{Y_n}| > \frac{\delta}{2}) + P(|\overline{Z_n}| > \frac{\delta}{2})$$

Commençons par contrôler la valeur de $P(|\overline{X_n}| > \frac{\delta}{2})$.

Pour tout $\varepsilon > 0$ et pour tout n , la variable $T_n = |X_n| \mathbb{1}_{|X_n| > M}$ est d'espérance inférieure à $\frac{\delta \varepsilon}{8}$ quand M est assez grand.

Ce résultat s'obtient en analysant la croissance de $E(|X_n| \mathbb{1}_{|X_n| \leq M})$ vers $E(|X_n|)$ quand M tend vers l'infini.

On obtient ainsi, en appliquant les inégalités triangulaire et de Markov, pour cette valeur de M assez grande :

$$P(|\overline{Y_n}| > \frac{\delta}{2}) \leq \frac{2E(|\overline{Y_n}|)}{\delta} \leq \frac{2E(|Y_n|)}{\delta} \leq \frac{4E(T_n)}{\delta} \leq \frac{\varepsilon}{2}$$

Maintenant, on peut contrôler la valeur de $P(|\overline{Z_n}| > \frac{\delta}{2})$. Comme la variable est centrée et est L^2 , on peut utiliser le lemme 9.2.2 pour conclure. \square

9.2.3 Un contre exemple

Soit une suite de variables aléatoires indépendantes X_n à valeurs dans $\{0; 1\}$ telles que $P(X_n = 0) = 1 - \frac{1}{n}$ et $P(X_n = 1) = \frac{1}{n}$.

Cette suite converge en probabilité vers 0. En effet, pour tout $\varepsilon \in]0; 1[$,

$$P(X_n > \varepsilon) = \frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0$$

Pourtant cette suite ne converge pas presque sûrement vers 0.

Soit ω un événement tel que $\lim_{n \rightarrow +\infty} X_n(\omega) = 0$.

Il existe N tel que pour tout $n \geq N$, $|X_n(\omega) - 0| < \frac{1}{2}$. Cela signifie que, pour tout $n \geq N$, $X_n(\omega) = 0$.

Montrons maintenant que cet événement appartient à un ensemble de probabilité nulle.

Considérons $B_N = \bigcap_{n \geq N} \{\omega / X_n(\omega) = 0\}$.

Pour tout M , la probabilité de $\bigcap_{M+N \leq n \leq N} \{\omega / X_n(\omega) = 0\}$ vaut $\prod_{N+M \leq n \leq N} \left(1 - \frac{1}{n}\right)$. Or ce produit est télescopique et vaut :

$$\prod_{N+M \leq n \leq N} \left(1 - \frac{1}{n}\right) = \frac{N-1}{N+M}$$

Pour M tendant vers l'infini, la probabilité de $\bigcap_{M+N \leq n \leq N} \{\omega / X_n(\omega) = 0\}$ tend vers 0 et, par convergence monotone, on peut conclure quant à la probabilité de B_N : elle est nulle.

9.3 Loi forte des grands nombres

9.3.1 Version L^2

Théorème :

Soit (X_n) une suite de variables aléatoires indépendantes identiquement distribuées et de carré intégrables.

Alors la suite définie pour tout $n \geq 1$ par $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ converge presque sûrement vers $E(X_1)$.

Dans toute la suite, posons $m = E(X_1)$.

(Remarque 5) Remarquons que $E(\bar{X}_n) = m$ et $V(\bar{X}_n) = \frac{1}{n^2} n V(X_1) = \frac{\sigma^2}{n}$ avec $\sigma^2 = V(X_1)$, car les X_n sont indépendants.

(Remarque 6) On peut centrer le problème en posant, pour tout n , $Z_n = X_n - m$. En effet, (\bar{X}_n) converge vers m si et seulement si (\bar{Z}_n) converge vers 0 et d'autre part les X_n sont indépendants si et seulement si les Z_n le sont (on peut le montrer à l'aide des classes monotones par exemple). Enfin, l'hypothèse « identiquement distribuée et de carré intégrable » est également équivalente.

La preuve se fait donc sur des variables centrées en exploitant éventuellement l'inégalité de Bienaymé-Tchebychev.

Démonstration. Pour tout n et pour tout $m > 0$, on a $P(|\bar{X}_{n^2}| > m) \leq \frac{\sigma^2}{n^2 m^2}$.

En particulier, $\sum_{n \in \mathbb{N}} P(|\bar{X}_{n^2}| > m) < +\infty$, ce qui donne, d'après le lemme de Borel-Cantelli, par passage au complémentaire :

$$P\left(\bigcup_p \bigcap_{n \geq p} \{|\bar{X}_{n^2}| \leq m\}\right) = 1$$

En exploitant le théorème de convergence monotone, on en déduit :

$$P\left(\bigcap_{\substack{m>0 \\ m \in \mathbb{Q}}} \bigcup_p \bigcap_{n \geq p} \{|\bar{X}_{n^2}| \leq m\}\right) = 1$$

Ainsi, la suite extraite (\bar{X}_{n^2}) converge presque sûrement vers 0.

Il s'agit de montrer que (\bar{X}_n) ne s'éloigne jamais trop de cette suite extraite.

Soit ainsi $n \in \mathbb{N}^*$, et $k = \lfloor \sqrt{n} \rfloor$, de sorte que $k^2 \leq n < (k+1)^2$. On va contrôler $E((\bar{X}_n - \bar{X}_{(k+1)^2})^2)$:

$$\begin{aligned} E((\bar{X}_n - \bar{X}_{(k+1)^2})^2) &= E(\bar{X}_n^2) + E(\bar{X}_{(k+1)^2}^2) - 2E(\bar{X}_n \bar{X}_{(k+1)^2}) \\ &= E(\bar{X}_n^2) + E(\bar{X}_{(k+1)^2}^2) - 2 \frac{1}{n(k+1)^2} \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq (k+1)^2}} E(X_i X_j) \end{aligned}$$

Toujours en raison de l'indépendance des variables, pour $i \neq j$, $E(X_i X_j) = 0$. On obtient donc :

$$E\left((\overline{X}_n - \overline{X}_{(k+1)^2})^2\right) = \left(\frac{1}{n} + \frac{1}{(k+1)^2} - \frac{2n}{n(k+1)^2}\right)\sigma^2 = \left(\frac{1}{n} - \frac{1}{(k+1)^2}\right)\sigma^2$$

Majorons maintenant, $\frac{1}{n} - \frac{1}{(k+1)^2} \leq \frac{1}{k^2} - \frac{1}{(k+1)^2} = \frac{2k+1}{k^2(k+1)^2} \leq \frac{2}{k^3}$. Or, on a $\frac{2}{k^3} \sim \frac{2}{n^{3/2}}$.

Fixons un rationnel $m > 0$. En notant, pour tout n , $A_n = \{|\overline{X}_n - \overline{X}_{(k+1)^2}| > m\}$, on a, par application de l'inégalité de Bienaymé-Tchebychev et du lemme de Borel-Cantelli :

$$P\left(\bigcap_p \bigcup_{n \geq p} A_n\right) = 0$$

En particulier, par passage au complémentaire et application du théorème de convergence monotone, on en déduit que :

$$P\left(\bigcap_{\substack{m>0 \\ m \in \mathbb{Q}}} \bigcup_p \bigcap_{n \geq p} \{|\overline{X}_n - \overline{X}_{(k+1)^2}| \leq m\}\right) = 1$$

L'intersection des événements $\bigcap_{\substack{m>0 \\ m \in \mathbb{Q}}} \bigcup_p \bigcap_{n \geq p} \{|\overline{X}_n - \overline{X}_{(k+1)^2}| \leq m\}$ et $\bigcap_{\substack{m>0 \\ m \in \mathbb{Q}}} \bigcup_p \bigcap_{n \geq p} \{|\overline{X}_{n^2}| \leq m\}$ est donc de probabilité

1, ce qui permet de conclure ! □

9.3.2 Version L^1 (Kolmogorov)

Contrairement à la version L^2 , la version de Kolmogorov établit une équivalence entre la convergence en moyenne et l'intégrabilité, et il précise que, en l'absence de convergence, la moyenne n'est presque sûrement pas bornée.

Théorème : Loi forte des grands nombres

Soit (X_n) une suite de variables aléatoires identiquement distribuées. On note, pour tout n , $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$.

Alors X_1 est intégrable si et seulement si la suite (\overline{X}_n) converge presque sûrement.

De plus, dans ce cas, la limite des (\overline{X}_n) est unique et vaut $E(X_1)$ et la convergence a aussi lieu dans L^1 .

Dans le cas contraire, la suite (\overline{X}_n) n'est presque sûrement pas bornée.

Avant d'aller plus loin, quelques remarques :

(Remarque 7) Les événements (\overline{X}_n) n'est pas bornée ou (\overline{X}_n) tend vers ℓ sont asymptotiques car ils ne peuvent s'écrire qu'en fonction d'un nombre infini de valeurs des (X_n) .

(Remarque 8) Dans le cas particulier où les (X_n) sont des variables de Bernoulli de paramètre p , cette loi entraîne que (\overline{X}_n) tend vers p . Ainsi la loi forte des grands nombres justifie qu'une probabilité correspond à la limite d'une fréquence.

(Remarque 9) Ce théorème ne nous précise pas avec quelle vitesse « probabiliste », la moyenne tend vers l'espérance. Pour avoir une idée de la distribution de la moyenne, il faut se référer au théorème centrale limite.

Nous allons maintenant démontrer ce théorème.

Démonstration. Commençons par le sens réciproque. Supposons que (\overline{X}_n) converge presque sûrement. Ici, nous allons prouver que $\left(\frac{|X_n|}{n-1}\right)_{n \geq 2}$ est bornée, ce qui permettra de contrôler la valeur de $E(|X_1|)$ par majoration. En effet :

$$E(|X_1|) \leq \sum_{n \in \mathbb{N}^*} nP(|X_1| \in [n-1; n[)$$

Or :

$$\begin{aligned} \sum_{n \in \mathbb{N}^*} n P(|X_1| \in [n-1; n]) &= \sum_{n \in \mathbb{N}^*} \sum_{k \in \mathbb{N}^*} \mathbb{1}_{k \leq n} P(|X_k| \in [n-1; n]) \text{ car les } (X_n) \text{ sont i.i.d.} \\ &= \sum_{k \in \mathbb{N}^*} \sum_{n \in \mathbb{N}^*} \mathbb{1}_{k \leq n} P(|X_k| \in [n-1; n]) = \sum_{k \in \mathbb{N}^*} P(|X_k| \in [k-1; +\infty[) \end{aligned}$$

Nous allons maintenant prouver que cette dernière somme n'est pas infinie en exploitant la loi du tout ou rien.

Notons que, pour tout $n \geq 2$, $\frac{X_n}{n-1} = \frac{n}{n-1} \bar{X}_n - \bar{X}_{n-1}$. Presque sûrement, on obtient ainsi $\lim_{n \rightarrow \infty} \frac{X_n}{n-1} = 0$

Considérons maintenant l'évènement asymptotique $\bigcap_{n \geq 2} \bigcup_{k \geq n} \left\{ \frac{|X_k|}{k-1} \geq 1 \right\}$. Sa probabilité est nécessairement nulle en raison de la limite de $\left(\frac{X_n}{n-1} \right)$ établie plus haut. La version améliorée du théorème de Borel-Cantelli (voir page 107) entraîne que $\sum_{k \in \mathbb{N}^*} P(|X_k| \in [k-1; +\infty[) < +\infty$ et, par suite :

$$E(|X_1|) < +\infty$$

On va maintenant prouver le sens direct. Supposons que X_1 est intégrable. On peut supposer que $E(X_1) = 0$ pour simplifier, quitte à centrer les variables (X_n) , ce qui ne change pas la généralité du problème. Notre objectif va être de prouver que, pour tout $\eta > 0$, $\limsup_n \sup_{k \geq n} \bar{X}_k \leq \eta$, ce qui permettra de prouver que $\limsup_n \bar{X}_n \leq 0$.

Pour ce faire, nous allons prouver par l'absurde que, pour tout $\eta > 0$, la suite $\left(\sum_{i=1}^k X_i - k\eta \right)$ est majorée. Fixons $\eta > 0$ et considérons, pour tout n , les suites

$$\begin{cases} A_{n, \eta} = \sup_{1 \leq k \leq n} \left(\sum_{i=1}^k X_i - k\eta \right) \\ B_{n, \eta} = \sup_{1 \leq k \leq n} \left(\sum_{i=2}^{k+1} X_i - k\eta \right) \end{cases}$$

Ces deux suites sont croissantes et, puisque les (X_n) sont i.i.d., elles ont la même loi. Remarquons maintenant que :

$$B_{n, \eta} + (X_1 - \eta) = \sup_{1 \leq k \leq n} \left(\sum_{i=1}^{k+1} X_i - (k+1)\eta \right) = \sup_{2 \leq k \leq n+1} \left(\sum_{i=1}^k X_i - k\eta \right)$$

Et ainsi, lorsque $(X_1 - \eta) > B_{n, \eta} + (X_1 - \eta)$, on a $A_{n+1, \eta} = X_1 - \eta$ et, dans le cas contraire, $A_{n+1, \eta} = B_{n, \eta} + (X_1 - \eta)$, ce qui permet d'écrire :

$$A_{n+1, \eta} = B_{n, \eta} + \max(X_1 - \eta; X_1 - \eta - B_{n, \eta})$$

En raison de la remarque précédente sur le sens de variation des suites et le fait qu'elles sont identiquement distribuées, on obtient, pour tout n , en passant à l'espérance :

$$E(\max(X_1 - \eta; X_1 - \eta - B_{n, \eta})) \geq 0$$

Remarquons que l'évènement « la suite $\left(\sum_{i=1}^k X_i - k\eta \right)$ n'est pas majorée » est asymptotique. Donc sa probabilité est nulle ou bien égale à 1. Supposons qu'elle est égale à 1. La suite $B_{n, \eta}$ tend donc presque sûrement vers $-\infty$, ce qui signifie que $\max(X_1 - \eta; X_1 - \eta - B_{n, \eta})$ tend presque sûrement vers $X_1 - \eta$. Or nous sommes dans les conditions d'application du théorème de convergence dominée puisque, pour tout n , $|\max(X_1 - \eta; X_1 - \eta - B_{n, \eta})| \leq \max(|X_1 - \eta|; |X_1 - \eta - B_{1, \eta}|)$. Finalement, on obtient, par passage à la limite sur n :

$$E(X_1) = 0 \geq \eta$$

Ce qui constitue une contradiction. On obtient que la suite $\left(\sum_{i=1}^k X_i - k\eta\right)$ est presque sûrement majorée et nous permet d'écrire :

$$\limsup_n \frac{1}{k} \sum_{i=1}^k X_i \leq \eta$$

Et comme $\eta > 0$ est arbitraire, on obtient :

$$\limsup_n \frac{1}{k} \sum_{i=1}^k X_i \leq 0$$

On prouve de même que $\liminf_n \frac{1}{k} \sum_{i=1}^k X_i \geq 0$, ce qui permet de conclure.

Reste à prouver que la convergence de (\overline{X}_n) vers $E(X_1)$ a également lieu dans L^1 .

Soit $\varepsilon > 0$ et $M > 0$ tel que $E(|X_1 - \mathbb{1}_{|X_1| \leq M} X_1|) < \frac{\varepsilon}{3}$.

On pose, pour tout n , $Z_n = \mathbb{1}_{|X_n| \leq M} X_n$ et $\overline{Z}_n = \frac{1}{n} \sum_{k=1}^n Z_k$. L'inégalité triangulaire donne :

$$E(|\overline{X}_n - X_1|) \leq E(|\overline{X}_n - \overline{Z}_n|) + E(|\overline{Z}_n - E(Z_1)|) + E(|Z_1 - X_1|) < \frac{2\varepsilon}{3} + E(|\overline{Z}_n - E(Z_1)|)$$

On obtient ainsi :

$$\limsup E(|\overline{X}_n - X_1|) < \varepsilon$$

Et comme ε est arbitraire, on obtient le résultat escompté. \square

Nous allons réaliser une seconde démonstration du sens direct.

Démonstration. Puisque les (X_n) sont intégrables, on peut mener le raisonnement indépendamment sur les parties positives et négatives. On peut donc supposer sans risque que les (X_n) sont positives.

Pour tout n , on va poser $Y_n = X_n \mathbb{1}_{[0; n[}$, $T_n = \sum_{k=1}^n Y_k$ et $\overline{Y}_n = \frac{1}{n} T_n$. Notons que les Y_n sont indépendantes et toutes L^∞ donc en particulier L^2 .

On va commencer par montrer que presque sûrement, $X_n = Y_n$ à partir d'un certain rang, c'est à dire que $P\left(\bigcup_{n \geq n} \bigcap \{X_k = Y_k\}\right) = 1$, ce qui revient à prouver, par passage au complémentaire, que

$$P\left(\bigcap_n \bigcup_{k \geq n} \{X_k \neq Y_k\}\right) = 0$$

À partir de ce premier résultat, on en déduira que $\overline{X}_n - \overline{Y}_n$ tend presque sûrement vers 0.

Exploitions le lemme de Borel-Cantelli pour arriver à nos fins. Pour tout k , $\{X_k \neq Y_k\} = \{X_k \in [k; +\infty[)\}$. En particulier :

$$\begin{aligned} \sum_n P(X_n \neq Y_n) &= \sum_n P(X_n \in [n; +\infty[)) \\ &= \sum_n \sum_k \mathbb{1}_{k \geq n} P(X_n \in [k; k+1[)) \\ &= \sum_n \sum_k \mathbb{1}_{k \geq n} P(X_k \in [k; k+1[)) \\ &= \sum_k \sum_n \mathbb{1}_{k \geq n} P(X_k \in [k; k+1[)) = \sum_k k P(X_k \in [k; k+1[)) \leq E(X_1) < +\infty \end{aligned}$$

Pour presque tout ω , on en déduit qu'il existe n_ω tel que, pour tout $k \geq n_\omega$, $X_k(\omega) = Y_k(\omega)$ et en particulier :

$$\overline{X}_k(\omega) - \overline{Y}_k(\omega) = \frac{1}{k} \sum_{i=1}^{n_\omega-1} (X_i(\omega) - Y_i(\omega)) \xrightarrow[k \rightarrow +\infty]{} 0$$

On va maintenant prouver :

- que l'on peut extraire une sous-suite de $(\bar{Y}_n - E(\bar{Y}_n))$ qui converge presque sûrement vers 0 ;
- que $\lim E(\bar{Y}_n) = E(X_1)$;
- que \bar{Y}_n tend en fait presque sûrement vers $E(X_1)$.

Exploitions la caractérisation sur la convergence presque sûre, page 110, pour fabriquer la suite extraite. Soit $\varepsilon > 0$. Chacune des Y_n est de carré intégrable et donc, par l'inégalité de Bienaymé-Chebychev :

$$P(|\bar{Y}_n - E(\bar{Y}_n)| \geq \varepsilon) \leq \frac{V(\bar{Y}_n)}{\varepsilon^2}$$

En raison de l'indépendance des (Y_n) , on a :

$$\begin{aligned} V(\bar{Y}_n) &= \frac{1}{n^2} \sum_{k=1}^n V(Y_k) \\ &\leq \frac{1}{n^2} \sum_{k=1}^n E(Y_k^2) = \frac{1}{n^2} \sum_{k=1}^n \int_{[0; k[} t^2 dP_X(t) \text{ avec } P_X \text{ la probabilité transportée par } X_1 \end{aligned}$$

Pour fabriquer la suite extraite, fixons $\alpha > 1$ et posons pour tout n , $p_n = [\alpha^n]$. En reprenant ce qui vient d'être fait, on a :

$$\sum_n P(|\bar{Y}_{p_n} - E(\bar{Y}_{p_n})| \geq \varepsilon) \leq \sum_n \frac{1}{p_n^2} \sum_{k=1}^{p_n} \int_{[0; k[} t^2 dP_X(t)$$

Exploitions maintenant les indicatrices pour poursuivre la majoration de cette somme :

$$\begin{aligned} \sum_n \frac{1}{p_n^2} \sum_{k=1}^{p_n} \int_{[0; k[} t^2 dP_X(t) &= \sum_n \sum_k \frac{1}{p_n^2} \mathbb{1}_{k \leq p_n} \int_{[0; k[} t^2 dP_X(t) \\ &= \sum_n \sum_k \sum_i \frac{1}{p_n^2} \mathbb{1}_{k \leq p_n} \mathbb{1}_{i \leq k} \int_{[i-1; i[} t^2 dP_X(t) \\ &\leq \sum_n \sum_k \sum_i \frac{1}{p_n^2} \mathbb{1}_{k \leq p_n} \mathbb{1}_{i \leq k} i^2 P_X([i-1; i]) = \sum_i i^2 P_X([i-1; i]) \sum_n \mathbb{1}_{i \leq p_n} \frac{p_n - i + 1}{p_n^2} \end{aligned}$$

Pour conclure, on va travailler sur la dernière somme :

$$\sum_n \mathbb{1}_{i \leq p_n} \frac{p_n - i + 1}{p_n^2} \leq \sum_n \mathbb{1}_{i \leq p_n} \frac{1}{p_n} = \frac{1}{i} \sum_n \mathbb{1}_{i \leq p_n} \frac{i}{p_n}$$

Soit maintenant n_0 le plus petit entier tel que $\alpha^{n_0} \geq i$. On remarque que, pour tout n , $p_n \geq i$ entraîne $n \geq n_0$. Par construction des p_n , on a aussi $\frac{1}{p_n} \leq \frac{1}{\alpha^n}$. Ces deux réflexions permettent d'écrire :

$$\sum_n \mathbb{1}_{i \leq p_n} \frac{p_n - i + 1}{p_n^2} \leq \frac{1}{i} \sum_{n \geq n_0} \frac{1}{\alpha^{n-n_0}} = \frac{1}{i} c_\alpha \text{ avec } c_\alpha = \frac{1}{1 - \frac{1}{\alpha}}$$

Revenons à notre majoration de départ et rassemblons les morceaux :

$$\sum_n P(|\bar{Y}_{p_n} - E(\bar{Y}_{p_n})| \geq \varepsilon) \leq \sum_i i^2 P_X([i-1; i]) \frac{1}{i} c_\alpha = c_\alpha \sum_i i P_X([i-1; i]) \leq c_\alpha (E(X_1) + 1) < +\infty$$

Cela nous prouve que, presque sûrement, $\lim_n (\bar{Y}_{p_n} - E(\bar{Y}_{p_n})) = 0$.

Par le théorème de Césaro, on prouve facilement que $\lim_n E(\bar{Y}_n) = E(X_1)$

Reste à prouver que \bar{Y}_n ne s'éloigne pas trop de \bar{Y}_{p_n} . Considérons ainsi un entier n suffisamment grand. Il existe ainsi un unique l tel que $p_l \leq n < p_{l+1}$. Sachant que les Y_n sont des v.a.r. positives, on obtient l'encadrement :

$$\frac{T_{p_l}}{p_{l+1}} \leq \bar{Y}_n \leq \frac{T_{p_{l+1}}}{p_l} \iff \frac{p_l}{p_{l+1}} \bar{Y}_{p_l} \leq \bar{Y}_n \leq \frac{p_{l+1}}{p_l} \bar{Y}_{p_{l+1}}$$

Par passage à la limite sur n , cela donne, presque sûrement :

$$\frac{1}{\alpha} E(X_1) \leq \liminf \bar{Y}_n \leq \limsup \bar{Y}_n \leq \alpha E(X_1)$$

Et puisque α est arbitraire, on en déduit que (\bar{Y}_n) tend presque sûrement vers $E(X_1)$.

Mais puisqu'on a prouvé précédemment que $(\bar{X}_n - \bar{Y}_n)$ tend presque sûrement vers 0, on obtient le résultat escompté. \square

9.4 Fonctions caractéristiques

9.4.1 Définition et premières propriétés

Définition : Fonction caractéristique

Soit X une variable aléatoire réelle. On pose :

$$\psi[X] : \omega \mapsto E[e^{i\omega X}]$$

C'est donc la transformée de Fourier de la mesure P_X .

(Remarque 10) D'après ce que l'on sait des transformées de Fourier, on sait que la fonction caractéristique d'une v.a.r. est \mathcal{C}_0 .

(Remarque 11) Sa valeur en 0 est 1 (masse totale).

Proposition : Fonction caractéristique d'une somme de v.a.r. indépendantes

Soient X et Y deux v.a.r. indépendantes. Alors :

$$\psi[X + Y] = \psi[X]\psi[Y]$$

Démonstration. On sait que la probabilité transportée par $X + Y$ est la convolution de P_X et P_Y . Le résultat final s'obtient à partir de ce que l'on sait des transformées de Fourier. \square

Proposition : Fonction caractéristique d'une loi normale

Si N suit une loi normale centrée réduite alors $\psi[N] = g$ où $g : \omega \mapsto e^{-\omega^2/2}$.

Démonstration. On a déjà prouvé ce résultat mais on va refaire une démonstration en passant par l'analyse complexe. Pour tout ω :

$$\begin{aligned} \psi[N](\omega) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2 + i\omega x} d\lambda(x) \\ &= \frac{1}{\sqrt{2\pi}} \lim_{L \rightarrow +\infty} \int_{\gamma_L} e^{-z^2/2 + i\omega z} dz \end{aligned}$$

On a posé γ_L le chemin orienté qui relie les points d'affixe $-L$ et L dans le plan complexe. Travaillons sur cette dernière intégrale et réalisons un changement de variable :

$$\begin{aligned} \int_{\gamma_L} e^{-z^2/2 + i\omega z} dz &= \int_{\gamma_L} e^{-\frac{1}{2}(z - i\omega)^2 - \frac{1}{2}\omega^2} dz \\ &= e^{-\frac{1}{2}\omega^2} \int_{\tilde{\gamma}_L} e^{-s^2/2} ds \end{aligned}$$

On a posé $\tilde{\gamma}_L$ le chemin qui relie les points d'affixe $-L - i\omega$ et $L - i\omega$. On va maintenant décomposer ce chemin en trois :

- un chemin entre les points $-L - i\omega$ et $-L$, noté γ_L^- , et de longueur $|\omega|$;
- le chemin entre les points $-L$ et L , noté γ_L ;
- le chemin entre les points L et $L - i\omega$, noté γ_L^+ , et de longueur $|\omega|$.

Ainsi :

$$\int_{\gamma_L} e^{-z^2/2+i\omega z} dz = e^{-\frac{1}{2}\omega^2} \left[\int_{\gamma_L^-} e^{-s^2/2} ds + \int_{\gamma_L} e^{-s^2/2} ds + \int_{\gamma_L^+} e^{-s^2/2} ds \right]$$

On obtient aisément les majorations suivantes, par un raisonnement sur les modules et l'inégalité triangulaire :

$$\left| \int_{\gamma_L^-} e^{-s^2/2} ds \right| \leq |\omega| e^{-L^2/2} \quad \text{et} \quad \left| \int_{\gamma_L^+} e^{-s^2/2} ds \right| \leq |\omega| e^{-L^2/2}$$

En particulier, pour $L \rightarrow +\infty$, ces deux intégrales tendent vers 0. Quant à l'intégrale sur γ_L , elle tend vers $\sqrt{2\pi}$, ce qui donne bien :

$$\psi[N](\omega) = e^{-\frac{1}{2}\omega^2}$$

□

9.4.2 Théorème Central Limite

Proposition : Fonctions caractéristiques et moments

Soit X une variable aléatoire de classe L^p avec $p \in \mathbb{N}$. Alors, pour tout $k \leq p$, $\psi[X]^{(k)}$ existe, est uniformément continue, et pour tout $\omega \in \mathbb{R}$:

$$\psi[X]^{(k)}(\omega) = i^k \int t^k e^{it\omega} dP_X(t)$$

En particulier, $\psi[X]^{(k)}(0) = i^k E(X^k)$

Démonstration. L'existence est une conséquence assez simple du théorème de convergence dominée de Lebesgue et du théorème de transfert.

Considérons ω_0 un réel et montrons que $\psi[X]^{(k)}$ est continue en ω_0 . Soit $\varepsilon > 0$.

On choisit $A > 0$ de telle sorte que $\int_{\mathbb{R} \setminus [-A; A]} |t|^k dP_X(t) < \frac{\varepsilon}{3}$.

Pour tout ω_1 réel, on a :

$$\left| \psi[X]^{(k)}(\omega_1) - \psi[X]^{(k)}(\omega_0) \right| < \frac{2\varepsilon}{3} + A|\omega_1 - \omega_0| \int_{[-A; A]} |t|^k dP_X(t) \leq \frac{2\varepsilon}{3} + A|\omega_1 - \omega_0| E(|X|^k)$$

En effet, $\left| \int_{[-A; A]} t^k (e^{it\omega_1} - e^{it\omega_0}) dP_X(t) \right| \leq \int_{[-A; A]} |t|^k |e^{it\omega_1} - e^{it\omega_0}| dP_X(t)$ et, pour tout $t \in [-A; A]$, $|e^{it\omega_1} - e^{it\omega_0}| \leq A|\omega_1 - \omega_0|$. En considérant maintenant ω_1 tel que $|\omega_1 - \omega_0| < \frac{\varepsilon}{3AE(|X|^k)}$, on obtient :

$$\left| \psi[X]^{(k)}(\omega_1) - \psi[X]^{(k)}(\omega_0) \right| < \varepsilon$$

Notons que le choix de A et donc de la jauge $\frac{\varepsilon}{3AE(|X|^k)}$ ne dépend pas de ω_0 , ce qui prouve l'uniforme continuité de $\psi[X]^{(k)}$. □

Corollaire : Développement limité en 0 et moments

On reprend les mêmes hypothèses. Alors $\psi[X]$ possède en 0 un développement limité d'ordre n .

Pour tout ω :

$$\psi[X](\omega) = \sum_{k=0}^n \frac{i^k E(X^k)}{k!} \omega^k + o(\omega^n)$$

Démonstration. Ce résultat est un corollaire de ce qui précède, de la formule de Taylor-Lagrange, en exploitant le fait que $\psi[X]^{(n)}$ est continue. \square

Écrivons maintenant le théorème central limite.

Théorème : Théorème de la limite centrale

Soit $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables aléatoires de classe L^2 , indépendantes et identiquement distribuées. On pose $\mu = E(X_1)$ et $\sigma = \sqrt{V(X_1)}$.

Alors la suite $\left(\frac{S_n - n\mu}{\sqrt{n}\sigma}\right)$ converge (en loi) vers une loi normale centrée réduite, en posant, pour tout $n \geq 1$,

$$S_n = \sum_{k=1}^n X_k$$

On va exploiter la caractérisation de la convergence étroite par les transformées de Fourier.

Démonstration. Considérons les variables aléatoires (Z_k) définies pour tout $k \in \mathbb{N}^*$ par $Z_k = \frac{X_k - \mu}{\sigma}$. Ces variables sont centrées réduites, indépendantes et identiquement distribuées. De plus, pour tout $n \geq 1$,

$$\sum_{k=1}^n Z_k = \frac{S_n - n\mu}{\sigma}.$$

On peut maintenant supposer que les (X_k) sont également centrées et réduites sans nuire à la généralité. Il s'agit de montrer que $\left(\frac{S_n}{\sqrt{n}}\right)$ converge vers une loi normale centrée réduite. Or la transformée de Fourier de cette variable vaut, pour tout ω , d'après le théorème du transfert :

$$\psi\left[\frac{S_n}{\sqrt{n}}\right](\omega) = \int e^{i\omega \frac{t}{\sqrt{n}}} dP_{S_n}(t) = \psi[S_n]\left(\frac{\omega}{\sqrt{n}}\right)$$

D'après ce que nous savons des fonctions caractéristiques de sommes de variables aléatoires indépendantes, on obtient ainsi, sachant que les variables sont identiquement distribuées :

$$\psi\left[\frac{S_n}{\sqrt{n}}\right](\omega) = \prod_{k=1}^n \psi[X_k]\left(\frac{\omega}{\sqrt{n}}\right) = \psi[X_1]\left(\frac{\omega}{\sqrt{n}}\right)^n$$

Exploitions maintenant ce qui précède sur le développement limité en 0. Il existe une fonction à valeurs complexes ε , continue en 0 et telle que $\varepsilon(0) = 0$ et qui vérifie :

$$\psi\left[\frac{S_n}{\sqrt{n}}\right](\omega) = \left(1 - \frac{\omega^2}{2n} + \frac{\omega}{n}\varepsilon\left(\frac{\omega}{\sqrt{n}}\right)\right)^n$$

Considérons maintenant la série entière définie sur le disque ouvert de convergence de rayon 1, $\ln(1+z) = \sum_{n \in \mathbb{N}^*} \frac{(-1)^{n+1} z^n}{n}$.

Pour tout ω , il existe N tel que, pour tout $n \geq N$, $-\frac{\omega^2}{2n} + \frac{\omega}{n}\varepsilon\left(\frac{\omega}{\sqrt{n}}\right) \in D$, où D est le disque de convergence. En particulier, on obtient alors :

$$\left(1 - \frac{\omega^2}{2n} + \frac{\omega}{n}\varepsilon\left(\frac{\omega}{\sqrt{n}}\right)\right)^n = e^{n \ln\left(1 - \frac{\omega^2}{2n} + \frac{\omega}{n}\varepsilon\left(\frac{\omega}{\sqrt{n}}\right)\right)} = e^{-\frac{\omega^2}{2} + \omega^2 \tilde{\varepsilon}\left(\frac{\omega}{\sqrt{n}}\right)}$$

Avec $\tilde{\varepsilon}$ une autre fonction à valeur complexes, continue en 0 avec $\tilde{\varepsilon}(0) = 0$. Par passage à la limite, cela donne :

$$\psi\left[\frac{S_n}{\sqrt{n}}\right](\omega) \xrightarrow{n \rightarrow +\infty} e^{-\frac{\omega^2}{2}}$$

On retrouve ici la fonction caractéristique d'une loi normale centrée réduite et on peut ainsi conclure. \square

9.5 Quelques lois

9.5.1 Lois discrètes

9.5.1.1 Fonctions génératrices

Définition : Fonction génératrice

Soit X une variable aléatoire prenant des valeurs entières.

On définit la fonction génératrice de X comme étant la série formelle :

$$g_x[S] = E(S^X) = \sum_{k \in \mathbb{N}} P(X = k) \times S^k$$

On vérifie $g_x[1] = 1$.

En particulier le rayon de convergence de cette série est au moins égal à 1.

Démonstration. Évident d'après ce que l'on sait sur les séries entières. □

Proposition : Espérance, variance et fonction génératrice

Avec les notations précédentes :

$$\begin{aligned} E(X) &= g'_x[1] \\ V(X) &= g''_x[1] + g'_x[1] - g'_x[1]^2 \end{aligned}$$

Démonstration.

$$\begin{aligned} g'_x[S] &= \sum_{k \in \mathbb{N}^*} k P(X = k) S^{k-1} \\ g''_x[S] &= \sum_{k \geq 2} k(k-1) P(X = k) S^{k-2} \\ &= \sum_{k \geq 2} k^2 P(X = k) S^{k-1} - \sum_{k \geq 2} k P(X = k) S^{k-1} \\ &= \sum_{k \geq 1} k^2 P(X = k) S^{k-1} - \sum_{k \geq 1} k P(X = k) S^{k-1} \end{aligned}$$

On en déduit que $g'_x[1] = E(X)$ et $g''_x[1] = E(X^2) - E(X)$.

Ce qui donne bien $g''_x[1] + g'_x[1] - g'_x[1]^2 = E(X^2) - E(X)^2 = V(X)$. □

Enfin, une proposition portant sur les sommes de variables aléatoires et les sommes à nombre de termes aléatoires.

Proposition : somme de variables aléatoires discrètes indépendantes

Soient $(X_i)_{i \in \mathbb{N}^*}$ des variables aléatoires discrètes de mêmes lois et indépendantes.

Soit $m \in \mathbb{N}^*$ un nombre entier fixé et N une variable aléatoire à valeurs dans \mathbb{N}^* indépendante des $(X_i)_{i \in \mathbb{N}^*}$.

On note g_x la fonction génératrice de X_1 et g_n la fonction génératrice de N .

Alors la variable $S_m = \sum_{1 \leq k \leq m} X_k$ a pour fonction génératrice g_x^m .

Et la variable aléatoire $T = \sum_{1 \leq k \leq N} X_k$ a pour fonction génératrice $g_n \circ g_x$.

Démonstration. Soit $k \in \mathbb{N}$ quelconque.

$P(S_m = k) = P\left(\sum_{1 \leq i \leq m} X_i = k\right)$. Étudions l'évènement

$$E_{km} = \sum_{1 \leq i \leq m} X_i = k.$$

On a $E_{km} = \bigcup_{\substack{i_1, \dots, i_m \\ i_1 + \dots + i_m = k}} \{X_1 = i_1\} \cap \dots \cap \{X_m = i_m\}$.

On en déduit :

$$\begin{aligned} P(S_m = k) &= \sum_{\substack{i_1, \dots, i_m \\ i_1 + \dots + i_m = k}} P(X_1 = i_1) \times \dots \times P(X_m = i_m) \\ &= \sum_{\substack{i_1, \dots, i_m \\ i_1 + \dots + i_m = k}} a_{i_1} \times \dots \times a_{i_m} \end{aligned}$$

avec a_i qui sont les coefficients de g_x .

Donc, par définition, $P(S_m = k)$ correspond au terme de degré k de la série g_x^m .

Reste à prouver le dernier résultats.

$P(T = k) = P\left(\sum_{1 \leq i \leq N} X_i = k\right)$. On utilise la formule des probabilités totales :

$$\begin{aligned} P(T = k) &= \sum_{n \in \mathbb{N}} P\left(\sum_{1 \leq i \leq N} X_i = k \mid N = n\right) \times P(N = n) \\ &= \sum_{n \in \mathbb{N}} P\left(\sum_{1 \leq i \leq n} X_i = k\right) \times P(N = n) \end{aligned}$$

Notons (b_i) les coefficients de g_N et remarquons que $P\left(\sum_{1 \leq i \leq n} X_i = k\right)$ correspond au coefficient de degré k de la série g_x^n , noté $a_k^{(n)}$.

Nous avons donc :

$$P(T = k) = \sum_{n \in \mathbb{N}} a_k^{(n)} b_n \times P(N = n)$$

Cette expression correspond effectivement bien au coefficient de degré k de la série $g_n \circ g_x$. □

9.5.1.2 Loïs de Bernoulli et binomiale

Définition : Loi de Bernoulli, loi binomiale

On dit qu'une variable B suit une loi de Bernoulli de paramètre $p \in]0; 1[$ lorsque $P(X = 0) = 1 - p$ et $P(X = 1) = p$.

On dit qu'une variable X suit une loi binomiale de paramètres n et p lorsque X est la somme de n variables de Bernoulli indépendantes de paramètre p .

En particulier, on a alors, pour tout $k \in \llbracket 0; n \rrbracket$, $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Démonstration. La fonction génératrice de B est $g_b[S] = 1 - p + pS$ donc la fonction génératrice de X est

$$g_x[S] = (1 - p + pS)^n = \sum_{0 \leq k \leq n} \binom{n}{k} p^k (1 - p)^{n-k} S^k.$$

On obtient ainsi $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. □

Proposition : Espérance et variance

On reprend les mêmes notations.

$$E(B) = p$$

$$V(B) = p(1 - p)$$

$$E(X) = np$$

$$V(X) = np(1 - p)$$

Démonstration. Pour B , on applique la formule. Pour X , on utilise la linéarité de l'espérance et la linéarité de la variance dans le cas de somme de variables aléatoires indépendantes.

On peut aussi utiliser les fonctions génératrices... □

9.5.1.3 Loi géométrique et loi de poisson**Définition : Loi géométrique**

On dit qu'une variable T suit une loi géométrique de paramètre $\theta \in]0; 1[$ lorsque pour tout $k \in \mathbb{N}^*$, $P(T = k) = \theta \times (1 - \theta)^{k-1}$.



Cette loi donne le rang d'obtention du premier succès dans le cas d'une suite infinie de Bernoulli de paramètres θ .

Elle se prolonge dans le continu en loi exponentielle.

Proposition : Fonction génératrice, espérance, variance d'une loi géométrique

Avec les mêmes notations et hypothèses :

$$g_T[S] = \frac{\theta S}{1 - (1 - \theta)S}$$

$$E(T) = \frac{1}{\theta}$$

$$V(T) = \frac{1}{\theta^2} - \frac{1}{\theta}$$

Démonstration. On écrit la formule de la fonction génératrice :

$$\begin{aligned} g_t[S] &= \sum_{k \geq 1} \theta(1 - \theta)^{k-1} S^k \\ &= \theta S \sum_{k \geq 1} \theta((1 - \theta)S)^{k-1} \\ &= \frac{\theta S}{1 - (1 - \theta)S} \end{aligned}$$

L'espérance de T est donnée par :

$$\begin{aligned} E(T) &= \sum_{k \geq 1} k\theta(1 - \theta)^{k-1} \\ &= \theta \sum_{k \geq 1} k(1 - \theta)^{k-1} \end{aligned}$$

On reconnaît dans la somme, l'opposée de la dérivée de la série $\sum_{k \geq 0} (1 - \theta)^k = \frac{1}{\theta}$. On obtient ainsi

$$E(T) = \theta \times \frac{1}{\theta^2} = \frac{1}{\theta}$$

On calcule maintenant :

$$E(T^2) = \sum_{k \geq 1} k^2 \theta (1 - \theta)^{k-1}$$

En particulier :

$$\begin{aligned} E(T^2) - E(T) &= \sum_{k \geq 1} k(k-1) \theta (1 - \theta)^{k-1} \\ &= \theta(1 - \theta) \sum_{k \geq 2} k(k-1) \theta (1 - \theta)^{k-2} \end{aligned}$$

On reconnaît dans la somme la dérivée seconde de la série $\sum_{k \geq 0} (1 - \theta)^k = \frac{1}{\theta}$. On obtient donc

$$E(T^2) - E(T) = \theta(1 - \theta) \times \frac{2}{\theta^3} = \frac{2}{\theta^2} - \frac{2}{\theta}$$

Finalement, $E(T^2) = \frac{2}{\theta^2} - \frac{1}{\theta}$ et donc $V(T) = E(T^2) - E(T)^2 = \frac{1}{\theta^2} - \frac{1}{\theta}$. □

Définition : Loi de Poisson

On dit qu'une variable N suit une loi de Poisson de paramètre $\lambda > 0$ lorsque, pour tout $k \in \mathbb{N}$, $P(N = k) = e^{-\lambda} \frac{\lambda^k}{k!}$.

Proposition : Fonction génératrice, espérance, variance d'une loi de Poisson

Avec les mêmes notations et hypothèses :

$$\begin{aligned} g_n[S] &= e^{-\lambda + \lambda S} \\ E(N) &= \lambda \\ V(N) &= \lambda \end{aligned}$$

Démonstration. On calcule la fonction génératrice :

$$\begin{aligned} g_n[S] &= \sum_{k \geq 1} e^{-\lambda} \frac{\lambda^k}{k!} S^k \\ &= e^{-\lambda} \sum_{k \geq 1} \frac{(\lambda S)^k}{k!} S^k \\ &= e^{-\lambda} \times e^{\lambda S} = e^{-\lambda + \lambda S} \end{aligned}$$

On obtient $g'_n[S] = \lambda g_n[S]$ et $g''_n[S] = \lambda^2 g_n[S]$.

Ainsi, $E[N] = g'_n[1] = \lambda$ et $V[N] = g''_n[1] + E[N] - E[N]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$. □

Proposition : Convergence de binomiales vers Poisson

Soit $\lambda > 0$

Pour $n > [\lambda]$, on pose X_n une variable aléatoire suivant une loi binomiale de paramètres n et $\frac{\lambda}{n}$.
Alors la suite X_n converge en loi vers une variable N suivant une loi de Poisson de paramètre λ .

Démonstration. Pour tous entiers k et n :

$$\begin{aligned} P(X_n = k) &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \left(1 - \frac{\lambda}{n}\right)^{n-k} \times \frac{\lambda^k}{k!} \times \frac{n \times (n-1) \times \dots \times (n-k+1)}{n^k} \end{aligned}$$

On utilise un développement limité pour $\frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0$:

$$\left(1 - \frac{\lambda}{n}\right)^{n-k} = e^{(n-k)\ln(1-\frac{\lambda}{n})} = e^{-\lambda+o(1)}$$

D'autre part :

$$\frac{n \times (n-1) \times \dots \times (n-k+1)}{n^k} = 1 \times \left(1 - \frac{1}{n}\right) \times \left(1 - \frac{2}{n}\right) \times \dots \times \left(1 - \frac{k-1}{n}\right) \xrightarrow{n \rightarrow +\infty} 1$$

Finalement, $P(X_n = k) \xrightarrow{n \rightarrow +\infty} e^{-\lambda} \times \frac{\lambda^k}{k!}$, ce qui achève de prouver la convergence en loi vers une $\mathcal{P}(\lambda)$. \square

9.5.2 Lois gamma**9.5.2.1 Fonction gamma****Définition : Fonction Gamma**

La fonction Γ définie sur \mathbb{R}_*^+ associe à tout x

$$\Gamma(x) = \int_{\mathbb{R}^+} e^{-t} t^{x-1} d\lambda(t)$$

De plus, cette fonction est \mathcal{C}^∞ .

Démonstration. La fonction $\phi : (t, r) \mapsto e^{-t} t^{r-1}$ est \mathcal{C}^∞ sur $\mathbb{R}^+ \times \mathbb{R}_*^+$.

De plus, cette fonction est intégrable par rapport à t sur \mathbb{R}^+ . En effet, on peut utiliser deux arguments :

- $t^2 \phi(t) \xrightarrow{t \rightarrow +\infty} 0$ qui prouve l'intégrabilité sur $[1; +\infty[$;
- pour tout $0 < t \leq 1$, $0 \leq \phi(t, r) \leq t^{r-1}$ qui prouve l'intégrabilité sur $]0; 1]$.

On va maintenant encore utiliser des arguments de domination pour prouver la dérivabilité de Γ .

Pour tout n , $\frac{\partial^n}{\partial r^n} \phi(t, r) = \ln(t)^n \times \phi(t, n)$.

Soit $[a; b] \subset \mathbb{R}_*^+$. Pour tout $r \in [a; b]$, si $t \geq 1$, $t^{b-1} \geq t^{r-1}$ et si $t \leq 1$, $t^{a-1} \geq t^{r-1}$.

Ainsi, en posant $\psi_n : t \mapsto \begin{cases} \ln(t)^n e^{-t} t^{b-1} & \text{si } t \geq 1 \\ |\ln(t)|^n e^{-t} t^{a-1} & \text{si } 0 \leq t \leq 1 \end{cases}$, on a pour tout $t \geq 0$:

$$\left| \frac{\partial^n}{\partial r^n} \phi(t, r) \right| \leq \psi_n(t)$$

Or, pour tout $0 < t \leq 1$, $\psi_n(t) = t^{a/2} |\ln(t)|^n e^{-t} t^{a/2-1}$, et on sait que $t^{a/2} |\ln(t)|^n = o(1)$ ce qui prouve l'intégrabilité de ψ_n sur $]0; 1]$.

D'autre part, pour tout $t \geq 1$, $t^2 \psi_n(t) = |\ln(t)|^n e^{-t} t^{b+1} \xrightarrow{t \rightarrow +\infty} 0$, ce qui prouve l'intégrabilité de ψ_n sur $[1; +\infty[$.

Finalement, par le théorème de convergence dominée, Γ est n -fois dérivable. Donc Γ est bien \mathcal{C}^∞ . \square

Proposition : Lien entre Γ et factorielle

Pour tout $x > 0$:

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

En particulier, pour tout entier n , $\Gamma(n) = (n-1)!$

Démonstration. On procède par intégration par parties, sachant que, dans ce cas, l'intégrale de Riemann impropre et l'intégrale de Lebesgue sont confondues (on intègre une fonction positive) :

$$\begin{aligned}\Gamma(x+1) &= \int_{\mathbb{R}^+} e^{-t} t^x d\lambda(t) \\ &= \left[-t^x e^{-t} \right]_0^{+\infty} + x \int_{\mathbb{R}^+} e^{-t} t^{x-1} d\lambda(t) \\ &= x\Gamma(x)\end{aligned}$$

On peut vérifier aisément que $\Gamma(1) = 1 = 0!$ d'où $\Gamma(2) = 1 \times \Gamma(1) = 1!$ puis $\Gamma(3) = 2 \times \Gamma(2) = 2!$ et ainsi de suite par une récurrence immédiate. \square

9.5.2.2 Loi gamma**Définition : Variable suivant une loi gamma**

Soit X une variable aléatoire et soit $r > 0$ un nombre.

On dit que X suit une loi gamma de paramètre r lorsque X admet une fonction densité

$$\phi_r : t \mapsto \frac{t^{r-1} e^{-t} \mathbb{1}_{\mathbb{R}^+}(t)}{\Gamma(r)}$$

On note $X \sim \gamma(r)$

Démonstration. Il est clair que la fonction ϕ_r est normée et positive par définition de la fonction Γ \square

Proposition : Cas particuliers de la loi gamma

X suit une loi exponentielle de paramètre λ si et seulement si λX suit une loi gamma de paramètre 1.

Si U suit une loi normale centrée réduite alors $\frac{U^2}{2}$ suit une loi gamma de paramètre $\frac{1}{2}$.

Démonstration. Soit $X \sim \exp(\lambda)$

On va utiliser la caractérisation de la densité par des fonctions continues bornées.

Soit f une fonction continue bornée. Par la formule du transfert et un changement de variable, on obtient :

$$\begin{aligned}E[f(\lambda X)] &= \int \lambda f(\lambda t) e^{-\lambda t} \mathbb{1}_{\mathbb{R}^+}(t) dt \\ &= \int f(u) e^{-u} du\end{aligned}$$

Donc la densité de λX est bien la densité d'une variable suivant une loi $\gamma(1)$.

Soit $U \sim \mathcal{N}(0; 1)$. On utilise la même technique ainsi que la parité :

$$\begin{aligned} E \left[f \left(\frac{U^2}{2} \right) \right] &= \frac{1}{\sqrt{2\pi}} \int f \left(\frac{t^2}{2} \right) e^{-t^2/2} dt \\ &= \frac{2}{\sqrt{2\pi}} \int f \left(\frac{t^2}{2} \right) e^{-t^2/2} \mathbb{1}_{\mathbb{R}^+}(t) dt \\ &= \frac{2}{\sqrt{2\pi}} \int f(u) e^{-u} \mathbb{1}_{\mathbb{R}^+}(u) \frac{\sqrt{2} du}{2\sqrt{u}} \\ &= \frac{1}{\sqrt{\pi}} \int f(u) u^{-1/2} e^{-u} \mathbb{1}_{\mathbb{R}^+}(u) du \end{aligned}$$

On retrouve que $\frac{U^2}{2}$ suit une loi $\gamma\left(\frac{1}{2}\right)$ et, au passage, on trouve la valeur de $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. \square

Proposition : Fonction caractéristique d'une loi gamma

Soit $r > 0$ et X une variable aléatoire suivant une loi $\gamma(r)$.

Alors la fonction caractéristique de X est :

$$\psi_X : \omega \mapsto \frac{1}{(i\omega - 1)^r}$$

Démonstration. Cela se prouve avec un changement de variable dans l'intégrale. \square

Proposition : Somme de deux lois gamma

Si X et Y sont indépendantes et suivent des lois gamma de paramètres r et s alors $X + Y$ suit une loi gamma de paramètre $r + s$.

Démonstration. Cela se prouve très facilement en utilisant la fonction caractéristique. \square

Définition : Loi du χ^2 à n degrés de libertés

Soient U_1, U_2, \dots, U_n , n variables indépendantes qui suivent toutes des lois normales centrées réduites.

On dit que $U_1^2 + U_2^2 + \dots + U_n^2$ suit une loi du χ^2 à n degrés de liberté.

En particulier, une variable X suit une loi du χ^2 à n degré de libertés si et seulement si $\frac{X}{2}$ suit une loi $\gamma\left(\frac{n}{2}\right)$.

Démonstration. Évident d'après ce qui précède concernant l'étude des lois gamma.

En effet, par linéarité, $\frac{X}{2}$ est une somme de n variables indépendantes suivant des lois γ de paramètre $\frac{1}{2}$. Il suit donc une loi $\gamma\left(\frac{n}{2}\right)$.

La réciproque est évidente, par construction. \square

Chapitre 10

Séries de Fourier

Dans ce document, on va exploiter les résultats portant sur la convolution ainsi que sur les espaces de Hilbert. Remarquons que les fonctions L^2 2π -périodiques à valeurs dans \mathbb{C} , que l'on notera $L^2_{2\pi}$ dans la suite, forment un espace de Hilbert pour cette forme sesquilinéaire définie positive :

$$\begin{aligned} \varphi : (L^2_{2\pi})^2 &\rightarrow \mathbb{C} \\ (f; g) &\mapsto \langle f|g \rangle = \int_{[0; 2\pi]} f(t) \overline{g(t)} d\lambda(t) \end{aligned} \quad \text{avec } \lambda \text{ la mesure de Lebesgue sur } \mathbb{R}$$

De plus, la famille $(e_n)_{n \in \mathbb{Z}}$ avec $e_n : t \mapsto \frac{1}{\sqrt{2\pi}} e^{int}$ est une famille orthonormée. On va montrer que cette famille est une base de notre espace.

10.1 Polynômes trigonométriques

10.1.1 Propriétés

On rappelle quelques propriétés des fonctions $\mathbb{R} \rightarrow \mathbb{C}$. Soit ainsi φ définie sur un intervalle réel ouvert I et à valeurs dans \mathbb{C} . On note $f = \Re(\varphi)$ et $g = \Im(\varphi)$.

- φ est continue lorsque f et g sont continues ;
- φ est dérivable lorsque f et g sont dérivables ; dans ce cas

$$\varphi' = f' + ig'$$

- les formules de dérivée d'une combinaison linéaire et d'un produit sont les mêmes ;
- la dérivée des fonctions usuelles polynômes, fonctions rationnelles est identique à la dérivée des polynômes et fonctions rationnelles à coefficients réels.

De plus, pour toute fonction holomorphe h , $h \circ \varphi$ est dérivable lorsque φ est dérivable et on a

$$(h \circ \varphi)' = \varphi' \times h' \circ \varphi$$

En particulier, pour tout complexe a , $t \mapsto e^{at}$ est dérivable et de dérivée $t \mapsto ae^{at}$.

Définition : Polynôme trigonométrique de degré au plus n

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ une fonction réelle. On dit que f est un polynôme trigonométrique de degré n lorsqu'il existe $2n + 1$ coefficients complexes $(a_k)_{-n \leq k \leq n}$ tels que, pour tout t :

$$f(t) = \sum_{-n \leq k \leq n} a_k e^{ikt} \quad \text{avec } a_n \neq 0 \text{ ou } a_{-n} \neq 0$$

En particulier f est périodique et sa période est un diviseur de 2π .

On peut faire quelques remarques :

- les polynômes trigonométriques sont de classe \mathcal{C}^∞ ;
- la dérivée d'un polynôme trigonométrique est un polynôme trigonométrique ;

- un produit, une combinaison linéaire de deux polynômes trigonométriques est un polynôme trigonométrique.
- les coefficients (a_k) sont définis de manière unique (proposition plus bas).

Proposition : Base orthonormée des polynômes trigonométriques de degré au plus n

La famille $(e_k)_{k \in [-n; n]}$ avec $e_k : t \mapsto \frac{1}{\sqrt{2\pi}} e^{ikt}$ constitue une base orthonormée des polynômes trigonométriques pour le produit scalaire défini en préambule.

En particulier, les coefficients (a_k) d'un polynôme trigonométrique f valent

$$a_k = \langle f | e_k \rangle$$

Démonstration. Assez évident. □

Proposition : Convolution de fonctions L^1 $2 - \pi$ périodiques

Dans toute la suite, on définit une convolution sur l'ensemble des fonctions $2 - \pi$ périodiques intégrables sur $[-\pi; \pi]$ par :

$$f * g : x \mapsto \int_{[-\pi; \pi]} f(x-t)g(t) d\lambda(t)$$

Ce produit est :

- commutatif ;
- distributif sur l'addition ;
- la convolée de deux fonctions L^1 et $2 - \pi$ périodiques donne une fonction L^1 et $2 - \pi$ périodique.
- la convolée d'une fonction non périodique, L^1 sur $[-\pi; \pi]$ et d'une fonction L^1 périodique est périodique ;
- la convolée d'une fonction non périodique, L^1 sur $[-\pi; \pi]$ et d'un polynôme trigonométrique est un polynôme trigonométrique.

Démonstration. Montrons uniquement le dernier point :

$$x \mapsto \int_{[-\pi; \pi]} e^{ik(x-t)} g(t) d\lambda(t) = e^{ikx} \int_{[-\pi; \pi]} e^{-ikt} g(t) d\lambda(t) \quad \text{est un polynôme trigonométrique}$$

□

10.1.2 Convergence uniforme d'un polynôme trigonométrique vers une fonction continue

On considère une fonction f , 2π -périodique et continue.

Soit c la fonction définie par $c(x) = \frac{\cos(x) + 1}{2}$. Notons que c est paire, positive, de maximum 1, atteint pour $x = 2k\pi$ avec $k \in \mathbb{Z}$.

Pour tout entier naturel n , on pose

$$g_n = \frac{c^n}{\int_{[-\pi; \pi]} c^n}, \text{ de sorte que } \int_{[-\pi; \pi]} g_n = 1.$$

Enfin, on considère la suite de polynômes trigonométriques $f_n = g_n * f$.

Alors f_n converge uniformément vers f . Pourquoi ?

Notons tout d'abord que, pour tout x ,

$$f_n(x) - f(x) = \int_{[-\pi; \pi]} (f(x-t) - f(x))g_n(t) d\lambda(t). \text{ Cela provient de l'égalité}$$

$$f(x) = f(x) \int_{[-\pi; \pi]} g_n = \int_{[-\pi; \pi]} f(x)g_n(t) d\lambda(t).$$

On considère maintenant $\varepsilon > 0$ quelconque et on utilise l'uniforme continuité de f .

On sait qu'il existe $\eta > 0$ indépendant de x tel que, pour tout $|t| < \eta$, $|f(x-t) - f(x)| \leq \frac{\varepsilon}{2}$.

On en déduit :

$$\begin{aligned} |f_n(x) - f(x)| &\leq \int_{[-\pi; \pi]} |f(x-t) - f(x)| g_n(t) d\lambda(t) \\ |f_n(x) - f(x)| &\leq \frac{\varepsilon}{2} \int_{[-\pi; \pi] \cap \{|t| < \eta\}} g_n(t) d\lambda(t) + 2 \|f\|_\infty \int_{[-\pi; \pi] \cap \{|t| \geq \eta\}} g_n(t) d\lambda(t) \end{aligned}$$

Il reste maintenant à prouver que $\lim_{n \rightarrow +\infty} \int_{[-\pi; \pi] \cap \{|t| \geq \eta\}} g_n(t) d\lambda(t) = 0$.

On peut considérer que $0 < \eta \leq \pi$. Notons alors que $c\left(\frac{\eta}{2}\right) > c(\eta)$.

Revenons maintenant à l'expression :

$$\int_{[-\pi; \pi] \cap \{|t| \geq \eta\}} g_n(t) d\lambda(t) = \frac{\int_{[-\pi; \pi] \cap \{|t| \geq \eta\}} c^n(t) d\lambda(t)}{\int_{[-\pi; \pi]} c^n(t) d\lambda(t)} \leq \frac{c^n(\eta) \lambda([- \pi; \pi] \cap \{|t| \geq \eta\})}{c^n\left(\frac{\eta}{2}\right) \lambda([- \pi; \pi] \cap \{|t| \leq \frac{\eta}{2}\})} \leq \left(\frac{c(\eta)}{c\left(\frac{\eta}{2}\right)}\right)^n \times \frac{\lambda([- \pi; \pi] \cap \{|t| \geq \eta\})}{\lambda([- \pi; \pi] \cap \{|t| \leq \frac{\eta}{2}\})}$$

Cette dernière inégalité permet de conclure.

10.2 Séries de Fourier, noyau de Dirichlet

10.2.1 Introduction, convergence en norme deux

Soit f une fonction de classe L^2 sur $[-\pi; \pi]$ et 2π périodique.

On appelle série de Fourier de f de degré n la projection de f sur l'ensemble des polynômes trigonométriques de degré au plus n . On notera ainsi :

$$S_n[f] = \sum_{-n \leq k \leq n} \langle f | e_k \rangle e_k$$

D'après le paragraphe précédent, pour toute fonction continue et 2π -périodique f et pour tout polynôme trigonométrique de degré au plus n noté g_n , on a

$$\|g_n - f\|_2 \leq 2\pi \|g_n - f\|_\infty$$

Par suite, l'ensemble des polynômes trigonométriques est dense dans l'ensemble des fonctions continues pour la norme deux. Et comme l'ensemble des fonctions continues est dense dans l'ensemble des fonctions L^2 et $2-\pi$ périodiques, on en déduit que l'ensemble des polynômes trigonométriques est dense dans l'ensemble des fonctions L^2 et $2-\pi$ périodiques pour la norme deux.

D'autre part, on sait que $\|S_n[f] - f\|_2$ réalise la distance minimale entre les polynômes trigonométriques de degré au plus n et les f .

On en déduit que $S_n[f]$ converge vers f en norme deux. En particulier, l'ensemble des $(e_k)_{k \in \mathbb{Z}}$ constitue une base hilbertienne des fonctions L^2 $2-\pi$ périodiques.

10.2.2 Noyau de Dirichlet : définition et étude

Pour tout x , on a

$$\begin{aligned} S_n[f](x) &= \sum_{-n \leq k \leq n} \int_{[-\pi; \pi]} f \overline{e_k} e_k(x) \\ &= \int_{[-\pi; \pi]} f(t) \times \frac{1}{2\pi} \sum_{-n \leq k \leq n} e^{ik(x-t)} dt \\ &= f * d_n(x) \end{aligned}$$

En posant $d_n : t \mapsto \frac{1}{2\pi} \sum_{-n \leq k \leq n} e^{ikt}$.

Ainsi, la série de Fourier de f de degré n s'analyse comme la convolution de f et de la fonction d_n qu'on appelle le noyau de Dirichlet de degré n .

On « lit » assez rapidement que d_n est une fonction à valeurs réelles. Utilisons nos connaissances sur les sommes de termes de suite géométrique pour déterminer plus précisément l'expression de d_n .

Ainsi, pour $t = 0$, on a $d_n(0) = 2n + 1$. Et pour $t \notin 2\pi\mathbb{Z}$:

$$\begin{aligned} 2\pi d_n(t) &= e^{-int} \sum_{0 \leq k \leq 2n} e^{ikt} \\ &= e^{-int} \times \frac{1 - e^{(2n+1)it}}{1 - e^{it}} \\ &= e^{-int} \times \frac{e^{i(n+1/2)t}}{e^{it/2}} \times \frac{e^{-i(n+1/2)t} - e^{i(n+1/2)t}}{e^{-it/2} - e^{it/2}} \\ &= e^{-int} \times e^{int} \times \frac{\sin((n+1/2)t)}{\sin(\frac{t}{2})} \\ &= \frac{\sin\left(\frac{(2n+1)t}{2}\right)}{\sin\left(\frac{t}{2}\right)} \end{aligned}$$

On obtient donc :

$$d_n : t \mapsto \begin{cases} \frac{2n+1}{2\pi} & \text{si } t \in 2\pi\mathbb{Z} \\ \frac{1}{2\pi} \frac{\sin\left(\frac{(2n+1)t}{2}\right)}{\sin\left(\frac{t}{2}\right)} & \text{sinon} \end{cases}$$

La dérivée de d_n s'obtient après de longs calculs :

$$2\pi d'_n : t \mapsto \frac{n \sin((n+1)t) - (n+1) \sin(nt)}{2 \sin^2\left(\frac{t}{2}\right)}$$

Il faut étudier le numérateur de cette fonction. On pose donc :

$$\varphi : t \mapsto n \sin((n+1)t) - (n+1) \sin(nt)$$

La dérivée de cette fonction est, après calculs, $\varphi' : t \mapsto -2n(n+1) \sin\left(\frac{(2n+1)t}{2}\right) \sin\left(\frac{t}{2}\right)$.

Sur $]0; \pi]$, la dérivée s'annule et change de signe en $(\alpha_k)_{1 \leq k \leq n}$ avec $\alpha_k = \frac{2k\pi}{2n+1}$.

La fonction φ possède donc des extrema en $(\alpha_k)_{0 \leq k \leq n}$. On calcule

$$\varphi(\alpha_k) = (-1)^k (2n+1) \sin\left(\frac{\pi k}{2n+1}\right) \quad \text{qui est du signe de } (-1)^k$$

Par suite, d'_n change de signe entre chacun des α_k .

On note ainsi $(\beta_k)_{0 \leq k \leq n-1}$ les lieux des extrema de d_n . On a $\beta_0 = 0$, et pour tout $1 \leq k \leq n-1$, $\beta_k \in]\alpha_k; \alpha_{k+1}[$. Là encore l'analyse du signe de φ nous indique que pour k pair, le lieu est un maximum et pour k impair le lieu est un minimum.

On va maintenant chercher à majorer $|d_n(\beta_k)|$ pour $1 \leq k \leq n-1$, sachant que

- $d_n(\beta_0) = 2n + 1$
- entre α_k et α_{k+1} la fonction $t \mapsto \left| \sin\left(\frac{(2n+1)t}{2}\right) \right|$ est maximale en $\frac{\alpha_k + \alpha_{k+1}}{2} = \frac{(2k+1)\pi}{(2n+1)}$ et son maximum est 1 ;
- sur $[0; \pi]$ la fonction $t \mapsto \sin\left(\frac{t}{2}\right)$ est positive et croissante.
De plus, cette fonction est minorée sur $[0; \pi]$ par sa corde $t \mapsto \frac{x}{\pi}$.

On a ainsi le contrôle suivant :

$$2\pi |d_n(\beta_k)| \leq \frac{1}{\sin\left(\frac{\alpha_k}{2}\right)} = \frac{1}{\sin\left(\frac{k\pi}{2n+1}\right)} \leq \frac{2n+1}{k}$$

Ce dernier résultat nous indique la forme de la courbe de d_n sur $[-\pi; \pi]$ qui possède

- un maximum en 0 ;
- des extrema locaux secondaires d'amplitudes atténuées plus on se rapproche des extrémités du segment $[-\pi; \pi]$.

10.2.3 Le cas C^1

10.2.3.1 Cas simple

On suppose maintenant que f est dérivable et que sa dérivée est de classe L^2 , ce qui bien sûr inclut le cas C^1 que l'on retrouvera en pratique.

On va montrer dans ce paragraphe que la série de Fourier de f converge normalement vers f .

Proposition : Coefficient de Fourier des dérivées n-èmes

Soit f une fonction de classe C^n . Alors, pour tout $k \leq n$ et pour tout $m \in \mathbb{Z}$, on a

$$\langle f^{(k)} | e_m \rangle = (im)^k \langle f | e_m \rangle$$

Démonstration. Très facile à prouver par IPP et récurrence. □

Proposition : Convergence normale de la série de Fourier dans le cas d'une fonction de classe C^1

Tout est dans le titre ! La série de Fourier converge normalement vers f .

Démonstration. On raisonne en norme deux. On sait, d'après l'inégalité de Parseval que :

$$\sum_{k \in \mathbb{Z}} |\langle f' | e_k \rangle|^2 < +\infty$$

Or, $|\langle f' | e_k \rangle|^2 = k^2 |\langle f | e_k \rangle|^2$.

On exploite maintenant l'inégalité de Cauchy-Schwarz, pour tout $n \in \mathbb{N}^*$:

$$\sum_{\substack{-n \leq k \leq n \\ k \neq 0}} \left| \frac{1}{k} \right| |k \langle f | e_k \rangle| \leq \sqrt{\sum_{k \in \mathbb{Z}^*} |k|^2 |\langle f | e_k \rangle|^2} \sqrt{\sum_{k \in \mathbb{Z}^*} |k|^2 \left| \frac{1}{k^2} \right|} < +\infty$$

On en déduit que $\sum_{k \in \mathbb{Z}} |\langle f | e_k \rangle|$ est convergente, ce qui achève la démonstration. □

10.2.3.2 Théorème de Dini-Dirichlet et applications

On va maintenant essayer d'établir sous des conditions légèrement moins strictes un résultat de convergence simple de la série de Fourier d'une fonction f de classe L^1 en un point x vers une limite ℓ .

On cherche donc à étudier la limite de :

$$\begin{aligned} Sn[f](x) - \ell &= \int_{[-\pi; \pi]} f(x-t) d_n(t) dt - \ell \\ &= \int_{[-\pi; \pi]} (f(x-t) - \ell) d_n(t) dt \\ &= \int_{[0; \pi]} (f(x-t) - \ell) d_n(t) dt + \int_{[0; \pi]} (f(x+t) - \ell) d_n(-t) dt \\ &= \int_{[0; \pi]} (f(x-t) + f(x+t) - 2\ell) d_n(t) dt \text{ avec un changement de variable et par parité de } d_n \\ &= \frac{1}{2\pi} \int_{[0; \pi]} \frac{f(x-t) + f(x+t) - 2\ell}{\sin\left(\frac{t}{2}\right)} \sin\left(\frac{(2n+1)t}{2}\right) dt \\ &= \frac{1}{2\pi} \text{Im} \left(\widehat{g} \left(\frac{(2n+1)}{2} \right) \right) \end{aligned}$$

En posant $g : t \mapsto \frac{f(x-t) + f(x+t) - 2\ell}{\sin(\frac{t}{2})} \mathbb{1}_{[0; \pi]}$ et en notant \widehat{g} la transformée de Fourier de g .

Il suffit donc de montrer $\widehat{g}\left(\frac{(2n+1)}{2}\right) \xrightarrow{n \rightarrow +\infty} 0$.

Mais pour cela, nous savons qu'il suffit que g soit L^1 . Cela nous offre un critère de convergence.

Théorème : Critère de Dirichlet-Dini

On reprend les mêmes hypothèses.

Si $\int_{[0; \pi]} \left| \frac{f(x-t) + f(x+t) - 2\ell}{t} \right| dt < +\infty$ alors

$$S_n[f](x) \xrightarrow{n \rightarrow +\infty} \ell$$

Démonstration. C'est le critère de Riemann. En effet, $g(t) \underset{0}{\sim} 2 \times \left(\frac{f(x-t) + f(x+t) - 2\ell}{t} \right)$. Ainsi, si la fonction $t \mapsto \frac{f(x-t) + f(x+t) - 2\ell}{t}$ est L^1 sur $[0; \pi]$ alors g l'est aussi. Or on a vu dans le développement précédent que :

$$S_n[f](x) - \ell = \frac{1}{2\pi} \text{Im} \left(\widehat{g} \left(\frac{(2n+1)}{2} \right) \right)$$

On en déduit le résultat escompté en raison de la limite de la transformée de Fourier d'une fonction L^1 en $+\infty$. \square

Nous en déduisons un corollaire très pratique.

Corollaire : Convergence simple de la série de Fourier dans le cas \mathcal{C}^1 par morceaux

Soit f une fonction 2π périodique et \mathcal{C}^1 par morceaux.

On pose $\tilde{f} : x \mapsto \begin{cases} f(x) & \text{si } x \text{ n'est pas un saut} \\ \frac{f(x^+) + f(x^-)}{2} & \text{si } x \text{ est un saut} \end{cases}$

Alors $S_n[f]$ converge simplement vers \tilde{f} .

Démonstration. Soit $x \in]-\pi; \pi]$.

D'après le critère de Dini-Dirichlet, il suffit de vérifier si $g : t \mapsto \frac{f(x-t) + f(x+t) - 2\tilde{f}(x)}{t}$ est L^1 sur $[0; \pi]$.

Dans le cas où x est un saut, on a $g(t) = \frac{f(x+t) + f(x-t) - f(x^+) - f(x^-)}{t}$. Quand t tend vers 0 cette fonction tend vers $f'_d(x) + f'_g(x)$ avec f'_d et f'_g les dérivées à droite et à gauche, ce qui permet de conclure. La fonction g est bien intégrable.

Le cas où x n'est pas un saut est encore plus simple.

Dans les deux cas, le critère de Dini s'applique et cela permet de conclure. \square

10.2.3.3 Deux exemples en dents de scie

Soit la fonction s , 2π périodique telle que $s|_{[-\pi; \pi]} = \text{id}|_{[-\pi; \pi]}$.

On note pour tout $k \in \mathbb{Z}$, c_k les coefficients de Fourier des éléments e_k , $c_k = \langle s | e_k \rangle$.

On a $c_0 = 0$ car s est impaire. De plus, pour tout $k \neq 0$, on obtient après calcul :

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_{[-\pi; \pi]} t e^{-ikt} d\lambda(t) \\ &= \dots \\ &= \frac{i(-1)^k}{k} \end{aligned}$$

On en déduit aisément l'expression de la série de Fourier associée à s . Pour tout $n \in \mathbb{N}^*$:

$$S_n[s] : x \mapsto \sum_{1 \leq k \leq n} \frac{2 \times (-1)^{k+1} \sin(kx)}{k}$$

L'égalité de Parseval donne :

$$\lim_{n \rightarrow +\infty} \sum_{k=-n}^n |c_k|^2 = \|s\|_2^2 = \frac{1}{2\pi} \int_{]-\pi; \pi]} t^2 d\lambda(t)$$

Un peu de calcul permet d'obtenir l'identité :

$$\sum_{k \in \mathbb{N}^*} \frac{1}{k^2} = \frac{\pi^2}{6}$$

Considérons également la fonction u , 2π périodique telle que, pour tout $t \in]-\pi; \pi]$, $u(t) = |t|$. On note (b_k) les coefficients de Fourier associés.

On a $b_0 = \frac{\pi}{2}$ (valeur moyenne) et pour tout $k \in \mathbb{Z}^*$, après calcul :

$$b_k = \begin{cases} \frac{-2}{k^2\pi} & \text{si } k \text{ est impair} \\ 0 & \text{sinon} \end{cases}$$

La série de Fourier associée à u est donc, pour tout $n \in \mathbb{N}^*$:

$$S_{2n+1}[u] : x \mapsto \frac{\pi}{2} - \frac{4}{\pi} \sum_{0 \leq k \leq n} \frac{\cos((2k+1)x)}{(2k+1)^2}$$

La convergence simple de cette série en 0 vers $s(0) = 0$ donne l'identité :

$$\sum_{k \in \mathbb{N}^*} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$$

Chapitre 11

Intégrale de Kurtzweil-Henstock

Ce document définit une intégrale qui améliore considérablement la formulation du théorème fondamental de l'analyse, qui possède une propriété de convergence monotone, de convergence dominée. En outre, la classe de fonctions ainsi intégrables est plus vaste que la classe des fonctions Lebesgue-intégrables.

11.1 Définition et premières propriétés de l'intégrale de Kurzweil-Henstock

11.1.1 Définitions

Définition : Subdivision pointée

Soit $[a; b]$ un intervalle.

Une subdivision pointée D est la donnée de $N + 1$ points $a = a_0 < a_1 < \dots < a_N = b$ et de N points x_1, x_2, \dots, x_N tels que pour tout $i \in \llbracket 0; N - 1 \rrbracket$, $x_i \in [a_i; a_{i+1}]$.

On pourra également désigner une subdivision pointée par le un N -uplet de couples $D = ([a_i; a_{i+1}]; x_i)_{0 \leq i \leq N-1}$.

Comme dans le cas des subdivisions classiques, on définit l'inclusion d'une subdivision dans une autre de manière « naturelle. »

Ainsi, pour deux subdivisions pointées $D = ([a_i; a_{i+1}]; x_i)_{0 \leq i \leq N-1}$ et $D' = ([\alpha_i; \alpha_{i+1}]; t_i)_{0 \leq i \leq Q-1}$, on écrira $D \subset D'$ lorsque :

- pour tout $i \in \llbracket 1; N - 1 \rrbracket$, il existe $j \in \llbracket 1; Q - 1 \rrbracket$ tel que $a_i = \alpha_j$;
- pour tout $i \in \llbracket 0; N - 1 \rrbracket$, il existe $j \in \llbracket 0; Q - 1 \rrbracket$ tel que $x_i = t_j$;

Définition : Jauge, subdivision δ -fine

Une jauge sur $[a; b]$ est une fonction sur $[a; b]$ à valeurs strictement positives.

On dit qu'une subdivision pointée $D = ([a_i; a_{i+1}]; x_i)_{0 \leq i \leq N-1}$ sur $[a; b]$ est δ -fine lorsque pour tout $0 \leq i \leq N - 1$:

$$a_{i+1} - a_i \leq \delta(x_i)$$

Lemme : Existence de subdivision pointée δ -fine

Soit δ une jauge sur $[a; b]$.

Alors il existe une subdivision δ -fine de $[a; b]$.

Démonstration. Considérons l'ensemble des intervalles $]x - \delta(x); x + \delta(x)[$ pour $x \in [a; b]$. Leur union recouvre $[a; b]$ qui est compact.

Il existe ainsi une famille finie $(x_i)_{0 \leq i \leq n-1}$ telle que :

- $\bigcup_{1 \leq i \leq n}]x_i - \delta(x_i); x_i + \delta(x_i)[$ recouvre $[a; b]$;
- les x_i sont distincts deux à deux, rangés par ordre croissants ;
- aucun des sous-intervalles $]x_i - \delta(x_i); x_i + \delta(x_i)[$ n'est inclus dans un autre.

On pose ensuite $a_0 = a$, $a_n = b$ et pour tout $1 \leq i \leq n-1$, a_n un nombre quelconque de $]x_i - \delta(x_i); x_{i+1} + \delta(x_{i+1})[$ tel que $x_i > a_n > x_{i-1}$.

On obtient ainsi, par construction, une distribution pointée δ -fine. \square

Définition : Somme de Riemann associée à une subdivision pointée

Soit $D = ([a_i; a_{i+1}]; x_i)_{0 \leq i \leq N-1}$ une subdivision pointée et f une fonction définie sur $[a; b]$ et à valeurs dans \mathbb{R} . On note

$$S_D[f] = \sum_{0 \leq k \leq N-1} (a_{i+1} - a_i) f(x_i)$$

Notons que, pour une subdivision pointée D donnée, l'application $S_D : f \mapsto S_D[f]$ est une forme linéaire (et positive si $E = \mathbb{R}$).

Définition : Intégrale de Kurtzweil-Henstock

Soit $f : [a; b] \rightarrow \mathbb{R}$. On dit que f est intégrable au sens de Kurtzweil-Henstock lorsqu'il existe un nombre note $\int_a^b f$ tel que, pour tout $\varepsilon > 0$, il existe une jauge δ telle que pour toute subdivision D , δ -fine, on a :

$$\left\| S_D[f] - \int_a^b f \right\| < \varepsilon$$

De plus, le nombre $\int_a^b f$ est unique.

Démonstration. Il faut montrer l'unicité. Soient ainsi $A \neq B$ deux nombres vérifiant la définition. Pour $\varepsilon = \frac{|B-A|}{3}$, il existe deux jauges δ_A et δ_B telles que pour toutes subdivisions pointées D_A et D_B respectivement δ_A et δ_B fines, on a :

$$\|S_{D_A}[f] - A\| < \frac{|B-A|}{3} \quad \text{et} \quad \|S_{D_B}[f] - B\| < \frac{|B-A|}{3}$$

En particulier, en posant $\delta = \min(\delta_A; \delta_B)$, et pour toute subdivision E δ -fine, on a E qui est à la fois δ_A et δ_B fine. En particulier, l'inégalité triangulaire donne :

$$|B-A| \leq \|S_E[f] - A\| + \|S_E[f] - B\| < \frac{2|B-A|}{3}$$

Ceci constitue une absurdité! \square

Cette intégrale est bien une généralisation de l'intégrale de Riemann comme l'explique la proposition qui suit.

Proposition : Lien avec l'intégrale de Riemann

Si f est une fonction Riemann-intégrable sur $[a; b]$ alors f est KH intégrable sur $[a; b]$ et dans ce cas l'intégrale de f sur $[a; b]$ de Riemann et de Kurtzweil-Henstock sont confondues.

La réciproque est fausse.

La preuve s'appuie sur l'équivalence suivante que nous rappelons et qui est valable pour les fonctions à valeurs dans \mathbb{R}^d avec $d \in \mathbb{N}^*$

Définition : Deux manières de définir l'intégrale de Riemann pour les fonctions à valeurs dans \mathbb{R}^d

Soit $[a; b]$ un intervalle et f une fonction définie sur $[a; b]$ et à valeurs dans \mathbb{R}^d . Alors les deux définitions suivantes de l'intégrale de Riemann sont équivalentes :

Définition classique On dit que f est Riemann intégrable sur $[a; b]$ lorsqu'il existe deux suites de fonctions en escaliers φ_n et ψ_n à valeurs respectivement dans \mathbb{R}^d et \mathbb{R}^+ telles que :

- pour tout n , $\|f - \varphi_n\| \leq \psi_n$;
- $\lim_{n \rightarrow +\infty} \int_a^b \psi_n = 0$.

Dans ce cas $\int_a^b \varphi_n$ converge et cette limite ne dépend pas du choix des deux suites. On note $\int_a^b f$ cette limite.

Définition avec les subdivisions On dit que f est Riemann intégrable sur $[a; b]$ lorsqu'il existe un nombre A tel que pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que, pour toute subdivision pointée D de pas inférieur à η , on a

$$\|S_D[f] - A\| < \varepsilon$$

Dans ce cas, on appelle intégrale de f sur $[a; b]$ le nombre $\int_a^b f = A$.

Dans le cas d'une fonction f à valeurs dans un espace de Banach quelconque, la première définition entraîne la seconde. En revanche, je n'ai pas réussi à prouver que la seconde entraîne la première.

En revanche, pour \mathbb{R}^d , c'est possible et cela peut se faire par exemple à l'aide des sommes de Darboux sur chacune des coordonnées.

Démonstration. Revenons à la preuve que la Riemann intégrabilité entraîne la KH intégrabilité.

Cela est trivial. Supposons f Riemann intégrable. Pour tout $\varepsilon > 0$, nous savons qu'il existe $\eta > 0$ telle que pour toute subdivision pointée D de pas inférieur à η , on a

$$\left\| S_D[f] - \int_a^b f \right\| < \varepsilon$$

Pour montrer la KH intégrabilité, il suffit donc de considérer la jauge constante égale à η . □

On va maintenant trouver un contre exemple de fonction KH intégrable qui n'est pas Riemann intégrable : c'est la fonction de Dirichlet $\mathbb{1}_{\mathbb{Q}}$ sur $[0; 1]$. Tout le problème ici, c'est de trouver la bonne jauge.

Soit ainsi $\varepsilon > 0$. On pose $(x_n)_{n \in \mathbb{N}}$ les rationnels de $[0; 1]$.

Pour tout $n \in \mathbb{N}$, on pose $\delta(x_n) = \frac{\varepsilon}{2^{n+1}}$ et $\delta(x) = 1$ si $x \notin \mathbb{Q}$.

Soit $D = ([a_i; a_{i+1}]; t_i)_{0 \leq i \leq n-1}$ une distribution pointée δ -fine. Alors :

$$\begin{aligned} S_D[f] &= \sum_{0 \leq i \leq n-1} (a_{i+1} - a_i) \mathbb{1}_{\mathbb{Q}}(t_i) \\ &= \sum_{\substack{0 \leq i \leq n-1 \\ t_i \in \mathbb{Q}}} (a_{i+1} - a_i) \end{aligned}$$

Par construction de la jauge, on en déduit :

$$0 \leq S_D[f] < \varepsilon$$

Cela permet de prouver que la fonction de Dirichlet est KH-intégrable et d'intégrale nulle.

11.1.2 Les propriétés habituelles de l'intégrale

Proposition : Linéarité de l'intégrale

Soient f et g deux fonctions à valeurs dans \mathbb{R}^d et KH-intégrables sur $[a; b]$. Soit $\lambda \in \mathbb{R}$.

Alors $f + \lambda g$ est KH-intégrable et on a

$$\int_a^b (f + \lambda g) = \int_a^b f + \lambda \int_a^b g$$

Démonstration. On va bien sûr exploiter la linéarité des sommes de Riemann.

Soit $\varepsilon > 0$. Il existe deux jauges δ_f et δ_g telles que pour toutes subdivisions pointées D_f et D_g respectivement δ_f et δ_g fines, on a

$$\left\| S_{D_f}[f] - \int_a^b f \right\| < \frac{\varepsilon}{2} \quad \text{et} \quad \left\| S_{D_g}[g] - \int_a^b g \right\| < \frac{\varepsilon}{2(|\lambda| + 1)}$$

Considérons maintenant la jauge $\delta = \min(\delta_f; \delta_g)$. Pour toute subdivision pointée E δ -fine, E est à la fois δ_f et δ_g fine.

En particulier, on a :

$$\left\| S_E[f + \lambda g] - \left(\int_a^b f + \lambda \int_a^b g \right) \right\| = \left\| S_E[f] - \int_a^b f + \lambda \left(S_E[g] - \int_a^b g \right) \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon|\lambda|}{2(|\lambda| + 1)} \leq \varepsilon$$

Cela achève la démonstration. □

La proposition qui suit porte sur les fonctions à valeurs réelles.

Proposition : Positivité de l'intégrale

Soit f une fonction positive, définie sur $[a; b]$, et KH-intégrable.

Alors :

$$\int_a^b f \geq 0$$

En particulier, si g et h sont deux fonctions KH intégrables sur $[a; b]$ telles que pour tout x de $[a; b]$ $g(x) \geq h(x)$ alors :

$$\int_a^b g \geq \int_a^b h$$

Démonstration. Si f est positive, toutes les sommes de Riemann sont positives. En particulier, il est impossible que l'intégrale de f soit strictement négative (il suffirait de prendre $\varepsilon = -\int_a^b f$ pour aboutir à une absurdité).

La seconde partie de la proposition se montre avec la linéarité de l'intégrale et la première partie. En effet, dans ce cas $g - h$ est une fonction positive. □

Un corollaire concerne bien sûr l'inégalité de la moyenne dans le cas où f est bornée.

11.1.3 Critère de Cauchy et relation de Chasles

La notation KHC n'est pas officielle.

Théorème : Critère de Kurzweil-Henstock-Cauchy

Soit f une fonction définie sur un intervalle $[a; b]$.

On dit que f vérifie le critère de Kurzweil-Henstock-Cauchy lorsque pour tout $\varepsilon > 0$, il existe une jauge δ telle que pour toutes subdivisions pointées D et D' δ -fines, on ait :

$$\|S_D[f] - S_{D'}[f]\| < \varepsilon$$

De plus, une fonction est KH intégrable si et seulement si elle vérifie le critère de KHC.

Démonstration. Le sens direct est évident et se fait avec les techniques de contrôle habituelles en utilisant l'inégalité triangulaire sur

$$\|S_D[f] - S_{D'}[f]\| \leq \left\| S_D[f] - \int_a^b f \right\| + \left\| S_{D'}[f] - \int_a^b f \right\|$$

Montrons le sens réciproque. Pour ce faire, il faut fabriquer une suite de Cauchy de subdivisions « de plus en plus fines ».

On pose ainsi, pour tout entier n , $\varepsilon_n = \frac{1}{n+1}$.

Soit une suite de jauges (δ_n) telles que, pour tout n , et pour toutes subdivisions D_n et D'_n δ_n -fines on a :

$$\|S_{D_n}[f] - S_{D'_n}[f]\| < \varepsilon$$

On fabrique maintenant, par récurrence, une suite décroissante de jauges (η_n) en posant $\eta_0 = \delta_0$ et, pour tout $n \in \mathbb{N}^*$, $\eta_n = \min(\delta_n; \eta_{n-1})$.

Par construction pour tout n et pour tout p et pour toute subdivision D on a les propriétés suivantes :

- D est η_n -fine entraîne D est δ_n -fine ;
- D est η_{n+p} -fine entraîne D est η_n -fine (et donc δ_n -fine d'après le point précédent).

Finalement, soit une suite (D_n) de subdivisions telle que, pour tout n , D_n est η_n -fine. On obtient ainsi, pour tout $p \in \mathbb{N}$:

$$\|S_{D_{n+p}}[f] - S_{D_n}[f]\| < \varepsilon_n$$

En particulier, la suite $(S_{D_n}[f])$ est de Cauchy dans \mathbb{R}^d qui est complet et donc converge.

On pose A la limite de cette suite. Pour tous entiers n et p et pour toute subdivision E_n qui est η_n -fine, on a

$$\|S_{D_{n+p}}[f] - S_{E_n}[f]\| < \varepsilon_n$$

Par passage à la limite sur p , cela donne :

$$\|A - S_{E_n}[f]\| \leq \varepsilon_n$$

Cela permet donc de conclure : A correspond bien à l'intégrale KH de f puisque ε_n peut être arbitrairement petit. \square

On va maintenant se doter de la relation de Chasles.

Définition : Deux conventions pour l'intégrale

Soit f une fonction KH intégrable sur $[a; b]$. On pose :

$$\int_b^a f = - \int_a^b f$$

De même, pour tout $c \in [a; b]$, on pose $\int_c^c f = 0$.

On va également énoncer deux lemmes.

Lemme : Intégrabilité sur un sous-intervalle

Soient $a \leq c < d \leq b$ quatre nombres et f une fonction définie sur $[a; b]$ et à valeurs dans \mathbb{R}^d .

Si f est intégrable sur $[a; b]$ et si f est KH intégrable sur $[a; b]$ alors f est KH intégrable sur $[c; d]$.

Démonstration. On va montrer que si f vérifie le critère de Cauchy sur $[a; b]$ alors f vérifie aussi le critère de Cauchy sur $[c; d]$.

Soit ainsi $\varepsilon > 0$ un nombre. On sait qu'il existe une jauge δ sur $[a; b]$ telle que pour toutes subdivisions pointées D et D' δ -fines de $[a; b]$, on a

$$\|S_D[f] - S_{D'}[f]\| < \varepsilon$$

En particulier, pour toutes subdivisions E et E' de $[c; d]$ qui sont δ -fines, et pour toutes subdivisions F et G respectivement de $[a; c]$ et $[d; b]$ là encore δ -fines, on pose $D = F \cup E \cup G$ et $D' = F \cup E' \cup G$.

Par construction, on a D et D' qui sont δ -fines et de plus, on vérifie aisément que

$$\|S_D[f] - S_{D'}[f]\| = \|S_E[f] - S_{E'}[f]\| < \varepsilon$$

On en déduit que f vérifie bien le critère de Cauchy sur $[c; d]$, ce qui permet de conclure.

Ici, on n'a pas traité le cas éventuel où $c = a$ ou $d = b$ qui s'élimine facilement en considérant une subdivision F ou G vide ! \square

Lemme : Relation de Chasles, première version

Soient $a < b < c$ trois nombres. Si f est KH intégrable sur $[a; b]$ et $[b; c]$ alors f est KH intégrable sur $[a; c]$ et on a

$$\int_a^b f + \int_a^c f = \int_a^c f$$

Démonstration. Tout le problème ici consiste à fabriquer la jauge sur $[a; b]$ à partir des jauges sur $[a; b]$ et $[b; c]$.

Soit ainsi $\varepsilon > 0$. On suppose qu'on a deux jauges δ_b sur $[a; b]$ et δ_c sur $[b; c]$ telles que pour toutes subdivisions pointées D et E respectivement de $[a; b]$ et $[b; c]$ qui soient respectivement δ_b et δ_c fines on ait

$$\left\| S_D[f] - \int_a^b f \right\| < \frac{\varepsilon}{3} \quad \text{et} \quad \left\| S_E[f] - \int_b^c f \right\| < \frac{\varepsilon}{3}$$

Tout le problème maintenant c'est de fabriquer une jauge δ sur $[a; c]$ de telle sorte que chaque subdivision D δ -fine soit δ_b -fine sur $[a; b]$ et δ_c -fine sur $[b; c]$. Il se pose alors le problème de la jonction en b . En effet, on veut que les sous-intervalles débutant sur $[a; b]$ ne s'achèvent pas sur $[b; c]$ et que les sous-intervalles de $[b; c]$ ne débutent pas sur $[a; b]$ car les jauges δ_b et δ_c ne portent que sur des subdivisions respectivement propres à $[a; b]$ et $[b; c]$.

Pour circonvier cette contrainte, on fabrique δ de sorte que, pour chaque x appartenant à $[a; c] \setminus \{b\}$, $\delta(x)$ soit inférieur à la distance entre x et b , c'est à dire $|b - x|$. Il faut également régler le problème éventuel du saut en b .

On pose ainsi, pour tout $x \in [a; c] \setminus \{b\}$, $\delta(x) = \min(\delta_b(x); \delta_c(x); |b - x|)$ et $\delta(b) = \min(\delta_b(b); \delta_c(b))$

Ainsi, un sous-intervalle qui « pointerait » sur b ne serait pas non plus problématique car il serait l'union de deux sous-intervalles compatibles avec les jauges δ_b (à gauche) et δ_c (à droite).

Pour une telle jauge δ et pour toute subdivision pointée F δ -fine, on a bien, par construction :

$$\left| S_F[f] - \int_a^b f - \int_b^c f \right| < \frac{2\varepsilon}{3}$$

\square

À l'aide des conventions énoncées plus haut et des deux lemmes, on obtient la relation de Chasles « classique »

Théorème : Relation de Chasles

Soit E une fonction définie et HK-intégrable sur un intervalle I . Alors, pour tout $(a; b, c) \in I^3$, on a

$$\int_a^b f + \int_b^c f = \int_a^c f$$

Pour prouver ce résultat, on analyse les cas possibles. C'est un peu fastidieux mais faisable.

11.1.4 Le théorème fondamental de l'analyse

11.1.4.1 Énoncé et preuve

Dans le cadre de l'intégrale de Kurzweil-Henstock, le théorème fondamental de l'analyse prend une forme bien plus souple et agréable.

Théorème : Théorème fondamental de l'analyse

Soit f une fonction définie et dérivable sur $[a; b]$.

Alors f' est KH-intégrable sur $[a; b]$ et :

$$\int_a^b f' = f(b) - f(a)$$

Avant de se lancer dans la preuve, on rappelle qu'une fonction f est dérivable en x lorsque pour tout h suffisamment petit

$$f(x+h) - f(x) = hf'(x) + h\varepsilon(h) \text{ avec } \varepsilon(0) = 0 \text{ et } \varepsilon \text{ continue en } 0$$

Remarquons également que si $a = a_0 < a_1 < \dots < a_n = b$ sont $n+1$ nombres, on a

$$f(b) - f(a) = \sum_{0 \leq k \leq n-1} (f(a_{k+1}) - f(a_k))$$

Ces deux remarques nous donnent un moyen assez simple de contrôler $|S_D[f'] - (f(b) - f(a))|$ et nous permettent de faire la démonstration sans trop de difficulté.

Démonstration. Soit $\varepsilon > 0$. Pour tout $x \in [a; b]$, il existe $\delta(x) > 0$ tel que pour tout $t \in]x - \delta(x); x + \delta(x)[\cap [a; b]$, on a

$$|f'(x)(t-x) - (f(t) - f(x))| \leq \frac{\varepsilon}{b-a} \times |t-x|$$

On a ainsi construit une jauge δ . Soit ainsi une subdivision pointée $D = ([a_i; a_{i+1}]; x_i)_{0 \leq i \leq n-1}$ qui est δ -fine.

On va maintenant contrôler

$$\begin{aligned} |S_D[f'] - (f(b) - f(a))| &= \left| S_D[f'] - \sum_{0 \leq i \leq n-1} (f(a_{i+1}) - f(a_i)) \right| \\ &= \left| \sum_{0 \leq i \leq n-1} [f'(x_i)(a_{i+1} - a_i) - (f(a_{i+1}) - f(a_i))] \right| \\ &\leq \sum_{0 \leq i \leq n-1} |f'(x_i)(a_{i+1} - a_i) - (f(a_{i+1}) - f(a_i))| \end{aligned}$$

Pour tout $0 \leq i \leq n-1$, travaillons sur le i -ème terme :

$$\begin{aligned} |f'(x_i)(a_{i+1} - a_i) - (f(a_{i+1}) - f(a_i))| &= |f'(x_i)(a_{i+1} - x_i) + f'(x_i)(x_i - a_i) - (f(a_{i+1}) - f(x_i)) - (f(x_i) - f(a_i))| \\ &\leq |f'(x_i)(a_{i+1} - x_i) - (f(a_{i+1}) - f(x_i))| + |f'(x_i)(x_i - a_i) - (f(x_i) - f(a_i))| \\ &\leq \frac{\varepsilon}{b-a} (a_{i+1} - x_i) + \frac{\varepsilon}{b-a} (x_i - a_i) = \frac{\varepsilon}{b-a} (a_{i+1} - a_i) \end{aligned}$$

Finalement, on obtient :

$$|S_D[f'] - (f(b) - f(a))| \leq \frac{\varepsilon}{b-a} \sum_{0 \leq i \leq n-1} (a_{i+1} - a_i) = \varepsilon$$

□

11.1.4.2 Un exemple pathologique

En considérant l'exemple de $f : x \mapsto x^2 \sin\left(\frac{1}{x^2}\right)$ prolongée en 0 par $f(0) = 0$; on peut prouver qu'on dispose d'une fonction continue et dérivable sur \mathbb{R} .

Cependant la dérivée de f sur \mathbb{R}^* vaut :

$$f' : x \mapsto \frac{-2 \cos\left(\frac{1}{x^2}\right)}{x} + 2x \sin\left(\frac{1}{x^2}\right)$$

Cette dérivée n'est pas bornée sur un intervalle fermé contenant 0 donc f' n'est pas Riemann intégrable sur cet intervalle. Pourtant, f' est KH intégrable d'après le théorème qui précède.

Pire que cela, on peut montrer que la fonction $x \mapsto \left| \frac{-2 \cos\left(\frac{1}{x^2}\right)}{x} \right|$ n'est pas Lebesgue intégrable sur cet intervalle alors que la fonction $x \mapsto 2x \sin\left(\frac{1}{x^2}\right)$ l'est.

En particulier, cela implique que f' en tant que somme de fonction convergente et divergente n'est pas Lebesgue-intégrable.

11.2 Des résultats de convergence

11.2.1 Un premier exemple

On sait que l'intégrale de Riemann impropre $\int_0^1 \frac{dt}{2\sqrt{t}}$ vaut 1.

En particulier cela signifie que la fonction $f : t \mapsto \frac{\mathbb{1}_{[0; 1]}(t)}{2\sqrt{t}}$ prolongée par exemple par $f(0) = 0$ est Lebesgue-intégrable sur $[0; 1]$.

Mais est-elle KH intégrable ? La réponse est oui et nous allons le prouver « à la main » afin d'essayer de découvrir des propriétés sympathiques de convergence.

Soit ainsi $\varepsilon > 0$ un nombre.

On veut fabriquer une jauge δ sur $[0; 1]$ telle que pour toute subdivision pointée D δ -fine on ait

$$|S_D[f] - 1| < \varepsilon$$

C'est ici que l'on exploite l'intégrale impropre en tant que limite d'intégrales de Riemann « classiques. »

Ainsi, il existe $\eta > 0$ tel que $0 \leq \int_0^\eta f < \frac{\varepsilon}{2}$. Soit maintenant $\rho > 0$, tel que pour tout subdivision pointée D de $[\eta; 1]$ de pas inférieur à ρ , on a :

$$\int_\eta^1 f - S_D[f] < \frac{\varepsilon}{2}$$

On fabrique ainsi δ en posant pour tout $x \in]0; \eta]$, $\delta(x) = x \frac{\rho}{\eta}$ et pour tout $x \geq \eta$, $\delta(x) = \rho$ et enfin $\delta(0) = \eta$.

Par construction, cette jauge vérifie le cahier des charges.

11.2.2 Lemme de Henstock

Ce lemme permet de contrôler la valeur de $\int_a^b f$ sur un ensemble d'intervalles disjoints contrôlés par une jauge.

Lemme : Henstock

Soit $[a; b]$ un intervalle et f une fonction KH intégrable.

Soit $\varepsilon > 0$ et δ une jauge associée à ε .

Alors, pour tout ensemble d'intervalles fermés bornés deux à deux disjoints, inclus dans $[a; b]$, $([\alpha_i; \beta_i])_{0 \leq i \leq n-1}$ associés à des points (x_i) et vérifiant pour tout i , $\beta_i - \alpha_i \leq \delta(x_i)$, on a

$$\left| \sum_{0 \leq i \leq n-1} \left(f(x_i)(\beta_i - \alpha_i) - \int_{\alpha_i}^{\beta_i} f \right) \right| \leq \varepsilon$$

$$\sum_{0 \leq i \leq n-1} \left| f(x_i)(\beta_i - \alpha_i) - \int_{\alpha_i}^{\beta_i} f \right| \leq 2\varepsilon$$

Démonstration. On va commencer par prouver la première inégalité. On commence par compléter les trous formés par les $F_i = [\alpha_i; \beta_i]$. Soient ainsi les intervalles fermés $(G_j)_{0 \leq j \leq p-1} = ([\mu_j; \nu_j])_{0 \leq j \leq p-1}$ inclus dans $[a; b]$ qui complètent les F_i , c'est à dire tels que :

•

$$\left(\bigcup_{0 \leq j \leq p-1} G_j \right) \cup \left(\bigcup_{0 \leq i \leq n-1} F_i \right) = [a; b]$$

• $\forall (i; j) \in \llbracket 0; n-1 \rrbracket \times \llbracket 0; p-1 \rrbracket$, $\overset{\circ}{F}_i \cap \overset{\circ}{G}_j = \emptyset$

Soit $\eta > 0$ quelconque. Pour tout $0 \leq j \leq p-1$, soit δ_j une jauge sur G_j telle $\delta_j \leq \delta$ et telle que pour toute subdivision D_j δ_j -fine on a $\left| \int_{\mu_j}^{\nu_j} f - S_{D_j}[f] \right| < \frac{\eta}{p}$.

Considérons enfin la subdivision pointée D obtenue par réunion des $(D_j)_{0 \leq j \leq p-1}$ et des $(x_i; [\alpha_i; \beta_i])_{0 \leq i \leq n-1}$. Par construction D est δ -fine.

De plus, en notant $\Delta_F = \sum_{0 \leq i \leq n-1} \left(f(x_i)(\beta_i - \alpha_i) - \int_{\alpha_i}^{\beta_i} f \right)$ et $\Delta_G = \sum_{0 \leq j \leq p-1} \left(S_{D_j}[f] - \int_{\mu_j}^{\nu_j} f \right)$, on a par

construction, $\Delta_F + \Delta_G = S_D[f] - \int_a^b f = \Delta$.

En particulier, on sait que $|\Delta| < \varepsilon$ et, par l'inégalité triangulaire, $|\Delta_G| < \eta$.

On en déduit que :

$$|\Delta_F| = |\Delta - \Delta_G| < \varepsilon + \eta$$

Et comme $\eta > 0$ est arbitraire, on en déduit la première inégalité du lemme de Henstock.

Pour la seconde inégalité, on sépare les termes positifs et négatifs de Δ_F . La somme des termes positifs est majorée par ε , de même que la somme des termes négatifs est minorée par $-\varepsilon$. La somme des valeurs absolues est donc majorée par 2ε . \square

Un corollaire facile du lemme.

Corollaire : Fonction définie par une intégrale

Soit f une fonction KH intégrable sur un intervalle $[a; b]$. Alors :

$$\varphi : x \mapsto \int_a^x f \text{ est continue}$$

Démonstration. Soit $x \in [a; b]$. Soit $\varepsilon > 0$ et δ la jauge associée à $\frac{\varepsilon}{2}$.

Pour tout y de l'intervalle $[x - \delta; x + \delta] \cap [a; b]$, on a

$$(y - x)f(x) - \frac{\varepsilon}{2} \leq f(y) - f(x) = \int_x^y f \leq (y - x)f(x) + \frac{\varepsilon}{2}$$

Et ainsi, il existe $0 < \eta < \delta(x)$ tel que pour tout y de l'intervalle $[x - \eta; x + \eta] \cap [a; b]$, on a

$$\varepsilon < f(y) - f(x) < \varepsilon$$

\square

11.2.3 Intégrales généralisées

On va maintenant s'intéresser dans un premier temps aux résultats étendant l'intégrale de Kurzweil-Henstock sur des intervalles ouverts. Il s'agit donc déjà de définir une jauge pour des intervalles de longueur potentiellement infinie.

Définition : Jauge pour des intervalles de longueur infinie, somme de Riemann associée et intégrale KH

On s'intéresse ici à un intervalle $[a; +\infty[$ avec $a \in \mathbb{R}$.

Une jauge δ sur cet intervalle est une fonction strictement positive définie sur $[a; +\infty[$.

Une subdivision pointée D δ -fine est un $n+1$ -uplet $(x_i; I_i)_{0 \leq i \leq n}$ avec :

- pour tout $i \leq n-1$, $x_i \in [a; +\infty[$ et I_i et $I_i = [\alpha_i; \alpha_{i+1}]$ tel que $\alpha_{i+1} - \alpha_i \leq \delta(x_i)$;
- $x_n = +\infty$ et $I_n = [\alpha_n; +\infty[$ avec $\alpha_n \geq \frac{1}{\delta(+\infty)}$;
- $\alpha_0 = a$.

De plus, on pose, par convention, que la somme de Riemann de cette subdivision est

$$S_D[f] = \sum_{0 \leq i \leq n-1} f(x_i)(\alpha_{i+1} - \alpha_i)$$

Enfin, on dira qu'une fonction est KH intégrable sur $[a; +\infty[$ lorsqu'il existe un nombre A tel que, pour tout $\varepsilon > 0$, il existe une jauge δ sur $[a; +\infty[$ telle que, pour tout subdivision pointée D δ -fine on ait :

$$|S_D[f] - A| < \varepsilon$$

On dira que A est l'intégrale de f .

Définition : Jauge pour un intervalle $[a; b[$

Si $b \in \mathbb{R}$, on définit une subdivision pointée δ -fine sur $[a; b[$ de la même manière que sur $[a; b]$ en imposant $f(b) = 0$.

Proposition : Propriétés de cette extension de l'intégrale KH

Les propriétés suivantes de l'intégrale KH sur un segment reste inchangées :

- Elle est linéaire, positive ;
- Le critère de Cauchy s'énonce de manière identique et reste une condition nécessaire et suffisante de KH intégrabilité ;
- Pour tout intervalle $[c; d] \subset [a; +\infty[$, si f est intégrable sur $[a; +\infty[$ alors f est intégrable sur $[c; d]$;
- Si f est intégrable sur $[a; c]$ et sur $[c; +\infty[$ alors elle est intégrable sur $[a; +\infty[$.
- Les conventions associées à la relation de Chasles et la relation de Chasles reste identique.
- Le lemme de Henstock reste valable.

Avec cette convention, la notion d'intégrale généralisée de Riemann prend une forme beaucoup plus agréable.

Proposition : Intégrale généralisée : deux formes

Soit un intervalle $[a; b[$ avec éventuellement $b = +\infty$. On suppose que, pour tout $a \leq c < b$, f est KH-intégrable sur $[a; c]$ et que $\lim_{c \nearrow b} \int_a^c f$ existe. Alors :

$$f \text{ est KH intégrable sur } [a; b[\text{ et } \int_a^b f = \lim_{c \nearrow b} \int_a^c f$$

Réciproquement, si f est KH intégrable sur $[a; b[$ alors

$$\int_a^b f = \lim_{c \nearrow b} \int_a^c f$$

Démonstration. On va supposer $b \in \mathbb{R}$. Soit $A = \lim_{c \nearrow b} \int_a^c f$. Soit enfin $\varepsilon > 0$.

Il existe $a < c_0 < b$ tel que pour tout $x \in [c_0; b[$ on ait $|A - \int_a^x f| < \frac{\varepsilon}{2}$. Soit également la jauge δ_{-1} définie sur $[a; c_0]$ telle que pour toute subdivision D_{-1} δ_{-1} -fine de $[a; c_0]$, on ait :

$$\left| S_{D_{-1}}[f] - \int_a^{c_0} f \right| < \frac{\varepsilon}{4}$$

Soit maintenant une suite strictement croissante (c_n) telle que $\lim \uparrow c_n = b$. On pose par exemple $c_n = \frac{n}{n+1}b + \frac{1}{n+1}c_0$.

Ensuite, on fabrique une suite de jauges (δ_n) des intervalles $[c_n; c_{n+1}]$ telles que, pour tout n , et pour toute subdivision pointée D_n δ_n fine, on ait :

$$\left| S_{D_n}[f] - \int_{c_n}^{c_{n+1}} f \right| < \frac{\varepsilon}{2^{n+3}}$$

Cette construction permet de fabriquer une jauge δ , en posant $c_{-1} = a$.

Pour tout $x \in [a; b]$, on distingue ainsi quatre cas :

- ou bien il existe un unique entier $n \in \mathbb{N} \cup \{-1\}$ tel que $x \in]c_n; c_{n+1}[$ et on pose alors $\delta(x) = \min(\delta_n(x); (c_{n+1} - x); (x - c_n))$;
- ou bien il existe un entier $n \in \mathbb{N}$ tel que $x = c_n$ et on pose alors $\delta(x) = \min(\delta_n(c_n); \delta_{n-1}(c_n))$;
- ou bien $x = a$ et on pose $\delta(x) = \delta_{-1}(a)$;
- ou bien $x = b$ et on pose $\delta(x) = (b - c_0)$.

Soit maintenant une subdivision $D = (x_i; [\alpha_i; \alpha_{i+1}])_{0 \leq i \leq n-1}$ de $[a; b]$ δ -fine. Par construction, on a forcément $x_{n-1} = \alpha_n = b$ et il existe donc un unique p tel que $\alpha_{n-1} \in [c_p; c_{p+1}[$. On va maintenant examiner la valeur de

$$|S_D[f] - A| \leq \left| S_D[f] - \int_a^{\alpha_{n-1}} f \right| + \left| \int_a^{\alpha_{n-1}} f - A \right|$$

On sait que $\left| \int_a^{\alpha_{n-1}} f - A \right| < \frac{\varepsilon}{2}$ par définition de c_0 .

On réarrange ensuite la somme $S_D[f]$ en $p+1$ sous-sommes étant chacune relatives à des subdivisions $(D_i)_{-1 \leq i \leq p-1}$ δ_i -fines des intervalles $[c_i; c_{i+1}]$ avec un résidu $S_{D_p}[f]$ δ_p -fin relatif à l'intervalle $[c_p; \alpha_p]$. On exploite également la relation de Chasles sur l'intégrale $\int_a^{\alpha_{n-1}} f$.

On obtient ainsi :

$$\left| S_D[f] - \int_a^{\alpha_{n-1}} f \right| \leq \sum_{-1 \leq i \leq p} \left| S_{D_i}[f] - \int_{c_i}^{c_{i+1}} f \right| + \left| S_{D_p}[f] - \int_{c_p}^{\alpha_{n-1}} f \right|$$

Par construction pour les $p+1$ premiers termes et en exploitant le lemme de Henstock pour le dernier terme, on obtient :

$$\left| S_D[f] - \int_a^{\alpha_{n-1}} f \right| < \sum_{-1 \leq i \leq p} \frac{\varepsilon}{2^{p+3}} < \frac{\varepsilon}{2}$$

Finalement, on a bien :

$$|S_D[f] - A| < \varepsilon$$

Le cas où $b = +\infty$ se traite de manière identique.

On va maintenant examiner la réciproque. Cette fois-ci, on suppose $b = +\infty$ (pour changer) et on suppose donc que f est KH intégrable sur $[a; +\infty[$ et pour alléger, on pose $A = \int_a^{+\infty} f$

Soit ainsi $\varepsilon > 0$. On part sur la même approche. Soit δ la jauge de $[a; +\infty]$ $\frac{\varepsilon}{2}$ -fine. Soit $c = \frac{1}{\delta(+\infty)}$. Pour tout $x \geq c$, examinons $|\int_a^x f - A|$.

On sait qu'il existe une jauge δ_x de $[a; x]$ plus fine que δ et telle que pour toute subdivision D_x δ_x -fine on ait :

$$\left| S_{D_x}[f] - \int_a^x f \right| < \frac{\varepsilon}{2}$$

Soit maintenant la subdivision D obtenue en ajoutant à D_x l'intervalle $[x; +\infty[$ qui pointe sur $+\infty$.

D est δ -fine et on a $S_{D_x}[f] = S_D[f]$. Par conséquent, l'inégalité triangulaire donne

$$\left| \int_a^x f - A \right| \leq \left| \int_a^x f - S_{D_x}[f] \right| + |S_{D_x}[f] - A| < \varepsilon$$

□

Dans l'intégrale sur un intervalle semi-ouvert $[a; b[$ avec $b \in \mathbb{R}$, on a fixé $f(b) = 0$. Il est donc naturel de se demander si cette convention est un biais potentiel pour la théorie, en particulier, dans le cas d'une fonction « classiquement intégrable » sur $[a; b]$ avec $f(b) \neq 0$.

Cela change-t-il quelque chose de modifier ponctuellement des valeurs de f ?

Lemme : Modification d'un ensemble dénombrable de points

Soit une fonction f définie sur un intervalle $[a; b]$ et KH intégrable sur ce même intervalle.

Soit $(c_n)_{n \in \mathbb{N}}$ un ensemble dénombrable de points distincts deux à deux de $[a; b]$ et $(m_n)_{n \in \mathbb{N}}$ une suite de nombres.

On pose \tilde{f} la fonction modifiée $\tilde{f} : x \mapsto \begin{cases} f(x) & \text{si } x \notin (c_n)_{n \in \mathbb{N}} \\ m_n & \text{si } x = c_n \end{cases}$.

Alors \tilde{f} est KH intégrable et son intégrale est identique à celle de f .

Démonstration. Là encore, il s'agit pour un $\varepsilon > 0$ de construire une jauge adaptée. Soit δ_0 la jauge $\frac{\varepsilon}{2}$ adaptée pour f .

Il faut maintenant s'assurer que les sauts potentiels sur les points de (c_n) soient contrôlables.

Soit $n \in \mathbb{N}$, pour tout intervalle $[\alpha; \beta]$ contenant c_n , on a

$$|(\beta - \alpha)\tilde{f}(c_n) - (\beta - \alpha)f(c_n)| \leq (\beta - \alpha)|m_n - f(c_n)| \leq \delta(c_n)|m_n - f(c_n)|$$

Cela nous donne donc l'idée pour la jauge « prolongée ». On pose ainsi, pour tout $n \in \mathbb{N}$,

$$\delta(c_n) = \min \left(\delta_0(c_n); \frac{\varepsilon}{2^{n+2}\delta(c_n)(|m_n - f(c_n)| + 1)} \right).$$

Et pour tout $x \notin (c_n)_{n \in \mathbb{N}}$, on pose $\delta(x) = \delta_0(x)$.

Soit ainsi une subdivision pointée D δ -fine. On a

$$\left| S_D[\tilde{f}] - \int_a^b f \right| \leq |S_D[\tilde{f}] - S_D[f]| + \left| S_D[f] - \int_a^b f \right|$$

Par construction de la jauge au niveau des sauts, on a $|S_D[\tilde{f}] - S_D[f]| < \sum_{n \in \mathbb{N}} \frac{\varepsilon}{2^{n+2}} = \frac{\varepsilon}{2}$. Finalement, on a bien

$$\left| S_D[\tilde{f}] - \int_a^b f \right| < \varepsilon$$

□

11.2.4 Convergence monotone

Théorème : Théorème de convergence monotone

Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions croissantes KH-intégrables sur un intervalle $[a; b]$.

On suppose que :

- la suite $\lim \uparrow \int_a^b f_n < +\infty$;
- $\lim \uparrow f_n$ ne prend que des valeurs finies.

Alors $\lim \uparrow f_n$ est KH intégrable et

$$\lim \uparrow \int_a^b f_n = \int_a^b (\lim \uparrow f_n)$$

Démonstration. Encore une fois, c'est sur la jauge qu'il faut travailler.

Soit $A = \lim \uparrow \int_a^b f_n$ et soit $f = \lim \uparrow f$.

Soit également $\varepsilon > 0$.

On sait qu'il existe un rang p tel que

$$0 \leq A - \int_a^b f_p < \frac{\varepsilon}{4}$$

On suppose d'autre part que l'on a construit une suite de jauges $(\delta_n)_{n \in \mathbb{N}}$ telles que, pour tout n , δ_n est $\frac{\varepsilon}{2^{n+3}}$ -adaptée à (f_n) .

On va maintenant fabriquer notre jauge δ :

Pour tout x , il existe un rang $n \geq p$ tel que $f(x) - f_n(x) < \frac{\varepsilon}{4(b-a)}$. On pose alors $\delta(x) = \delta_n(x)$.

Considérons maintenant une subdivision $D = (x_i; [\alpha_i; \alpha_{i+1}])_{0 \leq i \leq N-1}$ δ -fine. À chaque x_i , on associe le rang n_i qui a servi à fabriquer la jauge de x_i . On a donc :

$$|S_D[f] - A| \leq \sum_{0 \leq i \leq N-1} (f(x_i) - f_{n_i}(x_i))(\alpha_{i+1} - \alpha_i) + \left| \sum_{0 \leq i \leq N-1} f_{n_i}(x_i)(\alpha_{i+1} - \alpha_i) - A \right|$$

Par construction, on a $\sum_{0 \leq i \leq N-1} (f(x_i) - f_{n_i}(x_i))(\alpha_{i+1} - \alpha_i) < \frac{\varepsilon}{4}$.

Reste à contrôler le second terme... Pour ce faire, on va considérer la fonction \tilde{f} dont la restriction sur chaque $] \alpha_i; \alpha_{i+1}[$ est f_{n_i} et qui vaut par exemple 0 sur les α_i . Par construction, on a

$$\sum_{0 \leq i \leq N-1} f_{n_i}(x_i)(\alpha_{i+1} - \alpha_i) = S_D[\tilde{f}]$$

De plus, \tilde{f} est KH intégrable (comme somme de fonctions KH intégrables) et on a $0 \leq A - \int_a^b \tilde{f} \leq A - \int_a^b f_p < \frac{\varepsilon}{4}$ en raison de la croissance de (f_n) .

Ainsi, on a une première inégalité sur le second terme :

$$\left| \sum_{0 \leq i \leq N-1} f_{n_i}(x_i)(\alpha_{i+1} - \alpha_i) - A \right| \leq \left| S_D[\tilde{f}] - \int_a^b \tilde{f} \right| + \left| \int_a^b \tilde{f} - A \right| < \frac{\varepsilon}{4} + \left| S_D[\tilde{f}] - \int_a^b \tilde{f} \right|$$

Travaillons enfin sur le dernier terme à contrôler

$$S_D[\tilde{f}] - \int_a^b \tilde{f} = \sum_{0 \leq i \leq N-1} \left(f_{n_i}(x_i)(\alpha_{i+1} - \alpha_i) - \int_{\alpha_i}^{\alpha_{i+1}} f_{n_i} \right)$$

On réalise une partition sur les différentes valeurs des $(n_i)_{0 \leq i \leq N-1}$. Supposons par exemple que les (n_i) prennent les valeurs deux à deux distinctes $(m_k)_{0 \leq k \leq q-1}$.

Pour tout $0 \leq k \leq q-1$, on pose I_k l'ensemble d'indices tels que pour tout $i \in I_k$, $n_i = m_k$, de sorte que les $(I_k)_{0 \leq k \leq q-1}$ réalisent une partition de $\llbracket 0; N-1 \rrbracket$.

On réécrit ainsi la somme précédente en regroupant les termes selon cette partition :

$$\sum_{0 \leq i \leq N-1} \left(f_{n_i}(x_i) (\alpha_{i+1} - \alpha_i) - \int_{\alpha_i}^{\alpha_{i+1}} f_{n_i} \right) = \sum_{0 \leq k \leq q-1} \sum_{i \in I_k} \left(f_{m_k}(x_i) (\alpha_{i+1} - \alpha_i) - \int_{\alpha_i}^{\alpha_{i+1}} f_{m_k} \right)$$

On peut alors appliquer le lemme de Henstock à chacun des regroupements de la partition. On obtient donc :

$$\left| S_D[\tilde{f}] - \int_a^b \tilde{f} \right| \leq \sum_{0 \leq k \leq q-1} \left| \sum_{i \in I_k} \left(f_{m_k}(x_i) (\alpha_{i+1} - \alpha_i) - \int_{\alpha_i}^{\alpha_{i+1}} f_{m_k} \right) \right| \leq \sum_{0 \leq k \leq q-1} \frac{\varepsilon}{2^{m_k+3}} < \frac{\varepsilon}{4}$$

Finalement, on a

$$|S_D[f] - A| < \frac{3\varepsilon}{4} < \varepsilon$$

□

11.2.5 Convergence dominée

Avant toute chose, remarquons que lorsque f et $|f|$ sont KH intégrables alors l'inégalité triangulaire s'applique :

$$\int |f| \geq \left| \int f \right|$$

Lemme : Critère de convergence pour la valeur absolue

Soit f une fonction KH intégrable sur un intervalle $[a; b]$ (ou $[a; b[$ ou $]a; b]$ ou $]a; b[$).

On note \mathcal{D} l'ensemble des subdivisions (non pointées) de cet intervalle.

Alors $|f|$ est KH intégrable si et seulement si l'ensemble $\left\{ \sum_{I \in D} \left| \int_I f \right| ; \text{ avec } D \in \mathcal{D} \right\}$ est majoré.

De plus, dans ce cas :

$$\int_a^b |f| = \sup_{D \in \mathcal{D}} \sum_{I \in D} \left| \int_I f \right|$$

Montrons ce lemme qui nous servira à prouver très facilement la convergence monotone.

Démonstration. Le sens direct est évident. En effet, supposons que $|f|$ est KH intégrable. La relation de Chasles et l'inégalité triangulaire donnent, pour tout $D \in \mathcal{D}$:

$$\sum_{I \in D} \left| \int_I f \right| \leq \sum_{I \in D} \int_I |f| = \int_a^b |f|$$

Attaquons-nous maintenant à la réciproque. Soit ainsi la borne supérieure de l'ensemble $\left\{ \sum_{I \in D} \left| \int_I f \right| ; \text{ avec } D \in \mathcal{D} \right\}$ que l'on notera A et soit $\varepsilon > 0$.

En raison de la définition de borne supérieure, il existe ainsi une subdivision $D = (I_k)_{0 \leq k \leq N-1} \in \mathcal{D}$ telle que :

$$A - \frac{\varepsilon}{2} < \sum_{0 \leq k \leq N-1} \left| \int_{I_k} f \right| \leq A$$

Pour tout $k \in \llbracket 0; N-1 \rrbracket$, il existe une jauge δ_k de l'intervalle I_k telle que pour tout subdivision pointée D_k de I_k qui est δ_k -fine on a :

$$\left| S_{D_k}[f] - \int_{I_k} f \right| < \frac{\varepsilon}{4N}$$

On fabrique maintenant une jauge δ sur $[a; b]$ (ou $[a; b[$ ou $]a; b]$ ou $]a; b[$) en recollant toutes ces jauges, avec la technique déjà utilisée pour prouver la relation de Chasles. Soit ainsi E une subdivision pointée δ -fine.

Par construction, E « imbriquée » dans D , en ce sens que E est formée de la réunion de subdivisions pointées E_k δ_k -fines des intervalles I_k .

Pour tout $k \in \llbracket 0; N-1 \rrbracket$, on note ainsi $E_k = (x_i^{(k)}; [\alpha_i^{(k)}; \alpha_{i+1}^{(k)}])_{0 \leq i \leq n_k-1}$

Montrons maintenant que la quantité $S_E [|f|]$ n'est pas trop éloignée de A . En raison de la relation de Chasles, de l'inégalité triangulaire et de la définition de borne supérieure, on a :

$$A \geq \sum_{0 \leq k \leq N-1} \sum_{0 \leq i \leq n_k-1} \left| \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right| \geq \sum_{0 \leq k \leq N-1} \left| \int_{I_k} f \right| > A - \frac{\varepsilon}{2} \quad (11.1)$$

On va donc étudier :

$$\begin{aligned} \left| S_E [|f|] - \sum_{0 \leq k \leq N-1} \sum_{0 \leq i \leq n_k-1} \left| \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right| \right| &= \left| \sum_{0 \leq k \leq N-1} \sum_{0 \leq i \leq n_k-1} \left[f(x_i^{(k)}) (\alpha_{i+1}^{(k)} - \alpha_i^{(k)}) - \left| \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right| \right] \right| \\ &\leq \sum_{0 \leq k \leq N-1} \sum_{0 \leq i \leq n_k-1} \left| f(x_i^{(k)}) (\alpha_{i+1}^{(k)} - \alpha_i^{(k)}) - \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right| \end{aligned}$$

On exploite ensuite le lemme de Henstock. On a la majoration suivante, pour chacune des sommes sur les intervalles I_k :

$$\sum_{0 \leq i \leq n_k-1} \left| f(x_i^{(k)}) (\alpha_{i+1}^{(k)} - \alpha_i^{(k)}) - \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right| \leq 2 \times \frac{\varepsilon}{4N} = \frac{\varepsilon}{2N}$$

On en déduit un contrôle de la différence entre $S_E [|f|]$ et $\sum_{0 \leq k \leq N-1} \sum_{0 \leq i \leq n_k-1} \left| \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right|$.

$$\left| S_E [|f|] - \sum_{0 \leq k \leq N-1} \sum_{0 \leq i \leq n_k-1} \left| \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right| \right| \leq \sum_{0 \leq k \leq N-1} \sum_{0 \leq i \leq n_k-1} \left| f(x_i^{(k)}) (\alpha_{i+1}^{(k)} - \alpha_i^{(k)}) - \int_{\alpha_i^{(k)}}^{\alpha_{i+1}^{(k)}} f \right| \leq N \times \frac{\varepsilon}{2N} \quad (11.2)$$

Ce qui permet d'obtenir, en exploitant les inégalités 11.1 et 11.2, on obtient :

$$A + \frac{\varepsilon}{2} \geq S_E [|f|] > A - \varepsilon$$

On a donc bien prouvé que $|f|$ était intégrable et d'intégrale A . □

Un corollaire important : le principe de domination.

Corollaire : Domination

Soient f et g deux fonctions KH intégrables sur un intervalle.

On suppose que sur cet intervalle on a $|f| \leq g$. Alors $|f|$ est KH intégrable.

Démonstration. Très simple. On reprend les hypothèses et notations du critère de convergence. Ainsi, il est immédiat que si $|f|$ est intégrable alors l'ensemble $\left\{ \sum_{I \in D} \left| \int_I f \right| \right\}$ avec $D \in \mathcal{D}$ est majoré par $\int_a^b g$ en raison de l'inégalité triangulaire et de l'hypothèse de domination. Ainsi $|f|$ est intégrable. □

On peut également retrouver des résultats de l'intégration de Lebesgue comme le lemme de Fatou qui permettent d'obtenir un théorème de convergence dominée. Nous ne le ferons pas ici.

En revanche, nous allons maintenant montrer un moyen de définir rapidement la mesure de Lebesgue à l'aide de l'intégrale KH sur un intervalle fermé borné. Le passage aux intervalles quelconques s'obtiendra par convergence monotone sur la suite croissante d'intervalles $[-n; n]$ et ne sera pas détaillé ici.

Avant cela, deux résultats qui portent sur le maximum de deux fonctions.

Lemme : Intégrabilité de f^+ et f^-

Soit f une fonction KH intégrable sur un intervalle.

Alors $|f|$ si et seulement si f^+ et f^- sont KH intégrables.

Démonstration. C'est évident en remarquant que $|f| = f^+ + f^-$ et que $f^+ = \frac{|f| + f}{2}$ et $f^- = \frac{|f| - f}{2}$ □

Proposition : Intégrabilité du maximum et du minimum de deux fonctions

Soient f et g deux fonctions KH intégrables sur un intervalle. On suppose que $|f - g|$ est KH intégrable. Alors $\max(f; g)$ et $\min(f; g)$ sont intégrables.

Démonstration. C'est évident d'après le lemme qui précède. En effet, $\max(f; g) = g + (f - g)^+$ et $\min(f; g) = f - (f - g)^+$ □

Théorème : Mesure de Lebesgue : une définition à l'aide des intégrales KH

On se place sur un intervalle fermé borné $[a; b]$.

Alors, pour tout borélien A de $[a; b]$, $\mathbb{1}_A$ est intégrable sur $[a; b]$ et de plus $\lambda(A) = \int_a^b \mathbb{1}_A$ définit la mesure de Lebesgue de A .

Pour prouver ce théorème, on va utiliser des arguments de classes monotones (pour l'unicité) et de stabilité (pour l'existence).

Démonstration. Il est clair que si A est un intervalle de $[a; b]$, $\mathbb{1}_A$ est KH intégrable et que dans ce cas l'intégrale et la mesure de Lebesgue coïncident.

Il est clair également que si $\mathbb{1}_A$ est KH intégrable alors $1 - \mathbb{1}_A$ est également KH intégrable. Cela montre la stabilité par complémentaire.

On va maintenant montrer la stabilité par intersection. Mais cela est simple compte-tenu de ce qui précède. En effet, $\mathbb{1}_{A \cap B} = \min(\mathbb{1}_A; \mathbb{1}_B)$.

Comme on a stabilité par complémentaire et par intersection alors on a stabilité par union et différence.

Reste à prouver la stabilité par passage à la limite supérieure pour achever cette démonstration. Mais cela est évident en raison du théorème de convergence monotone. Ainsi, si (A_n) est une suite croissante d'ensembles KH intégrables, on sait que $\int_a^b \mathbb{1}_{A_n}$ est majorée par $b - a$, ce qui permet de conclure.

Les boréliens sont donc KH intégrables.

De plus, il est clair que l'intégrable est σ -additive.

Enfin, comme l'intégrale et la mesure de Lebesgue coïncident sur tous les intervalles de $[a; b]$, ils coïncident sur les boréliens par application du théorème des classes monotones (voir le premier chapitre). □

Table des matières

1	Introduction à la théorie de la mesure	1
1.1	Cardinaux, sommes dénombrables, nouveau formalisme sur les ensembles	1
1.2	Tribu, mesure	3
2	L'intégrale de Lebesgue	9
2.1	Fonctions mesurables	9
2.2	Fonctions échelonnées positives	14
2.3	Intégrale d'une fonction positive	15
2.4	Intégrale d'une fonction à valeurs réelles ou complexe	19
3	La mesure de Lebesgue	25
3.1	Théorème de Carathéodory	25
3.2	Applications	29
3.3	Le théorème de Riesz	35
4	L'intégrale de Riemann	37
4.1	Définition	37
4.2	Applications	41
5	Méthodes de quadratures	49
5.1	Autour de la formule d'Euler Mac Laurin	49
5.2	Premiers résultats sur les quadratures	57
5.3	Description et propriétés des méthodes	61
6	Compléments sur l'intégration	65
6.1	Tribu, mesure produit, théorème de Fubini-Tonelli	65
6.2	Inégalités, espaces \mathbf{L}^p , dérivation	69
7	Convolution et transformée de Fourier	73
7.1	Convolution	73
7.2	Transformée de Fourier : définitions et premières propriétés	80
8	Espace de Schwartz, transformée de Fourier et mesures	85
8.1	Généralités sur l'espace de Schwartz	85
8.2	Transformée de Fourier dans \mathcal{S}	88
8.3	Étude des suites de mesures finies	92
9	Probabilités	101
9.1	Quelques résultats préliminaires	101
9.2	Convergences presque sûrement, convergence \mathcal{L}^p , convergence en probabilité	110
9.3	Loi forte des grands nombres	114
9.4	Fonctions caractéristiques	119
9.5	Quelques lois	122

10 Séries de Fourier	129
10.1 Polynômes trigonométriques	129
10.2 Séries de Fourier, noyau de Dirichlet	131
11 Intégrale de Kurtzweil-Henstock	137
11.1 Définition et premières propriétés de l'intégrale de Kurzweil-Henstock	137
11.2 Des résultats de convergence	144