

Introduction générale et rappels

Après quelques rappels, le cours proprement dit comprendra quatre chapitres : l'intégrale de Riemann sur \mathbf{R} , l'intégrale de Lebesgue sur \mathbf{R} , les séries de Fourier et l'intégrale de Lebesgue sur \mathbf{R}^d . Dans cette introduction j'essaie d'expliquer quelques uns des objectifs mathématiques de ce cours.

Les principes que j'essaie de mettre en œuvre dans ce cours tiennent en quatre points :

- 1) Dégager les idées essentielles.*
- 2) Donner des énoncés simples et efficaces.*
- 3) Fournir de multiples exemples pour chaque notion étudiée.*
- 4) Faire des dessins partout (sauf dans ce polycopié!).*

0.1 Qu'est-ce que l'intégration ?

0.1.1 L'objectif de l'intégrale

Si l'on revient aux sources, l'objectif d'une théorie de l'intégration est le suivant. On a des fonctions $y = f(x)$, définies sur un ensemble X , qui peut être *a priori* assez général (une partie de \mathbf{R} , de \mathbf{R}^d , etc.) et à valeurs réelles ou complexes. On cherche à mesurer la taille d'une fonction f (disons positive) par un nombre (une "norme" de f). Il y a plusieurs possibilités. La plus simple est de regarder le maximum de f : $\|f\|_\infty = \sup_{x \in X} f(x)$ (la norme de la convergence uniforme). Mais c'est une mesure de f dans une seule direction, celle des y . L'intégrale $\int_X f(x) dx = \|f\|_1$, au contraire, est un nombre qui mesure de la taille de la fonction dans les deux directions : en x et en y (c'est la norme de la convergence "en moyenne"). L'idée intuitive qui

nous guide c'est bien sûr le cas où X est un intervalle de \mathbf{R} et où l'intégrale c'est simplement **l'aire** sous la courbe $y = f(x)$.

On peut retenir cela sous forme d'une maxime : *Une intégrale a deux façons d'être grande, si la fonction prend des grandes valeurs sur un domaine pas trop petit, ou si elle prend des valeurs pas trop petites sur un grand domaine : les grands maigres et les petits gros.*

0.1.2 Propriétés attendues

Selon Lebesgue lui-même, une intégrale digne de ce nom doit vérifier plusieurs propriétés :

- 1) Elle doit être **linéaire** : $\int_X (\lambda f + \mu g) = \lambda \int_X f + \mu \int_X g$.
- 2) Elle doit être **positive** (donc croissante) : si on a $f \geq 0$, on a $\int_X f \geq 0$ (ce qui implique, avec 1), que si on a $f \leq g$, on a $\int_X f \leq \int_X g$).
- 3) Elle doit avoir de bonnes propriétés de **convergence** ; on attend un théorème du genre : sous certaines hypothèses, si on a $f = \lim f_n$, on a $\int_X f = \lim \int_X f_n$ (autrement dit, on peut intervertir les signes \lim et \int).

C'est cette troisième propriété qui fait toute la différence entre la théorie de Riemann et celle de Lebesgue : tout est dans les hypothèses. Évidemment on souhaite qu'elles soient les plus faibles possible et on verra au chapitre 2 que l'intégrale de Lebesgue permet un grand progrès à cet égard avec les théorèmes de convergence monotone et de convergence dominée.

Cette propriété de passage à la limite admet des variantes avec des séries au lieu de suites : $\int_X \sum_n f_n = \sum_n \int_X f_n$ (intersion des signes \int et Σ). Elle permet aussi d'étudier les intégrales dépendant d'un paramètre $F(t) = \int_X f(t, x) dx$. En particulier on obtiendra un théorème de "dérivation sous le signe somme" : $F'(t) = \int_X \frac{\partial f(x, t)}{\partial t} dx$ (sous des hypothèses minimales).

Parmi les innombrables applications de ces résultats de convergence on étudiera au chapitre 3 les séries de Fourier :

$$\sum_{n=0}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx$$

qui sont un outil essentiel pour étudier les fonctions périodiques et donc les phénomènes de vibrations et oscillations de la physique.

0.1.3 Construction de l'intégrale

Les fonctions quelconques

Pour obtenir les théorèmes de Lebesgue, il faut passer par une construction de l'intégrale assez sophistiquée. Notons déjà que l'essentiel, pour savoir définir des intégrales, est de savoir le faire pour fonctions réelles positives. En effet, si f est une fonction réelle de signe quelconque on peut toujours l'écrire $f = f^+ - f^-$ avec $f^+(x) = \max(f(x), 0)$ et $f^-(x) = \max(-f(x), 0)$ (on notera qu'on a aussi $|f| = f^+ + f^-$). On définit alors, par linéarité, $\int f = \int f^+ - \int f^-$. Si f est à valeurs complexes, on l'écrit $f = \operatorname{Re} f + i \operatorname{Im} f$ et on est ainsi ramené par linéarité au cas des fonctions à valeurs réelles.

Intégrer les constantes

Pour les fonctions positives, on commence par le commencement : les constantes. Là, en dimension 1, il y a une formule, sans doute la plus importante de ce cours, mais que les étudiants instruits ont peut-être déjà rencontrée :

l'aire d'un rectangle est le produit de sa longueur par sa largeur.

On généralise cette formule en dimensions plus grandes en une formule "du cylindre". Si on a une fonction constante et égale à b sur une partie A d'un ensemble X , son intégrale est donnée par la formule *base* \times *hauteur* : $\mu(A)b$ où $\mu(A)$ désigne la "mesure" de A . Dans le cas où la fonction constante est égale à 1 sur A et nulle ailleurs (ce qu'on appelle la fonction caractéristique de A , notée χ_A ou encore $\mathbf{1}_A$) l'intégrale est exactement égale à $\mu(A)$.

Bien entendu, avant d'écrire une telle formule, encore faut-il savoir ce qu'est cette mesure. C'est assez clair pour un intervalle $[a, b]$: $\mu([a, b]) = b - a$, voire pour un pavé, mais pour une partie quelconque, c'est une autre histoire et un autre aspect de la théorie : la théorie de la mesure.

Mesures

Au sens mathématique du terme, une mesure sur un ensemble X est une application, définie sur certaines parties de X et à valeurs réelles ≥ 0 , qui associe à une partie A un nombre $\mu(A)$: sa mesure. Une telle mesure doit vérifier l'axiome essentiel suivant :

Si A et B sont deux parties disjointes, on a $\mu(A \cup B) = \mu(A) + \mu(B)$.

On dit que μ est (simplement) additive. Il y a une variante de cette condition où l'on impose une additivité dénombrable (ou σ -additivité), toujours pour des parties disjointes : $\mu\left(\bigcup_{n \in \mathbf{N}} A_n\right) = \sum_{n \in \mathbf{N}} \mu(A_n)$. Cette propriété d'additivité dénombrable est l'un des avantages essentiels de la théorie de Lebesgue sur celle de Riemann. En voici un exemple : les rectangles de longueur 1 et de largeur $1/2^n$, $n \geq 1$ posés les uns sur les autres et qui forment le carré

unité. Avec Riemann, on ne peut pas écrire *a priori* que l'aire du carré est la somme des aires des rectangles. La notion de dénombrabilité sera rappelée ci-dessous.

La construction

Après les fonctions constantes, en pensant toujours à la linéarité, on peut intégrer les combinaisons linéaires de ces fonctions : $\int \sum_{i=1}^n \lambda_i \chi_{A_i} = \sum_{i=1}^n \lambda_i \mu(A_i)$. Si les A_i sont des intervalles (voire des pavés), ce sont les fonctions en escalier (cf. Ch. 1), s'ils peuvent être plus compliqués, on parle de fonctions étagées (cf. Ch. 4). Il reste à passer aux fonctions quelconques.

C'est l'objectif d'une théorie de l'intégration et donc de ce cours ! Il y a pour cela plusieurs outils essentiels : au chapitre 1 (resp. 2) nous utiliserons des techniques d'encadrement (resp. de passage à la limite), à partir des fonctions en escalier, au chapitre 4, des méthodes de passage à la borne supérieure à partir des fonctions étagées générales.

On obtient ainsi des fonctions intégrables (celles que l'on sait intégrer). Il y a en fait deux notions : intégrable au sens de Riemann et au sens de Lebesgue et nous les caractériserons. On notera qu'il y a beaucoup plus de fonctions intégrables au sens de Lebesgue, y compris des fonctions un peu méchantes comme la fonction caractéristique de \mathbf{Q} ou celles des ensembles de Cantor, mais c'est un des avantages de la théorie.

0.1.4 D'autres théorèmes

Hormis les résultats de convergence évoqués ci-dessus, d'autres résultats sont essentiels. D'abord, en dimension 1, ceux que le lecteur a déjà vu en DEUG (pardon, en L1-L2) : le lien intégrale-primitive, l'intégration par parties et le changement de variables, cf. Ch. 1.

Dans le chapitre 3 on abordera le problème des séries de Fourier avec un autre outil, de nature géométrique : le produit scalaire et on en verra la puissance.

En dimension plus grande, deux résultats se détachent, cf. Ch. 4 : le théorème de Fubini, qui permet de ramener une intégrale multiple à une intégrale simple, et la formule de changement de variables.

0.2 Rappels

0.2.1 Propriétés des nombres réels

0.2.1 Proposition-Définition. *Il existe un unique corps ordonné archimédien appelé corps des nombres réels et noté \mathbf{R} qui vérifie l'un au choix des axiomes suivants :*

- 1) *Toute partie non vide et majorée $A \subset \mathbf{R}$ admet une borne supérieure.*
- 2) *Toute suite de Cauchy de nombres réels est convergente.*
- 3) *La propriété de Bolzano-Weierstrass : si (x_n) est une suite bornée de nombres réels il existe une sous-suite (x_{n_k}) convergente.*
- 4) *Si (x_n) et (y_n) sont deux suites adjacentes de nombres réels elles ont une limite commune $x \in \mathbf{R}$. Ce nombre est l'unique x qui vérifie $x_n \leq x \leq y_n$ pour tout n .*

Le corps \mathbf{R} contient le corps \mathbf{Q} des rationnels.

Borne supérieure

Rappelons ce qu'est la borne supérieure¹ d'un ensemble A .

D'abord un majorant de A est un nombre M qui vérifie : $\forall x \in A, x \leq M$. Ensuite la borne supérieure M de A , notée $\sup A$, est le plus petit des majorants. Cela signifie :

- 1) C'est un majorant : $\forall x \in A, x \leq M$.
- 2) $\forall \epsilon > 0, \exists x \in A$, tel que $M - \epsilon < x \leq M$.

Cette deuxième condition exprime que M est le plus petit des majorants. En effet, pour tout $\epsilon > 0, M - \epsilon$ n'en est pas un puisqu'il y a un x de A qui est plus grand.

Voici un exemple de manipulation de cette notion :

0.2.2 Proposition. *Soient A et B deux parties majorées de \mathbf{R} , avec $A \subset B$. Alors, on a $\sup A \leq \sup B$.*

Démonstration. Comme $\sup B$ est un majorant de B , on a, pour tout $b \in B, b \leq \sup B$. Comme A est inclus dans B , on a, *a fortiori*, $a \leq \sup B$ pour tout $a \in A$. On voit que $\sup B$ est un majorant de A . Comme $\sup A$ est le plus petit majorant de A on a bien $\sup A \leq \sup B$.

Suites de Cauchy

Rappelons la notion de suite de Cauchy dont nous ferons un usage constant tout au long de ce cours. Une suite (disons de réels) (x_n) est dite de Cauchy si elle vérifie :

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall p, q \in \mathbf{N}, (p, q \geq N \implies |x_p - x_q| \leq \epsilon).$$

¹Il y a évidemment une notion de borne inférieure, duale de celle-ci.

Il est clair qu'une suite convergente (disons vers l) est de Cauchy à cause de l'inégalité triangulaire : $|x_p - x_q| \leq |x_p - l| + |l - x_q|$, mais la réciproque n'est pas évidente (c'est un théorème dans le cas de \mathbf{R} , voire un axiome). Ainsi, cette assertion est fautive si l'on se limite aux rationnels. En effet, si l'on considère la suite de rationnels :

$$x_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!},$$

on sait qu'elle converge vers e , de sorte que cette suite est de Cauchy dans \mathbf{R} (et donc aussi dans \mathbf{Q}), mais elle n'est pas convergente dans \mathbf{Q} car e est irrationnel. Un espace dans lequel toute suite de Cauchy converge est dit **complet**. C'est le cas de \mathbf{R} , mais pas de \mathbf{Q} (on sent bien qu'il "manque" des choses dans \mathbf{Q} , par exemple $\sqrt{2}$, e , π , etc.).

0.2.2 Dénombrabilité

Si E est un ensemble fini la notion de cardinal de E , qui n'est rien d'autre que le nombre des éléments de E , est bien connue. Cette notion est liée de façon fondamentale à celle de bijection, puisque c'est en mettant deux ensembles en bijection (les moutons et les cailloux par exemple) que s'est dégagé le concept même de nombre. Dans le cas des ensembles infinis, on fonde encore la notion de cardinal sur celle de bijection mais les choses sont plus difficiles.

0.2.3 Définition. *On dit que deux ensembles E et F ont même cardinal s'il existe une bijection f de E sur F . On dit qu'un ensemble E est dénombrable s'il a le même cardinal que \mathbf{N} , c'est-à-dire s'il existe une bijection $f : \mathbf{N} \rightarrow E$.*

0.2.4 Exemples.

- 1) L'ensemble $\mathbf{N} - \{0\}$ a même cardinal que \mathbf{N} (bien qu'il lui manque 0). En effet, on a une bijection $f : \mathbf{N} \rightarrow \mathbf{N} - \{0\}$ donnée par $f(n) = n + 1$. On montre de même que \mathbf{N} est en bijection avec $\mathbf{N} - \{0, 1, \dots, n\}$.
- 2) L'ensemble $2\mathbf{N}$ des nombres pairs a même cardinal que \mathbf{N} bien qu'il lui manque un ensemble infini (les impairs). En effet, on a la bijection $n \mapsto 2n$ de \mathbf{N} sur $2\mathbf{N}$.

Plus généralement on a la proposition suivante :

0.2.5 Proposition. *Soit E un sous-ensemble infini de \mathbf{N} . Alors E est dénombrable.*

Démonstration. On construit par récurrence une bijection $f : \mathbf{N} \rightarrow E$ en utilisant le fait que toute partie non vide de \mathbf{N} admet un plus petit élément. On prend pour $f(0)$ le plus petit élément de E , puis pour $f(1)$ le plus petit élément de $E - \{f(0)\}$. Supposons $f(0), f(1), \dots, f(n)$ construits. On définit $f(n+1)$ comme le plus petit élément de $E - \{f(0), f(1), \dots, f(n)\}$ (cet ensemble est non vide sinon E serait fini). Je dis que l'application f ainsi construite est bijective. En effet, deux éléments distincts de \mathbf{N} ont, par définition, des images distinctes. Par ailleurs, tout élément de E est atteint par f (précisément, si n est un élément de E , on a $n = f(p)$ où p est le cardinal de $E \cap \{0, 1, \dots, n-1\}$).

0.2.6 Corollaire. *L'ensemble des nombres premiers est dénombrable.*

0.2.7 Corollaire. *Tout ensemble infini contenu dans un ensemble dénombrable est dénombrable. Si E est infini et si $f : E \rightarrow \mathbf{N}$ est une application injective, E est dénombrable.*

Jusqu'ici on a vu que certains sous-ensembles de \mathbf{N} étaient "aussi gros" que \mathbf{N} . Voici maintenant des ensembles contenant \mathbf{N} mais pourtant pas plus gros.

0.2.8 Proposition. *L'ensemble \mathbf{Z} est dénombrable.*

Démonstration. On construit une bijection de \mathbf{Z} sur \mathbf{N} en posant $f(n) = 2n$ si $n \geq 0$ et $f(n) = -2n - 1$ si $n < 0$.

0.2.9 Remarque. L'idée de la démonstration précédente est très simple : il s'agit de ranger \mathbf{Z} dans \mathbf{N} . Pour cela on fabrique, avec les pairs et les impairs, deux " tiroirs", qui sont aussi gros que \mathbf{N} en vertu de 0.2.5, et dans lesquels on range d'un côté les entiers positifs et de l'autre les négatifs. On peut d'ailleurs, en utilisant les congruences modulo n faire n tiroirs au lieu de deux.

0.2.10 Proposition. *Soit E un ensemble. On suppose que E s'écrit comme réunion disjointe*

$$E = \bigcup_{i \in I} E_i$$

où I et les E_i sont tous dénombrables. Alors, E est dénombrable.

Démonstration. En vertu de 0.2.7, il suffit de construire une application injective de E dans \mathbf{N} . Pour cela on considère l'ensemble P des nombres premiers que l'on numérote (il est en bijection avec \mathbf{N}) : $P = \{p_0, p_1, p_2, \dots\}$. Par

ailleurs, on a une bijection de I sur \mathbf{N} et on peut donc supposer $I = \mathbf{N}$. On a aussi une bijection de E_i avec \mathbf{N}^* : $E_i = \{a_{i,1}, \dots, a_{i,n}, \dots\}$. On définit alors $f : E \rightarrow \mathbf{N}$ en posant $f(a_{i,n}) = p_i^n$. On voit que f est une bijection de E sur son image, donc une injection, cqfd.

0.2.11 Remarque. Dans cette démonstration, l'idée, comme pour 0.2.8, est de fabriquer des tiroirs dans \mathbf{N} , mais en nombre infini pour y ranger tous les ensembles E_i . Cette fois, ce sont les puissances des nombres premiers qui constituent nos tiroirs.

0.2.12 Corollaire. *L'ensemble \mathbf{Q} est dénombrable.*

Démonstration. On peut écrire \mathbf{Q} comme la réunion (dénombrable) des ensembles \mathbf{Q}_b des rationnels écrits sous forme de fractions irréductibles a/b avec $b \in \mathbf{N}^*$. Il reste donc à montrer que les \mathbf{Q}_b sont dénombrables. Or, ces ensembles sont clairement infinis et contenus dans \mathbf{Z} , donc dénombrables.

0.2.13 Corollaire. *Le produit de deux ensembles dénombrables est dénombrable.*

Démonstration. En effet, on a $A \times B = \bigcup_{a \in A} (\{a\} \times B)$.

Jusqu'ici, les ensembles infinis que nous avons considérés ($\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q}$) étaient tous dénombrables, c'est-à-dire de même cardinal que \mathbf{N} . Le corps \mathbf{R} , en revanche, a un cardinal strictement plus grand, puisque nous allons montrer qu'il n'est pas dénombrable.

0.2.14 Théorème. *Le corps des réels n'est pas dénombrable.*

Démonstration. Nous raisonnons par l'absurde en supposant \mathbf{R} dénombrable. *A fortiori*, l'ensemble des réels compris entre 0 et 1 (1 exclu) serait lui aussi dénombrable comme sous-ensemble infini d'un ensemble dénombrable, voir 0.2.5. Cela signifie qu'on pourrait numéroter les réels de $[0, 1[: x_1, x_2, \dots, x_n, \dots$ cette suite les épuisant tous. Nous allons montrer que ceci est impossible en construisant un réel de $[0, 1[$, distinct de tous ceux de la suite (x_n) .

Pour cela on considère les développements décimaux illimités de x_1, \dots, x_n, \dots et on construit x en se donnant son développement décimal illimité (propre) $x = 0, a_1 a_2 \dots a_n \dots$ comme suit. On choisit a_1 différent du premier chiffre après la virgule de x_1 et de 9 (pour éviter les développements impropres). C'est possible puisque on doit juste éviter deux chiffres parmi dix. On choisit ensuite a_2 distinct de 9 et du deuxième chiffre après la virgule de x_2 , et ainsi de suite, on choisit a_n distinct de 9 et du n -ième chiffre après la virgule de x_n . On a construit ainsi un nombre réel, qui est dans $[0, 1[$ puisque le

développement est propre. De plus, il est différent de tous les termes de la suite. En effet, il est différent de x_1 car ils n'ont pas le même premier chiffre après la virgule, de x_2 à cause du deuxième chiffre, de x_n à cause du n -ième chiffre : cqfd.

0.2.3 Normes

0.2.15 Définition. Soit E un espace vectoriel. Une application $N : E \rightarrow \mathbf{R}^+$ est une norme sur E si elle vérifie les trois propriétés suivantes :

- 1) on a $N(x) = 0 \iff x = 0$,
 - 2) on a $N(\lambda x) = |\lambda|N(x)$ pour tout $x \in E$ et tout $\lambda \in \mathbf{R}$,
 - 3) on a $N(x + y) \leq N(x) + N(y)$ pour tous $x, y \in E$.
- On note souvent $N(x) = \|x\|$.

0.2.16 Exemples.

- 1) Si $E = \mathbf{R}^n$ on a trois normes classiques pour $x = (x_1, \dots, x_n) : \|x\|_\infty = \sup |x_i|$, $\|x\|_1 = \sum_i |x_i|$, $\|x\|_2 = \sqrt{\sum_i x_i^2}$.
- 2) Si E est l'espace vectoriel des fonctions continues sur $[a, b]$ on a, entre autres, $\|f\|_\infty = \sup_x |f(x)|$ et $\|x\|_1 = \int_a^b |f(x)|dx$.

Quand on dispose d'une norme sur un espace vectoriel E on définit aussitôt une distance sur E en posant $d(x, y) = \|x - y\|$. On a aussi des notions de convergence, de continuité, etc. On les écrit exactement comme sur \mathbf{R} en remplaçant les valeurs absolues par les normes. Par exemple, si on a une norme sur l'espace vectoriel E , une suite (x_n) d'éléments de E converge vers $x \in E$ si l'on a :

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n \in \mathbf{N}, (n \geq N \implies \|x_n - x\| \leq \epsilon).$$

0.2.4 Convergence uniforme

Rappelons la définition :

0.2.17 Définition. Soit X un ensemble et soit (f_n) une suite de fonctions de X dans \mathbf{R} ou \mathbf{C} . On dit que la suite (f_n) converge uniformément (sur X) vers une fonction $f : X \rightarrow \mathbf{R}$ (ou \mathbf{C}) si on a :

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n \in \mathbf{N}, n \geq N \implies (\forall x \in X, |f(x) - f_n(x)| \leq \epsilon).$$

C'est exactement la convergence pour la norme $\|\cdot\|_\infty$ car la dernière expression équivaut à $\|f - f_n\|_\infty \leq \epsilon$.

Rappelons aussi le théorème :

0.2.18 Théorème. *On suppose que X est muni d'une distance d (on peut penser à une partie de \mathbf{R} ou \mathbf{R}^d). Si la suite (f_n) converge uniformément vers f sur X et si les f_n sont continues en $x_0 \in X$, il en est de même de f .*

Démonstration. On se donne $\epsilon > 0$. On veut majorer $|f(x) - f(x_0)|$. Pour cela on décompose :

$$f(x) - f(x_0) = f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)$$

et on applique l'inégalité triangulaire :

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|.$$

Comme la convergence est uniforme², il existe un N tel que le premier et le troisième termes soient $\leq \epsilon$ pour tout $n \geq N$ et pour tout $x \in X$. Fixons un tel n . Comme f_n est continue en x_0 le second terme est $\leq \epsilon$ pour $d(x, x_0) < \eta$. On a gagné.

0.2.19 Corollaire. *Dans la situation précédente, si les f_n sont continues sur tout X il en est de même de f .*

²C'est une réflexion intéressante de se demander pourquoi la convergence simple ne marche pas. On se rend compte assez vite qu'il faut, soit pouvoir majorer le premier terme indépendamment de x (convergence uniforme), soit le second indépendamment de n (équicontinuité).

Chapitre 1

L'intégrale de Riemann

Dans tout ce qui suit on travaille avec une fonction $f : [a, b] \rightarrow \mathbf{R}$, définie sur un intervalle **fermé borné** $[a, b]$ (on supposera $a \leq b$, sauf mention expresse du contraire) et on suppose que f est **bornée**.

1.1 Fonctions intégrables

Pour définir l'intégrale d'une fonction on pense à l'aire sous la courbe représentative. Pour une fonction constante positive égale à λ sur $[a, b]$ l'intégrale va être l'aire du rectangle : $\lambda(b - a)$. On convient de prendre aussi cette définition si λ est < 0 . Si on a une fonction en escalier sur $[a, b]$, c'est-à-dire une fonction constante et égale à λ_i sur les intervalles $]a, s_1[$, $]s_1, s_2[$, ..., $]s_{n-1}, b[$, son intégrale sera la somme des aires des rectangles : $\sum_{i=0}^{n-1} \lambda_i (s_{i+1} - s_i)$. Pour une fonction f quelconque, le principe de l'intégrale de Riemann est d'encadrer f par des fonctions en escalier.

1.1.1 Subdivisions, fonctions en escalier, sommes de Darboux

1.1.1 Définition.

- 1) On appelle *subdivision* du segment $I = [a, b]$ un ensemble fini S de points de I , contenant les extrémités. On peut toujours ordonner $S : s_0 = a < s_1 < s_2 < \dots < s_{n-1} < s_n = b$.
- 2) Le *pas* de la subdivision est la quantité $|S| = \max(s_{i+1} - s_i)$.
- 3) On dit qu'une subdivision T est *plus fine* que S si on a $S \subset T$.

1.1.2 Exemple. Si n est un entier > 0 , la subdivision formée des points $s_i = a + \frac{i(b-a)}{n}$ pour $i = 0, \dots, n$, est appelée **subdivision régulière** d'ordre n .

1.1.3 Définition. Soit $f : [a, b] \rightarrow \mathbf{R}$. On dit que f est une fonction **en escalier** s'il existe une subdivision $S = \{a = s_0, s_1, \dots, s_n = b\}$ telle que f soit constante (égale à λ_i) sur chaque intervalle ouvert $]s_i, s_{i+1}[$, pour $i = 0, \dots, n - 1$. On note $f = \text{esc}(S, \lambda_i)$.

Une fonction $f : \mathbf{R} \rightarrow \mathbf{R}$ est dite en escalier s'il existe un segment $[a, b]$ tel que f soit en escalier en restriction à $[a, b]$ et nulle en dehors de $[a, b]$.

1.1.4 Remarque. On notera qu'on ne dit rien des valeurs de f en les points de la subdivision. En particulier, la notation $f = \text{esc}(S, \lambda_i)$ désigne plusieurs fonctions.

1.1.5 Proposition. La somme, le produit de deux fonctions en escalier sont des fonctions en escalier. Si f est en escalier, $g \circ f$ aussi, quelle que soit g . Si f est une fonction en escalier, l'ensemble des x tels que $f(x) \geq k$ (resp. $> k$, $\leq k$, $< k$) est une réunion finie d'intervalles (peut-être réduits à un point).

Démonstration. La première assertion est facile en utilisant la réunion des subdivisions. Pour la seconde, $g \circ f$ est associée à S et aux $g(\lambda_i)$. Pour l'image réciproque, il suffit, pour chaque intervalle de la subdivision (et chaque point), de regarder si f est $\geq k$ sur cet intervalle.

Pour l'instant, nous n'avons pas encore défini l'intégrale de quelque fonction que ce soit, mais, moralement, l'intégrale d'une telle fonction en escalier

est $\int_a^b f(t)dt = \sum_{i=0}^{n-1} (s_{i+1} - s_i)\lambda_i$. C'est ce qui nous guide pour la suite.

Sommes de Darboux

On reprend une fonction $f : [a, b] \rightarrow \mathbf{R}$, bornée quelconque. Si $S = (s_i)$ est une subdivision de $[a, b]$, on peut encadrer f par deux fonctions en escalier g et h associées à cette subdivision. Pour cela, pour $i = 0, 1, \dots, n - 1$, on pose

$$m_i = \inf_{]s_i, s_{i+1}[} f(x) \quad \text{et} \quad M_i = \sup_{]s_i, s_{i+1}[} f(x).$$

1.1.6 Remarques.

- 1) Les bornes m_i et M_i existent car on a supposé f bornée. Point n'est besoin qu'elle soit continue.
- 2) On associe à f et S les fonctions en escalier $g(x)$ et $h(x)$ respectivement égales à m_i et M_i sur les intervalles $]s_i, s_{i+1}[$ et égales à f en les points s_i . Il est clair qu'on a $g(x) \leq f(x) \leq h(x)$.

La définition que nous avons en vue pour l'intégrale d'une fonction en escalier rend naturelle la définition suivante :

1.1.7 Définition. On appelle **sommes de Darboux** (inférieure et supérieure) associées à S et f les deux sommes

$$\sigma(S, f) = \sum_{i=0}^{n-1} (s_{i+1} - s_i) m_i \quad \text{et} \quad \Sigma(S, f) = \sum_{i=0}^{n-1} (s_{i+1} - s_i) M_i.$$

En effet, on a, "moralement", $\sigma(S, f) = \int_a^b g(t) dt$ et $\Sigma(S, f) = \int_a^b h(t) dt$.

Si S_n est la subdivision régulière d'ordre n , les sommes correspondantes sont notées $\sigma_n(f)$ et $\Sigma_n(f)$.

1.1.8 Remarque. On notera que les valeurs de m_i, M_i et les sommes $\sigma(S, f)$ et $\Sigma(S, f)$ ne changent pas si on modifie f en les points s_i .

1.1.9 Proposition. Soient S, T des subdivisions.

- 1) On a $\sigma(S, f) \leq \Sigma(S, f)$.
- 2) Si T est plus fine que S on a $\sigma(S, f) \leq \sigma(T, f)$ et $\Sigma(S, f) \geq \Sigma(T, f)$.
- 3) On a $\sigma(S, f) \leq \Sigma(T, f)$.

Démonstration. Le point 1) est clair. Pour le point 2), faisons le cas de σ , l'autre est analogue. On raisonne par récurrence sur le nombre de points supplémentaires introduits dans T . S'il y a un seul point t ajouté entre s_i et s_{i+1} , la quantité $(s_{i+1} - s_i) m_i$ est remplacée par $(t - s_i) m'_i + (s_{i+1} - t) m''_i$ où m'_i et m''_i sont respectivement les minimums de f sur $]s_i, t[$ et $]t, s_{i+1}[$. Comme on a $m_i \leq m'_i$ et $m_i \leq m''_i$ on a la conclusion.

Enfin, le point 3) résulte des précédents en considérant la subdivision $S \cup T$:

$$\sigma(S, f) \leq \sigma(S \cup T, f) \leq \Sigma(S \cup T, f) \leq \Sigma(T, f).$$

1.1.2 Fonctions intégrables : définition

La proposition précédente admet une conséquence importante :

1.1.10 Proposition. Les bornes $\sup_S \sigma(S, f)$ et $\inf_S \Sigma(S, f)$ existent et vérifient $\sup_S \sigma(S, f) \leq \inf_S \Sigma(S, f)$.

Démonstration. Soit T une subdivision quelconque. Pour toute subdivision S on a $\sigma(S, f) \leq \Sigma(T, f)$. On voit que les nombres $\sigma(S, f)$ sont majorés par $\Sigma(T, f)$. Ils ont donc une borne supérieure et, comme celle-ci est le plus petit

des majorants, elle vérifie $\sup_S \sigma(S, f) \leq \Sigma(T, f)$. Comme cela vaut pour tout T , on voit que les nombres $\Sigma(T, f)$ sont minorés par $\sup_S \sigma(S, f)$. Ils ont donc une borne inférieure qui est le plus grand des minorants, donc vérifie $\sup_S \sigma(S, f) \leq \inf_S \Sigma(S, f)$.

La définition suivante est alors bien naturelle :

1.1.11 Définition. On dit qu'une fonction bornée f est **Riemann-intégrable** si l'on a $\sup_S \sigma(S, f) = \inf_S \Sigma(S, f)$. La valeur commune de ces bornes est appelée **intégrale (de Riemann) de a à b de f** et on la note $\int_a^b f(t) dt$. On a, pour toutes subdivisions S, T : $\sigma(S, f) \leq \int_a^b f(t) dt \leq \Sigma(T, f)$.

1.1.12 Proposition. La fonction f est Riemann-intégrable si et seulement si, pour tout $\epsilon > 0$, il existe une subdivision T telle que $\Sigma(T, f) - \sigma(T, f) < \epsilon$.

Démonstration. Si on a la condition précédente, on a : $\delta = \inf_S \Sigma(S, f) - \sup_S \sigma(S, f) \leq \Sigma(T, f) - \sigma(T, f) < \epsilon$. On conclut alors grâce à la remarque suivante, triviale, mais essentielle :

1.1.13 Remarque. Si un nombre réel $\delta \geq 0$ vérifie $\delta \leq \epsilon$ pour tout $\epsilon > 0$, on a $\delta = 0$.

La réciproque vient de la définition d'une borne supérieure ou inférieure.

La propriété suivante fait intervenir le pas de la subdivision :

1.1.14 Proposition. Pour que f soit intégrable il faut et il suffit que la quantité $\Sigma(S, f) - \sigma(S, f)$ tende vers 0 lorsque le pas de la subdivision S tend vers 0. L'intégrale est alors la limite commune de $\Sigma(S, f)$ et $\sigma(S, f)$.

Démonstration. Si $\Sigma(S, f) - \sigma(S, f)$ tend vers 0 quand $|S|$ tend vers 0, on a évidemment $\inf_S \Sigma(S, f) - \sup_S \sigma(S, f) = 0$, de sorte que f est intégrable et il est clair que l'intégrale est la limite de $\Sigma(S, f)$ et $\sigma(S, f)$.

Supposons maintenant f intégrable. Nous allons montrer que $\sigma(S, f)$ tend vers $\int_a^b f(t) dt$ quand $|S|$ tend vers 0. Cela montrera le résultat (car le raisonnement est analogue avec Σ). On note m et M les bornes de f sur $[a, b]$ et on pose $h = M - m$. Soit $\epsilon > 0$. Comme f est intégrable, il existe une subdivision T , à p points, telle que $\int_a^b f - \sigma(T, f) \leq \epsilon$. Soit S une subdivision quelconque et posons $S' = S \cup T$. On a a fortiori $\int_a^b f - \sigma(S', f) \leq \epsilon$. Nous allons montrer que, si le pas de S est assez petit, on a $\sigma(S', f) - \sigma(S, f) \leq \epsilon$. On suppose d'abord $|S| < \text{Min}(t_{j+1} - t_j) = \tau$, de sorte que chaque intervalle $[s_i, s_{i+1}]$ contient au plus un point de T . Les intervalles de S qui ne contiennent pas de

points de T apportent une même contribution aux sommes $\sigma(S', f)$ et $\sigma(S, f)$. Si l'intervalle $[s_i, s_{i+1}]$ contient le point t , la différence entre les deux contributions, sur cet intervalle, est égale, avec les notations de la preuve de 1.1.9, à $(t - s_i)(m'_i - m_i) + (s_{i+1} - t)(m''_i - m_i)$, donc majorée par $(s_{i+1} - s_i)(M - m)$, donc par $|S|h$. Pour les p intervalles qui contiennent des points de T elle est donc majorée par $p|S|h$. Il suffit alors d'avoir $|S| < \text{Min}(\epsilon/ph, \tau)$ pour que $\sigma(S', f) - \sigma(S, f)$ soit $\leq \epsilon$, donc $\int_a^b f - \sigma(S, f) \leq 2\epsilon$.

1.1.15 Corollaire. *Pour que f soit intégrable, il faut et il suffit que la différence $\Sigma_n(f) - \sigma_n(f)$ tende vers 0 quand n tend vers l'infini. Les deux suites ont alors pour limite commune l'intégrale.*

Chasles

1.1.16 Notation. Soit $f : [a, b] \rightarrow \mathbf{R}$. On note $\sigma_{[a,b]}(f) = \sup_S \sigma(S, f)$ et $\Sigma_{[a,b]}(f) = \sup_S \Sigma(S, f)$.

1.1.17 Proposition. *Soient a, b, c trois réels avec $a < b < c$ et soit f une fonction bornée sur $[a, c]$. Les restrictions f_a et f_c de f à $[a, b]$ et $[b, c]$ sont Riemann-intégrables si et seulement si f l'est et, si l'on note $\int_a^b f(t)dt$ et $\int_b^c f(t)dt$ les intégrales de f_a et f_c , on a la relation de Chasles :*

$$\int_a^c f(t)dt = \int_a^b f(t)dt + \int_b^c f(t)dt.$$

Démonstration. C'est ici qu'on voit l'intérêt de prendre des subdivisions non régulières. Soient S et T des subdivisions quelconques de $[a, b]$ et $[b, c]$ respectivement, de sorte que $S \cup T$ est une subdivision de $[a, c]$. On a $\sigma(S, f_a) + \sigma(T, f_c) = \sigma(S \cup T, f) \leq \sigma_{[a,c]}(f)$. En passant à la borne supérieure, d'abord sur S , puis sur T , on en déduit : $\sigma_{[a,b]}(f_a) + \sigma(T, f_c) \leq \sigma_{[a,c]}(f)$ puis $\sigma_{[a,b]}(f_a) + \sigma_{[b,c]}(f_c) \leq \sigma_{[a,c]}(f)$. Avec la relation analogue sur Σ on en déduit que si f_a et f_c sont intégrables, f l'est aussi et qu'on a la relation.

Réciproquement, supposons f intégrable. Si U est une subdivision quelconque de $[a, c]$, on considère les restrictions S et T de U à $[a, b]$ et $[b, c]$ (en leur ajoutant au besoin le point b). On a alors $\sigma(U, f) \leq \sigma(S, f_a) + \sigma(T, f_c)$. En effet, tous les termes des sommes sont égaux, sauf peut-être pour les intervalles qui contiennent b où l'on a seulement l'inégalité. En passant aux sup, on a $\sigma_{[a,c]}(f) \leq \sigma_{[a,b]}(f_a) + \sigma_{[b,c]}(f_c)$. On a, de même, $\Sigma_{[a,b]}(f_a) + \Sigma_{[b,c]}(f_c) \leq \Sigma_{[a,c]}(f)$ et l'intégrabilité de f implique $\sigma_{[a,c]}(f) = \Sigma_{[a,c]}(f)$, donc aussi $\sigma_{[a,b]}(f_a) = \Sigma_{[a,b]}(f_a)$ et $\sigma_{[b,c]}(f_c) = \Sigma_{[b,c]}(f_c)$, donc l'intégrabilité de f_a et f_c .

1.1.18 Remarque. Si on a $a > b$, on convient de définir $\int_a^b f(t)dt$ comme $-\int_b^a f(t)dt$. La relation de Chasles est alors vérifiée quelle que soit la position des points a, b, c .

1.1.3 Fonctions intégrables : exemples

Les exemples suivants montrent que la définition de l'intégrale adoptée ci-dessus redonne bien ce qu'on attend pour les fonctions constantes et les fonctions en escalier.

1.1.19 Proposition. Une fonction constante, égale à λ , sur $[a, b]$, est intégrable et on a $\int_a^b f(t)dt = \lambda(b - a)$. Il en est de même si la fonction est constante égale à λ , sur $]a, b[$ et prend des valeurs arbitraires en a et b .

Démonstration. En effet, on a $\sigma(S, f) = \Sigma(S, f) = \lambda(b - a)$ pour toute subdivision.

1.1.20 Remarque. On a là un premier exemple du fait que la modification d'une fonction sur un ensemble "négligeable" (ici l'ensemble $\{a, b\}$) ne change pas l'intégrale.

1.1.21 Proposition. Une fonction en escalier $f = \text{esc}(S, \lambda_i)$ est intégrable et on a :

$$\int_a^b f(t)dt = \sigma(S, f) = \Sigma(S, f) = \sum_{i=0}^{n-1} (s_{i+1} - s_i)\lambda_i.$$

La formule $\int_a^b f(t)dt = \sigma(T, f) = \Sigma(T, f)$ est encore vraie si T est plus fine que S .

Démonstration. On applique 1.1.17 et 1.1.19 plusieurs fois.

Maintenant qu'on sait ce qu'est l'intégrale d'une fonction en escalier, on peut reformuler 1.1.12 :

1.1.22 Proposition. Une fonction f est Riemann-intégrable si et seulement si, pour tout $\epsilon > 0$, il existe deux fonctions en escalier g et h vérifiant $g \leq f \leq h$ et $\int_a^b h - \int_a^b g < \epsilon$ et on a alors $\int_a^b g \leq \int_a^b f \leq \int_a^b h$.

Démonstration. Soit S une subdivision. Si g et h sont les fonctions en escalier associées à S et aux valeurs m_i et M_i respectivement (cf. 1.1.6.2), les sommes de Darboux $\sigma(S, f)$ et $\Sigma(S, f)$ sont égales à $\int_a^b g(t)dt$ et $\int_a^b h(t)dt$, d'où le résultat par 1.1.12. Réciproquement, si f est encadrée par les fonctions en

escalier $g = \text{esc}(S, \lambda_i)$ et $h = \text{esc}(T, \mu_i)$ on peut supposer $S = T$ (quitte à les remplacer toutes deux par $S \cup T$) et on a alors $\lambda_i \leq m_i \leq M_i \leq \mu_i$, d'où il résulte :

$$\int_a^b g(t) dt \leq \sigma(S, f) \leq \Sigma(S, f) \leq \int_a^b h(t) dt$$

ce qui permet de conclure, toujours par 1.1.12.

1.1.23 Proposition. *Une fonction¹ monotone est Riemann-intégrable.*

Démonstration. Supposons par exemple f croissante et considérons la subdivision régulière d'ordre n . On a $m_i \geq f(s_i)$ et $M_i \leq f(s_{i+1})$, d'où :

$$\sigma_n(f) \geq \frac{b-a}{n}(f(a) + f(s_1) + f(s_2) + \cdots + f(s_{n-1})) \quad \text{et}$$

$$\Sigma_n(f) \leq \frac{b-a}{n}(f(s_1) + f(s_2) + \cdots + f(s_{n-1}) + f(b))$$

d'où $\Sigma_n(f) - \sigma_n(f) \leq \frac{(f(b) - f(a))(b-a)}{n}$ et la conclusion par 1.1.15.

1.1.24 Proposition. *Une fonction continue sur $[a, b]$ est Riemann-intégrable.*

Démonstration. Soit $\epsilon > 0$. Comme la fonction f est continue sur $[a, b]$ elle est uniformément continue :

$$\forall \epsilon > 0, \exists \eta > 0, \forall x, y \in [a, b] \quad (|x - y| \leq \eta \implies |f(x) - f(y)| \leq \epsilon).$$

On prend alors la subdivision régulière associée à un entier n qui vérifie $(b-a)/n < \eta$. Sur chaque intervalle de la subdivision on a $M_i - m_i \leq \epsilon$ et donc $\Sigma_n(f) - \sigma_n(f) \leq \epsilon(b-a)$.

1.1.25 Exemple. La fonction χ caractéristique des rationnels de $[0, 1]$ n'est pas intégrable. En effet, sur chaque intervalle d'une subdivision le maximum de χ vaut 1 (car il y a des rationnels dans tout intervalle non vide) et le minimum vaut 0 (car il y a des irrationnels dans tout intervalle non vide). On en déduit que les sommes $\sigma(S, f)$ valent toutes 0 tandis que les $\Sigma(S, f)$ valent toutes 1.

¹Bien entendu, les fonctions sont toujours supposées définies sur un segment $[a, b]$ et bornées.

1.2 Propriétés de l'intégrale

1.2.1 Linéarité et croissance

1.2.1 Théorème.

1) L'ensemble E des fonctions Riemann-intégrables $f : [a, b] \rightarrow \mathbf{R}$ est un sous-espace vectoriel de l'espace vectoriel de toutes les fonctions de $[a, b]$ dans \mathbf{R} et l'application $I : f \mapsto \int_a^b f(t)dt$ est une forme linéaire sur E .

2) La fonction f est intégrable si et seulement si $f^+ = \text{Max}(f, 0)$ et $f^- = \text{Max}(-f, 0)$ le sont. Si f est intégrable, $|f|$ l'est aussi².

3) L'application I est croissante : si on a $f \leq g$ on a $I(f) \leq I(g)$. En particulier, on a l'inégalité $|\int_a^b f(t)dt| \leq \int_a^b |f(t)|dt$.

Démonstration. 1) Seul le cas de l'addition de deux fonctions pose problème.

a) Notons d'abord que la formule $\int_a^b (f + g) = \int_a^b f + \int_a^b g$ est vraie dans le cas des fonctions en escalier. En effet, on peut supposer que ces fonctions sont attachées à une même subdivision en vertu de 1.1.21 et le résultat est alors évident.

b) Si on a deux fonctions intégrables f_1 et f_2 et un nombre $\epsilon > 0$, en vertu de 1.1.22, il existe des fonctions en escalier g_1, h_1, g_2, h_2 , avec $g_i \leq f_i \leq h_i$ et $\int_a^b h_i - \int_a^b g_i < \epsilon$ pour $i = 1, 2$. Mais alors, les fonctions en escalier $g_1 + g_2$ et $h_1 + h_2$ encadrent $f_1 + f_2$ et on a $\int_a^b (h_1 + h_2) - \int_a^b (g_1 + g_2) < 2\epsilon$ en vertu du a). Cela montre déjà que $f_1 + f_2$ est intégrable. De plus, $\int_a^b (f_1 + f_2)$ et $\int_a^b f_1 + \int_a^b f_2$ sont toutes deux comprises entre $\int_a^b (g_1 + g_2)$ et $\int_a^b (h_1 + h_2)$, donc leur différence est, en valeur absolue, plus petite que 2ϵ . Comme cela vaut pour tout $\epsilon > 0$, cette différence est nulle et on a bien $\int_a^b (f_1 + f_2) = \int_a^b f_1 + \int_a^b f_2$.

2) Si f^+ et f^- sont intégrables, $f = f^+ - f^-$ l'est aussi. Réciproquement, si f est intégrable, comme on a $f^- = (-f)^+$ il suffit de traiter le cas de f^+ . Soit $\epsilon > 0$ et soit S une subdivision qui vérifie $\Sigma(S, f) - \sigma(S, f) \leq \epsilon$. Sur chaque intervalle de la subdivision, si l'on note M_i et m_i les extrema de f , ceux de f^+ sont $M_i^+ = \max(M_i, 0)$ et $m_i^+ = \max(0, m_i)$. On a donc $M_i^+ - m_i^+ \leq M_i - m_i$ (distinguer selon que M_i est positif ou négatif) donc $\Sigma(S, f^+) - \sigma(S, f^+) \leq \epsilon$ et la conclusion s'ensuit. Le cas de $|f|$ en résulte par linéarité car on a $|f| = f^+ + f^-$.

Le point 3) est clair.

²Dans le cas de Riemann, la réciproque est fautive : $|f|$ intégrable n'implique pas f intégrable, exemple, la fonction qui vaut 1 sur les rationnels de $[0, 1]$ et -1 sur les irrationnels. Avec Lebesgue, le résultat sera vrai, au moins si les fonctions sont mesurables.

1.2.2 Proposition. *Si f et g sont intégrables il en est de même de fg .*

Démonstration. On se ramène au cas $f, g \geq 0$ par utilisation des fonctions du type f^+ et linéarité. Dans le cas positif, on choisit une subdivision qui donne un écart de ϵ à la fois pour f et g . Si les extrema de f, g, fg sur les intervalles de la subdivision sont $M_i, m_i, N_i, n_i, P_i, p_i$ respectivement, on a $P_i \leq M_i N_i$ et $p_i \geq m_i n_i$. On écrit alors, par la vieille ruse, $P_i - p_i \leq M_i N_i - m_i n_i = M_i(N_i - n_i) + n_i(M_i - m_i)$ et cela permet de majorer $\Sigma(S, fg) - \sigma(S, fg)$ par $(M + N)\epsilon$ où M et N désignent les maxima de f et g sur $[a, b]$.

1.2.2 Fonctions à valeurs complexes

Le fait que l'intégrale soit linéaire permet d'étendre sa définition aux fonctions de $[a, b]$ dans \mathbf{C} .

1.2.3 Définition. *Soit $f : [a, b] \rightarrow \mathbf{C}$ une fonction. On dit que f est Riemann-intégrable si sa partie réelle et sa partie imaginaire le sont et on pose :*

$$\int_a^b f(t)dt = \int_a^b \operatorname{Re} f(t)dt + i \int_a^b \operatorname{Im} f(t)dt.$$

Les propriétés vues en 1.2.1 restent valables à l'exception de 2) et 3) dans lesquelles subsiste seulement l'inégalité : $|\int_a^b f(t)dt| \leq \int_a^b |f(t)|dt$ (où $|z|$ désigne le module du nombre complexe z).

1.3 Primitives et intégrales

Les deux paragraphes qui suivent concernent des choses connues : le lien intégrale-primitive, l'intégration par parties, le changement de variables, les intégrales impropres. Ils sont juste là pour vous les remettre en mémoire. **Attention**, cela ne veut pas dire que ce n'est pas important : on ne peut rien faire sans cela et on utilisera ces techniques dans tous les exemples pratiques.

1.3.1 Le théorème fondamental

1.3.1 Théorème. *Soit f une fonction Riemann-intégrable sur $[a, b]$. Alors la fonction F définie par $F(x) = \int_a^x f(t)dt$ est continue. De plus, si f est continue, F est une primitive de f .*

Démonstration. Montrons d'abord la continuité de F . Soit M un majorant de $|f|$ sur $[a, b]$. On a, pour $h \geq 0$, $|F(x+h) - F(x)| = \left| \int_x^{x+h} f(t) dt \right| \leq \int_x^{x+h} |f(t)| dt \leq hM$ et cette quantité tend vers 0 quand h tend vers 0. Le raisonnement est analogue pour $h < 0$.

Écrivons que f est continue en x . On a :

$$\forall \epsilon > 0, \exists \eta > 0, \forall t \in \mathbf{R}, (x - \eta \leq t \leq x + \eta \implies f(x) - \epsilon \leq f(t) \leq f(x) + \epsilon).$$

Si h est positif, on en déduit, pour $h < \eta$: $(f(x) - \epsilon)h \leq F(x+h) - F(x) \leq (f(x) + \epsilon)h$. On voit ainsi que le rapport $\frac{F(x+h) - F(x)}{h}$ tend vers $f(x)$ quand h tend vers 0. Le raisonnement est analogue pour $h < 0$.

1.3.2 Applications : changement de variables, intégration par parties, etc.

On se contente d'énoncer les deux résultats et on renvoie le lecteur à ses cours de DEUG pour toutes précisions sur ces sujets.

1.3.2 Proposition. Soient u, v deux fonctions définies sur $[a, b]$ et de classe C^1 . On a la formule (dite d'intégration par parties) :

$$\int_a^b u(t)v'(t)dt = [uv]_a^b - \int_a^b u'(t)v(t)dt.$$

Démonstration. La notation $[uv]_a^b$ signifie $u(b)v(b) - u(a)v(a)$. Le résultat vient de la formule $(uv)' = u'v + uv'$.

1.3.3 Proposition. Soit I un intervalle de \mathbf{R} , $f : I \rightarrow \mathbf{R}$ une fonction continue et $\varphi : [a, b] \rightarrow I$ une fonction de classe C^1 . On a la formule :

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

Démonstration. On regarde les fonctions $F(y) = \int_{\varphi(a)}^y f(u) du$, $F(\varphi(x))$ et $G(x) = \int_a^x f(\varphi(t)) \varphi'(t) dt$.

1.3.4 Remarque. On notera qu'on ne suppose ni que φ est injective, ni que φ' ne s'annule pas.

1.3.3 Intégrales impropres

Rappelons (voir les cours de DEUG) qu'on définit, à partir de l'intégrale de Riemann, des intégrales impropres dans deux cas :

- lorsque l'intervalle d'intégration n'est pas borné,
- lorsque la fonction f n'est pas bornée.

Dans les deux cas, que la borne b soit infinie ou qu'elle soit finie mais que la fonction f soit infinie en b (par exemple), on définit $\int_a^b f(t)dt = \lim_{x \rightarrow b} \int_a^x f(t)dt$ si cette limite existe. On dit alors que l'intégrale est convergente.

Par exemple, l'intégrale $\int_1^{+\infty} \frac{dt}{t^2}$ et l'intégrale $\int_0^1 \frac{dt}{\sqrt{t}}$ sont convergentes, mais ni $\int_1^{+\infty} \frac{dt}{t}$, ni $\int_0^1 \frac{dt}{t}$.

Rappelons qu'on a une condition suffisante de convergence qui est la convergence absolue (c'est-à-dire celle de l'intégrale de $|f(t)|$). Un exemple où cette condition n'est pas remplie est celui de l'intégrale $\int_1^{+\infty} \frac{\sin t}{t} dt$. On parle alors de semi-convergence.

1.4 Intégrale et limites

Dans un espace vectoriel (de points ou de fonctions), pour étudier la convergence de suites, la continuité, etc. on a besoin de distances ou de normes. Sur l'espace vectoriel des fonctions Riemann-intégrables sur $[a, b]$ il y a *a priori* deux "normes" intéressantes :

- la norme de la convergence uniforme définie par $\|f\|_\infty = \sup_{[a,b]} |f(x)|$,
- la semi-norme de la convergence en moyenne (ou convergence L^1) définie par $\|f\|_1 = \int_a^b |f(t)| dt$.

1.4.1 Remarques.

- 1) La semi-norme L^1 n'est pas une norme car on peut avoir $\|f\|_1 = 0$ sans que f soit nulle. C'est le cas, par exemple, d'une fonction qui est nulle sauf sur un ensemble fini.
- 2) En vertu de 1.2.1.3, on a l'inégalité $\|f\|_1 \leq \|f\|_\infty (b - a)$. Cela montre que la convergence uniforme implique la convergence en moyenne.

Dans ce paragraphe nous établissons les propriétés de l'intégrale de Riemann vis-à-vis de la convergence uniforme.

1.4.1 La convergence uniforme : complétude

1.4.2 Théorème. Si (f_n) est une suite de fonctions Riemann-intégrables sur $[a, b]$ qui converge uniformément vers une fonction f , alors la fonction f est Riemann-intégrable et on a $\int_a^b f = \lim \int_a^b f_n$.

Démonstration. On écrit la convergence :

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n \in \mathbf{N} (n \geq N \implies \forall x \in I, |f(x) - f_n(x)| \leq \epsilon).$$

Cela montre déjà que f est bornée (par $m_N - \epsilon$ et $M_N + \epsilon$, si m_N et M_N désignent les bornes de f_N).

Fixons un nombre $\epsilon > 0$ arbitraire (qui fournit un N comme ci-dessus) et soit S une subdivision. Si on note $m_{n,i}$ et $M_{n,i}$ (resp. m_i, M_i) les inf et sup de f_n (resp. f) sur $[s_i, s_{i+1}]$, on voit que l'on a $m_{N,i} - \epsilon \leq m_i$ et $M_{N,i} + \epsilon \geq M_i$. On en déduit en additionnant le tout :

$$\sigma(S, f_N) - \epsilon(b-a) \leq \sigma(S, f) \leq \Sigma(S, f) \leq \Sigma(S, f_N) + \epsilon(b-a).$$

En passant au sup et inf sur les subdivisions on obtient :

$$\sup \sigma(S, f_N) - \epsilon(b-a) \leq \sup \sigma(S, f) \leq \inf \Sigma(S, f) \leq \inf \Sigma(S, f_N) + \epsilon(b-a),$$

mais, comme f_N est intégrable, on a $\sup \sigma(S, f_N) = \inf \Sigma(S, f_N) = \int_a^b f_N(t)dt$ et, finalement :

$$\int_a^b f_N(t)dt - \epsilon(b-a) \leq \sup \sigma(S, f) \leq \inf \Sigma(S, f) \leq \int_a^b f_N(t)dt + \epsilon(b-a). \quad (*)$$

On a donc $\inf \Sigma(S, f) - \sup \sigma(S, f) \leq 2\epsilon(b-a)$, mais comme ϵ est arbitraire, cela n'est possible que si $\inf \Sigma(S, f) = \sup \sigma(S, f)$, ce qui signifie que f est intégrable. La relation (*) (qui vaut en remplaçant N par un $n \geq N$) montre alors que l'intégrale de f est la limite des intégrales des f_n .

1.4.3 Corollaire. L'espace vectoriel E muni de la norme uniforme est complet.

Démonstration. Soit (f_n) une suite de Cauchy de fonctions Riemann-intégrables (au sens de la norme uniforme). On a donc :

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall p, q \in \mathbf{N}, (p, q \geq N \implies \|f_p - f_q\| \leq \epsilon).$$

La dernière inégalité signifie encore qu'on a $\forall x \in [a, b], |f_p(x) - f_q(x)| \leq \epsilon$, inégalité que nous noterons (*). Si on fixe un $x \in [a, b]$, on en déduit que la suite $(f_n(x))$ est une suite de Cauchy de réels. Elle converge donc vers un réel que l'on appelle $f(x)$. Si l'on fixe $p \geq N$ et qu'on fait tendre q vers l'infini dans la relation (*) on en déduit qu'on a, pour tout $x \in [a, b], |f_p(x) - f(x)| \leq \epsilon$, donc que (f_n) converge uniformément vers f . On conclut avec 1.4.2.

1.4.2 Fonctions réglées

Parmi les limites uniformes de fonctions Riemann-intégrables sur un segment (dont on vient de voir qu'elles étaient encore Riemann-intégrables) il y a en particulier les limites de fonctions en escalier qui constituent une classe intéressante. On peut même se demander, *a priori*, si ce ne sont pas toutes les fonctions Riemann-intégrables. On verra qu'il n'en est rien.

1.4.4 Définition. Soit $f : [a, b] \rightarrow \mathbf{C}$ une fonction. On dit que f est réglée s'il existe une suite (f_n) de fonctions en escalier telle que f soit limite uniforme de la suite (f_n) sur $[a, b]$.

Le théorème suivant (que nous ne démontrerons pas) donne une caractérisation des fonctions réglées en terme de régularité.

1.4.5 Théorème. Une fonction $f : [a, b] \rightarrow \mathbf{R}$ est réglée si et seulement si elle admet une limite à gauche et une limite à droite en chaque point.

1.4.6 Exemples.

- 1) Une fonction continue est réglée.
- 2) Une fonction monotone est réglée. (En effet, il est clair qu'elle admet en chaque point une limite à droite et une limite à gauche.)
- 3) Nous verrons au paragraphe suivant un exemple de fonction intégrable (Riemann) non réglée.

1.4.7 Proposition. Une fonction réglée sur $[a, b]$ est bornée et n'a qu'un nombre fini ou dénombrable de points de discontinuité.

Démonstration. Écrivons que f est limite uniforme d'une suite de fonctions en escalier (f_n) :

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n \in \mathbf{N}, (n \geq N \implies \forall x \in [a, b], |f(x) - f_n(x)| \leq \epsilon).$$

Si l'on se donne ϵ et qu'on fixe N , il est clair que f est bornée par $\epsilon + \max |f_N|$.

Chacune des fonctions f_n est continue sauf aux points d'un ensemble fini A_n . Appelons A la réunion des A_n . C'est un ensemble dénombrable. Si x_0 est en dehors de A , le raisonnement usuel montre que la fonction f est continue en x_0 .

1.4.8 Exemple. Notons qu'il peut y avoir une infinité de discontinuités pour une fonction réglée (même monotone). Un exemple est fourni par la fonction définie sur $[0, 1]$ par $f(0) = 0$ et $f(x) = 1/n$ sur l'intervalle $]1/(n+1), 1/n]$.

1.5 Caractérisation des fonctions Riemann-intégrables

1.5.1 Notation. Si $I = (a, b)$ est un intervalle borné de \mathbf{R} (on ne précise pas s'il est fermé, ouvert, etc.) on note $l(I)$ sa longueur, c'est-à-dire $b - a$. On utilisera aussi la même notation $l(A)$ si A est une réunion finie d'intervalles disjoints. Si on a $a_1 < b_1 \leq a_2 < b_2 \leq \dots \leq a_n < b_n$, on pose :

$$l((a_1, b_1) \cup (a_2, b_2) \cup \dots \cup (a_n, b_n)) = \sum_{i=1}^n b_i - a_i.$$

1.5.1 Les ensembles de Cantor

La construction des ensembles de Cantor, c'est comme la fabrication du gruyère : on fait des trous ! Expliquons d'abord le principe. On part du segment $[0, 1]$, on lui enlève un intervalle (ouvert) au centre (c'est-à-dire un intervalle symétrique par rapport à $1/2$). Il reste deux segments sur les bords. On leur enlève à chacun un intervalle ouvert au centre. Il reste quatre segments, on leur enlève un intervalle au centre et on recommence. On fait ça une infinité de fois. L'ensemble de Cantor c'est le fromage, c'est-à-dire ce qui reste quand on a enlevé tous les trous (s'il reste quelque chose !)

Pour formaliser cette construction, on se donne une suite de nombres > 0 , $\lambda_0, \lambda_1, \dots, \lambda_n, \dots$ qui vont être les longueurs des intervalles enlevés à chaque étape : à l'étape 0 l'intervalle enlevé à $[0, 1]$ est de longueur λ_0 , à l'étape 1 les deux intervalles enlevés sont de longueur λ_1 , etc. À l'étape n on enlève donc 2^n intervalles, tous de longueur λ_n . Si l'on veut qu'il reste quelque chose à la fin, il faut qu'on ait :

$$(*) \quad \sum_0^{+\infty} 2^n \lambda_n \leq 1.$$

Précisément, si la suite (λ_n) vérifie cette condition, on pose $K_0 = [0, 1]$ et on définit par récurrence un ensemble K_n , qui est réunion de 2^n segments, et qu'on obtient en enlevant à K_{n-1} (qui est formé, lui, de 2^{n-1} segments) les 2^{n-1} intervalles centraux de longueur λ_{n-1} de ces segments. L'ensemble de Cantor est alors défini ainsi :

1.5.2 Définition. Soit $\lambda = (\lambda_n)$ une suite vérifiant la condition (*). On note K_λ l'intersection $K_\lambda = \bigcap_{n \in \mathbf{N}} K_n$. L'ensemble K_λ est appelé **ensemble de**

Cantor relatif à λ . Dans le cas $\lambda_n = \frac{1}{3^{n+1}}$ on parle d'**ensemble triadique de Cantor**.

La proposition suivante résume les principales propriétés des K_n et de K_λ :

1.5.3 Proposition.

1) L'ensemble K_n est réunion de 2^n segments non vides et disjoints. Sa longueur totale (c'est-à-dire la somme des longueurs des 2^n intervalles qui le constituent) est égale à $1 - \sum_{i=0}^{n-1} 2^i \lambda_i$. Cette quantité est $> 2^n \lambda_n$ (donc > 0)

et a pour limite $1 - \sum_{n=0}^{+\infty} 2^n \lambda_n$.

2) L'ensemble K_λ est compact, non vide, contenu dans $[0, 1]$. Il ne contient aucun intervalle non réduit à un point.

Démonstration. Le point 1) est évident par construction (les segments sont non vides car on retire λ_n qui est $< \left(1 - \sum_{i=0}^{n-1} 2^i \lambda_i\right) / 2^n$ par définition de la suite λ). Il est clair que K_λ est compact (il est fermé borné). Il est non vide car il contient 0 et 1. Montrons qu'il ne contient pas d'intervalle non réduit à un point. Si K_λ contenait un intervalle de largeur $\epsilon > 0$, cet intervalle serait contenu dans K_n . Or, la longueur des segments constitutifs de K_n tend vers 0 avec n et elle est $< \epsilon$ pour n assez grand.

1.5.4 Remarque. Quand on aura défini une notion de mesure un peu plus générale, on montrera que la mesure de K_λ est exactement la limite des longueurs des K_n : $1 - \sum_{n=0}^{+\infty} 2^n \lambda_n$. En particulier, dans le cas du vrai Cantor (triadique) où l'on a exactement $\sum_0^{+\infty} 2^n \lambda_n = 1$ cette mesure sera nulle.

1.5.2 Ensembles négligeables

Nous définissons maintenant les ensembles négligeables de \mathbf{R} qui vont jouer un rôle essentiel dans la théorie de Riemann, et surtout dans celle de Lebesgue :

1.5.5 Définition. Soit A une partie de \mathbf{R} . On dit que A est **négligeable** (on dira plus tard que A est de mesure nulle, au sens de Lebesgue) si pour

tout $\epsilon > 0$ il existe une suite d'intervalles I_n (nécessairement bornés) tels que :

- 1) A est contenu dans la réunion des I_n ,
- 2) la somme des longueurs des I_n (somme d'une série à termes positifs) est $\leq \epsilon$.

Le lemme suivant, d'apparence anodine, est essentiel dans ces questions :

1.5.6 Lemme. (Lemme de Borel-Lebesgue) Soit K une réunion finie disjointe d'intervalles. On suppose que K est contenu dans une réunion dénombrable d'intervalles $\bigcup_{n \in \mathbf{N}} J_n$. Alors, on a $l(K) \leq \sum_{n \in \mathbf{N}} l(J_n)$.

Démonstration. 1) On note d'abord que le résultat est évident si K est contenu dans un unique intervalle J . En effet, on écrit $K = (a_1, b_1) \cup \dots \cup (a_r, b_r)$, avec $a_1 < b_1 \leq a_2 < b_2 \leq \dots \leq a_r < b_r$ et $J = (a, b)$. On a $a \leq a_1$ et $a_r \leq b$, d'où $b - a \geq \sum_i (b_i - a_i)$.

2) Montrons ensuite que si J_1, \dots, J_r sont des intervalles de \mathbf{R} , on a $l(J_1 \cup \dots \cup J_r) \leq \sum_{k=1}^r l(J_k)$.

C'est évident si les intervalles sont disjoints. Sinon, on se ramène par récurrence au cas de deux intervalles I, J . On pose $I = (a, b)$, $J = (c, d)$ et on suppose que a est le plus petit des quatre. Comme l'union est un intervalle, c'est que l'on a $c \leq b$ et $I \cup J = (a, \text{Max}(b, d))$. On a alors $l(I \cup J) = \text{Max}(b - a, d - a)$. Par ailleurs on a $l(I) + l(J) = (b - a) + (d - c)$ et cette quantité est $\geq b - a$ (c'est clair) et $\geq d - a$ (car $b - c \geq 0$).

3) Montrons le lemme dans le cas particulier (dit compact-ouvert) où K est réunion disjointe de segments K_1, \dots, K_q et où tous les J_n sont ouverts. Comme K est compact, on peut supposer que les J_n sont en nombre fini r . Considérons la réunion $J_1 \cup \dots \cup J_r$. C'est un ouvert de \mathbf{R} . Ses composantes connexes C_1, \dots, C_p sont des intervalles ouverts et chacune est réunion de certains des J_k . Chaque composante contient un certain nombre des segments K_j et la somme des longueurs de ces segments est $\leq l(C_i)$ par le point 1). Au total, on a donc $l(K) \leq \sum l(C_i) = l(C_1 \cup \dots \cup C_p) = l(J_1 \cup \dots \cup J_r)$ et cette quantité est $\leq \sum_{k=1}^r l(J_k)$ en vertu du point 2).

4) Passons maintenant au cas général. Si on avait $\sum_n l(J_n) < l(K)$, soit ϵ un nombre > 0 plus petit que la différence entre ces nombres. On remplace $K = (a_1, b_1) \cup \dots \cup (a_r, b_r)$ par $\overline{K} = [a_1, b_1] \cup \dots \cup [a_r, b_r]$ et on remplace la famille des J_n ($n \in \mathbf{N}$) par des intervalles ouverts J'_n qui débordent chaque J_n de $\epsilon/2^{n+2}$, auxquels on ajoute, en chaque extrémité des intervalles $]a_i, b_i[$, des intervalles ouverts L'_i et L''_i de largeur $\epsilon/(4r)$. La longueur totale des intervalles J'_n, L'_i, L''_i est égale à

$$\sum_n l(J'_n) + \sum_n \frac{\epsilon}{2^{n+2}} + 2r \frac{\epsilon}{4r} = \sum_n l(J_n) + \frac{\epsilon}{2} + \frac{\epsilon}{2} < l(K).$$

Mais, comme \overline{K} est contenu dans $\bigcup_n J'_n \bigcup_i (L'_i \cup L''_i)$ et que ces intervalles sont ouverts, on a $l(\overline{K}) \leq \sum l(J'_n) + \sum l(L'_i) + \sum l(L''_i)$, en vertu du cas compact-ouvert, et c'est absurde.

1.5.7 Proposition.

- 1) Un sous-ensemble d'un ensemble négligeable est négligeable.
- 2) Un point est un ensemble négligeable.
- 3) Une réunion finie ou dénombrable d'ensembles négligeables est négligeable. En particulier, un ensemble fini (resp. dénombrable) est négligeable.
- 4) L'ensemble triadique de Cantor est non dénombrable mais négligeable.
- 5) Un intervalle non réduit à un point n'est pas négligeable.
- 6) Avec les notations de 1.5.2, si la suite λ vérifie $\sum_i 2^i \lambda_i < 1$, l'ensemble de Cantor K_λ n'est pas négligeable.

Démonstration. Les points 1) et 2) sont évidents. Montrons directement la conséquence de 3). Soit $A = \{a_1, \dots, a_n, \dots\}$ un ensemble dénombrable. Soit ϵ un réel positif. On considère les intervalles $I_n = [a_n - \frac{\epsilon}{2^n}, a_n + \frac{\epsilon}{2^n}]$. Il est clair que A est contenu dans la réunion des I_n et la somme de leurs longueurs est ϵ .

Pour montrer 3), on considère une réunion dénombrable d'ensembles négligeables $A = \bigcup_{n \in \mathbf{N}^*} A_n$. Soit $\epsilon > 0$. Pour chaque n il existe une suite d'intervalles $I_{n,k}$ tels que $A_n \subset \bigcup_{k \in \mathbf{N}} I_{n,k}$ avec $\sum_{k \in \mathbf{N}} l(I_{n,k}) \leq \epsilon/2^n$. Alors, l'ensemble de tous les $I_{n,k}$ est un ensemble dénombrable d'intervalles que l'on peut numéroter comme une suite et leur réunion contient A . De plus, la somme des longueurs des $I_{n,k}$ est $\leq \sum_{n=1}^{+\infty} \frac{\epsilon}{2^n} = \epsilon$.

4) Dans le cas du Cantor triadique, l'ensemble K_n est réunion disjointe de 2^n intervalles chacun de longueur $1/3^n$, soit un total de $(2/3)^n$. Comme K est l'intersection des K_n , il est clair que K est négligeable.

Pour voir que K n'est pas dénombrable, on montre que K est l'ensemble des x de $[0, 1]$ qui admettent une écriture triadique : $x = \sum_{n=1}^{+\infty} \frac{a_n}{3^n}$ avec des a_n égaux à 0 ou à 2. Le même argument que celui utilisé pour montrer que \mathbf{R} n'est pas dénombrable montre que K ne l'est pas non plus.

5) C'est le lemme de Borel-Lebesgue, voir 1.5.6.

6) Choisissons un nombre ϵ tel que $0 < \epsilon < 1 - \sum_i 2^i \lambda_i$. Si K_λ était négligeable il serait contenu dans la réunion U d'une suite d'intervalles A_n de somme des longueurs plus petite que ϵ . Quitte à augmenter un peu le ϵ , on peut supposer ces intervalles ouverts (si on prend $\epsilon' > \epsilon$ on pourra augmenter éventuellement la longueur du n -ième intervalle de $(\epsilon' - \epsilon)/2^n$ pour le rendre ouvert). Si F désigne le complémentaire de U , on a alors $\bigcap_n (K_n \cap F) = \emptyset$. Par

compacité, cela signifie que l'un des $K_n \cap F$ est vide, donc que K_n est contenu dans U . Mais, comme la longueur de K_n est plus grande que $1 - \sum_i 2^i \lambda_i$, c'est absurde (lemme de Borel-Lebesgue, 1.5.6).

1.5.3 Caractérisation des fonctions intégrables

1.5.8 Théorème. *Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction bornée. Alors, f est Riemann-intégrable si et seulement si l'ensemble des points de discontinuité de f est négligeable.*

Démonstration. Nous montrons seulement ici que si f est intégrable l'ensemble de ses points de discontinuité est négligeable et nous admettrons la réciproque.

Dire qu'une fonction f est discontinue en x_0 c'est dire :

$$\exists h > 0, \forall \eta > 0, \exists x \in [a, b], |x - x_0| < \eta \text{ et } |f(x) - f(x_0)| \geq h.$$

Pour $h > 0$, posons :

$$\Omega_h = \{x_0 \in [a, b] \mid \forall \eta > 0, \exists x \in [a, b], \text{ avec } |x - x_0| < \eta \text{ et } |f(x) - f(x_0)| \geq h\}.$$

L'ensemble \mathcal{D} des points de discontinuité de f est égal à $\bigcup_{h>0} \Omega_h$ ou même à $\bigcup_{n \in \mathbf{N}^*} \Omega_{1/n}$. En vertu de 1.5.7 on aura gagné si on montre que Ω_h est négligeable pour tout h .

Soit $\epsilon > 0$. Nous allons montrer que Ω_h est inclus dans une réunion dénombrable d'intervalles dont la somme des longueurs est $\leq 2\epsilon$. Écrivons que f est intégrable. Cela signifie qu'il existe une subdivision S telle que l'on ait $\Sigma(S, f) - \sigma(S, f) < \epsilon h$.

Si $]s_i, s_{i+1}[$ rencontre Ω_h , on a $M_i - m_i \geq h$. La condition :

$$\sum_{i=1}^n (s_{i+1} - s_i)(M_i - m_i) \leq \epsilon h$$

assure que la somme des quantités $h(s_{i+1} - s_i)$, pour tous les intervalles $]s_i, s_{i+1}[$ qui rencontrent Ω_h , est plus petite que ϵh . Il en résulte que $\Omega_h - \{s_0, s_1, \dots, s_n\}$ est inclus dans une réunion finie d'intervalles $]s_i, s_{i+1}[$ dont la somme des longueurs est $< \epsilon$. Mais on peut aussi englober $\{s_0, s_1, \dots, s_n\}$ dans une telle réunion finie d'intervalles, et donc Ω_h est contenu dans une réunion d'intervalles de longueur $\leq 2\epsilon$.

On peut appliquer ce qui précède aux fonctions caractéristiques. La remarque fondamentale est la suivante :

1.5.9 Proposition. Soit A une partie de \mathbf{R}^d (ou plus généralement d'un espace métrique) et χ_A sa fonction caractéristique. Les points de discontinuité de χ_A sont les points de la frontière de A : $\text{Fr } A = \overline{A} - A^\circ$.

Démonstration. Si x n'est pas frontière, il est dans A° ou n'est pas dans \overline{A} et la fonction χ_A est constante et égale à 1 ou 0 au voisinage de x , donc elle est continue. Au contraire, si x est frontière, il y a, dans tout voisinage de x , à la fois des points qui sont dans A et des points qui n'y sont pas, donc des points où χ_A vaut 1 et d'autres où elle vaut 0.

1.5.10 Exemples.

1) On retrouve le fait que la fonction caractéristique des rationnels de $[0, 1]$ n'est pas Riemann-intégrable. En effet, elle est discontinue en tous les points de $[0, 1]$.

2) Le théorème permet de donner un exemple de fonction Riemann-intégrable et non réglée : la fonction caractéristique de l'ensemble triadique de Cantor K . En effet, l'ensemble triadique de Cantor est négligeable, mais non dénombrable. De plus, cet ensemble est fermé et d'intérieur vide, donc égal à sa frontière. Donc les points de discontinuité de χ_K sont tous les points de K . Comme K est négligeable, la fonction χ_K est Riemann-intégrable, comme il n'est pas dénombrable, elle n'est pas réglée (cf. 1.4.7).

3) En revanche la fonction caractéristique d'un ensemble de Cantor K_λ vérifiant $\sum_n 2^n \lambda_n < 1$ n'est pas Riemann-intégrable (car son ensemble de points de discontinuité K_λ n'est pas négligeable). Si on note χ_λ (resp. χ_n) la fonction caractéristique de K_λ (resp. de K_n) on vérifie les faits suivants :

- Les fonctions χ_n forment une suite décroissante de fonctions en escalier.
- La suite (χ_n) est de Cauchy au sens de L^1 . En effet, pour $p \leq q$, la

différence $\chi_p - \chi_q$ est une fonction en escalier d'intégrale $\sum_{n=p}^{q-1} 2^n \lambda_n$ qui tend vers 0 quand p tend vers l'infini.

- La fonction χ_λ est la limite (simple) de χ_n et le candidat à être sa limite L^1 , mais n'est pas Riemann-intégrable.

1.5.11 Remarque. La caractérisation de l'intégrabilité permet de retrouver facilement le fait que le produit de deux fonctions intégrables est intégrable, que la valeur absolue ou le module l'est. Elle donne aussi le résultat suivant :

1.5.12 Proposition. Soient f, g deux fonctions bornées, $f : [a, b] \rightarrow [c, d]$ et $g : [c, d] \rightarrow \mathbf{R}$. On suppose f Riemann-intégrable et g continue. Alors $g \circ f$ est Riemann-intégrable.

Démonstration. En effet, en un point où f est continue, $g \circ f$ l'est aussi. Il en résulte que $g \circ f$ est continue en dehors d'un ensemble négligeable, donc intégrable.

1.5.13 Remarque. L'assertion analogue avec seulement f, g intégrables est fausse. Voici un contre-exemple. Soit f la fonction définie sur $[0, 1]$ par $f(x) = 0$ si x est irrationnel et $f(p/q) = 1/q$ si p/q est rationnel (écrit sous forme irréductible). Cette fonction est continue sur $[0, 1]$, sauf en les rationnels. Elle est donc Riemann-intégrable. Si g est la fonction égale à 1 partout, sauf en 0 où elle vaut 0, g est intégrable aussi, mais $g \circ f$ n'est autre que la fonction caractéristique de $\mathbf{Q} \cap [0, 1]$ dont on a vu qu'elle n'est pas Riemann-intégrable.

1.6 Annexe : Une remarque sur la définition des sommes de Darboux

1.6.1 Une variante des sommes de Darboux

Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction bornée quelconque. Si $S = (s_i)$ est une subdivision de $[a, b]$, on a posé, pour $i = 0, 1, \dots, n-1$:

$$m_i = \inf_{]s_i, s_{i+1}[} f(x) \quad \text{et} \quad M_i = \sup_{]s_i, s_{i+1}[} f(x).$$

Mais on aurait tout aussi bien pu considérer les extrema sur les segments fermés :

$$m'_i = \inf_{[s_i, s_{i+1}] } f(x) \quad \text{et} \quad M'_i = \sup_{[s_i, s_{i+1}] } f(x).$$

Cela permet de définir deux types de sommes de Darboux :

$$\begin{aligned} \sigma(S, f) &= \sum_{i=0}^{n-1} (s_{i+1} - s_i) m_i \quad \text{et} \quad \Sigma(S, f) = \sum_{i=0}^{n-1} (s_{i+1} - s_i) M_i, \\ \sigma'(S, f) &= \sum_{i=0}^{n-1} (s_{i+1} - s_i) m'_i \quad \text{et} \quad \Sigma'(S, f) = \sum_{i=0}^{n-1} (s_{i+1} - s_i) M'_i. \end{aligned}$$

Bien entendu, on a $m_i \geq m'_i$ et $M_i \leq M'_i$. Si f est continue, on a égalité. En effet, on a $f(s_i) = \lim_{x \rightarrow s_i} f(x) \geq m_i$ et de même en s_{i+1} . Mais, en général, les bornes sont distinctes et donc les sommes de Darboux avec et sans prime ne sont pas égales. Heureusement on a le théorème suivant :

1.6.1 Théorème. *On a les égalités :*

$$\sup_S \sigma(S, f) = \sup_S \sigma'(S, f) \quad \text{et} \quad \inf_S \Sigma(S, f) = \inf_S \Sigma'(S, f).$$

Démonstration. Traitons le cas des sommes Σ , l'autre est analogue. Notons déjà qu'on peut supposer la fonction f positive. En effet, comme f est bornée, la fonction $f + k$, où k est une constante positive, est positive pour k assez grand. Sur chaque intervalle, les maxima de $f + k$ sont obtenus en ajoutant k à ceux de f . On a donc $\Sigma(S, f + k) = \Sigma(S, f) + k(b - a)$ et $\Sigma'(S, f + k) = \Sigma'(S, f) + k(b - a)$ et la même relation vaut en prenant les bornes inférieures, de sorte qu'il suffit de montrer l'égalité pour $f + k$.

Soit S une subdivision. On a $M_i \leq M'_i$ et donc $\Sigma(S, f) \leq \Sigma'(S, f)$. On en déduit $I = \inf_S \Sigma(S, f) \leq \inf_S \Sigma'(S, f) = I'$.

Montrons l'inégalité en sens inverse. Pour cela, il suffit de montrer que, pour tout $\epsilon > 0$, il existe une subdivision T telle que l'on ait $\Sigma'(T, f) \leq I + 2\epsilon$. Comme I est la borne inférieure des sommes Σ , il existe une subdivision $S = s_0, s_1, \dots, s_n$ telle que l'on ait $I \leq \Sigma(S, f) \leq I + \epsilon$. On considère une subdivision $T = t_0, t_1, \dots, t_{2n+1}$, avec $t_0 = s_0 = a$, $t_1 = t_0 + \eta$, $t_{2n+1} = s_n = b$, $t_{2n} = b - \eta$ et, pour chaque $i = 1, \dots, n - 1$, $t_{2i} = s_i - \eta$ et $t_{2i+1} = s_i + \eta$. On choisit $\eta > 0$, suffisamment petit pour que la suite des t_i soit strictement croissante. On considère la somme $\sum_{i=0}^{2n} (t_{i+1} - t_i) M'_i$ où M'_i est la borne supérieure de f sur l'intervalle $[t_i, t_{i+1}]$ et on scinde cette somme selon la parité des indices. L'intervalle $[t_{2i+1}, t_{2i+2}]$, pour i variant de 0 à $n - 1$ est contenu dans $]s_i, s_{i+1}[$ par construction. On a donc $M'_{2i+1} \leq M_i$ et on en déduit l'inégalité :

$$\sum_{i=0}^{n-1} (t_{2i+2} - t_{2i+1}) M'_{2i+1} \leq \sum_{i=0}^{n-1} (s_{i+1} - s_i) M_i = \Sigma(S, f) \leq I + \epsilon.$$

(Attention, c'est ici qu'il est important d'avoir supposé $f \geq 0$. En effet, on a bien $t_{2i+2} - t_{2i+1} \leq s_{i+1} - s_i$ et $M'_{2i+1} \leq M_i$, mais on ne peut conclure que les produits sont dans le même ordre que si ces quantités sont positives.)

Pour majorer les autres termes, on introduit la borne supérieure M de f sur $[a, b]$. On a alors $\sum_{i=0}^n (t_{2i+1} - t_{2i}) M'_{2i} \leq 2\eta(n + 1)M$. On choisit η pour que cette quantité soit $\leq \epsilon$ et on a alors $\Sigma'(T, f) \leq I + 2\epsilon$ comme annoncé.

1.6.2 Sommes de Riemann

1.6.2 Définition. Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction bornée et soit $S = s_0, s_1, \dots, s_n$ une subdivision de $[a, b]$. On appelle **somme de Riemann** associée à f et S toute somme de la forme $\sum_{i=0}^{n-1} (s_{i+1} - s_i) f(\xi_i)$ où ξ_i est un point quelconque de l'intervalle $[s_i, s_{i+1}]$.

1.6.3 Remarque. Si s est une somme de Riemann associée à la fonction f et à la subdivision S , on a l'encadrement : $\sigma'(S, f) \leq s \leq \Sigma'(S, f)$.

1.6.4 Théorème. Avec les notations précédentes on suppose la fonction f intégrable sur $[a, b]$. Alors, les sommes de Riemann convergent vers $\int_a^b f(t)dt$ lorsque le pas de la subdivision tend vers 0. Précisément, cela signifie que pour tout $\epsilon > 0$ il existe $\eta > 0$ tel que, pour toute subdivision de pas $\leq \eta$ et pour tout choix des points $\xi_i \in [s_i, s_{i+1}]$ on a :

$$\left| \sum_{i=0}^{n-1} (s_{i+1} - s_i) f(\xi_i) - \int_a^b f(t)dt \right| < \epsilon.$$

Démonstration. On a montré (voir 1.1.12) que, si $|S|$ est assez petit, on a $\Sigma(S, f) - \sigma(S, f) \leq \epsilon$. Par ailleurs, on sait qu'on a $\sigma(S, f) \leq I = \int_a^b f(t)dt \leq \Sigma(S, f)$. Si $s = \sum_{i=0}^{n-1} (s_{i+1} - s_i) f(\xi_i)$ est une somme de Riemann associée à f et S , on a aussi $\sigma'(S, f) \leq s \leq \Sigma'(S, f)$ en vertu de la remarque ci-dessus. On a donc bien $|I - s| \leq \epsilon$.

1.6.5 Corollaire. Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction intégrable. Soit n un entier positif, S_n la subdivision régulière d'ordre n et soit $\xi_i(n)$ un point quelconque de l'intervalle $\left[a + \frac{i(b-a)}{n}, a + \frac{(i+1)(b-a)}{n} \right]$. Alors, la suite

$$I_n = \frac{b-a}{n} \sum_{i=0}^{n-1} f(\xi_i(n)) \text{ converge vers } \int_a^b f(t)dt.$$

Chapitre 2

L'intégrale de Lebesgue

2.1 Intégrales et suites, premier épisode : la convergence uniforme

Parmi les défauts de l'intégrale de Riemann, qui vont motiver l'introduction de l'intégrale de Lebesgue, on peut citer le fait que Riemann requière que l'intervalle et les fonctions soient bornés ou le fait que certaines fonctions, somme toute assez simples, comme la fonction caractéristique des rationnels de $[0, 1]$, ne soient pas intégrables. Le point essentiel sur lequel nous allons porter notre attention n'est pas l'un de ceux-là, c'est le lien entre intégrales et suites. L'objectif est d'obtenir des théorèmes du genre de celui-ci :

2.1.1 Théorème. *Soit f_n une suite de fonctions intégrables (en un sens à préciser), qui converge (en un sens à préciser) vers une fonction f . Alors f est intégrable (en un sens à préciser) et $\int f$ est la limite des $\int f_n$.*

Les applications en analyse de ce type de théorèmes sont innombrables.

2.1.1 Riemann et la norme de la convergence uniforme : un exemple

Dans le cas des fonctions Riemann-intégrables, nous avons vu un seul théorème de passage à la limite du type de 2.1.1 : le théorème 1.4.2 qui utilise la norme de la convergence uniforme. Rappelons que, si (f_n) converge uniformément vers f , la convergence des intégrales vient de l'inégalité :

$$(*) \quad \left| \int_a^b f_n - \int_a^b f \right| \leq \int_a^b |f - f_n| \leq (b - a) \|f - f_n\|_\infty.$$

Il est clair que cette méthode n'est pas du tout un outil raisonnable pour étudier des suites d'intégrales. En effet, dire qu'une suite (f_n) est petite au sens de la convergence uniforme c'est dire que les valeurs des f_n sont partout petites et qu'on peut donc coincer l'aire sous la courbe dans un rectangle horizontal aplati. Or, son intégrale peut être petite aussi si la fonction est grande (pas trop) sur un petit intervalle, donc coincée dans un rectangle vertical étroit. Un exemple très simple illustre ces faits, il s'agit des intégrales de Wallis : $I_n = \int_0^{\pi/2} \sin^n x dx$. Comme la suite (décroissante) $\sin^n x$ tend vers 0 (car on a $\sin x < 1$), sauf au point $x = \pi/2$ où elle est constante et égale à 1, on conjecture que la suite I_n tend vers 0.

Cependant, comme la limite est discontinue il n'y a pas convergence uniforme et on ne peut appliquer directement 1.4.2. On peut toutefois montrer le résultat en utilisant uniquement l'outil convergence uniforme. Si on examine le graphe, on voit que la courbe de $\sin^n x$ tend vers le coude formé par le segment $[0, \pi/2]$ de l'axe horizontal et le segment $[0, 1]$ de la verticale $x = \pi/2$. On va donc majorer l'aire par celle de deux rectangles, l'un allongé, l'autre debout. Précisément, on se donne $\epsilon > 0$ et on écrit

$$I_n = \int_0^{\pi/2} \sin^n x dx = \int_0^{\pi/2-\epsilon} \sin^n x dx + \int_{\pi/2-\epsilon}^{\pi/2} \sin^n x dx.$$

Il est clair que le deuxième morceau est $\leq \epsilon$. Quant au premier, sur l'intervalle $[0, \pi/2 - \epsilon]$, la convergence vers 0 de la suite $\sin^n x$ est uniforme (ces fonctions sont majorées par $\sin^n(\pi/2 - \epsilon)$). L'intégrale tend donc vers 0 et elle est $\leq \epsilon$ pour n assez grand. Il en résulte que I_n est $\leq 2\epsilon$.

Ce raisonnement n'est pas difficile, mais il demande déjà une bonne manipulation des ϵ . Quand on disposera des théorèmes de Lebesgue on aura deux arguments simples pour conclure. On dira, au choix :

- La suite (f_n) est majorée par la fonction intégrable 1. Elle converge presque partout vers 0 (ici, presque partout signifie : sauf en $\pi/2$), donc son intégrale converge vers l'intégrale de 0, soit 0. (Convergence dominée)
- La suite (f_n) est décroissante. Elle converge presque partout vers 0, donc son intégrale converge vers 0. (Convergence monotone)

2.2 Intégrales et suites, deuxième épisode : la convergence en moyenne

Il y a évidemment une (semi)-norme mieux adaptée pour mesurer la taille des intégrales : la norme de la convergence en moyenne définie par

$\|f\|_1 = \int_a^b |f(t)| dt = \int |f|$. La norme L^1 possède encore la propriété qui nous intéresse : si une suite de fonctions (f_n) converge vers f au sens de L^1 , les intégrales $I_n = \int f_n$ convergent vers $\int f$. Cela résulte de l'inégalité analogue à l'inégalité (*) du cas uniforme, mais où l'on s'arrête avant la fin :

$$\left| \int_a^b f - \int_a^b f_n \right| \leq \int_a^b |f - f_n| = \|f - f_n\|_1.$$

2.2.1 Remarque. Attention, il faut se débarrasser tout de suite d'une idée fautive : si les intégrales $I_n = \int f_n$ convergent vers $\int f$, ce n'est pas pour autant que la suite (f_n) converge vers f au sens de L^1 . En effet, dire que les intégrales convergent c'est dire que $|\int f - \int f_n| = |\int(f - f_n)|$ tend vers 0, tandis que dire que (f_n) tend vers f au sens de L^1 c'est dire que $\int |f - f_n|$ tend vers 0 et on a seulement l'inégalité : $|\int(f - f_n)| \leq \int |f - f_n|$.

2.2.2 Exemple. On considère les fonctions f_n définies sur $[0, 1]$, qui sont des pics triangulaires à support dans $[0, 2/n]$ avec pointe de hauteur n en $1/n$ et la fonction $f = 1$. Comme les intégrales des f_n valent toutes 1, $\int f_n$ converge vers $\int f = 1$. Cependant, (f_n) ne converge pas du tout vers f , ni au sens de L^1 , ni simplement. En fait, la suite (f_n) converge simplement¹ vers 0, mais pas au sens L^1 puisque les intégrales des f_n sont toutes égales à 1.

Il y a tout de même un cas où la convergence des intégrales implique la convergence au sens de la norme L^1 : lorsque les fonctions $f - f_n$ sont de signe constant, par exemple si la suite est monotone, cf. 2.5.7.

2.2.1 Les suites de Cauchy

La convergence pour la norme L^1 permet d'approcher le théorème idéal, avec une hypothèse plus faible que la convergence uniforme, mais il reste à savoir comment montrer la convergence de (f_n) vers f au sens de L^1 . En tous cas, l'exemple ci-dessus montre que la convergence simple est insuffisante.

Pour aborder cette question nous allons utiliser les suites de Cauchy. Cela est justifié par les deux remarques suivantes :

2.2.3 Remarques.

1) Si une suite de fonctions (f_n) converge vers f au sens de L^1 c'est une suite de Cauchy, en raison de l'inégalité :

$$\|f_p - f_q\|_1 \leq \|f_p - f\|_1 + \|f - f_q\|_1.$$

¹Le lecteur méditera cet exemple à la lumière de 2.5.12.

2) Si (f_n) est de Cauchy au sens de L^1 , la suite des intégrales $I_n = \int f_n$ est une suite de Cauchy de \mathbf{R} . En effet, on a l'inégalité $|\int f_p - \int f_q| \leq \int |f_p - f_q|$. Comme \mathbf{R} est complet, il en résulte que I_n est convergente.

3) **Attention**, comme pour la convergence, le fait que la suite des intégrales soit de Cauchy n'implique pas que la suite de fonctions est de Cauchy au sens de L^1 , cf. 2.2.1 et 2.2.2. Cela sera cependant le cas si la suite (f_n) est monotone, cf. 2.5.7.

La réciproque de la remarque 1) est une question fondamentale en analyse : une suite de Cauchy d'éléments d'un espace E est-elle automatiquement convergente dans E ? Ou encore, l'espace E est-il complet? Ce n'est pas toujours vrai car il peut "manquer" des fonctions. Ainsi, si on travaille avec les fonctions Riemann-intégrables, on peut trouver une suite de fonctions qui est de Cauchy au sens de la norme L^1 et qui converge simplement vers une fonction f , mais telle que f ne soit pas Riemann-intégrable. Nous donnons ci-dessous deux exemples de cette situation.

2.2.4 Exemple. On considère les fonctions $f_n : [0, 1] \rightarrow \mathbf{R}$ définies par $f_n(x) = 0$ pour $x < 1/n$ et $f_n(x) = 1/\sqrt{x}$ pour $x \geq 1/n$. On vérifie que la suite (f_n) est de Cauchy au sens de L^1 . Elle converge vers $1/\sqrt{x}$ qui n'est pas une fonction Riemann-intégrable.

Au vu de cet exemple, on pourrait se demander s'il ne suffit pas de rajouter les fonctions, comme ici $1/\sqrt{x}$, qui admettent une intégrale de Riemann impropre. L'exemple suivant montre qu'il n'en est rien, puisqu'il concerne une fonction bornée sur un intervalle borné :

2.2.5 Exemple. On a vu au chapitre 1 que la fonction caractéristique d'un ensemble de Cantor K_λ non négligeable n'est pas Riemann-intégrable. On a vu aussi que, si on note χ_λ (resp. χ_n) la fonction caractéristique de K_λ (resp. de K_n) les fonctions χ_n forment une suite décroissante de fonctions en escalier et que la suite (χ_n) est de Cauchy au sens de L^1 . La fonction χ_λ , qui est la limite (simple) de χ_n et le candidat à être sa limite L^1 , n'est pas Riemann-intégrable.

L'espace des fonctions Riemann-intégrables n'est donc pas complet pour la norme L^1 . La théorie de Lebesgue va consister essentiellement à le compléter.

2.2.2 Convergence L^1 et convergence simple

On sait que la convergence uniforme implique la convergence simple, mais que la réciproque est inexacte. Dans le cas de la convergence L^1 , ces notions

sont encore plus indépendantes, la même suite pouvant converger vers des fonctions différentes selon le sens du mot convergence.

2.2.6 Exemple. Dans l'exemple de Wallis, la suite (f_n) tend vers 0 au sens de L^1 , mais ne tend pas simplement vers 0 (elle tend vers 1 au point $\pi/2$).

2.2.7 Exemple. On numérote les rationnels de $[0, 1]$: $r_0, r_1, \dots, r_n, \dots$ et on prend pour f_n la fonction caractéristique de l'ensemble $\{r_0, r_1, \dots, r_n\}$. Comme les intégrales $\int_0^1 |f_n|$ sont toutes nulles, la suite (f_n) converge vers la fonction nulle au sens de L^1 . En revanche, au sens de la convergence simple, il est clair que la suite (f_n) converge vers la fonction caractéristique χ des rationnels de $[0, 1]$. Elle ne converge pas vers χ au sens de L^1 car cette fonction n'est pas Riemann-intégrable.

Dans les deux exemples précédents, les limites aux sens des convergences simple et L^1 ne diffèrent que sur des ensembles assez petits : un point ou un ensemble dénombrable. Cela nous conduit à introduire la notion de presque partout, qui va permettre de négliger ... les ensembles négligeables. Précisément :

2.2.8 Définition. On dit qu'une propriété $P(x)$ du nombre réel x est vraie **presque partout** sur \mathbf{R} (ou sur une partie de \mathbf{R}) si l'ensemble des points où P est fautive est négligeable.

2.2.9 Exemple. Par exemple une fonction est continue (ou nulle, ou dérivable, etc.) presque partout si elle l'est sauf sur un ensemble négligeable. Ainsi, la fonction caractéristique de \mathbf{Q} est presque partout nulle. De même, la fonction caractéristique de l'ensemble triadique de Cantor ordinaire est presque partout continue. De la même façon, la suite $(\sin^n x)$ converge presque partout vers la fonction nulle. On utilise des variantes de ce terme : presque toujours, presque sûrement, etc.

On peut maintenant se demander si la convergence au sens L^1 entraîne la convergence simple presque partout. L'exemple suivant montre qu'il n'en est rien !

2.2.10 Exemple. L'exemple horrible On considère la suite de fonctions (f_n) où f_0 est la fonction caractéristique de $[0, 1/2]$, f_1 celle de $[1/2, 1]$, f_2 celle de $[0, 1/3]$, f_3 de $[1/3, 2/3]$, f_4 de $[2/3, 1]$, etc. Les fonctions f_n sont Riemann-intégrables (ce sont des fonctions en escalier), positives, définies sur $[0, 1]$. La suite (f_n) tend vers 0 au sens de L^1 (de sorte qu'elle est de Cauchy). En effet, l'intégrale de $|f_n|$ tend vers 0 quand n tend vers l'infini

car les f_n sont majorées par la constante 1 (la longueur !) mais la largeur de leur support tend vers 0. Pourtant, la suite (f_n) ne tend pas simplement vers 0, même presque partout (en fait, elle ne tend vers 0 en aucun point !). En effet, en chaque point $x \in [0, 1]$, il y a une infinité de fonctions f_n non nulles (c'est le cas chaque fois que le support de f_n repasse par x).

2.2.3 Notre objectif

Notre objectif est le suivant : considérer les suites de Cauchy au sens de L^1 de fonctions Riemann-intégrables (f_n) (ou même des suites de fonctions en escalier, car on peut approcher toute fonction Riemann-intégrable par des fonctions en escalier au sens de L^1) et, comme elles ne convergent pas en général vers des fonctions Riemann-intégrables, cf. 2.2.7 ou 2.2.5, définir de nouvelles fonctions intégrables f qui seront les limites de ces suites.

2.2.11 Remarque. Dans la théorie de Riemann, les intégrales sont prises sur des intervalles compacts $[a, b]$, mais, **attention**, si on a une suite (f_n) (même de Cauchy) de fonctions, les f_n ne sont pas nécessairement toutes définies sur un même segment. Par exemple, la suite des fonctions $f_n = \frac{1}{n^2} \chi_{[-n, n]}$ est de Cauchy au sens de L^1 (elle converge vers la fonction nulle), mais les supports de ses termes augmentent avec n . Il va donc être commode de considérer des intégrales sur \mathbf{R} tout entier. Pour cela il y a un moyen simple : si f est définie sur $I = [a, b]$ et si elle est Riemann-intégrable, on la prolonge en une fonction définie sur \mathbf{R} en posant $f(x) = 0$ pour $x \notin I$. On pose alors $\int_{\mathbf{R}} f(x) dx = \int_a^b f(x) dx$. On note même cette intégrale $\int_{\mathbf{R}} f$. Inversement, si f est définie sur \mathbf{R} , on peut considérer $\int_a^b f(x) dx$, c'est, par définition, l'intégrale de la restriction à $[a, b]$ de f . Il n'y a donc pas d'inconvénient à considérer que les fonctions sont définies sur \mathbf{R} tout entier. Bien entendu, pour le moment, les seules fonctions réputées intégrables sont les fonctions à support borné qui sont Riemann-intégrables. C'est le cas des fonctions en escalier dont l'intégrale est définie dans la théorie de Riemann et le reste dans celle de Lebesgue.

2.3 Le lemme d'Egorov

Nous pouvons maintenant entrer dans le vif du sujet en commençant par un lemme technique, pas facile, mais fondamental.

2.3.1 Le lemme d'Egorov

Le problème est le suivant. On a dit qu'on allait s'intéresser aux suites de Cauchy de fonctions f_n Riemann-intégrables (voire en escalier), au sens de L^1 . L'exemple horrible 2.2.10 nous a montré qu'une telle suite peut ne pas converger simplement vers une fonction f , même presque partout. Face à cette difficulté, il va falloir se contenter d'un pis-aller. Dans l'exemple horrible, la suite ne converge pas simplement vers 0, mais elle contient des sous-suites qui convergent simplement, et toutes vers la fonction nulle, par exemple la sous-suite des fonctions caractéristiques des intervalles $[0, 1/n]$. Ce phénomène est général :

2.3.1 Théorème. (Lemme d'Egorov) Soit $f_n : \mathbf{R} \rightarrow \mathbf{R}$ une suite de fonctions en escalier. On suppose que (f_n) est de Cauchy au sens de L^1 .

1) Il existe une sous-suite f_{n_k} qui converge simplement presque partout vers une fonction f . De plus, pour tout $\epsilon > 0$, la convergence est uniforme en dehors d'un ensemble A_ϵ de "mesure" plus petite que ϵ , précisément, A_ϵ est contenu dans une réunion d'intervalles $\bigcup_{n \in \mathbf{N}} J_n$ telle que l'on ait $\sum_{n \in \mathbf{N}} l(J_n) < \epsilon$.

2) On suppose que $\|f_n\|_1 = \int_a^b |f_n(t)| dt$ converge vers 0. Alors, il existe une sous-suite (f_{n_k}) de (f_n) qui converge vers 0 presque partout.

3) Réciproquement, s'il existe une sous-suite (f_{n_k}) de (f_n) qui converge vers 0 presque partout, la suite $\|f_n\|_1 = \int_a^b |f_n(t)| dt$ converge vers 0.

Preuve du point 1

Il s'agit de montrer que la suite $(f_n(x))$ (ou une de ses sous-suites) converge pour presque tout x . La première idée de cette preuve c'est que noter qu'une suite $(f_n(x))$ converge si et seulement si la série $(f_{n+1}(x) - f_n(x))$ converge. En effet, cela résulte de la formule :

$$S_N(x) = \sum_{n=0}^{N-1} (f_{n+1}(x) - f_n(x)) = f_N(x) - f_0(x).$$

On sait qu'il suffit pour cela que cette série converge absolument, ou mieux encore qu'on ait, pour n assez grand :

$$(*_n) \quad |f_{n+1}(x) - f_n(x)| \leq \frac{1}{2^n}.$$

On va donc s'intéresser à la suite $f_{n+1} - f_n$. La condition de Cauchy nous dit que la suite des normes $L^1 : \|f_{n+1} - f_n\|_1$ tend vers 0, mais on préférerait que la série converge (ce n'est pas pareil!). L'exemple horrible montre que ça n'est pas vrai en général (la suite en question est en $1/n$). C'est là qu'il faut extraire une sous-suite :

2.3.2 Lemme. *Si (f_n) est une suite de Cauchy au sens de L^1 il existe une sous-suite (f_{n_k}) qui vérifie $\|f_{n_{k+1}} - f_{n_k}\|_1 \leq \frac{1}{2^{2k}}$.*

Démonstration. On construit la sous-suite par récurrence sur k . On sait qu'il existe N_0 tel que, pour $p, q \geq N_0$ on ait $\|f_p - f_q\|_1 \leq 1$. On choisit alors $n_0 \geq N_0$. Puis on a un N_1 tel que, pour $p, q \geq N_1$, on ait $\|f_p - f_q\|_1 \leq 1/4$. On choisit alors $n_1 \geq N_1$ et $n_1 > n_0$. Au rang k on a un N_k avec, pour $p, q \geq N_k$, $\|f_p - f_q\|_1 \leq 1/2^{2k}$ et on prend $n_k \geq N_k$ et $n_k > n_{k-1}$.

Pour simplifier les notations, maintenant qu'on a construit cette sous-suite, on peut faire comme si c'était la suite initiale, autrement dit supposer que (f_n) vérifie $\|f_{n+1} - f_n\|_1 \leq 1/2^{2n}$. On va montrer que la suite $(f_n(x))$ converge alors presque partout, avec l'idée d'utiliser la série $(f_{n+1}(x) - f_n(x))$. Cette série converge, en tous cas, si elle vérifie $(*_n)$ pour n assez grand. Autrement dit, x est un "bon point" (un point où la suite $f_n(x)$ converge) s'il existe p tel que, pour tout $n \geq p$, on ait $(*_n)$ et donc x est un "mauvais point" si pour tout p , il existe $n \geq p$, tel que $(*_n)$ soit fausse.

Pour étudier cette condition on considère l'ensemble : $A_n = \{x \in \mathbf{R} \mid |f_{n+1}(x) - f_n(x)| > 1/2^n\}$. L'ensemble B des mauvais points est alors $B = \bigcap_{p \in \mathbf{N}^*} B_p$ avec $B_p = \bigcup_{n \geq p} A_n$, et il s'agit de montrer que B est négligeable.

Comme $f_{n+1} - f_n$ est en escalier, A_n est une réunion finie d'intervalles et la condition de norme L^1 impose que la somme $l(A_n)$ de leurs longueurs vérifie $l(A_n) \leq 1/2^n$ (c'est l'aire du rectangle!). Il en résulte que B_p est une réunion

dénombrable d'intervalles dont la somme des longueurs est $\leq \sum_{n=p}^{+\infty} \frac{1}{2^n} = \frac{1}{2^{p-1}}$

et on peut rendre cette longueur $< \epsilon$ pour p assez grand. Comme B est contenu dans B_p , il est bien négligeable, par définition.

On note que la convergence est uniforme (et même normale) sur $\mathbf{R} - B_p$, ce qui montre l'assertion supplémentaire.

Preuve des points 2 et 3

Pour les deux assertions, on peut supposer les f_n positives (il suffit de remplacer f_n par $|f_n|$). On pose $I_n = \int_{\mathbf{R}} f_n$.

Montrons le point 2). Le raisonnement est analogue à celui utilisé pour le point 1). Comme la suite I_n tend vers 0 on peut extraire une sous-suite (f_{n_k}) telle que l'on ait $I_{n_k} \leq 1/2^{2k}$. Pour simplifier les notations on suppose désormais que l'on a $I_n \leq 1/2^{2n}$ et il s'agit de montrer que $(f_n(x))$ tend vers 0 presque partout.

On pose : $A_n = \{x \in \mathbf{R} \mid f_n(x) > 1/2^n\}$. Comme f_n est en escalier, A_n est une réunion finie d'intervalles et on a $I_n \geq \frac{l(A_n)}{2^n}$, d'où $l(A_n) \leq \frac{1}{2^n}$. On considère ensuite $B_p = \bigcup_{n \geq p} A_n$ et $B = \bigcap_{p \in \mathbf{N}} B_p$.

Si x n'est pas dans B , il existe p tel que $x \notin B_p$, ce qui signifie que, pour tout $n \geq p$, $x \notin A_n$, ou encore qu'on a, pour $n \geq p$, $f_n(x) < \frac{1}{2^n}$, ce qui montre que $f_n(x)$ converge vers 0. Il reste à montrer que B est négligeable. En effet, il est contenu dans B_p , lequel est contenu dans une réunion d'intervalles dont la somme des longueurs est $s_p = \sum_{n \geq p} \frac{1}{2^n} = \frac{1}{2^{p-1}}$, d'où le résultat.

Montrons maintenant le point 3).

L'inégalité $|\|f_p\|_1 - \|f_q\|_1| \leq \|f_p - f_q\|_1$ montre que la suite $\|f_n\|_1$ est une suite de Cauchy de \mathbf{R} , donc elle converge. Pour montrer que c'est vers 0, il suffit de le faire pour une sous-suite et on choisit la suite (f_{n_k}) donnée par l'énoncé qui converge simplement vers 0 presque partout. On se ramène ainsi au cas d'une suite (f_n) qui converge presque partout vers 0. On se donne alors $\epsilon > 0$ et on va montrer que $\|f_q\|_1$ est $\leq 4\epsilon$ pour q assez grand.

On applique la condition de Cauchy. Il existe N tel que, pour $p, q \geq N$ on ait $\|f_p - f_q\|_1 \leq \epsilon$. On fixe un $p \geq N$. Soit M le maximum de f_p et soit I la réunion d'intervalles où f_p est non nulle. En vertu de la première phase, on peut trouver une suite d'intervalles (J_n) , avec $\sum_n l(J_n) \leq \epsilon/M$, telle que la suite (f_n) converge uniformément vers 0 sur le complémentaire de $J = \bigcup_{n \in \mathbf{N}} J_n$. Cela montre que, si q est assez grand, on a $f_q(x) \leq \epsilon/l(I)$ pour $x \notin J$. Soit K la réunion d'intervalles sur laquelle f_q est $> \epsilon/l(I)$. On a $K \subset J$, donc $l(K) \leq \sum_n l(J_n) \leq \epsilon/M$ par le lemme de Borel-Lebesgue (cf. 1.5.6).

On met tout cela ensemble. On décompose l'intégrale

$$\|f_q\|_1 = \int_{\mathbf{R}} f_q = \int_{\mathbf{R}-I} f_q + \int_{I-K} f_q + \int_K f_q.$$

Sur $\mathbf{R}-I$, comme f_p est nulle, on a $\int_{\mathbf{R}-I} f_q = \int_{\mathbf{R}-I} |f_q - f_p| \leq \|f_q - f_p\|_1 \leq \epsilon$. Sur $I-K$, on a $f_q \leq \epsilon/l(I)$ et donc $\int_{I-K} f_q \leq l(I-K)\epsilon/l(I) \leq \epsilon$. Enfin, sur K , on a $\int_K f_q \leq \int_K f_p + \int_K |f_q - f_p|$. Mais on a $\int_K f_p \leq l(K)M \leq \epsilon$ et comme l'autre terme est $\leq \|f_q - f_p\|_1 \leq \epsilon$, on a bien majoré $\int_{\mathbf{R}} f_q$ par 4ϵ .

2.3.3 Corollaire. *Sous les hypothèses du lemme d'Egorov (point 1), si deux sous-suites (f_{n_k}) et $(f_{n'_k})$ de (f_n) convergent l'une vers f et l'autre vers f' presque partout, alors les fonctions f et f' sont égales presque partout.*

Démonstration. La suite $g_k = f_{n_k} - f_{n'_k}$ est de Cauchy au sens de L^1 , elle tend vers $f - f'$ presque partout et $\int |g_k|$ tend vers 0 (car f_n est de Cauchy). On en déduit qu'une sous-suite de g_k tend vers 0 presque partout en vertu de 2.3.1.2, ce qui impose $f - f' = 0$ presque partout.

2.4 Définition et propriétés de l'intégrale

Voilà, tout est en place pour définir l'intégrale de Lebesgue. L'idée est de définir de nouvelles fonctions intégrables comme limites de suites de Cauchy (au sens de L^1) de fonctions en escalier. Grâce à Egorov, on peut supposer que ces suites convergent presque partout, quitte à extraire au besoin une sous-suite.

2.4.1 Fonctions intégrables

2.4.1 Définition. *On dit qu'une fonction $f : \mathbf{R} \rightarrow \mathbf{R}$ est Lebesgue-intégrable s'il existe une suite (f_n) de fonctions en escalier, qui est de Cauchy au sens de L^1 , et qui converge simplement vers f presque partout. L'ensemble des fonctions Lebesgue-intégrables sur \mathbf{R} est noté $\mathcal{L}(\mathbf{R})$.*

2.4.2 Proposition-Définition. *Avec les notations précédentes, les intégrales des f_n admettent une limite I qui ne dépend que de f et pas du choix de la suite (f_n) . On l'appelle **intégrale** de f (au sens de Lebesgue) et on note $I = \int_{\mathbf{R}} f$ ou $I = \int_{\mathbf{R}} f(t)dt$.*

Démonstration. Que les intégrales convergent résulte de 2.2.3. Il faut montrer que la limite des intégrales ne dépend pas du choix de la suite. Si (f_n) et (g_n) sont deux suites de Cauchy qui convergent vers f presque partout, la suite $(f_n - g_n)$ est encore de Cauchy et converge simplement presque partout vers 0. En vertu d'Egorov (2.3.1.3), $\|f_n - g_n\|_1$ tend vers 0 donc, *a fortiori*, $\int (f_n - g_n) = \int f_n - \int g_n$ tend vers 0.

2.4.3 Remarques.

1) Il résulte de la définition que si f est dans \mathcal{L} et si g est égale à f presque partout, g est dans \mathcal{L} et on a $\int_{\mathbf{R}} f = \int_{\mathbf{R}} g$. En particulier, si f est nulle presque partout, elle est intégrable et d'intégrale nulle. Ainsi, la fonction caractéristique des rationnels $\chi_{\mathbf{Q}}$ (qui n'est pas Riemann-intégrable) est

Lebesgue-intégrable et d'intégrale nulle. Plus généralement, si A est négligeable, la fonction χ_A est Lebesgue-intégrable et d'intégrale nulle.

2) Si f est dans \mathcal{L} il en est de même² de $|f|$ (car elle est limite de la suite $|f_n|$).

La remarque ci-dessus permet de définir la (semi)³-norme L^1 :

2.4.4 Définition. Soit $f \in \mathcal{L}$. On pose $\|f\|_1 = \int_{\mathbf{R}} |f(t)| dt$.

Il est clair que cette quantité vérifie bien l'inégalité triangulaire.

2.4.5 Proposition. Avec les notations de 2.4.1 et 2.4.2, la fonction f est limite de la suite (f_n) au sens de la norme L^1 , autrement dit la suite de réels $\|f - f_n\|_1 = \int_{\mathbf{R}} |f - f_n|$ tend vers 0 quand n tend vers l'infini. En particulier, si f est Lebesgue-intégrable, pour tout $\epsilon > 0$, il existe une fonction en escalier g qui vérifie $\|f - g\|_1 < \epsilon$.

Démonstration. C'est un peu plus subtil qu'il n'y paraît ! Pour p fixé on considère la suite (g_q) de fonctions en escalier définie par $g_q = |f_p - f_q|$. C'est une suite de Cauchy au sens de L^1 . Cela résulte de l'inégalité $\|g_q - g_r\|_1 \leq \|f_q - f_r\|_1$. Par ailleurs, g_q tend vers $g = |f_p - f|$ presque partout. Par définition de l'intégrale et de la norme on en déduit que $\|f_p - f\|_1 = \|g\|_1$ est la limite de $\|f_p - f_q\|_1 = \|g_q\|_1$ quand q tend vers l'infini. (Autrement dit "on peut passer à la limite dans les normes L^1 ".) Si on fixe $\epsilon > 0$, la condition de Cauchy pour la suite (f_n) assure qu'on a $\|f_p - f_q\|_1 \leq \epsilon$ pour $p, q \geq N$. En passant à la limite sur q on en déduit $\|f_p - f\|_1 \leq \epsilon$ pour $p \geq N$, ce qui est la conclusion cherchée.

2.4.6 Notation. Désormais, les mots intégrable, intégrale sous-entendent "au sens de Lebesgue". Au contraire, on précisera si l'on veut parler d'intégrale au sens de Riemann.

2.4.2 Exemples

Les fonctions Riemann-intégrables

Pour faire le lien avec l'intégrale de Riemann nous définissons d'abord l'intégrabilité sur une partie :

²La réciproque est presque vraie, mais pas tout à fait, voir le paragraphe sur les fonctions mesurables.

³Comme dans le cas de Riemann, il s'agit d'une semi-norme et pas d'une norme car une fonction peut être de "norme" nulle sans être nulle. C'est le cas, par exemple, de la fonction caractéristique d'un ensemble fini, voire négligeable.

2.4.7 Définition. Si f est définie sur une partie $A \subset \mathbf{R}$, on dit qu'elle est **intégrable sur A** si la fonction \tilde{f} obtenue en prolongeant f par 0 en dehors de A est intégrable sur \mathbf{R} . On pose $\int_A f := \int_{\mathbf{R}} \tilde{f}$ (ou $\int_a^b f(t)dt = \int_{\mathbf{R}} \tilde{f}$ si A est le segment $[a, b]$).

On a ainsi, en particulier, une notion de fonction intégrable sur un intervalle.

2.4.8 Proposition. Soit f une fonction Lebesgue-intégrable sur \mathbf{R} et soit I un intervalle. La fonction f "tronquée", égale à f sur I et nulle en dehors de I est Lebesgue-intégrable.

Démonstration. En effet, les restrictions aux intervalles de fonctions en escalier en sont encore.

Le premier exemple de fonctions intégrables au sens de Lebesgue ce sont les fonctions Riemann-intégrables : tout ce qui a été fait au chapitre 1 demeure donc valable.

2.4.9 Proposition. Si f est Riemann-intégrable sur $[a, b]$ et si \tilde{f} est la fonction obtenue en prolongeant f par 0 en dehors de $[a, b]$, \tilde{f} est Lebesgue-intégrable et les intégrales $\int_a^b f(t)dt$ au sens de Riemann et de Lebesgue sont égales.

Démonstration. Il résulte de 1.1.22, appliqué avec $\epsilon = 1/n$, qu'il existe deux suites (g_n) et (h_n) de fonctions en escalier définies sur $[a, b]$, qui vérifient $g_n \leq f \leq h_n$ et telles que l'intégrale $\int_a^b (h_n - g_n)(t)dt$ tende vers 0. La suite (g_n) converge vers f au sens de L^1 dans l'espace des fonctions Riemann-intégrables, donc est de Cauchy⁴ pour L^1 . De plus, comme les intégrales $\int |h_n - g_n|$ tendent vers 0, il existe une sous-suite⁵ $(h_{n_k} - g_{n_k})$ qui converge presque partout vers 0 (c'est Egorov 2.3.1.2 appliqué à $\tilde{h}_n - \tilde{g}_n$). Mais alors, on a $0 \leq f - g_{n_k} \leq h_{n_k} - g_{n_k}$ de sorte que (g_{n_k}) converge presque partout vers f . Comme cette suite est encore de Cauchy, il en résulte que f est bien Lebesgue-intégrable. Son intégrale au sens de Lebesgue est la limite de $\int_a^b g_{n_k}(t)dt$, c'est bien l'intégrale de Riemann.

Des intégrales impropres

⁴On peut aussi utiliser les inégalités : $|g_p - g_q| \leq (f - g_p) + (f - g_q) \leq (h_p - g_p) + (h_q - g_q)$.

⁵L'exemple suivant : f est nulle, h_n est la suite de l'exemple horrible et $g_n = -h_n$ montre qu'il peut être nécessaire d'extraire une sous-suite.

2.4.10 Exercice.

- 1) Montrer que la fonction qui vaut $1/x^2$ pour $x \geq 1$ et 0 ailleurs est Lebesgue-intégrable (on exhibera une suite de fonctions en escalier convenable⁶ qui l'approche).
- 2) Montrer que la fonction qui vaut $1/\sqrt{x}$ sur $]0, 1]$ et 0 ailleurs est Lebesgue-intégrable.

Nous verrons, après les théorèmes de convergence, le résultat général concernant les intégrales impropres : une intégrale impropre au sens de Riemann qui converge **absolument** est aussi une intégrale au sens de Lebesgue.

D'autres fonctions intégrables

Nous avons vu ci-dessus que la fonction caractéristique des rationnels de $[0, 1]$, qui n'est pas Riemann-intégrable, l'est au sens de Lebesgue. Voici un autre exemple :

2.4.11 Proposition. *Soit K un compact de \mathbf{R} . Alors χ_K est intégrable au sens de Lebesgue.*

Démonstration. Soit n un entier > 0 . On recouvre le compact K par un nombre fini d'intervalles $[x - 1/n, x + 1/n]$, avec $x \in K$, et on note J'_n la réunion de ces intervalles et $J_n = J'_1 \cap J'_2 \cap \dots \cap J'_n$. L'ensemble J_n est encore une réunion finie de segments, de sorte que sa fonction caractéristique f_n est une fonction en escalier, la suite (J_n) décroît, donc aussi la suite (f_n) . Montrons qu'on a $K = \bigcap_n J_n$. En effet, il est clair que K est contenu dans l'intersection. Réciproquement, si x n'est pas dans K , sa distance à K , $\delta = d(x, K) = \inf_{y \in K} |x - y|$, est > 0 , donc x n'est pas dans J_n pour $n > 1/\delta$ (car tout point de J_n est à une distance $\leq 1/n$ d'un point de K par construction). Il en résulte que χ_K est limite simple des f_n . Enfin, la suite (f_n) est de Cauchy au sens de L^1 . En effet, comme la suite (f_n) est décroissante, on a, pour $p \leq q$, $\int |f_p - f_q| = \int (f_p - f_q) = \int f_p - \int f_q$. Or, la suite $\int f_n$ est une suite décroissante de nombres ≥ 0 . Elle a donc une limite, donc est de Cauchy, donc aussi (f_n) (la difficulté évoquée en 2.2.1 ne se produit pas).

2.4.12 Exemple. Si K est un "faux Cantor" non négligeable, on a vu que χ_K n'est pas Riemann-intégrable. En revanche, cette fonction l'est au sens de Lebesgue.

⁶Avec des supports de plus en plus grands.

2.4.3 Propriétés de l'intégrale

L'intégrale de Lebesgue a essentiellement les mêmes propriétés que l'intégrale de Riemann :

2.4.13 Théorème.

1) L'ensemble $\mathcal{L}(\mathbf{R})$ des fonctions Lebesgue-intégrables $f : \mathbf{R} \rightarrow \mathbf{R}$ est un sous-espace vectoriel de l'espace vectoriel de toutes les fonctions de \mathbf{R} dans \mathbf{R} .

2) L'application $I : f \mapsto \int_{\mathbf{R}} f(t)dt$ est une forme linéaire sur $\mathcal{L}(\mathbf{R})$.

3) L'application I est croissante : si on a $f \leq g$ on a $I(f) \leq I(g)$. En particulier, si f est intégrable, $|f|$ l'est aussi et on a l'inégalité $|\int_{\mathbf{R}} f(t)dt| \leq \int_{\mathbf{R}} |f(t)|dt$.

4) Si f et g sont intégrables, $\text{Max}(f, g)$ et $\text{Min}(f, g)$ le sont aussi.

Démonstration. Montrons ensemble 1) et 2). Soient $f, g \in \mathcal{L}$. On écrit f (resp. g) comme limite presque partout d'une suite de Cauchy (f_n) (resp. (g_n)) de fonctions en escalier. On note d'abord que $(f_n + g_n)$ converge presque partout vers $f + g$ (si (f_n) converge en dehors de A et (g_n) en dehors de B avec A, B négligeables, $(f_n + g_n)$ converge en dehors de $A \cup B$ qui est négligeable). Comme la suite $(f_n + g_n)$ est encore une suite de Cauchy de fonctions en escalier, $f + g$ est intégrable et son intégrale est la limite de $\int(f_n + g_n) = \int f_n + \int g_n$, c'est donc $\int f + \int g$.

Pour le point 3) on écrit f, g comme limites de (f_n) et (g_n) . Comme on a $f \leq g$, f et g sont encore limites, respectivement, de $f'_n = \text{Min}(f_n, g_n)$ et de $g'_n = \text{Max}(f_n, g_n)$, qui sont encore des suites de Cauchy de fonctions en escalier. Comme on a $f'_n \leq g'_n$, on a $\int f'_n \leq \int g'_n$, donc aussi $\int f \leq \int g$ à la limite⁷.

Le point 4) résulte de formules du type : $\text{Max}(f, g) = \frac{1}{2}(f + g + |f - g|)$ et de 1) et 3).

2.4.14 Remarques.

1) La fonction f constante et égale à $a > 0$ n'est pas intégrable. En effet, si on note f_n la fonction égale à a sur $[-n, n]$ et nulle ailleurs, c'est une fonction en escalier, d'intégrale $2an$ et on a $f_n \leq f$. Si f est intégrable on a $\int f \geq \int f_n = 2an$ en vertu du point 3), ce qui est absurde (on dira plus loin que f est mesurable et d'intégrale infinie). Un argument analogue montre que la fonction qui vaut $1/x$ sur $]0, 1]$ et 0 ailleurs n'est pas intégrable.

2) Attention, en général, si f et g sont intégrables (à valeurs réelles) :

⁷Exercice : écrire les détails de cette preuve.

- la composée $g \circ f$ ne l'est pas. Exemple : sur $]0, 1]$, $g(x) = x^2$ et $f(x) = 1/\sqrt{x}$.
- le produit fg ne l'est pas. Exemple : sur $]0, 1]$, $f = g = 1/\sqrt{x}$.

2.4.4 Fonctions à valeurs complexes

Comme dans le cas de l'intégrale de Riemann, une fonction $f : \mathbf{R} \rightarrow \mathbf{C}$ est dite Lebesgue-intégrale si ses parties réelles et imaginaires le sont et on pose $\int_{\mathbf{R}} f = \int_{\mathbf{R}} \operatorname{Re} f + i \int_{\mathbf{R}} \operatorname{Im} f$. On vérifie que, si f est intégrable, son module $|f|$ l'est aussi et qu'on a $|\int_{\mathbf{R}} f| \leq \int_{\mathbf{R}} |f|$.

2.5 Les théorèmes de convergence

2.5.1 Complétude

Nous montrons d'abord que le plan annoncé a été réalisé : l'espace \mathcal{L} est complet. C'est là que sert de manière essentielle le point 1) du lemme d'Egorov.

2.5.1 Théorème. *L'espace \mathcal{L} est complet pour la semi-norme de la convergence en moyenne : si (f_n) est une suite de Cauchy de \mathcal{L} au sens de L^1 , il existe $f \in \mathcal{L}$ telle que (f_n) converge vers f au sens de L^1 .*

Démonstration. Soit (f_n) une suite de Cauchy de \mathcal{L} . Comme f_n est intégrable, il existe une fonction en escalier g_n qui vérifie $\|f_n - g_n\|_1 \leq 2^{-n}$ (cf. prop. 2.4.5).

La suite (g_n) est de Cauchy au sens de L^1 . En effet, cela résulte de l'inégalité :

$$\|g_p - g_q\|_1 \leq \|g_p - f_p\|_1 + \|f_p - f_q\|_1 + \|f_q - g_q\|_1.$$

En vertu d'Egorov, il y a une sous-suite g_{n_k} qui converge simplement presque partout vers une fonction f . En vertu de la définition, la fonction f est dans \mathcal{L} et en vertu de 2.4.5, la suite converge vers f au sens de L^1 . L'inégalité $\|f - f_{n_k}\|_1 \leq \|f - g_{n_k}\|_1 + \|g_{n_k} - f_{n_k}\|_1$ montre que (f_{n_k}) converge vers f au sens de L^1 et la condition de Cauchy assure alors que la suite (f_n) tout entière converge vers f au sens de L^1 .

2.5.2 Lemme d'Egorov pour les suites de fonctions Lebesgue-intégrables

La convergence de (f_n) vers f au sens de L^1 c'est bien, car elle implique la convergence des intégrales $\int f_n$ vers $\int f$. Il reste cependant une question : implique-t-elle la convergence simple presque partout ? L'exemple horrible montre que ce n'est pas vrai, même pour les fonctions en escalier, mais, comme dans ce cas, on a un pis-aller avec des sous-suites :

2.5.2 Proposition. (Egorov-Lebesgue) *Si la suite de fonctions Lebesgue-intégrables (f_n) converge au sens de L^1 vers la fonction intégrable f , il existe une sous-suite (f_{n_k}) de (f_n) qui converge simplement vers f presque partout.*

2.5.3 Remarque. Inversement, si (f_n) est une suite de Cauchy de fonctions Lebesgue-intégrables et si une sous-suite (f_{n_k}) de (f_n) converge presque partout vers une fonction f , alors f est intégrable et la suite (f_n) converge vers f au sens de L^1 . En effet, la complétude montre que f_n tend vers une fonction intégrable g au sens de L^1 , donc aussi la sous-suite f_{n_k} . Mais, en appliquant 2.5.2 à (f_{n_k}) , on voit que f est égale à g presque partout.

Démonstration. (de 2.5.2) Elle nécessite quelques préliminaires.

2.5.4 Proposition. *Soit f une fonction intégrable sur \mathbf{R} et soit $\epsilon > 0$. Il existe une fonction en escalier g et une partie A , réunion dénombrable d'intervalles I_n dont la somme des longueurs est $\leq \epsilon$, telles que l'on ait :*

- 1) $\|f - g\|_1 \leq \epsilon$,
- 2) $\|f - g\|_{A^c, \infty} = \sup_{x \notin A} |f(x) - g(x)| \leq \epsilon$.

Démonstration. Le fait que f soit Lebesgue-intégrable assure qu'il existe une suite de Cauchy (g_n) de fonctions en escalier qui converge presque partout vers f et on a $\|f - g_n\|_1 \leq \epsilon$ pour n assez grand, en vertu de 2.4.5. Par ailleurs, en vertu d'Egorov, et quitte à extraire une sous-suite, on peut supposer que la convergence de (g_n) vers f est uniforme en dehors d'une partie A comme ci-dessus de longueur $\leq \epsilon$. On a le résultat en prenant $g = g_n$ pour n assez grand.

Ce résultat nous donne un théorème de structure pour les suites de fonctions intégrables :

2.5.5 Corollaire. (Lemme de structure des suites de fonctions Lebesgue-intégrables) *Soit (f_n) une suite de fonctions intégrables. Il existe une suite de fonctions en escalier (g_n) qui vérifie les deux propriétés suivantes :*

- 1) $\|f_n - g_n\|_1$ tend vers 0 quand n tend vers $+\infty$,
- 2) la suite $f_n - g_n$ tend simplement vers 0 presque partout.

Démonstration. On utilise la proposition précédente en trouvant g_n en escalier qui approche f_n à $1/2^n$ près à la fois au sens de L^1 et au sens de la convergence uniforme en dehors d'un ensemble A_n de "longueur" $\leq 1/2^n$. Je dis que la suite (g_n) convient. En effet, on a clairement la condition sur la norme L^1 . Pour la convergence simple, on introduit, comme dans Egorov, l'ensemble $E = \bigcap_{p \in \mathbb{N}} B_p$ avec $B_p = \bigcup_{n \geq p} A_n$. On a deux propriétés :

- 1) E est négligeable,
- 2) si x n'est pas dans E , $f_n(x) - g_n(x)$ tend vers 0 quand n tend vers l'infini.

La première assertion est claire, car E est inclus dans B_p qui est de longueur $\leq 1/2^{p-1}$. Pour la deuxième, on procède comme dans Egorov : si x n'est pas dans E il existe p tel que $x \notin \bigcup_{n \geq p} A_n$, c'est-à-dire que x n'est dans aucun A_n pour $n \geq p$. Par définition de A_n cela signifie qu'on a $|f_n(x) - g_n(x)| \leq 1/2^n$ pour $n \geq p$ et on a gagné.

Retour à Egorov-Lebesgue

En considérant $f_n - f$ on se ramène au cas $f = 0$. C'est alors essentiellement le lemme de structure des suites. On associe à (f_n) une suite (g_n) de fonctions en escalier comme dans 2.5.5. Le fait que (f_n) converge vers 0 au sens de L^1 et que $\|f_n - g_n\|_1$ tende vers 0 implique que (g_n) converge vers 0 au sens de L^1 . Par Egorov, il y a une sous-suite (g_{n_k}) qui converge presque partout vers 0. La deuxième condition de 2.5.5 montre que (f_{n_k}) converge aussi vers 0.

2.5.6 Corollaire. *Si une fonction intégrable f vérifie $\|f\|_1 = 0$, f est nulle presque partout.*

Démonstration. On considère la suite de fonctions constantes $f_n = 0$. Elle converge vers f au sens L^1 car on a $\|f - f_n\|_1 = \|f\|_1 = 0$. En vertu d'Egorov-Lebesgue, il existe une sous-suite f_{n_k} qui converge vers f presque partout. Mais, comme les f_{n_k} sont nulles, on a $f = 0$ presque partout.

2.5.3 Convergence monotone : Beppo-Levi

Le théorème de complétude, s'il est essentiel d'un point de vue théorique, reste difficilement utilisable dans la pratique. En revanche, les deux résultats suivants vont être essentiels dans toutes les applications.

2.5.7 Théorème. (de convergence monotone ou de Beppo-Levi) *Soit (f_n) une suite monotone de fonctions de \mathcal{L} à valeurs réelles. On suppose que*

la suite des intégrales $I_n = \int f_n$ est bornée. Alors, la suite (f_n) converge vers une fonction f Lebesgue-intégrable à la fois au sens de L^1 et au sens de la convergence simple presque partout, et la suite $\int f_n$ converge vers $\int f$.

Démonstration. Supposons par exemple la suite croissante. La suite I_n est croissante et bornée, donc elle converge vers un réel I . Montrons que (f_n) est de Cauchy au sens de L^1 . Cela signifie⁸ que $\int |f_q - f_p| = \int (f_q - f_p) = I_q - I_p$ (avec $p \leq q$) tend vers 0 quand p tend vers l'infini, c'est-à-dire que la suite des intégrales I_n est de Cauchy, ce qui est clair puisqu'elle est convergente. En vertu du théorème de complétude 2.5.1 la suite (f_n) converge vers une fonction $f \in \mathcal{L}$ au sens de L^1 . En vertu de 2.5.2, il y a une sous-suite f_{n_k} qui converge simplement vers f presque partout. Mais, comme (f_n) est croissante, la convergence d'une sous-suite au point x implique celle de la suite entière, cqfd.

2.5.8 Remarque. Dans la pratique, on montrera souvent la convergence presque partout de la suite (f_n) directement, ne serait-ce que pour identifier la fonction f .

2.5.9 Exemple. On peut reprendre l'exemple des intégrales de Wallis. Comme la suite $f_n(x) = \sin^n x$ est décroissante vers 0 sur $[0, \pi/2[$ et que les intégrales des f_n sont bornées par 1, on en déduit bien que les intégrales de Wallis convergent vers 0. Même argument pour $\int_1^{+\infty} e^{-x^n} dx$.

2.5.10 Corollaire. (Intégration terme à terme des séries positives)

Soit (u_n) une suite de fonctions intégrables positives. On suppose $\sum_{n=0}^{+\infty} \int_{\mathbf{R}} u_n < +\infty$. Alors, la série $\sum_{n=0}^{+\infty} u_n(x)$ converge presque partout et on a la formule :

$$\int_{\mathbf{R}} \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_{\mathbf{R}} u_n.$$

2.5.11 Exemple. En vertu du corollaire on a l'égalité :

$$\sum_{n=0}^{+\infty} \int_0^1 x^{2n}(1-x) = \int_0^1 \sum_{n=0}^{+\infty} x^{2n}(1-x)$$

qui donne $\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} = \ln 2$. (L'idée : on a $\ln 2 = \int_0^1 \frac{dx}{1+x}$ et on écrit $\frac{1}{1+x} = \frac{1-x}{1-x^2}$.)

⁸La difficulté évoquée en 2.2.1 ne se produit pas ici. En effet, il n'y a pas besoin de valeur absolue ici puisque la suite est monotone.

2.5.4 Convergence dominée : Lebesgue

2.5.12 Théorème. (de convergence dominée, ou de Lebesgue) Soit (f_n) une suite de fonctions de \mathcal{L} (à valeurs complexes). On suppose que (f_n) converge simplement presque partout vers une fonction f et qu'il existe une fonction Lebesgue-intégrable positive g telle que l'on ait, pour tout n , $|f_n| \leq g$ presque partout. Alors, f est intégrable et on a $\int f = \lim \int f_n$. De plus, la suite (f_n) converge vers f au sens de L^1 .

Démonstration. Notons que, quitte à modifier les f_n sur un ensemble négligeable, on peut supposer qu'on a $|f_n| \leq g$ partout. On commence par un lemme :

2.5.13 Lemme. Soit (g_n) une suite de fonctions intégrables positives, majorées par une fonction intégrable g . Alors la fonction $g' = \sup_{n \in \mathbb{N}} g_n$ est intégrable.

Démonstration. (du lemme) En vertu de 2.4.13.4, si f et g sont intégrables, $\text{Max}(f, g)$ l'est aussi et il en est de même, par récurrence, pour le maximum de n fonctions. On considère les fonctions intégrables $g'_n = \text{Max}(g_0, g_1, \dots, g_n)$. Elles sont majorées par g et on a $\sup_n g'_n = \sup_n g_n = g'$. Comme la suite g'_n est croissante, elle converge vers g' et les intégrales $\int g'_n$ sont majorées par $\int g$. On peut donc appliquer Beppo-Levi à la suite (g'_n) , ce qui montre que g' est intégrable.

Revenons à la convergence dominée. Montrons que la suite (f_n) est de Cauchy au sens de L^1 . Pour cela on considère la suite de fonctions (s_n) :

$$s_n = \sup_{p, q \geq n} |f_p - f_q|.$$

En vertu du lemme, ces fonctions sont intégrables (car on a $|f_p - f_q| \leq 2g$). De plus, la suite $s_n(x)$ tend vers 0 presque partout par hypothèse (car $f_n(x)$ est convergente, donc de Cauchy). La suite s_n étant décroissante, elle tend vers 0 au sens de L^1 par Beppo-Levi, ce qui implique que (f_n) est de Cauchy au sens de L^1 . Le théorème de complétude assure alors que (f_n) converge vers une fonction intégrable h au sens de L^1 , et cela implique que $\int f_n$ converge vers $\int h$. De plus, en vertu d'Egorov-Lebesgue, il y a une sous-suite (f_{n_k}) de (f_n) qui converge vers h presque partout. Mais, comme (f_n) converge vers f presque partout, cela montre que f et h sont égales presque partout et on a terminé.

2.5.14 Remarque. Si l'on sait d'avance que la fonction f est intégrable, on peut simplifier la preuve précédente en considérant la suite de fonctions $g_n = \sup_{p \geq n} |f - f_p|$, qui est décroissante et converge presque partout vers 0, donc converge au sens de L^1 par Beppo-Levi.

2.5.5 Une application des théorèmes de convergence : les intégrales généralisées

Dans ce paragraphe nous faisons le lien entre les intégrales de Riemann généralisées et l'intégrale de Lebesgue.

2.5.15 Théorème. Soit $f : [a, b[\rightarrow \mathbf{C}$ une fonction (b désigne une borne finie ou infinie). On suppose que, pour tout $c < b$, la fonction f est Riemann-intégrable sur $[a, c]$. Alors, f est Lebesgue-intégrable sur $[a, b[$ si et seulement si l'intégrale de Riemann généralisée $\int_a^b f(t)dt$ converge **absolument**. De plus l'intégrale au sens de Lebesgue et l'intégrale généralisée ci-dessus coïncident.

Démonstration. Traitons le cas $b = +\infty$, les autres sont analogues. Supposons d'abord que l'intégrale généralisée converge absolument. On considère la suite de fonctions g_n qui sont les tronquées de $|f|$ sur les intervalles $[a, n]$. La suite (g_n) converge vers $|f|$ en croissant et ses intégrales sont majorées par l'intégrale (généralisée) $\int_a^{+\infty} |f(t)|dt$. On peut donc lui appliquer le théorème de convergence monotone qui montre déjà que $|f|$ est Lebesgue-intégrable. On considère ensuite les fonctions f_n tronquées de f sur les mêmes intervalles. La suite (f_n) converge simplement vers f et les modules des f_n sont majorés par $|f|$ (qui est Lebesgue-intégrable). On peut donc appliquer le théorème de convergence dominée qui montre que f est Lebesgue-intégrable et que son intégrale est la limite des intégrales des f_n , donc égale à l'intégrale généralisée.

Réciproquement, si l'on suppose f Lebesgue-intégrable, on a vu que $|f|$ l'est aussi. On introduit la suite de fonctions (g_n) comme ci-dessus. On a $\int g_n = \int_a^n |f(t)|dt \leq \int_{\mathbf{R}} |f|$ et on conclut à la convergence absolue de l'intégrale impropre de f .

2.5.16 Remarque. Attention, la théorie ne s'applique pas pour les intégrales semi-convergentes. Par exemple, bien que l'intégrale $\int_0^{+\infty} \frac{\sin x}{x} dx$ soit semi-convergente, la fonction $\frac{\sin x}{x}$ n'est pas Lebesgue-intégrable sur \mathbf{R}^+ . En effet, si elle l'était, sa valeur absolue serait intégrable et l'intégrale convergerait absolument, ce qui n'est pas le cas (minorer la fonction sur les intervalles $[(2k+1)\frac{\pi}{2} - \frac{\pi}{4}, (2k+1)\frac{\pi}{2} + \frac{\pi}{4}]$).

2.5.17 Remarque. Certains résultats vus au chapitre 1 sont encore vrais lorsqu'on considère des intégrales sur des intervalles non bornés, ou des intégrales de fonctions non bornées, à condition de rajouter l'hypothèse que les fonctions considérées sont Lebesgue-intégrables. Précisément :

- Si f est Lebesgue-intégrable sur un intervalle I (éventuellement ouvert ou non borné) et si a est dans I , $F(x) = \int_a^x f$ est continue sur I . Si f est continue sur I , F est dérivable et de dérivée f .

- Si u, v sont de classe C^1 sur un intervalle $I = (a, b)$ (éventuellement ouvert ou infini) et si les fonctions $u'v$ et uv' sont dans $\mathcal{L}(I)$, on a la formule d'intégration par parties : $\int_I uv' = [uv]_I - \int_I u'v$ (où le symbole $[uv]_I$ désigne la limite de $u(x)v(x) - u(y)v(y)$ quand x tend vers a et y vers b).

2.5.6 Autres applications des théorèmes de convergence

1) Le cas des intégrales de Wallis devient trivial. En effet, les fonctions $\sin^n x$ sont toutes majorées par 1.

2) Voici un exemple un peu plus intéressant. On considère les intégrales $I_n = \int_0^n \left(1 - \frac{t}{n}\right)^n \ln t dt$, que l'on voit comme des intégrales sur \mathbf{R}^+ . Les fonctions $f_n(t) = \left(1 - \frac{t}{n}\right)^n \ln t$ sont toutes majorées par $e^{-t} \ln t$ qui est intégrable sur \mathbf{R}^+ . On peut donc appliquer le théorème de convergence dominée, de sorte que la limite de I_n est $\int_0^{+\infty} e^{-t} \ln t dt$. Cette remarque permet de montrer la formule suivante sur la constante d'Euler : $\gamma = -\int_0^{+\infty} e^{-t} \ln t dt$.

3) On peut reprendre le calcul de $\ln 2 = \int_0^1 \frac{dt}{1+t}$. Pour cela on développe :

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \dots + (-1)^n t^n + \dots$$

et, comme les fonctions $f_n(t) = 1 - t + t^2 - t^3 + \dots + (-1)^n t^n = \frac{1 + (-1)^{n+1} t^{n+1}}{1+t}$ sont majorées par 2, on peut intervertir les signes \int et \sum , d'où le résultat.

Parenthèse

Cette parenthèse est destinée aux étudiants qui envisagent de préparer le CAPES. J'espère que vous êtes convaincus de la puissance des théorèmes de Lebesgue. Attention, l'usage de ces théorèmes est prohibé à l'écrit du CAPES. Il faut donc se débrouiller sans cela (mais Lebesgue montre où aller). Dans le cas du résultat du point 3) précédent c'est facile, il suffit d'établir les deux points suivants :

1) On a, pour $t \neq -1$, $1 - t + t^2 \dots + (-1)^n t^n = \frac{1 + (-1)^{n+1} t^{n+1}}{1+t}$.

2) On a $0 \leq \int_0^1 \frac{t^{n+1}}{1+t} dt \leq \frac{1}{n+2}$.

2.5.7 Intégrales dépendant d'un paramètre

Il s'agit d'étudier les fonctions données par des intégrales : $F(\lambda) = \int_{\mathbf{R}} f(x, \lambda) dx$.

Limite et continuité

2.5.18 Théorème. Soit Λ un espace métrique quelconque (par exemple une partie de \mathbf{R}^d) et soit $f : \mathbf{R} \times \Lambda \rightarrow \mathbf{C}$ une fonction vérifiant les conditions suivantes :

- 1) $\forall \lambda \in \Lambda$, la fonction $x \mapsto f(x, \lambda)$ est intégrable,
- 2) Il existe un ensemble négligeable $A \subset \mathbf{R}$ tel que :
 - a) pour $x \notin A$, $f(x, \lambda)$ tend vers $g(x)$ quand λ tend vers λ_0 ,
 - b) il existe une fonction intégrable $h \geq 0$ telle que l'on ait, pour tout $\lambda \in \Lambda$ et tout $x \notin A$, $|f(x, \lambda)| \leq h(x)$.

Alors, la fonction g est intégrable et $F(\lambda) = \int_{\mathbf{R}} f(x, \lambda) dx$ tend vers $\int_{\mathbf{R}} g(x) dx$ quand λ tend vers λ_0 .

En particulier, si, pour tout $x \notin A$, la fonction $\lambda \mapsto f(x, \lambda)$ est continue en λ_0 , la fonction $F(\lambda) = \int_{\mathbf{R}} f(x, \lambda) dx$ est continue en λ_0 .

Démonstration. Il suffit de montrer que si (λ_n) est une suite convergeant vers λ_0 , la suite $(F(\lambda_n))$ converge vers $\int_{\mathbf{R}} g(x) dx$. C'est exactement le théorème de convergence dominée appliqué aux fonctions $f_n(x) = f(x, \lambda_n)$. Le cas de la continuité s'obtient en appliquant le résultat avec $g(x) = f(x, \lambda_0)$.

Dérivabilité

2.5.19 Théorème. Dérivation sous le signe somme

Soit I un intervalle ouvert de \mathbf{R} et soit $f : \mathbf{R} \times I \rightarrow \mathbf{R}$ une fonction vérifiant les conditions suivantes :

- 1) $\forall \lambda \in I$, $x \mapsto f(x, \lambda)$ est intégrable.
- 2) Il existe un ensemble négligeable $A \subset \mathbf{R}$ tel que :
 - a) pour $x \notin A$, la fonction $\lambda \mapsto f(x, \lambda)$ est dérivable sur I ,
 - b) il existe une fonction intégrable $g \geq 0$ telle que l'on ait, pour tout $\lambda \in I$

et tout $x \notin A$, $\left| \frac{\partial f}{\partial \lambda}(x, \lambda) \right| \leq g(x)$.

Alors, la fonction $\lambda \mapsto F(\lambda) = \int_{\mathbf{R}} f(x, \lambda) dx$ est dérivable sur I et on a :

$$F'(\lambda) = \int_{\mathbf{R}} \frac{\partial f}{\partial \lambda}(x, \lambda) dx.$$

Démonstration. Fixons $\lambda \in I$. Il s'agit de calculer la limite, quand h_n tend vers 0, de $\frac{F(\lambda + h_n) - F(\lambda)}{h_n}$. Cette quantité est l'intégrale de la fonction $\varphi_n(x) = \frac{f(x, \lambda + h_n) - f(x, \lambda)}{h_n}$. Par définition de la dérivée, pour $x \notin A$, cette fonction tend vers $\frac{\partial f}{\partial \lambda}(x, \lambda)$ quand n tend vers $+\infty$. Par ailleurs, le théorème des accroissements finis montre qu'on a $\varphi_n(x) = \frac{\partial f}{\partial \lambda}(x, \theta_n)$ avec $\theta_n \in]\lambda, \lambda + h_n[$. On a donc $|\varphi_n(x)| \leq g(x)$ pour $x \notin A$. On conclut avec le théorème de convergence dominée.

2.5.20 Remarque. Si I n'est pas ouvert, le théorème vaut, *mutatis mutandis*, avec les dérivées à gauche et à droite aux bornes de I .

2.5.21 Exemple. La fonction Γ . On considère la fonction définie par $\Gamma(\lambda) = \int_0^{+\infty} e^{-x} x^{\lambda-1} dx$ pour $\lambda > 0$. Cette fonction est bien définie car l'intégrale converge absolument pour $\lambda > 0$. La fonction Γ est continue et dérivable sur $]0, +\infty[$. Il suffit de le montrer sur $]\epsilon, A[$ avec $0 < \epsilon < 1 < A$. Pour la continuité, on majore la fonction $x \mapsto e^{-x} x^{\lambda-1}$ par la fonction h qui vaut $e^{-x} x^{\epsilon-1}$ pour $x < 1$ et $e^{-x} x^{A-1}$ pour $x \geq 1$. Pour la dérivabilité, il faut regarder l'intégrale $\int_0^{+\infty} e^{-x} x^{\lambda-1} \ln x dx$ et la fonction à intégrer est alors majorée par la fonction h qui vaut $e^{-x} x^{\epsilon-1} \ln x$ (resp. $e^{-x} x^{A-1} \ln x$) pour $x < 1$ (resp. $x \geq 1$) et qui est intégrable.

La fonction Γ est intéressante car elle interpole la suite $n!$. En effet, on vérifie qu'on a $\Gamma(1) = 1$ et $\Gamma(\lambda + 1) = \lambda \Gamma(\lambda)$ pour $\lambda > 0$, d'où, pour $n \in \mathbf{N}^*$, $\Gamma(n) = (n - 1)!$. On peut utiliser cette fonction pour prouver la formule de Stirling : $n! \sim n^n e^{-n} \sqrt{2\pi n}$.

2.5.8 Quelques contre-exemples

Dans les exemples qui suivent, ni le théorème de convergence monotone, ni celui de convergence dominée ne s'appliquent.

Concentration

On a déjà rencontré ce type d'exemple. Le plus simple est de prendre la fonction f_n constante et égale à n sur l'intervalle $[0, 1/n]$ et nulle en dehors (un rectangle debout). Il est clair que la suite (f_n) converge simplement vers 0 presque partout (sauf en 0), mais les intégrales sont toutes égales à 1.

Si l'on veut un exemple qui converge partout, on peut utiliser les fonctions pics de 2.2.2.

Une variante C^∞ sur \mathbf{R}^+ de cet exemple est donnée par $f_n(x) = n^2 x e^{-nx}$ (si on pose $f(x) = x e^{-x}$, on a $f_n(x) = n f(nx)$).

Étalement

Dans cet exemple au contraire, le rectangle est couché : f_n est constante et égale à $1/n$ sur $[0, n]$ et nulle en dehors. Là encore, la suite converge simplement vers 0, mais l'intégrale est constamment égale à 1. Exemple C^∞ : $f_n(x) = \frac{x}{n^2} e^{-x/n}$. (Avec les notations précédentes, on a $f_n(x) = \frac{1}{n} f(\frac{x}{n})$).

La bosse glissante

Cette fois, le rectangle glisse ! On prend f_n égale à 1 sur $[n, n+1]$ et nulle ailleurs. La suite tend simplement vers 0, mais l'intégrale est constamment égale à 1.

2.6 Fonctions et ensembles mesurables

Nous donnons dans cette section les définitions et les propriétés des fonctions et des ensembles mesurables. Ces notions sont beaucoup moins importantes dans l'approche que nous avons adoptée que dans celle qui part de la notion de mesure, mais elles seront fondamentales dans le chapitre sur les fonctions de plusieurs variables. L'idée qui préside à la définition d'une fonction mesurable est toute simple : une fonction mesurable c'est une fonction qui serait intégrable si elle n'était pas aussi haute ou aussi large.

2.6.1 Définition

2.6.1 Définition.

1) Soit f une fonction de \mathbf{R} dans \mathbf{R} et soient a, b des réels > 0 . On appelle **fonction tronquée** associée à f, a, b et on note $f_{a,b}$ la fonction définie comme suit :

i) $f_{a,b}(x) = 0$ pour $|x| > a$,

ii) pour $|x| \leq a$, $f_{a,b}(x) = f(x)$ si $|f(x)| \leq b$ et $f_{a,b}(x) = b \frac{f(x)}{|f(x)|}$ sinon. Autrement dit, $f_{a,b}(x) = \text{Min}(b, \text{Max}(f(x), -b))$.

2) Soit f une fonction de \mathbf{R} dans \mathbf{C} et soient a, b des réels > 0 . On appelle **fonction tronquée** associée à f, a, b et on note $f_{a,b}$ la fonction définie comme suit :

i) $f_{a,b}(x) = 0$ pour $|x| > a$,

ii) pour $|x| \leq a$, $f_{a,b}(x) = f(x)$ si $|f(x)| \leq b$ et $f_{a,b}(x) = b \frac{f(x)}{|f(x)|}$ sinon.

Dit en français, $f_{a,b}$ c'est la fonction égale à f , sauf qu'elle est nulle si la variable dépasse a (en valeur absolue ou en module) et qu'on tronque son module à b si la fonction dépasse b en module.

2.6.2 Définition. Soit f une fonction de \mathbf{R} dans \mathbf{R} ou \mathbf{C} . On dit que f est **mesurable** si, pour tous $a, b > 0$, la fonction tronquée $f_{a,b}$ est Lebesgue-intégrable.

Voici deux exemples importants de fonctions mesurables :

2.6.3 Proposition.

- 1) Toute fonction continue (par exemple une fonction constante) est mesurable.
- 2) Toute fonction intégrable (par exemple une fonction en escalier) est mesurable.

Démonstration. Comme le maximum (resp. le minimum) de deux fonctions continues est continu, le point 1) résulte du fait que les fonctions continues sur un segment sont Riemann-intégrables. Pour le point 2) il s'agit de voir que si f est intégrable, ses tronquées le sont aussi. Cela résulte du fait que le maximum et le minimum de deux fonctions intégrables sont intégrables, cf. 2.4.13.

2.6.2 Propriétés des fonctions mesurables

La notion de fonction mesurable est la plus stable qui soit. À se demander comment on peut faire pour obtenir des fonctions non mesurables. En fait, on peut pratiquement faire comme si toutes les fonctions étaient mesurables, cf. 2.6.16. On commence par la stabilité par limite simple :

2.6.4 Proposition. Si f est limite simple presque partout d'une suite de fonctions mesurables elle est mesurable.

Démonstration. Supposons f limite de (f_n) avec les f_n mesurables. Soient $a, b > 0$. Il est clair que la suite des $(f_n)_{a,b}$ converge simplement vers $f_{a,b}$ (utiliser la formulation avec Min et Max). Comme toutes ces fonctions sont majorées par la constante b sur $[-a, a]$ qui est intégrable, $f_{a,b}$ est intégrable en vertu du théorème de convergence dominée, donc f est mesurable.

On en déduit une caractérisation des fonctions mesurables :

2.6.5 Corollaire. Une fonction est mesurable si et seulement si elle est limite simple presque partout d'une suite de fonctions en escalier.

Démonstration. Comme les fonctions en escalier sont mesurables, la proposition précédente montre qu'une limite simple de fonctions en escalier est mesurable. Réciproquement, si f est mesurable, on applique 2.5.5 à la suite des fonctions tronquées $f_n = f_{n,n}$. Comme ces fonctions sont intégrables, il existe une suite (g_n) de fonctions en escalier telle que $f_n - g_n$ tend vers 0 presque partout. Mais alors g_n tend vers f presque partout.

2.6.6 Remarque. Le critère précédent montre qu'une fonction $f : \mathbf{R} \rightarrow \mathbf{C}$ est mesurable si et seulement si sa partie réelle et sa partie imaginaire le sont. Cela permet de se limiter, la plupart du temps, au cas des fonctions à valeurs réelles.

2.6.7 Proposition.

- 1) La somme, le produit de deux fonctions mesurables sont des fonctions mesurables.
- 2) Si $g : \mathbf{C} \rightarrow \mathbf{C}$ est continue et $f : \mathbf{R} \rightarrow \mathbf{C}$ mesurable, $g \circ f$ est mesurable⁹.
- 3) Une fonction égale presque partout à une fonction mesurable est mesurable.
- 4) Si les f_n , $n \in \mathbf{N}$, à valeurs dans \mathbf{R} , sont mesurables, il en est de même de $\sup f_n$ et $\inf f_n$ (si ces fonctions sont finies).

Démonstration. Les points 1), 2) et 3) sont clairs à partir de 2.6.5. Enfin, le point 4) résulte, après troncature, de 2.5.13.

La notion de fonction mesurable permet d'énoncer un critère d'intégrabilité :

2.6.8 Proposition. Soit f une fonction mesurable. Alors f est intégrable si et seulement si il existe une fonction intégrable h telle que l'on ait $|f| \leq h$. En particulier, f est intégrable si et seulement si $|f|$ l'est.

Démonstration. Si f est intégrable, $|f|$ aussi et on prend $h = |f|$. Réciproquement, on considère les fonctions tronquées $f_n = f_{n,n}$ (elles sont intégrables puisque f est mesurable). La suite (f_n) converge simplement vers f et elle est majorée par h . Il en résulte que f est intégrable par convergence dominée.

2.6.9 Remarques.

- 1) Cette proposition n'est pas vraie pour l'intégrale de Riemann. Ainsi, la fonction caractéristique des rationnels de $[0, 1]$ n'est pas intégrable, bien que mesurable et majorée par 1.
- 2) La condition de mesurabilité est essentielle. Il suffit de considérer une partie non mesurable A contenue dans $[0, 1]$ (s'il en existe, cf. ci-dessous !) et la fonction f qui vaut 1 sur A , -1 sur $[0, 1] - A$ et 0 ailleurs.

⁹L'assertion est fautive si on suppose f continue et g mesurable.

On peut maintenant faire un bilan : comment montrer, finalement, qu'une fonction est Lebesgue-intégrable ? Le chemin peut être balisé ainsi :

- On montre que f est mesurable.

C'est généralement une formalité : c'est plutôt montrer qu'une fonction n'est pas mesurable qui est difficile.

- On montre que $|f|$ est intégrable (on se ramène donc au cas $f \geq 0$) et il suffit pour cela de montrer que f est majorée par une fonction intégrable. C'est évident si f est bornée et à support borné : elle est majorée par une fonction constante sur un segment et nulle en dehors. Il n'y a donc de problème que pour les fonctions qui vont à l'infini en un point $x \in \mathbf{R}$ et/ou ne sont pas nulles en dehors d'un segment. Dans les cas usuels, l'expérience montre qu'on finit toujours par se ramener au cas d'une intégrale de Riemann impropre absolument convergente.

2.6.3 Ensembles mesurables

Intégrale des fonctions mesurables positives

On peut généraliser aux fonctions mesurables positives la notion d'intégrale :

2.6.10 Définition. Soit f une fonction mesurable ≥ 0 . On définit son intégrale $\int_{\mathbf{R}} f$ comme :

- 1) l'intégrale de Lebesgue $\int_{\mathbf{R}} f$ si f est Lebesgue-intégrable,
- 2) $+\infty$ sinon.

Une fonction mesurable positive est donc intégrable si et seulement si son intégrale est finie.

Les propriétés de l'intégrale des fonctions mesurables sont les mêmes que celles de l'intégrale de Lebesgue (linéarité, croissance, etc.), à ceci près que certaines intégrales peuvent être infinies et que l'on adapte alors les règles de calcul : on a ainsi $a + (+\infty) = +\infty$, $a \times +\infty = +\infty$ pour $a > 0$, $0 \times +\infty = 0$. On a aussi un théorème de convergence croissante : si une suite croissante (f_n) de fonctions mesurables converge presque partout vers f , $\int f_n$ tend vers $\int f$.

Ensembles mesurables

On note A^c le complémentaire d'une partie A .

2.6.11 Définition. Soit A une partie de \mathbf{R} . On dit que A est **mesurable** si sa fonction caractéristique χ_A est mesurable et la mesure de A est alors la quantité (finie ou infinie) $\lambda(A) = \int_{\mathbf{R}} \chi_A(x) dx$. L'application λ qui à un

ensemble mesurable associe sa mesure est appelée la mesure de Lebesgue sur \mathbf{R} .

2.6.12 Proposition.

Une union dénombrable (resp. une intersection dénombrable) de parties mesurables est mesurable. Le complémentaire d'un ensemble mesurable est mesurable.

Démonstration. Si $A = \bigcup_{n \in \mathbf{N}} A_n$ (resp. $A = \bigcap_{n \in \mathbf{N}} A_n$), on a $\chi_A = \sup_{n \in \mathbf{N}} \chi_{A_n}$ (resp. $\chi_A = \inf_{n \in \mathbf{N}} \chi_{A_n}$) et on conclut en utilisant 2.6.7.4. Pour le complémentaire on utilise la formule $\chi_{A^c} = 1 - \chi_A$.

2.6.13 Théorème. La mesure de Lebesgue vérifie les axiomes d'une mesure :

1) $\lambda(\emptyset) = 0$; $\lambda(A \cup B) = \lambda(A) + \lambda(B)$ si A et B sont mesurables disjoints; $\lambda(A) \leq \lambda(B)$ si A, B sont mesurables avec $A \subset B$ ("le tout est plus grand que la partie" aurait dit Euclide).

2) Plus généralement, si (A_n) est une suite de parties mesurables disjointes, on a $\lambda\left(\bigcup_{n \in \mathbf{N}} A_n\right) = \sum_{n \in \mathbf{N}} \lambda(A_n)$.

Démonstration. Le point 1) résulte de la formule $\chi_{A \cup B} = \chi_A + \chi_B$ pour A, B disjointes. L'inégalité $\lambda(A) \leq \lambda(B)$ vient de l'écriture $B = A \cup (B - A)$ où l'on a posé $B - A = B \cap A^c$. Montrons le point 2). Posons $B_n = A_0 \cup A_1 \cup \dots \cup A_n$. On a $\lambda(B_n) = \sum_{i=0}^n \lambda(A_i)$. La fonction χ_A est la limite simple croissante des χ_{B_n} . Il y a deux cas. Si les $\lambda(B_n) = \int_{\mathbf{R}} \chi_{B_n}$ sont bornées, on peut appliquer le théorème de convergence monotone et on a $\lim \lambda(B_n) = \sum_{n=0}^{+\infty} \lambda(A_n) = \int_{\mathbf{R}} \chi_A = \lambda(A)$. Sinon, la suite croissante $\lambda(B_n)$ tend vers $+\infty$. On a donc $\sum_{n=0}^{+\infty} \lambda(A_n) = +\infty$. Mais, comme A contient B_n , on a $\lambda(A) \geq \lambda(B_n)$ pour tout n , donc $\lambda(A) = +\infty$ et le résultat s'ensuit.

2.6.14 Corollaire.

1) Si (A_n) est une suite de parties mesurables, on a $\lambda\left(\bigcup_{n \in \mathbf{N}} A_n\right) \leq \sum_{n \in \mathbf{N}} \lambda(A_n)$.

2) Si (A_n) est une suite croissante de parties mesurables, on a $\lambda\left(\bigcup_{n \in \mathbf{N}} A_n\right) =$

$\lim \lambda(A_n)$.

Démonstration. 1) On pose $B_n = A_n - (A_0 \cup A_1 \cup \dots \cup A_{n-1})$. Les B_n sont mesurables, disjoints et on a $A = \bigcup_n A_n = \bigcup_n B_n$, d'où $\lambda(A) = \sum_n \lambda(B_n)$ et on conclut avec $\lambda(B_n) \leq \lambda(A_n)$.

2) On pose $B_0 = A_0$ et, pour $n \geq 1$, $B_n = A_n - A_{n-1}$. Les B_n sont mesurables, disjoints, on a $A_n = B_0 \cup \dots \cup B_n$ et $A = \bigcup_{n \in \mathbf{N}} A_n = \bigcup_{n \in \mathbf{N}} B_n$, d'où $\lambda(A_n) = \sum_{k=0}^n \lambda(B_k)$, $\lambda(A) = \sum_{k=0}^{+\infty} \lambda(B_k)$ et donc $\lambda(A) = \lim \lambda(A_n)$.

2.6.15 Proposition.

- 1) Un intervalle $I = (a, b)$ avec $a \leq b$ est mesurable et de mesure (éventuellement infinie) $b - a$. On a donc $\lambda(I) = l(I)$ (cf. 1.5.1).
- 2) Un ouvert, un fermé (a fortiori un compact), sont mesurables.
- 3) Un sous-ensemble A de \mathbf{R} est négligeable si et seulement si il est mesurable et de mesure nulle. Plus généralement, si un ensemble B diffère¹⁰ d'un mesurable A par un ensemble de mesure nulle, il est mesurable et on a $\lambda(A) = \lambda(B)$.
- 4) L'image réciproque d'un ouvert ou d'un fermé par une application mesurable est mesurable.

Démonstration. 1) La fonction caractéristique d'un intervalle borné est une fonction en escalier. Elle est donc mesurable par 2.6.3.2. De plus, son intégrale est bien égale à $b - a$. Le cas d'un intervalle non borné I en résulte en écrivant I comme réunion de ses traces sur les intervalles $[n, n + 1[$ pour $n \in \mathbf{Z}$ et en utilisant 2.6.12 et 2.6.13.

2) Tout ouvert U est réunion dénombrable d'intervalles ouverts (il suffit de prendre tous les intervalles $]r, s[$ à extrémités rationnelles contenus dans U) et on conclut encore avec 2.6.12.

Passons à 3). Si A est négligeable, on a vu en 2.4.3.1 que χ_A est intégrable (donc mesurable) et d'intégrale nulle (donc A est de mesure nulle). Réciproquement, soit A un ensemble mesurable et de mesure nulle. La fonction χ_A est mesurable positive et d'intégrale nulle, donc intégrable, et on a $\int \chi_A = 0 = \|\chi_A\|_1$. On conclut par la proposition 2.5.6.

Si A et B sont "presque égaux", c'est-à-dire si l'on a $A = (A \cap B) \cup N$ et $B = (A \cap B) \cup M$ avec M, N négligeables, on note d'abord que $A \cap B$ est mesurable (c'est $A \cap N^c$), donc aussi B , puisque M est mesurable. Comme N et M sont de mesure nulle on a $\mu(A) = \mu(A \cap B) = \mu(B)$.

Pour 4) on se ramène au cas d'un intervalle $A = [a, b]$. La fonction χ_A est alors limite simple des fonctions continues χ_n , affines par morceaux, qui valent 1 sur A et sont nulles à l'extérieur de $[a - 1/n, b + 1/n]$. Il en résulte que $\chi_{f^{-1}(A)} = \chi_A \circ f$ est limite des $\chi_n \circ f$, qui sont mesurables par 2.6.7.2.

2.6.16 Remarque. Il n'est pas évident d'exhiber des ensembles non mesurables dans \mathbf{R} . Il n'y en a que si l'on admet un axiome de la théorie des ensembles appelé axiome du choix.

¹⁰C'est-à-dire si $A - (A \cap B)$ et $B - (A \cap B)$ sont tous deux négligeables.

2.6.17 Proposition. (Inégalité de Bienaymé-Tchebychev)

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction mesurable et posons, pour $k \in \mathbf{R}$, $E_k = \{x \in \mathbf{R} \mid f(x) \geq k\}$. Alors, E_k est mesurable. De plus, si f est ≥ 0 et $k > 0$ on a l'inégalité : $\lambda(E_k) \leq \frac{1}{k} \int_{\mathbf{R}} f(x)dx$.

Démonstration. C'est l'aire du rectangle ! L'ensemble E_k est mesurable comme $f^{-1}([k, +\infty[$. La fonction f est minorée par $k\chi_{E_k}$. On a donc bien $\int_{\mathbf{R}} f \geq k \int_{\mathbf{R}} \chi_{E_k} = k\lambda(E_k)$.

2.6.18 Remarque. Si f est une fonction mesurable ≥ 0 et si on a $\int_{\mathbf{R}} f = 0$, alors f est nulle presque partout. En effet, elle est intégrable et on peut donc appliquer la proposition 2.5.6. On peut aussi appliquer Bienaymé-Tchebychev. On a ainsi $\lambda(E_k) = 0$ pour tout $k > 0$. Mais, si E est l'ensemble des points où f est > 0 , on a $E = \bigcup_{n \in \mathbf{N}^*} E_{\frac{1}{n}}$, de sorte que E est bien de mesure nulle en vertu de 2.6.14.2.

Chapitre 3

Séries de Fourier

Dans ce chapitre, nous étudions les séries de Fourier, qui sont l'un des outils essentiels pour étudier les fonctions périodiques, donc, notamment, tout ce qui modélise les mouvements vibratoires de la physique. Il va sans dire que nous n'épuisons pas ce vaste sujet. Ce chapitre a une forte connotation géométrique, par l'usage qui est fait du produit scalaire dans l'espace \mathcal{L}^2 .

3.1 L'espace \mathcal{L}^2

3.1.1 Définition. Soit I un intervalle de \mathbf{R} et f une fonction de I dans \mathbf{C} . On dit que f est de carré intégrable si f est mesurable et si $|f|^2$ est intégrable. On note $\mathcal{L}^2(I)$ l'ensemble des fonctions de carré intégrable sur I .

3.1.2 Proposition-Définition. L'ensemble $\mathcal{L}^2(I)$ est un \mathbf{C} -espace vectoriel, stable par conjugaison. Si f, g sont dans $\mathcal{L}^2(I)$ leur produit fg est intégrable (donc aussi $f\bar{g}$) et on a l'inégalité de Cauchy-Schwarz :

$$\left| \int_I f(x)\bar{g}(x) dx \right| \leq \left(\int_I |f(x)|^2 dx \right)^{1/2} \left(\int_I |g(x)|^2 dx \right)^{1/2}.$$

De plus, l'égalité a lieu si et seulement si il existe $\lambda, \mu \in \mathbf{C}$, non tous deux nuls, tels que la fonction $\lambda f + \mu g$ soit nulle presque partout.

On pose (provisoirement) :

$$(f|g) = \int_I f(x)\bar{g}(x)dx \quad \text{et} \quad \|f\|_2 = \left(\int_I |f(x)|^2 dx \right)^{1/2} = \sqrt{(f|f)}.$$

On a donc l'inégalité : $|(f|g)| \leq \|f\|_2 \|g\|_2$.

Démonstration. La stabilité par conjugaison est évidente, car on a $|\overline{f}| = |f|$. Montrons que $f + g$ est dans \mathcal{L}^2 si f et g y sont¹. Bien entendu, cette fonction est mesurable. Ensuite, on a $|f + g|^2 \leq (|f| + |g|)^2 = |f|^2 + |g|^2 + 2|f||g| \leq 2(|f|^2 + |g|^2)$ car on a $2|f||g| \leq |f|^2 + |g|^2$. Il en résulte que $|f + g|^2$ est intégrable (elle est mesurable et majorée par une fonction intégrable) et cela vaut aussi pour fg .

Pour montrer Cauchy-Schwarz, il faut faire un peu d'algèbre. L'application Φ qui à (f, g) associe $(f|g)$ est ce qu'on appelle une forme hermitienne positive. Cela signifie que Φ est \mathbf{C} -linéaire en f , antilinéaire en g (c'est-à-dire qu'on a $\Phi(f, \lambda g) = \overline{\lambda}\Phi(f, g)$), qu'on a $(g|f) = \overline{(f|g)}$ et qu'enfin le réel $(f|f)$ est ≥ 0). C'est presque ce qu'on appelle un **produit scalaire**. La seule chose qui manque est le caractère **défini** positif, c'est-à-dire le fait que $(f|f)$ n'est nul que si f est nulle. Dans le cas présent on peut seulement conclure que f est nulle presque partout. L'inégalité de Cauchy-Schwarz résulte alors du lemme suivant :

3.1.3 Lemme. *Soit E un \mathbf{C} -espace vectoriel et $(x|y)$ une forme hermitienne positive sur E . Alors on a l'inégalité de Cauchy-Schwarz : $|(x|y)|^2 \leq (x|x)(y|y)$. Si l'on pose $\|x\| = \sqrt{(x|x)}$, cette inégalité devient $|(x|y)| \leq \|x\| \|y\|$. Il y a égalité si et seulement si il existe $\lambda, \mu \in \mathbf{C}$ non tous deux nuls avec $(\lambda x + \mu y|\lambda x + \mu y) = 0$.*

Démonstration. Pour $\lambda \in \mathbf{C}$, on calcule $p = (x + \lambda y|x + \lambda y) = (x|x) + \lambda(y|x) + \overline{\lambda}(x|y) + |\lambda|^2(y|y)$. Cette quantité est ≥ 0 pour tout λ . Posons $(x|y) = \rho e^{i\theta}$ et choisissons $\lambda = r e^{i\theta}$, avec r réel. Un calcul immédiat donne $p = (y|y)r^2 + 2\rho r + (x|x)$ et cette quantité est positive ou nulle pour tout r . Si $(y|y)$ est nul, cela impose que ρ est nul et l'inégalité cherchée est une égalité. Sinon, c'est que le discriminant du polynôme du second degré en r est ≤ 0 et on en déduit le résultat.

Le cas d'égalité correspond, soit à $(y|y) = 0$, soit à l'existence d'une racine double r de l'équation $(y|y)r^2 + 2\rho r + (x|x) = 0$. Mais alors, si on pose $\lambda = r e^{i\theta}$ (où θ est encore l'argument de $(x|y)$), on a $(x + \lambda y|x + \lambda y) = 0$ comme annoncé.

La conclusion, dans le cas des fonctions, vient de 2.5.6. En effet, si on a $(\lambda f + \mu g|\lambda f + \mu g) = \int |\lambda f + \mu g|^2 = 0$, la fonction $\lambda f + \mu g$ est nulle presque partout.

3.1.4 Remarque. Attention, il n'y a pas en général d'implication entre les deux propriétés : f intégrable et f de carré intégrable. Par exemple, sur $[1, +\infty[$ la fonction $1/x$ n'est pas intégrable, mais elle est de carré intégrable,

¹Le cas de λf est trivial et laissé au lecteur.

tandis que sur $]0, 1]$, $1/\sqrt{x}$ est intégrable, mais pas son carré. Toutefois, sur un intervalle I de longueur $\lambda(I)$ finie, f de carré intégrable implique f intégrable en vertu de l'inégalité de Schwarz : $\|f\|_1 = (\int |f| \chi_I) \leq \|f\|_2 \sqrt{\lambda(I)}$. Bien entendu, une fonction continue (ou plus généralement une fonction mesurable et bornée) sur un segment $[a, b]$ est de carré intégrable.

3.1.5 Corollaire. (Inégalité de Minkowski)

Soient $f, g \in \mathcal{L}^2(I)$. On a $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$.

Démonstration. En élevant au carré il s'agit de montrer qu'on a :

$$(f + g|f + g) \leq (f|f) + (g|g) + 2\|f\|_2 \|g\|_2.$$

Or on a $(f + g|f + g) = (f|f) + (f|g) + (g|f) + (g|g)$ et on conclut avec Cauchy-Schwarz.

3.1.6 Remarque. Le corollaire montre que la quantité $\|f\|_2$ mérite cette notation, c'est-à-dire qu'elle est une semi-norme sur l'espace $\mathcal{L}^2(I)$. Ce n'est pas vraiment une norme car on peut avoir $\|f\|_2 = 0$ sans que f soit nulle (elle est seulement nulle presque partout, cf. chapitre 2 2.5.6). Nous dirons tout de même parfois norme, par abus de langage.

3.1.7 Définition. La "norme" $\|f\|_2$ est appelée norme de la convergence en moyenne quadratique ou norme L^2 et on parlera de convergence au sens de L^2 à son propos.

3.1.8 Remarque. Là encore, il n'y a pas de rapport, en général, entre convergence en moyenne et en moyenne quadratique, sauf si I est de longueur finie, auquel cas on a l'inégalité : $\|f\|_1 \leq \|f\|_2 \sqrt{\lambda(I)}$ qui montre que la convergence en moyenne quadratique implique la convergence en moyenne.

3.2 Fischer-Riesz ou la complétude de $\mathcal{L}^2(I)$

3.2.1 Théorème. (de complétude ou de Fischer-Riesz) L'espace $\mathcal{L}^2(I)$ est complet pour la semi-norme $\|\cdot\|_2$. Cela signifie que, si (f_n) ($n \in \mathbf{N}^*$) est une suite de Cauchy au sens de L^2 , il existe une fonction f , de carré intégrable, telle que $\|f - f_n\|_2$ tende vers 0 quand n tend vers l'infini. De plus, il existe alors une sous-suite (f_{n_k}) qui converge presque partout vers f .

Démonstration. Écrivons la condition de Cauchy. On a donc :

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall p, q \in \mathbf{N}, (p, q \geq N \implies \|f_p - f_q\|_2 \leq \epsilon).$$

On peut extraire de la suite f_n une sous-suite f_{n_k} qui vérifie :

$$\|f_{n_{k+1}} - f_{n_k}\|_2 \leq 2^{-k}.$$

Il suffit maintenant de montrer que la suite (f_{n_k}) converge vers une fonction f à la fois au sens de L^2 et au sens de la convergence simple presque partout. En effet, comme (f_n) est de Cauchy au sens L^2 , si une de ses sous-suites converge vers f , elle converge elle aussi vers f .

Cela permet de se ramener au cas où la suite (f_n) elle-même vérifie l'inégalité ci-dessus : $\|f_{n+1} - f_n\|_2 \leq 2^{-n}$. L'idée de la démonstration (comme dans Egorov) est de transformer la suite (f_n) en série. Pour cela on pose $f_0 = 0$ et on considère les fonctions différences $\delta_n(x) = f_{n+1}(x) - f_n(x)$, pour $n \in \mathbf{N}$. On a la formule $f_n(x) = \sum_{k=0}^{n-1} \delta_k(x)$ et l'inégalité $\|\delta_n\|_2 \leq 2^{-n}$. La convergence (au moins ponctuelle) de la suite (f_n) revient donc à celle de la série (δ_n) . On sait que pour qu'une série converge, il suffit qu'elle converge absolument. Cela nous amène à considérer les fonctions $S_n(x) = \sum_{k=0}^{n-1} |\delta_k(x)| = \sum_{k \leq n-1} |f_{k+1}(x) - f_k(x)|$. Le gros avantage de la suite S_n c'est qu'il s'agit d'une suite croissante, qui est donc justiciable de Beppo-Levi. En fait, comme on travaille avec des fonctions de carré intégrable, ce n'est pas S_n qu'il faut regarder, mais $T_n = S_n^2$. Il est clair que c'est encore une suite croissante. Comme les f_n sont dans l'espace vectoriel \mathcal{L}^2 , il en est de même des δ_n , donc de leurs modules, donc aussi de S_n . Cela montre que les fonctions T_n sont intégrables et on a $\int_I T_n = \|S_n\|_2^2$. Mais, par l'inégalité triangulaire,

on a $\|S_n\|_2 \leq \sum_{k=0}^{n-1} \|\delta_k\|_2 \leq 2$. On a donc $\int_I T_n \leq 2$. En vertu de Beppo-Levi,

la suite (T_n) converge presque partout vers une fonction intégrable T et la suite des intégrales $\int T_n$ converge vers $\int T$. En prenant la racine carrée, il en résulte que S_n converge presque partout vers $S = \sqrt{T}$, de carré intégrable. Cela implique que la série $\delta_n(x)$ converge (absolument) presque partout vers une fonction f , ou encore que la suite f_n converge presque partout vers f .

Comme f est limite simple de fonctions mesurables elle est mesurable. Par ailleurs, on a $|f_n(x)| \leq S_n(x) \leq S(x)$, d'où $|f_n(x)|^2 \leq S(x)^2 = T(x)$. À la limite on en déduit $|f(x)|^2 \leq T(x)$, et comme T est intégrable, f est de carré intégrable. Enfin, il reste à voir que f_n converge vers f au sens L^2 . Cela signifie que l'intégrale $\int_I |f_n - f|^2$ tend vers 0. Mais les fonctions $|f_n - f|^2$ tendent vers 0 presque partout et elles sont majorées par $|f_n|^2 + |f|^2 + 2|f||f_n|$ donc par $4T$ qui est intégrable. On conclut grâce au théorème de convergence dominée.

3.3 Séries de Fourier : la théorie \mathcal{L}^2

3.3.1 Introduction

Les séries de Fourier sont un moyen essentiel pour étudier les fonctions périodiques. On sait que ces fonctions jouent un rôle capital en physique dans l'étude de tous les phénomènes vibratoires. Ces séries ont été introduites par Joseph Fourier en 1822². Le principe c'est de ramener l'étude d'une fonction, disons de période 2π , à celle des fonctions périodiques prototypes : $\cos x$, $\sin x$ et les fonctions qui s'en déduisent : $\cos nx$, $\sin nx$ ou, mieux, à leurs variantes complexes, e^{inx} . On va donc chercher à écrire une fonction de période 2π comme somme d'une série :

$$f(x) = \sum_{n \in \mathbf{Z}} c_n e^{inx} \quad \text{ou} \quad f(x) = \sum_{n \geq 0} a_n \cos nx + \sum_{n \geq 1} b_n \sin nx,$$

la nature de la convergence étant à préciser.

3.3.2 Coefficients de Fourier

On s'intéresse aux fonctions de période $T > 0$ et on pose $\omega = \frac{2\pi}{T}$, de sorte que les fonctions $\cos n\omega x$, $\sin n\omega x$ et $e^{in\omega x}$ sont de période T . Il suffit d'étudier ces fonctions sur $[0, T]$. L'espace $\mathcal{L}^2 = \mathcal{L}^2([0, T])$ est muni du produit scalaire défini par $(f|g) = \int_0^T f(t)\overline{g(t)}dt$. L'exemple de la géométrie ordinaire nous enseigne que les bases orthogonales sont un outil essentiel dans ce type de situation. Or, ici, on a la propriété suivante :

3.3.1 Proposition. 1) Les fonctions $e_n(x) = e^{in\omega x}$, pour $n \in \mathbf{Z}$, forment une famille orthogonale de \mathcal{L}^2 . De plus, on a $(e_n|e_n) = T$.
2) Les fonctions $\cos n\omega x$ ($n \in \mathbf{N}$) et $\sin n\omega x$ ($n \in \mathbf{N}^*$) forment aussi une famille orthogonale de \mathcal{L}^2 .

Démonstration. 1) Il s'agit de calculer $(e_p|e_q) = \int_0^T e^{i\omega(p-q)x} dx$. Si p et q sont distincts, une primitive de $e^{i\omega(p-q)x}$ est $F(x) = \frac{e^{i\omega(p-q)x}}{\omega(p-q)}$ et on a $F(0) = F(T) = 1$, de sorte que l'intégrale est nulle. Si on a $p = q$ l'intégrale vaut T .

2) Le résultat est immédiat en écrivant cosinus et sinus en fonction des exponentielles ou en utilisant les formules de trigonométrie.

La valeur du carré scalaire nous conduit à modifier la définition de $(f|g)$:

²Théorie analytique de la chaleur, voir Fourier Œuvres, pp. 201-235. Un premier mémoire sur le sujet avait été proposé à l'Académie en 1811.

3.3.2 Définition. On définit le produit scalaire $(f|g)$ sur $\mathcal{L}^2([0, T])$ par la formule :

$$(f|g) = \frac{1}{T} \int_0^T f(t)\bar{g}(t)dt.$$

La famille (e_n) est alors une famille orthonormale³.

3.3.3 Remarque. Rappelons que si f est intégrable et de période T on a les deux formules : $\int_a^b f(x)dx = \int_{a+T}^{b+T} f(u)du$ (poser $x = u + T$) et $\int_0^T f(x)dx = \int_a^{a+T} f(x)dx$ pour tout $a \in \mathbf{R}$ (écrire $\int_0^T = \int_0^a + \int_a^{a+T} + \int_{a+T}^T$).

Les calculs fondamentaux

Les calculs qui suivent sont exactement copiés sur ceux que l'on pratique dans l'espace euclidien ordinaire. On sait en effet, que pour écrire un vecteur f sur la famille orthonormale (e_n) , on doit calculer les produits scalaires $(f|e_n)$:

3.3.4 Proposition. Soit f un polynôme trigonométrique, c'est-à-dire une combinaison linéaire finie des fonctions e_n , $f = \sum_{p \in I} c_p e_p$ (où I désigne une partie finie de \mathbf{Z}). On convient de poser $c_p = 0$ si p n'est pas dans I . Alors on a $(f|e_n) = c_n$ pour tout $n \in \mathbf{Z}$.

Démonstration. Cela résulte de 3.3.1.

La définition suivante coule alors de source :

3.3.5 Définition. Soit f une fonction intégrable sur $[0, T]$ et soit $n \in \mathbf{Z}$. On appelle n -ième coefficient de Fourier de f la quantité :

$$c_n = \hat{f}(n) = (f|e_n) = \frac{1}{T} \int_0^T f(x)e^{-in\omega x} dx.$$

On notera que la fonction $f(x)e^{-in\omega x}$ est bien intégrable (elle est mesurable et son module est le même que celui de f).

La proposition 3.3.4 montre alors que, dans le cas d'un polynôme trigonométrique, les coefficients de Fourier sont les coefficients du polynôme : si on a $f = \sum_{p \in I} c_p e_p = \sum_{p \in I} c_p e^{ip\omega x}$, avec I fini, les coefficients de Fourier $\hat{f}(p)$ de f sont égaux à c_p pour $p \in I$ et à 0 sinon.

³En revanche, les fonctions $\cos n\omega x$ et $\sin n\omega x$ pour $n \in \mathbf{N}^*$ sont de norme $1/\sqrt{2}$.

3.3.6 Corollaire. Les fonctions $e^{in\omega x}$ (pour $n \in \mathbf{Z}$) sont linéairement indépendantes sur \mathbf{C} . Cela signifie que si la fonction $f(x) = \sum_{n=-N}^N c_n e^{in\omega x}$ est identiquement nulle (ou simplement nulle presque partout), alors tous les c_n sont nuls.

Démonstration. En effet, c_n est le n -ième coefficient de Fourier de f et il se calcule donc par l'intégrale de 3.3.5. Comme f est nulle presque partout, on a bien $c_n = 0$ pour tout n .

3.3.7 Remarque. En fait, plus généralement, si une série trigonométrique $S_N(x) = \sum_{n=-N}^N c_n e^{in\omega x}$ converge presque partout vers une fonction f et si les fonctions $|S_N|$ sont majorées, indépendamment de N , par une fonction intégrable g , on a $c_n = \widehat{f}(n)$ en vertu du théorème de convergence dominée.

3.3.8 Corollaire. (Parseval pour les polynômes trigonométriques)

Si $f = \sum_{p \in I} c_p e_p$ est un polynôme trigonométrique, on a la formule $\|f\|_2^2 = \sum_{p \in I} |c_p|^2 = \sum_{p \in I} |\widehat{f}(p)|^2$.

Démonstration. C'est la semi-linéarité du produit scalaire :

$$\|f\|_2^2 = (f|f) = (f | \sum_{p \in I} c_p e_p) = \sum_{p \in I} \bar{c}_p (f|e_p) = \sum_{p \in I} \bar{c}_p c_p = \sum_{p \in I} |c_p|^2.$$

Les variantes réelles

Un polynôme trigonométrique peut s'écrire soit avec les exponentielles complexes, soit avec les fonctions cosinus et sinus. Le lien entre les deux est donné par les formules : $e^{in\omega x} = \cos n\omega x + i \sin n\omega x$ et

$$\cos n\omega x = \frac{1}{2}(e^{in\omega x} + e^{-in\omega x}), \quad \sin n\omega x = \frac{1}{2i}(e^{in\omega x} - e^{-in\omega x}).$$

On en déduit la proposition suivante :

3.3.9 Proposition. Soit f un polynôme trigonométrique que l'on écrit sous les deux formes suivantes :

$$f(x) = a_0 + \sum_{n=1}^N (a_n \cos n\omega x + b_n \sin n\omega x) = \sum_{n=-N}^N c_n e^{in\omega x}.$$

On a les relations : $a_0 = c_0$, et, pour $n \geq 1$, $c_n = \frac{1}{2}(a_n - ib_n)$, $c_{-n} = \frac{1}{2}(a_n + ib_n)$, $a_n = c_n + c_{-n}$, $b_n = i(c_n - c_{-n})$.

Démonstration. En remplaçant les sinus et cosinus par leurs expressions en termes d'exponentielles, on peut identifier les coefficients en vertu du corollaire 3.3.6 et on obtient les expressions ci-dessus.

3.3.10 Remarque. Les fonctions $\cos n\omega x$ et $\sin n\omega x$ sont elles aussi linéairement indépendantes sur \mathbf{R} , autrement dit, si la fonction $\sum_{n=0}^N a_n \cos n\omega x + \sum_{n=1}^N b_n \sin n\omega x$ est identiquement nulle, alors les coefficients a_n et b_n sont nuls. Plus généralement, dans un espace vectoriel muni d'un produit scalaire, toute famille orthogonale est libre.

3.3.11 Définition. Si f est une fonction périodique de période T on définit ses coefficients de Fourier "réels"⁴ :

$$a_0 = \frac{1}{T} \int_0^T f(t) dt, \quad a_n = \frac{2}{T} \int_0^T f(t) \cos n\omega t dt, \quad b_n = \frac{2}{T} \int_0^T f(t) \sin n\omega t dt \quad (\text{pour } n \in \mathbf{N}^*)$$

3.3.12 Remarque. On notera que si f est paire (resp. impaire), les coefficients b_n (resp. a_n) sont nuls.

3.3.13 Corollaire. Si f est un polynôme trigonométrique :

$$f(x) = a_0 + \sum_{n=1}^N (a_n \cos n\omega x + b_n \sin n\omega x),$$

les coefficients a_n et b_n sont les coefficients de Fourier réels de f .

Démonstration. Cela résulte du théorème analogue pour les c_n .

On montre aussitôt la variante réelle de Parseval pour les polynômes trigonométriques :

3.3.14 Corollaire. Avec les notations de 3.3.13, on a :

$$\|f\|_2^2 = |a_0|^2 + \frac{1}{2} \sum_{n=1}^N |a_n|^2 + \frac{1}{2} \sum_{n=1}^N |b_n|^2.$$

⁴L'appellation n'est justifiée que pour les fonctions à valeurs réelles.

3.3.3 Fischer-Riesz pour les séries de Fourier

Notre objectif, pour une fonction f intégrable sur $[0, T]$, est donc d'étudier la convergence de la série $\sum_{n \in \mathbf{Z}} \widehat{f}(n) e^{in\omega x}$. Bien entendu, comme les coefficients de Fourier ne changent pas si on modifie f sur un ensemble de mesure nulle, on ne peut pas espérer que la série converge mieux que presque partout. Si on suppose seulement f intégrable, c'est faux⁵ (c'est un résultat de Kolmogorov, 1926). Nous allons d'abord étudier le cas de la convergence au sens de la norme L^2 , en faisant usage le plus possible de l'outil produit scalaire. Pour la convergence presque partout, d'autres hypothèses nous seront nécessaires.

Notons d'abord que, si f est périodique de période T , et si l'on pose $g(x) = f\left(\frac{Tx}{2\pi}\right)$, la fonction g est périodique de période 2π . Quitte à faire ce changement de variables, on peut supposer $T = 2\pi$ et donc $\omega = 1$. C'est ce que nous ferons désormais.

3.3.15 Notation. On désigne par f une fonction de \mathbf{R} dans \mathbf{C} , 2π -périodique, intégrable sur $[0, 2\pi]$. On note $\widehat{f}(n)$ son coefficient de Fourier d'indice $n \in \mathbf{Z}$ (défini car f est intégrable). On a donc, pour $n \in \mathbf{Z}$, $c_n = \widehat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt$. On pose $S_N f(x) = \sum_{n=-N}^{n=N} \widehat{f}(n) e^{inx}$. On a aussi les formules suivantes pour les coefficients de Fourier réels (pour $n \in \mathbf{N}^*$) :

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(t) dt, \quad a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt dt, \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt dt.$$

Le premier théorème essentiel que nous allons montrer est le suivant :

3.3.16 Théorème. *Si f est dans $\mathcal{L}^2([0, 2\pi])$, la suite $S_N f$ converge vers f en moyenne quadratique et on a l'égalité de Parseval :*

$$\|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = \sum_{n=-\infty}^{n=+\infty} |\widehat{f}(n)|^2 = |a_0|^2 + \frac{1}{2} \sum_{n=1}^{+\infty} |a_n|^2 + \frac{1}{2} \sum_{n=1}^{+\infty} |b_n|^2.$$

Démonstration. Elle comporte trois étapes. L'étape 1 est essentiellement de nature algébrique, avec des manipulations de produits scalaires. L'étape 2 contient le cœur du théorème, qui consiste à faire le calcul dans le cas particulier de la fonction caractéristique d'un intervalle. Enfin, l'étape 3 est un théorème d'approximation des fonctions de L^2 par les fonctions en escalier.

⁵Mais la divergence n'est pas grossière. En effet, on montre facilement, en approchant f par des fonctions en escalier, que les coefficients de Fourier $\widehat{f}(n)$ tendent vers 0 quand n tend vers $\pm \infty$ (exercice).

3.3.17 Corollaire. Si f est dans $\mathcal{L}^2([0, 2\pi])$ et si ses coefficients de Fourier sont tous nuls, f est nulle presque partout.

Démonstration. En effet, on a $\|f\|_2 = 0$ d'après Parseval.

3.3.18 Remarque. On peut montrer que le corollaire précédent vaut encore si f est dans \mathcal{L}^1 seulement.

3.3.19 Remarque. Attention, le théorème ci-dessus n'implique pas *a priori* que la série de Fourier $S_N f(x)$ converge vers $f(x)$ même presque partout, mais seulement que c'est vrai pour une sous-suite. Cependant, le résultat est vrai, mais c'est l'un des plus difficiles de toute l'analyse (théorème de Carleson, 1966).

3.3.4 Première étape : Parseval

Un lemme

L'idée de base, en pensant qu'en définitive on a $f = \sum_{n \in \mathbf{Z}} \widehat{f}(n) e_n$, est de décomposer f en deux morceaux, l'un qui correspond aux e_n d'indices vérifiant $|n| \leq N$, et qui n'est autre que $S_N f$, et l'autre qui correspond aux indices plus grands, et qui est $f - S_N f$. Comme la famille (e_n) est orthogonale, on espère bien que ces morceaux sont orthogonaux. Le lemme suivant montre que c'est bien le cas :

3.3.20 Lemme. Avec les notations 3.3.15 on a, pour tout N , les formules :

- 0) $\|S_N f\|_2^2 = \sum_{n=-N}^N |\widehat{f}(n)|^2$.
- 1) $(f - S_N f | S_N f) = 0$.
- 2) $\|f\|_2^2 = \|f - S_N f\|_2^2 + \|S_N f\|_2^2$.

Démonstration. Le point 0) a déjà été vu (cf. 3.3.8). Le point 1) vient de 0) et de la "linéarité" du produit scalaire :

$$(f | S_N f) = (f | \sum_{n=-N}^{n=N} \widehat{f}(n) e^{inx}) = \sum_{n=-N}^{n=N} \overline{\widehat{f}(n)} (f | e^{inx}) = \sum_{n=-N}^{n=N} \overline{\widehat{f}(n)} \widehat{f}(n).$$

Le point 2) n'est autre le théorème de Pythagore : on calcule le carré scalaire de f en écrivant $f = (f - S_N f) + S_N f$. On a donc $\|f\|_2^2 = \|f - S_N f\|_2^2 + \|S_N f\|_2^2 + (S_N f | f - S_N f) + (f - S_N f | S_N f)$. Mais, en vertu de 1), $f - S_N f$ et $S_N f$ sont orthogonaux d'où le résultat.

Et ses conséquences

3.3.21 Corollaire. La série $\sum_{n=-\infty}^{+\infty} |\widehat{f}(n)|^2$ converge et sa somme est $\leq \|f\|_2^2$.

Démonstration. En effet, les sommes partielles sont égales à $\|S_N f\|_2^2$ en vertu de 1), donc majorées par $\|f\|_2^2$ en vertu de 2).

3.3.22 Corollaire. L'application $f \mapsto S_N f$ est une application linéaire de \mathcal{L}^2 dans \mathcal{L}^2 qui vérifie $\|S_N f\|_2 \leq \|f\|_2$ (donc est continue).

Démonstration. C'est le point 2 du lemme.

3.3.23 Corollaire. La suite $(S_N f)$ est de Cauchy au sens L^2 (donc elle converge vers une fonction $g \in L^2$ en vertu de Fischer-Riesz 3.2.1).

Démonstration. En effet, on a, pour $p \leq q$:

$$\|S_q f - S_p f\|_2^2 = \sum_{n=-q}^{-p-1} |\widehat{f}(n)|^2 + \sum_{n=p+1}^q |\widehat{f}(n)|^2$$

(toujours le Parseval des polynômes) et, comme la série des $|\widehat{f}(n)|^2$ converge, elle est de Cauchy et les sommes ci-dessus tendent vers 0.

3.3.24 Commentaire. Attention, la difficulté c'est qu'*a priori*, on ne sait pas que la fonction g est égale à f . Cependant, comme on est sur un intervalle borné, la fonction g est dans \mathcal{L}^1 et la convergence a lieu aussi dans \mathcal{L}^1 . En particulier, cela implique qu'il y a une sous-suite (f_{n_k}) qui converge presque partout vers g . On voit que, pour prouver le théorème pour une fonction f il suffit de montrer que $S_N f$ converge presque partout vers f . En effet, cela montre qu'on a $f = g$ presque partout, d'où la convergence au sens de L^2 de $S_N f$ vers f . Cela implique que $\|S_N f\|_2$ tend vers $\|f\|_2$, donc aussi les carrés, ce qui donne Parseval.

3.3.5 Le théorème pour une fonction en escalier

Les coefficients de Fourier d'une fonction constante

Dans ce paragraphe et le suivant, f est la fonction caractéristique de l'intervalle $[0, a]$, avec $0 < a < 2\pi$.

3.3.25 Lemme. On a les formules :

$$0) \widehat{f}(0) = \frac{a}{2\pi},$$

$$1) \widehat{f}(n) = \frac{1 - e^{-ina}}{2i\pi n}, \text{ pour } n \neq 0.$$

Démonstration. C'est un calcul évident.

On en déduit le résultat suivant :

3.3.26 Corollaire. Pour $x \in \mathbf{R}$, on pose $\sigma_N(x) = \sum_{n=1}^N \frac{\sin nx}{n}$. On a alors :

$$S_N f(x) = \frac{a}{2\pi} + \frac{1}{\pi} \sigma_N(x) - \frac{1}{\pi} \sigma_N(x - a).$$

Le calcul fondamental

On cherche à calculer la limite de $S_N f(x)$ quand N tend vers $+\infty$. On commence par calculer celle de $\sigma_N(x)$. Pour cela on pose $\tau_N(x) = \sum_{n=1}^N \frac{e^{inx}}{n}$ et on a $\sigma_N(x) = \text{Im}(\tau_N(x))$. Le lemme fondamental est le suivant :

3.3.27 Lemme. Soit $x \in]0, 2\pi[$. On a :

$$\lim_{n \rightarrow +\infty} \tau_N(x) = \sum_{n=1}^{+\infty} \frac{e^{inx}}{n} = -\ln\left(2 \sin \frac{x}{2}\right) + i\left(\frac{\pi - x}{2}\right).$$

La limite de $\sigma_N(x)$ quand N tend vers $+\infty$ est égale à $\frac{\pi - x}{2}$.

Démonstration. La preuve qui suit n'est totalement convaincante que si l'on connaît un peu la théorie des fonctions de variable complexe. Voir en annexe le lemme 3.5.1 pour une preuve directe. Nous admettrons qu'on peut, sous certaines conditions, définir le logarithme d'un nombre complexe et que celui-ci continue à vérifier deux propriétés : il est une fonction réciproque de l'exponentielle, il vérifie l'équation fonctionnelle $\ln(zw) = \ln z + \ln w$. Comme tout nombre complexe z s'écrit $z = re^{i\theta}$, cela donne aussitôt, pour $z \neq 0$, $\ln z = \ln r + i\theta$, soit $\ln z = \ln |z| + i \text{Arg } z$. La difficulté c'est qu'il faut que l'argument soit bien défini, alors qu'il ne l'est, *a priori*, que modulo 2π . Pour éviter cette difficulté, on se contentera de regarder le logarithme d'un complexe en dehors de l'axe réel négatif. L'argument peut alors être choisi dans $] - \pi, \pi[$.

On part de la relation bien connue :

$$\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n.$$

Cette relation vaut pour z de module < 1 . On “intègre” cette relation et on obtient :

$$-\ln(1-z) = \sum_{n=1}^{+\infty} \frac{z^n}{n},$$

pour les mêmes z , mais aussi pour $z = e^{ix}$ avec $x \in]0, 2\pi[$ (on notera que la série converge : pour $|z| < 1$ c'est le critère de D'Alembert, pour $z = e^{ix}$ c'est la transformation d'Abel). On en déduit que la somme cherchée vaut $-\ln|1 - e^{ix}| - i\text{Arg}(1 - e^{ix})$. Un calcul trigonométrique immédiat donne :

$$1 - e^{ix} = 2 \sin \frac{x}{2} \left(\sin \frac{x}{2} - i \cos \frac{x}{2} \right) = 2 \sin \frac{x}{2} \left(\cos\left(\frac{x}{2} - \frac{\pi}{2}\right) + i \sin\left(\frac{x}{2} - \frac{\pi}{2}\right) \right).$$

Comme x est dans $]0, 2\pi[$, $x/2$ est dans $]0, \pi[$, de sorte que $2 \sin \frac{x}{2}$ est positif, donc c'est le module de la quantité considérée et $\frac{x}{2} - \frac{\pi}{2}$ en est l'argument (ce nombre est dans $] - \pi/2, \pi/2[$). On a bien le résultat.

3.3.28 Proposition. *Soit $x \in]0, 2\pi[$, $x \neq a$. La limite de $S_N f(x)$ quand N tend vers l'infini est égale à $f(x)$, c'est-à-dire à 1 pour $x < a$ et à 0 pour $x > a$.*

Démonstration. On écrit $S_N f(x) = \frac{a}{2\pi} + \frac{1}{\pi} \sigma_N(x) - \frac{1}{\pi} \sigma_N(x-a)$. Si x est $> a$, x et $x-a$ sont dans $]0, 2\pi[$ et la limite vaut $\frac{a}{2\pi} + \frac{\pi-x}{2\pi} - \frac{\pi-(x-a)}{2\pi} = 0$. Si x est $< a$, comme $\sigma_N(x)$ est de période 2π , on a $\sigma_N(x-a) = \sigma_N(2\pi+x-a)$ et, cette fois, $2\pi+x-a$ est dans $]0, 2\pi[$. La limite est donc égale à

$$\frac{a}{2\pi} + \frac{\pi-x}{2\pi} - \frac{\pi-(2\pi+x-a)}{2\pi} = 1.$$

3.3.29 Remarque. On notera que le résultat n'est vrai ni en a ni en 0. En effet, on a $f(0) = f(a) = 1$ (voire 0 si on avait pris la fonction caractéristique de l'intervalle ouvert ou semi-ouvert), tandis que $S_N f(0)$ et $S_N f(a)$ ont toutes deux pour limite $1/2$ (cf. 3.4.6).

Le résultat pour une fonction en escalier

3.3.30 Proposition. Soit f une fonction en escalier définie sur $[0, 2\pi[$ et prolongée en une fonction 2π -périodique sur \mathbf{R} . Alors, la suite $(S_N f(x))$ converge presque partout vers $f(x)$. Il en résulte qu'elle converge aussi vers f au sens de L^2 (cf. 3.3.24).

Démonstration. On écrit f comme combinaison linéaire de fonctions caractéristiques d'intervalles $J_i = [a_i, a_{i+1}[$, qui sont différences des fonctions caractéristiques de $[0, a_{i+1}[$ et de $[0, a_i[$, donc du type précédent. On utilise alors la linéarité de $f \mapsto S_N f$ et les calculs précédents. On notera que la convergence n'a lieu que presque partout à cause des discontinuités en 0 et a_i . Précisément, en a_i on vérifie aisément que $S_N f$ tend vers la moyenne des valeurs de f sur les intervalles qui entourent a_i , cf. 3.4.6.

3.3.6 Le résultat d'approximation

Cette dernière étape concerne exclusivement la convergence L^2 (et plus la convergence simple).

3.3.31 Lemme. Soient $f \in \mathcal{L}^2([0, 2\pi])$ et $\epsilon > 0$. Il existe une fonction en escalier h définie sur $[0, 2\pi]$ qui vérifie $\|f - h\|_2 \leq 2\epsilon$.

Démonstration. L'idée est simple : on sait approcher une fonction f par une fonction h en escalier au sens de L^1 (voir Chapitre 2, 2.4.5). Bien entendu, même sur un intervalle borné, l'approximation au sens de L^1 n'implique pas qu'elle vaut au sens de L^2 , mais si les fonctions f et h sont bornées par M on a :

$$\|f - h\|_2^2 = \int_0^{2\pi} |f(t) - h(t)|^2 dt \leq 2M \int_0^{2\pi} |f(t) - h(t)| dt = 2M \|f - h\|_1.$$

Or, on peut approcher, au sens de L^2 , une fonction par une fonction bornée :

3.3.32 Lemme. Soient $f \in \mathcal{L}^2([0, 2\pi])$ et $\epsilon > 0$. Il existe une fonction $k \in \mathcal{L}^2([0, 2\pi])$, bornée, qui vérifie $\|f - k\|_2 \leq \epsilon$. La fonction k est aussi dans $\mathcal{L}^1([0, 2\pi])$.

Démonstration. On considère les fonctions f_n obtenues en tronquant $|f|$ par $n \in \mathbf{N}$: $f_n(x) = f(x)$ si $|f(x)| \leq n$ et $f_n(x) = 0$ sinon. Elles sont dans \mathcal{L}^2 , donc mesurables, bornées, donc aussi dans \mathcal{L}^1 , et la suite (f_n) converge vers f au sens de L^2 . En effet, il s'agit de voir que la suite $\|f - f_n\|_2^2 =$

$\int_0^{2\pi} |f(x) - f_n(x)|^2 dx$ converge vers 0. Or, il y a convergence simple et on a $|f(x) - f_n(x)|^2 \leq |f(x)|^2$, d'où la conclusion par convergence dominée. On peut alors prendre pour k une fonction f_n pour n assez grand.

On peut maintenant finir 3.3.31. On approche f à ϵ près au sens de L^2 par une fonction bornée k comme ci-dessus. Si $|k|$ est bornée par M , on approche ensuite k par une fonction en escalier h_0 au sens de L^1 à $\epsilon/2M$ près : $\|k - h_0\|_1 \leq \epsilon/2M$ (cf. 2.4.5). Attention, *a priori*, h_0 n'est pas bornée par M . On considère alors la fonction h qui est égale à h_0 si $|h_0(x)| \leq M$, et à $M \frac{h_0(x)}{|h_0(x)|}$ sinon. La fonction h est encore en escalier et on vérifie qu'on a $|k - h| \leq |k - h_0|$, d'où $\|k - h\|_1 \leq \epsilon/2M$. Comme les fonctions k et h sont bornées par M on en déduit qu'on a $\|k - h\|_2 \leq \epsilon/2$, donc $\|f - h\|_2 \leq \epsilon$ comme annoncé.

3.3.7 La fin du théorème

Soit f une fonction de \mathcal{L}^2 et soit $\epsilon > 0$. Il s'agit de montrer qu'on a $\|f - S_N f\|_2 \leq \epsilon$ pour N assez grand. Pour cela on approche f par une fonction en escalier h à $\epsilon/3$ au sens de L^2 (cf. 3.3.31). On sait que $S_N h$ converge vers h dans L^2 . Pour N grand, on peut donc majorer $\|h - S_N h\|_2$ par $\epsilon/3$. Enfin, on a l'inégalité : $\|S_N f - S_N h\|_2 = \|S_N(f - h)\|_2 \leq \|f - h\|_2$ (cf. 3.3.22). On conclut en écrivant :

$$\|f - S_N f\|_2 \leq \|f - h\|_2 + \|h - S_N h\|_2 + \|S_N h - S_N f\|_2.$$

3.3.8 Applications

Voici deux conséquences immédiates du théorème dans le cas de la fonction $(\pi - x)/2$.

3.3.33 Proposition. *On a les formules :*

$$(1) \quad \frac{\pi}{4} = \sum_{p=0}^{+\infty} \frac{(-1)^p}{2p+1}, \quad (2) \quad \frac{\pi^2}{6} = \sum_{n=1}^{+\infty} \frac{1}{n^2}.$$

Démonstration. Pour montrer (1) on utilise la formule : $\sum_{n=1}^{+\infty} \frac{\sin nx}{n} = \frac{\pi - x}{2}$, appliquée en $x = \pi/2$. Pour montrer (2) on considère la fonction f périodique

de période 2π , qui vaut $(\pi - x)/2$ sur $]0, 2\pi[$ et 0 en 0. On constate que f est impaire, de sorte que ses coefficients de Fourier a_n sont nuls. Une intégration par parties permet de montrer qu'on a $b_n = 1/n$ pour $n > 0$. Comme f est dans \mathcal{L}^2 , on peut appliquer Parseval (variante réelle) qui donne :

$$\|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \frac{(\pi - x)^2}{4} dx = \frac{\pi^2}{12} = \frac{1}{2} \sum_{n=1}^{+\infty} \frac{1}{n^2},$$

d'où le résultat.

3.3.34 Remarque. La première formule est le développement en série de Taylor de $\text{Arctan } 1$ en 0, la seconde la formule classique qui donne la valeur au point 2 de la fonction zeta de Riemann : $\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$.

3.3.9 Une réciproque

Le résultat suivant montre que les suites de carré sommable correspondent toutes à des séries de Fourier de fonctions de \mathcal{L}^2 :

3.3.35 Théorème. Soit $(c_n)_{n \in \mathbf{Z}}$ une suite de nombres complexes tels que $\sum_{n=-\infty}^{+\infty} |c_n|^2 < +\infty$. Alors, il existe une fonction $f \in \mathcal{L}^2$, unique à égalité presque partout près, qui vérifie $\widehat{f}(n) = c_n$.

Démonstration. On pose $f_N = \sum_{n=-N}^N c_n e_n$. La suite f_N est de Cauchy au sens L^2 (même argument qu'en 3.3.23), donc elle converge vers une fonction f au sens L^2 en vertu de Fischer-Riesz (cf. 3.2.1). De plus, on a $\widehat{f}(n) = c_n$. En effet, on a $\widehat{f_N}(n) = c_n$ pour $-N \leq n \leq N$ et l'application $f \mapsto \widehat{f}(n)$ est linéaire continue (c'est un produit scalaire).

Si g est une autre fonction de \mathcal{L}^2 vérifiant $\widehat{g}(n) = c_n$, on a $\widehat{f-g}(n) = 0$ donc f et g sont égales presque partout en vertu de 3.3.17.

3.4 Le théorème de Dirichlet

C'est le deuxième théorème essentiel de la théorie. Si le théorème 3.3.16 assure la convergence au sens L^2 (avec Parseval comme sous-produit essentiel), on a vu qu'il ne dit rien sur la convergence presque partout. C'est l'objet du théorème de Dirichlet (mais avec une hypothèse de régularité nettement plus forte sur la fonction). Nous ne démontrerons ici que la version faible de ce résultat.

3.4.1 La version faible

Rappelons d'abord une définition :

3.4.1 Définition. Soit $f : [a, b] \rightarrow \mathbf{C}$ une fonction. On dit que f est de classe C^1 par morceaux s'il existe une subdivision $s_0 = a < s_1 < \dots < s_r = b$ de $[a, b]$ telle que, sur chaque intervalle ouvert $]s_i, s_{i+1}[$ pour $i = 0, \dots, r-1$, f soit la restriction d'une fonction de classe C^1 sur $]s_i, s_{i+1}[$.

3.4.2 Remarque. Attention, une telle fonction n'est pas nécessairement continue en les points de subdivision, mais ses discontinuités sont seulement "de première espèce", ce qui signifie que f admet en chaque point x une limite à gauche $f(x^-)$ et une limite à droite $f(x^+)$.

3.4.3 Théorème. Soit $f : \mathbf{R} \rightarrow \mathbf{C}$ une fonction continue 2π -périodique.

1) On suppose que la série de terme général $\widehat{f}(n)$ converge absolument, c'est-à-dire qu'on a $\sum_{n=-\infty}^{+\infty} |\widehat{f}(n)| < +\infty$. Alors, la série de Fourier de f converge

uniformément vers f : on a $f(x) = \sum_{n=-\infty}^{+\infty} \widehat{f}(n) e^{inx}$.

2) La condition est réalisée, en particulier, si la fonction f est de classe C^1 par morceaux.

Démonstration. Il est clair que la série de Fourier de f converge uniformément (et même normalement) vers une fonction continue g et qu'on a $\widehat{f}(n) = \widehat{g}(n)$, de sorte que l'on a $f = g$ presque partout. Mais comme ces fonctions sont continues, on a $f = g$. (Par exemple parce que l'intégrale de $|f - g|$ est nulle.)

Pour le point 2) la convergence des coefficients repose sur un calcul facile :

3.4.4 Lemme. Si $f : \mathbf{R} \rightarrow \mathbf{C}$ est continue, 2π -périodique, et de classe C^1 par morceaux sur $[0, 2\pi]$, on a, pour tout $n \in \mathbf{Z}$, $\widehat{f}'(n) = in\widehat{f}(n)$.

Démonstration. Le lecteur se convaincra que la formule d'intégration par parties est encore valable si l'on suppose seulement les fonctions de classe C^1 par morceaux. On a $\widehat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt$. On intègre par parties en

posant $u = f(t)$, $dv = e^{-int} dt$, d'où $du = f'(t) dt$ et $v = \frac{e^{-int}}{-in}$ et donc

$$\widehat{f}(n) = \frac{-1}{2\pi in} [f(t) e^{-int}]_0^{2\pi} + \frac{1}{2\pi in} \int_0^{2\pi} f'(t) e^{-int} dt.$$

Comme f est de période 2π la partie tout intégrée est nulle et on a le résultat.

On déduit de ce lemme qu'on a, pour $n \neq 0$, $|\widehat{f}(n)| = \frac{|\widehat{f}'(n)|}{n}$. Mais, pour des séries à termes positifs, on a $2a_n b_n \leq a_n^2 + b_n^2$, de sorte que, comme les séries $1/n^2$ et $|\widehat{f}'(n)|^2$ convergent (pour f' c'est Parseval!), il en est de même de $\frac{|\widehat{f}'(n)|}{n}$, donc de $|\widehat{f}(n)|$.

3.4.5 Remarque. Si f est plus régulière que C^1 cela se traduit par une convergence plus rapide des coefficients de Fourier vers 0 (si f est C^d ils sont majorés par une série en $1/n^d$). Réciproquement, on montre aisément que si l'on a une majoration des coefficients de Fourier de la forme $|\widehat{f}(n)| \leq A/n^d$, la fonction f est de classe au moins C^{d-2} .

3.4.2 Application : encore $\zeta(2)$

On considère la fonction f , périodique de période 2π , définie par $f(x) = \pi^2 - x^2$ pour $x \in [-\pi, +\pi]$. On vérifie que f est continue et paire⁶, de sorte que ses coefficients de Fourier b_n sont nuls. Un calcul immédiat donne $a_0 = \frac{2\pi^2}{3}$ et une double intégration par parties fournit les autres coefficients : $a_n = (-1)^{n+1} \frac{4}{n^2}$. On constate que la série des coefficients de Fourier est en $1/n^2$, donc converge absolument. On peut appliquer le théorème et on a : $\pi^2 - x^2 = \sum_{n=0}^{+\infty} a_n \cos nx$. On applique cette formule en $x = \pi$. On a $f(\pi) = 0$ et $\cos n\pi = (-1)^n$, ce qui donne $a_0 = \sum_{n=1}^{+\infty} \frac{4}{n^2}$, d'où le résultat.

Une autre méthode consiste à considérer le développement de $\pi^2 - x^2$ en $x = 0$.

3.4.3 La version forte

Nous admettrons le théorème suivant qui généralise 3.4.3 au cas non continu :

3.4.6 Théorème. *Soit $f : \mathbf{R} \rightarrow \mathbf{C}$ une fonction 2π -périodique de classe C^1 par morceaux (on rappelle qu'en chaque point, f admet alors une limite à gauche et une limite à droite, pas nécessairement égales). Alors, la série de Fourier converge presque partout vers f , précisément, on a, pour tout $x \in \mathbf{R}$:*

$$\lim_{N \rightarrow +\infty} S_N f(x) = \frac{f(x^+) + f(x^-)}{2}.$$

⁶Comme la fonction f est de classe C^1 sur $] -\pi, \pi[$, on pourrait aussi appliquer la version forte 3.4.6 du théorème.

3.4.7 Remarque. Le théorème de Dirichlet vaut aussi si f est réglée avec des dérivées à gauche et à droite en chaque point.

3.5 Annexe : la preuve du calcul de 3.3.27

Dans cette annexe nous prouvons le lemme suivant, qui est la partie de 3.3.27 qui nous est utile :

3.5.1 Lemme. Soit $x \in]0, 2\pi[$. On a la formule :

$$\sum_{n=1}^{+\infty} \frac{\sin nx}{n} = \frac{\pi}{2} - \frac{x}{2}.$$

Démonstration. Il suffit de montrer la formule pour $x \in]0, \pi[$. En effet, si x est dans $]\pi, 2\pi[$, on se ramène à l'autre cas en posant $x = 2\pi - u$. Comme la formule est évidente pour $x = \pi$ on peut supposer $x \in]0, \pi[$. On considère l'intégrale $I = \int_0^1 \frac{e^{ix}}{1 - e^{ix}t} dt$. Pour calculer cette intégrale, on développe en série la fonction à intégrer, pour $0 \leq t < 1$:

$$\frac{1}{1 - e^{ix}t} = \sum_{n=0}^{+\infty} e^{inx}t^n.$$

Considérons la somme partielle $S_{N,x}(t) = \sum_{n=0}^N e^{inx}t^n = \frac{1 - e^{i(N+1)x}t^{N+1}}{1 - e^{ix}t}$.

On a, en module, $|S_{N,x}(t)| \leq \frac{2}{|1 - e^{ix}t|} = h_x(t)$. Mais, comme x est différent de 0 modulo 2π , la fonction en dénominateur ne s'annule pas, de sorte que $h_x(t)$ est une fonction continue, donc intégrable sur $[0, 1]$.

On peut donc appliquer le théorème de convergence dominée et on obtient :

$$I = \sum_{n=0}^{+\infty} \int_0^1 e^{i(n+1)x} t^n dt = \sum_{n=0}^{+\infty} \frac{e^{i(n+1)x}}{n+1}.$$

On en déduit que la partie imaginaire de I est égale à

$$J = \sum_{n=1}^{+\infty} \frac{\sin nx}{n}$$

qui est la quantité cherchée.

Un calcul immédiat montre que la partie imaginaire de $\frac{e^{ix}}{1 - e^{ix}t}$ est $\frac{\sin x}{1 + t^2 - 2t \cos x}$.

On en déduit $J = \int_0^1 \frac{\sin x}{1 + t^2 - 2t \cos x} dt$ et on écrit cette intégrale sous la forme $J = \int_0^1 \frac{dt}{\sin x \left[1 + \left(\frac{t - \cos x}{\sin x}\right)^2\right]}$ que l'on calcule par changement de variables en posant $u = \frac{t - \cos x}{\sin x}$ (on notera que $\sin x$ est non nul). On a $J = [\text{Arctan } u]_{-\cotan x}^{\frac{1 - \cos x}{\sin x}}$. Un petit calcul trigonométrique donne $\frac{1 - \cos x}{\sin x} = \tan \frac{x}{2}$. Or, on a $\text{Arctan}(\tan \frac{x}{2}) = \frac{x}{2}$ (car $x/2$ est dans $]0, \pi/2[$) et $\text{Arctan}(-\cotan x) = \text{Arctan}(\tan(x - \frac{\pi}{2})) = x - \frac{\pi}{2}$ (car $x - \pi/2$ est dans $] -\pi/2, \pi/2[$). En définitive, on a bien $J = \frac{\pi}{2} - \frac{x}{2}$ comme annoncé.

Chapitre 4

L'intégrale de Lebesgue dans \mathbf{R}^d

Comme annoncé dans l'introduction générale, nous changeons d'approche pour l'intégration à plusieurs variables, revenant à la méthode originelle de Lebesgue. Cependant, une partie importante du travail a été faite. En effet, nous avons défini sur \mathbf{R} la mesure de Lebesgue, qui associe à une partie mesurable. A sa mesure $\lambda(A)$ (qui est un réel ≥ 0 ou $+\infty$) et nous allons utiliser cette notion et l'étendre à \mathbf{R}^d (en admettant toutefois l'existence de la mesure) et construire l'intégrale de Lebesgue à partir de cette mesure.

Dans tout ce qui suit, on désigne par d un entier ≥ 1 . Le complémentaire d'un ensemble A est noté A^c . Si A et B sont deux parties de \mathbf{R}^d (sans relation d'inclusion *a priori*), on pose $A - B = A \cap B^c$.

4.1 La mesure de Lebesgue

4.1.1 Ensembles élémentaires

4.1.1 Définition. *On appelle ensemble élémentaire une partie $A \subset \mathbf{R}^d$ de la forme $A_1 \times \dots \times A_d$ où les A_i sont des parties mesurables de \mathbf{R} . On appelle pavé une partie P de \mathbf{R}^d de la forme $I_1 \times I_2 \times \dots \times I_d$ où les I_k sont des intervalles bornés de \mathbf{R} .*

Notons que si A est un ensemble élémentaire non vide, les A_i sont déterminés de manière unique : ce sont les projections de A sur les différents facteurs.

4.1.2 Remarque. On notera que, si A et B sont élémentaires, il en est de même de $A \cap B$. En effet, si $A = A_1 \times \dots \times A_d$ et $B = B_1 \times \dots \times B_d$, on a $A \cap B = (A_1 \cap B_1) \times \dots \times (A_d \cap B_d)$ et on sait que les $A_i \cap B_i$ sont mesurables dans \mathbf{R}

(cf. 2.6.12). La propriété est encore vraie pour une intersection dénombrable d'ensembles élémentaires car on a la formule :

$$\bigcap_i (A_{1i} \times \cdots \times A_{di}) = \left(\bigcap_i A_{1i} \right) \times \cdots \times \left(\bigcap_i A_{di} \right)$$

et une intersection dénombrable d'ensembles mesurables est mesurable (*loc. cit.*). En revanche, l'union de deux ensembles élémentaires n'en est pas un en général. De même, le complémentaire d'un ensemble élémentaire n'est pas élémentaire, mais il est réunion finie disjointe d'ensembles élémentaires. Précisément on a :

$$(A_1 \times \cdots \times A_d)^c = \bigcup_{\epsilon} A_1^{\epsilon_1} \times \cdots \times A_d^{\epsilon_d},$$

où les symboles ϵ_i sont, soit une absence de symbole, soit le symbole c , l'union étant étendue à toutes les distributions $\epsilon = (\epsilon_1, \dots, \epsilon_d)$ où l'un au moins des ϵ_i vaut c .

4.1.3 Définition. Soit $A = A_1 \times \dots \times A_d$ un ensemble élémentaire. On définit sa mesure de Lebesgue $\mu(A) \in [0, +\infty]$ comme le produit $\lambda(A_1) \cdots \lambda(A_d)$, où λ désigne la mesure de Lebesgue sur \mathbf{R} , avec les conventions suivantes :

- 1) si l'un des $\lambda(A_i)$ est nul on a $\mu(A) = 0$ (même si certains des $\lambda(A_j)$ sont infinis),
- 2) si tous les $\lambda(A_i)$ sont non nuls, et si l'un d'eux est infini, $\mu(A)$ est égal à $+\infty$.

4.1.4 Remarque. La justification de la convention 1) est la suivante. Considérons par exemple un ensemble de la forme $A \times \mathbf{R} \subset \mathbf{R}^2$, avec A négligeable. On peut écrire \mathbf{R} comme réunion des intervalles $I_k = [-k, k]$, pour $k \in \mathbf{N}$. L'ensemble $A \times \mathbf{R}$ est alors la réunion dénombrable des ensembles $E_k = A \times I_k$. Les E_k sont des ensembles élémentaires de mesure $\lambda(A) \times \lambda(I_k) = 0 \times 2k = 0$. La convention est donc nécessaire pour qu'une réunion dénombrable d'ensembles de mesure nulle soit encore de mesure nulle.

4.1.5 Remarque. L'usage du symbole $+\infty$ est donc le suivant : on a $a + (+\infty) = +\infty$, $a \times (+\infty) = +\infty$ si a est > 0 et, différence avec l'usage habituel, $0 \times (+\infty) = 0$.

4.1.2 Ensembles négligeables

4.1.6 Définition. On dit qu'une partie A de \mathbf{R}^d est négligeable si pour tout $\epsilon > 0$ il existe une suite d'ensembles élémentaires E_n tels que :

- 1) A est contenu dans la réunion des E_n ,
- 2) la somme des mesures des E_n (qui est la somme d'une série à termes positifs) est $\leq \epsilon$.

4.1.7 Proposition.

- 1) Un sous-ensemble d'un ensemble négligeable est négligeable.
- 2) Une union dénombrable de négligeables est négligeable.

Démonstration. Le point 1) est évident et le point 2) se prouve comme en dimension 1).

4.1.8 Exemples.

1) Avec la convention 1) de 4.1.3, il est clair qu'un ensemble élémentaire $A_1 \times \cdots \times A_d$ est négligeable dès que l'un des A_i l'est. La réciproque viendra de 4.1.19.2.

2) En particulier, l'hyperplan H_k de coordonnées défini par $x_k = a_k$ est négligeable. En effet, on a $H_k = \mathbf{R}^{k-1} \times \{a_k\} \times \mathbf{R}^{d-k}$. Avec la formule de changement de variables 4.5.2 on en déduit que tout hyperplan est négligeable et donc aussi toute partie contenue dans un hyperplan. Un bon exercice est de montrer directement que la droite $y = x$ est négligeable dans \mathbf{R}^2 .

4.1.9 Remarque. On a maintenant une notion de propriété vraie presque partout : cela signifie que la propriété est vraie sauf sur un ensemble négligeable.

4.1.3 Ensembles mesurables

La définition des ensembles mesurables se fait de manière axiomatique (pas plus que sur \mathbf{R} on ne sait pas vraiment décrire les parties mesurables) :

4.1.10 Proposition-Définition. Il existe une plus petite partie \mathcal{M} de $\mathcal{P}(\mathbf{R}^d)$ qui contient les ensembles élémentaires et les ensembles négligeables et qui vérifie les propriétés suivantes :

- i) Si A est dans \mathcal{M} , son complémentaire A^c aussi.
- ii) Si $(A_n)_{n \in \mathbf{N}}$ est une suite de parties (non nécessairement disjointes) de \mathcal{M} , la réunion des A_n est dans \mathcal{M} .

Les éléments de \mathcal{M} sont appelées parties **mesurables** de \mathbf{R}^d .

Une partie \mathcal{M} de $\mathcal{P}(\mathbf{R}^d)$ qui vérifie les conditions i) et ii) est appelée une **tribu** ou une **σ -algèbre**.

Démonstration. Montrons l'existence de \mathcal{M} . On considère toutes les tribus \mathcal{T} de $\mathcal{P}(\mathbf{R}^d)$ qui contiennent à la fois les ensembles négligeables et les ensembles élémentaires. Il en existe, car $\mathcal{P}(\mathbf{R}^d)$ en est une. On prend alors pour \mathcal{M} l'intersection de toutes ces tribus. Il est clair que \mathcal{M} convient.

Propriétés

4.1.11 Définition. Deux ensembles $A, B \subset \mathbf{R}^d$ sont dits presque égaux si leur différence symétrique : $A\Delta B = (A\cup B) - (A\cap B) = (A-B)\cup(B-A) = (A - (A\cap B)) \cup (B - (A\cap B))$ est négligeable. On note $A \sim B$ dans ce cas.

4.1.12 Proposition.

- 1) Si les A_n sont dans \mathcal{M} leur intersection y est aussi.
- 2) Si A et B sont mesurables, $A - B = A - (A\cap B)$ l'est aussi.
- 3) Si A est mesurable et si B est presque égal à A , B est mesurable.
- 4) Si on a $A \subset E \subset B$ avec $A, B \in \mathcal{M}$ et $B - A$ négligeable, alors E est dans \mathcal{M} .

Démonstration. Le point 1) vient du fait que l'intersection est le complémentaire de l'union des complémentaires. Le point 2) est évident. Pour 3), on note d'abord que $A - (A\cap B)$ est négligeable, donc mesurable. Il en résulte que $A\cap B = A - (A - (A\cap B))$ est mesurable, puis que $B = (A\cap B)\cup(B - (A\cap B))$ l'est aussi. Enfin, 4) résulte de 3). En effet, on a $A \sim E$ car $E - A$ est négligeable (il est contenu dans $B - A$ qui est négligeable).

Parmi les parties mesurables, certaines sont appelées à jouer un rôle important :

4.1.13 Proposition-Définition. Toute réunion dénombrable d'ensembles élémentaires est mesurable. On note \mathcal{D} le sous-ensemble de $\mathcal{P}(\mathbf{R}^d)$ formé des réunions dénombrables d'ensembles élémentaires. L'ensemble \mathcal{D} est stable par union dénombrable et par intersection finie. Une partie de \mathbf{R}^d presque égale à un élément de \mathcal{D} est mesurable.

Démonstration. Seule l'assertion sur l'intersection n'est pas évidente. Si on a $A_1, \dots, A_r \in \mathcal{D}$, on écrit chaque A_i comme réunion dénombrable d'élémentaires : $A_i = \bigcup_{n \in \mathbf{N}} A_{i,n}$. Alors, l'intersection $A_1 \cap \dots \cap A_r$ est la réunion des $A_{1,n_1} \cap \dots \cap A_{r,n_r}$ pour tous les $(n_1, \dots, n_r) \in \mathbf{N}^r$: c'est bien une réunion dénombrable d'ensembles élémentaires.

4.1.14 Exemples.

- 1) Un ouvert, un fermé (et donc un compact) de \mathbf{R}^d sont mesurables. En effet, un ouvert U est réunion dénombrable de pavés ouverts qui sont des

ensembles élémentaires, de sorte que U est dans \mathcal{D} . Pour un fermé, il suffit de passer au complémentaire.

2) On notera qu'un ensemble aussi simple que le cercle unité Γ de \mathbf{R}^2 (qui est négligeable, cf. 4.4.6) n'est pas réunion dénombrable d'élémentaires. En effet, si on a $A \times B \subset \Gamma$, cela implique que A et B sont de cardinal ≤ 2 . Pour voir cela on fixe $a \in A$, on a $a \times B \subset \Gamma$, mais, comme il n'y a que deux points $(a, y) \in \Gamma$ au plus, cela montre que $|B|$ est ≤ 2 . De même pour A . Une union dénombrable d'élémentaires contenus dans Γ est donc dénombrable donc distincte de Γ .

3) On peut montrer qu'il existe des ensembles mesurables qui ne sont pas dans \mathcal{D} , ni même presque égaux à un ensemble de \mathcal{D} . Cependant, on verra en 4.1.21 que tout ensemble mesurable est égal, "à ϵ près", à un ensemble de \mathcal{D} .

4) Comme dans le cas de la dimension 1 il n'est pas évident d'exhiber des ensembles non mesurables. Il n'en existe que grâce à l'axiome du choix.

4.1.15 Remarque. Ce qui précède montre que \mathcal{D} n'est pas une tribu. En effet, elle contient tous les ouverts par 4.1.14.1, mais pas tous les fermés (cf. 4.1.14.2). Elle n'est donc pas stable par passage au complémentaire.

Le lemme suivant précise la structure des parties de \mathcal{D} :

4.1.16 Lemme. *Toute réunion dénombrable d'ensembles élémentaires est aussi réunion dénombrable d'ensembles élémentaires **disjoints**.*

Démonstration. Soit \mathcal{D}_0 l'ensemble des réunions dénombrables disjointes d'ensembles élémentaires. Il est clair que \mathcal{D}_0 est stable par réunion dénombrable disjointe. D'après 4.1.2, on sait que toute intersection dénombrable d'ensembles élémentaires en est un (donc est dans \mathcal{D}_0) et que le complémentaire d'un ensemble élémentaire est réunion disjointe finie d'ensembles élémentaires, donc est dans \mathcal{D}_0 .

Soit alors $A = \bigcup_{n \in \mathbf{N}} E_n$ où les E_n sont élémentaires. On pose, pour $n \in \mathbf{N}$: $F_n = E_n - \bigcup_{i < n} E_i$. On a $A = \bigcup_{n \in \mathbf{N}} F_n$. En effet, il est clair que l'union est contenue dans A car on a $F_n \subset E_n$. Elle est égale car si x est dans A , et si on appelle n le plus petit entier tel que $x \in E_n$, alors x est dans F_n . Cette union est disjointe car si x est dans $F_p \cap F_q$ avec $p < q$, il est dans E_p et E_q ce qui est absurde par définition de F_q . Enfin, les F_n sont dans \mathcal{D}_0 . En effet, ils sont intersections de E_n et des complémentaires des E_i , $i < n$ et comme les E_i^c sont union disjointes finies d'élémentaires, on vérifie que F_n est dans \mathcal{D}_0 .

Il en résulte que A est dans \mathcal{D}_0 comme annoncé.

4.1.4 La mesure

Nous **admettrons** l'existence de la mesure, c'est-à-dire le théorème suivant :

4.1.17 Théorème. *Il existe une unique application $\mu : \mathcal{M} \rightarrow \mathbf{R}^+ \cup \{+\infty\}$ qui vérifie les propriétés suivantes :*

i) *Si $E = A_1 \times \cdots \times A_d$ est un ensemble élémentaire, $\mu(E)$ est défini comme en 4.1.3 : $\mu(E) = \lambda(A_1) \cdots \lambda(A_d)$ où λ est la mesure de Lebesgue sur \mathbf{R} .*

ii) *La mesure μ est σ -additive : si les parties A_n sont **disjointes** on a :*

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

4.1.18 Remarque. L'usage du symbole $+\infty$ est celui vu en 4.1.5.

4.1.19 Corollaire.

1) *Si A, B sont mesurables et $A \subset B$ on a $\mu(A) \leq \mu(B)$ ("le tout est plus grand que la partie"). **Si $\mu(A)$ est fini** on a $\mu(B - A) = \mu(B) - \mu(A)$.*

2) *Si les parties A_n sont mesurables quelconques on a :*

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

3) *Si la suite de parties mesurables (A_n) est croissante on a*

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

4) *Si la suite de parties mesurables (A_n) est décroissante et si l'une des mesures $\mu(A_n)$ est finie, on a $\mu\left(\bigcap_{n \in \mathbf{N}} A_n\right) = \lim_n \mu(A_n)$.*

Démonstration. 1) On a $B = A \cup (B - A)$ et $B - A$ est mesurable, donc $\mu(B) = \mu(A) + \mu(B - A) \geq \mu(A)$. On en déduit $\mu(B - A) = \mu(B) - \mu(A)$ si $\mu(A)$ est fini (s'il est infini, on a une expression indéterminée de la forme $\infty - \infty$).

2) C'est trivial si l'une des parties est de mesure infinie. Sinon, on pose $B_n = A_n - \bigcup_{i < n} A_i$. On a $A = \bigcup_n A_n = \bigcup_n B_n$ mais les B_n sont disjoints. On a donc $\mu(A) = \sum_n \mu(B_n) \leq \sum_n \mu(A_n)$.

3) On pose $A = \bigcup_n A_n$. On a $A = A_0 \cup \bigcup_{i=0}^{\infty} (A_{i+1} - A_i)$ (union disjointe),

donc $\mu(A) = \mu(A_0) + \sum_{i=0}^{\infty} \mu(A_{i+1} - A_i)$ et, de même, $\mu(A_n) = \mu(A_0) +$

$\sum_{i=0}^{n-1} \mu(A_{i+1} - A_i)$, d'où le résultat.

4) On peut supposer $\mu(A_0)$ finie. On pose $A = \bigcap_{n \in \mathbf{N}} A_n$, $B_n = A_0 - A_n$ et $B = A_0 - A$. On a $B = \bigcup_n B_n$ et les B_n sont croissants, d'où $\mu(B) = \lim \mu(B_n)$ par le point 3). Mais, comme $\mu(A_0)$ est fini, on a $\mu(B_n) = \mu(A_0) - \mu(A_n)$ et $\mu(B) = \mu(A_0) - \mu(A)$, d'où le résultat.

4.1.20 Remarque. Attention, l'exemple des parties $A_n = [n, +\infty[$ de \mathbf{R} montre que l'hypothèse $\mu(A_n)$ finie dans 4) est essentielle.

Un théorème de structure

Le théorème suivant décrit les ensembles mesurables :

4.1.21 Théorème. *Soit A un ensemble mesurable. Pour tout $\epsilon > 0$, il existe un ensemble $B \in \mathcal{D}$ (union dénombrable d'ensembles élémentaires), contenant A , tel que l'on ait $\mu(B - A) \leq \epsilon$. Autrement dit, A est égal à un ensemble de \mathcal{D} , "à ϵ près".*

Démonstration. Nous allons montrer, plus précisément, que A vérifie la condition (*) suivante : Pour tout $\epsilon > 0$, il existe des ensembles $B, C \in \mathcal{D}$ tels que l'on ait : *i) $A \subset B$, ii) $B - A \subset C$, iii) $\mu(C) \leq \epsilon$.*

Appelons \mathcal{T} l'ensemble des parties de \mathbf{R}^d qui vérifient la condition (*). L'ensemble \mathcal{T} contient les ensembles élémentaires. En effet, si A est élémentaire, il suffit de prendre $B = A$ et $C = \emptyset$. Il contient aussi les ensembles négligeables. En effet, si A est négligeable, pour tout $\epsilon > 0$ il est contenu dans une réunion $\bigcup_n E_n$ avec les E_n élémentaires et $\sum_n \mu(E_n) \leq \epsilon$. On prend alors $B = C = \bigcup_n E_n$ et on a bien $\mu(C) \leq \sum_n \mu(E_n) \leq \epsilon$.

Nous allons montrer que \mathcal{T} est une tribu. Comme \mathcal{M} est la plus petite tribu contenant les ensembles élémentaires et les négligeables, cela montrera que \mathcal{M} est contenue dans \mathcal{T} , donc que tout ensemble mesurable est dans \mathcal{T} .

Montrons que \mathcal{T} est stable par union dénombrable. Soient $A_n \in \mathcal{T}$ (avec $n > 0$), $A = \bigcup_n A_n$ et soit $\epsilon > 0$. Par définition, il existe $B_n, C_n \in \mathcal{D}$ avec $A_n \subset B_n$, $B_n - A_n \subset C_n$ et $\mu(C_n) \leq \epsilon/2^n$. On considère alors $B = \bigcup B_n$, $C = \bigcup C_n$. Il est clair qu'on a $A \subset B$, $B - A \subset C$ et C qui est réunion dénombrable d'ensembles élémentaires est de mesure $\leq \sum_n \epsilon/2^n = \epsilon$.

Montrons ensuite qu'une partie A est dans \mathcal{T} si et seulement si, pour tout pavé (ou tout ensemble élémentaire borné) P , $A \cap P$ est dans \mathcal{T} . Si on a cette dernière propriété, on considère les traces de A sur les pavés $P_n = [-n, n]^d$. Elles sont dans \mathcal{T} , donc aussi A qui est leur réunion. Inversement, si A est dans \mathcal{T} et si P est un pavé, pour tout $\epsilon > 0$ il existe $B, C \in \mathcal{D}$ avec la propriété (*). On vérifie qu'on a aussi cette propriété pour $A \cap P$ en prenant les ensembles $B \cap P$ et $C \cap P$.

Montrons maintenant que si A est dans \mathcal{D} , son complémentaire A^c est dans \mathcal{T} . On peut écrire $A = \bigcup_n A_n$ comme union dénombrable d'ensembles élémentaires **disjoints** (cf. 4.1.16). On a $A^c = \bigcap_n A_n^c$. Soit P un pavé, il suffit de montrer que $A^c \cap P$ est dans \mathcal{T} , en vertu de la remarque précédente. Soit $\epsilon > 0$. On va montrer la propriété (*) pour $A^c \cap P$. Comme les A_n sont disjoints et que $\bigcup_n (A_n \cap P)$ est contenue dans P , on a $\mu(\bigcup_n (A_n \cap P)) = \sum_n \mu(A_n \cap P) \leq \mu(P)$, de sorte que la série $\sum_n \mu(A_n \cap P)$ est convergente. Il existe donc un entier N tel que l'on ait $\sum_{n>N} \mu(A_n \cap P) \leq \epsilon$. Alors, l'ensemble $B = \bigcap_{n \leq N} (A_n^c \cap P)$ convient pour montrer la propriété (*) pour

$A^c \cap P$. En effet, il est clair que $A^c \cap P$ est contenu dans B et B est dans \mathcal{D} comme intersection finie d'éléments de \mathcal{D} (cf. 4.1.13). De plus, si x est dans B et non dans $A^c \cap P$, c'est qu'il existe $n > N$ tel que $x \notin A_n^c$. Autrement dit, x est dans $A_n \cap P$. On peut donc prendre pour C l'ensemble $\bigcup_{n>N} (A_n \cap P)$, qui est bien dans \mathcal{D} et, par construction, on a $\mu(C) = \sum_{n>N} \mu(A_n \cap P) \leq \epsilon$.

Il reste à montrer que, si A est dans \mathcal{T} , il en est de même de A^c . On se donne $\epsilon > 0$. On sait qu'il existe $B, C \in \mathcal{D}$ avec $A \subset B$, $B - A \subset C$ et $\mu(C) \leq \epsilon$. On a donc $B^c \subset A^c$ et $A^c - B^c = B - A$. En vertu du paragraphe précédent, B^c est dans \mathcal{T} , de sorte qu'il existe $D, E \in \mathcal{D}$ avec $B^c \subset D$, $D - B^c \subset E$ et $\mu(E) \leq \epsilon$. Mais alors, on a $A^c \subset D \cup C \in \mathcal{D}$ et $(D \cup C) - A^c \subset (D \cup C) - B^c \subset C \cup (D - B^c) \subset C \cup E$ et il est clair que $\mu(C \cup E) \leq 2\epsilon$, ce qui achève la démonstration.

Le corollaire suivant sera notamment utile dans la preuve du théorème de Fubini :

4.1.22 Corollaire. *Soit A une partie mesurable. Il existe une suite de parties $B_n \in \mathcal{D}$, contenant A , telles que $\lim \mu(B_n - A) = 0$ et $\mu(A) = \lim \mu(B_n)$.*

Démonstration. On applique 4.1.21. Pour tout $n \in \mathbf{N}^*$ il existe $B_n \in \mathcal{D}$ avec $A \subset B_n$ et $\mu(B_n - A) \leq 1/n$. On a alors $\mu(A) \leq \mu(B_n) = \mu(A) + \mu(B_n - A) \leq \mu(A) + (1/n)$, d'où le résultat.

4.1.23 Proposition. *Un ensemble A est négligeable si et seulement si il est mesurable et de mesure nulle. Deux ensembles mesurables presque égaux sont de même mesure.*

Démonstration. Si A est négligeable, il est mesurable et pour tout $\epsilon > 0$, il est contenu dans une réunion d'ensembles élémentaires : $A \subset \bigcup_n E_n$ avec $\sum_n \mu(E_n) \leq \epsilon$. On a donc $\mu(A) \leq \sum_n \mu(E_n) \leq \epsilon$ (4.1.19.2). Comme cela vaut pour tout $\epsilon > 0$, $\mu(A)$ est nul. Si A et B sont presque égaux, ils sont aussi presque égaux à $A \cap B$ et les différences entre A (resp. B) et l'intersection

sont négligeables. On en déduit $\mu(A) = \mu(A \cap B) = \mu(B)$ par le premier point.

Réciproquement, si A est mesurable et de mesure nulle, en vertu de 4.1.21, il est contenu dans un ensemble $B \in \mathcal{D}$, avec $\mu(B) = \mu(B - A) + \mu(A) = \mu(B - A) \leq \epsilon$. De plus, en vertu de 4.1.16, on peut supposer que B est réunion disjointe d'ensembles élémentaires B_n , de sorte qu'on a $\sum_n \mu(B_n) = \mu(B) \leq \epsilon$ et A est négligeable, par définition.

4.1.5 Fonctions mesurables

La définition des fonctions mesurables est inspirée de l'une des caractérisations des fonctions mesurables sur \mathbf{R} :

4.1.24 Proposition-Définition. *Soit f une fonction de \mathbf{R}^d dans \mathbf{R} (resp. \mathbf{C}). On dit que la fonction f est mesurable si pour tout ouvert $U \subset \mathbf{R}$ (resp. \mathbf{C}), la partie $f^{-1}(U)$ est mesurable.*

Dans le cas réel, il suffit pour cela que, pour tous $a, b \in \mathbf{R}$, l'ensemble :

$$f^{-1}(]a, b[) = \{x \in \mathbf{R}^d \mid a < f(x) < b\}$$

soit mesurable, ou que, pour tout $a \in \mathbf{R}$, $f^{-1}(]a, +\infty[)$ le soit.

Dans le cas complexe, f est mesurable si et seulement si ses parties réelle et imaginaire le sont.

Démonstration. Si la propriété est vraie pour les ouverts, elle l'est, *a fortiori*, pour les intervalles $]a, b[$. La réciproque résulte du fait que tout ouvert de \mathbf{R} est réunion dénombrable d'intervalles ouverts. Pour passer du cas $]a, +\infty[$ au cas $]a, b[$ on utilise les formules :

$$]a, b[=]a, +\infty[- [b, +\infty[\quad \text{et} \quad [b, +\infty[= \bigcap_{n \in \mathbf{N}^*}]b - \frac{1}{n}, +\infty[.$$

4.1.25 Exemples. Les constantes sont mesurables ; une fonction caractéristique χ_A est mesurable si et seulement si A l'est.

Comme dans le cas d'une variable, on montre que beaucoup de fonctions sont mesurables :

4.1.26 Proposition.

- 1) *Toute fonction continue est mesurable. Plus généralement, si f est continue sauf sur un ensemble négligeable A , f est mesurable.*
- 2) *Si les f_n , $n \in \mathbf{N}$, sont mesurables à valeurs dans \mathbf{R} , il en est de même*

de $\sup f_n$ et $\inf f_n$ (s'ils sont finis). En particulier, si f, g sont mesurables, il en est de même de $\text{Max}(f, g)$, $\text{Min}(f, g)$, f^+ , f^- et $|f|$.

3) Si f est limite simple d'une suite de fonctions mesurables, elle est mesurable.

4) Si $F : \mathbf{C} \rightarrow \mathbf{C}$ est continue et $f : \mathbf{R}^d \rightarrow \mathbf{C}$ mesurable, $F \circ f$ est mesurable.

5) La somme, le produit de deux fonctions mesurables sont des fonctions mesurables.

6) Une fonction égale presque partout à une fonction mesurable est mesurable.

Démonstration. Le point 1), dans le cas continu, vient du fait que l'image réciproque d'un ouvert par une fonction continue est un ouvert, donc une partie mesurable. Si f est continue sur $\mathbf{R}^d - A$, on considère sa restriction \bar{f} à $\mathbf{R}^d - A$. Si U est un ouvert de \mathbf{C} , $\bar{f}^{-1}(U)$ est un ouvert de $\mathbf{R}^d - A$ pour la topologie induite, donc de la forme $\Omega \cap (\mathbf{R}^d - A)$ avec Ω ouvert de \mathbf{R}^d . C'est donc une partie mesurable de \mathbf{R}^d . Comme on a $\bar{f}^{-1}(U) \subset f^{-1}(U) \subset \bar{f}^{-1}(U) \cup A$, on conclut par 4.1.12.3.

Pour 2), si $f = \sup f_n$, la propriété vient de la formule $f^{-1}(]a, +\infty[) = \bigcup_n f_n^{-1}(]a, +\infty[)$. Pour le point 3), on considère $g_n = \sup_{i \geq n} f_i$. Les g_n sont mesurables en vertu de 2), la suite (g_n) est décroissante et converge vers f , de sorte qu'on a $f = \inf g_n$, et cette fonction est mesurable, toujours par 2). Le point 4) est évident car on a $(F \circ f)^{-1}(U) = f^{-1}(F^{-1}(U))$ et, comme F est continue, $F^{-1}(U)$ est un ouvert.

Montrons que $f + g$ est mesurable si f et g le sont. Si f et g sont réelles, cela vient de la formule :

$$(f + g)^{-1}(]a, +\infty[) = \bigcup_{r \in \mathbf{Q}} \left(f^{-1}(]r, +\infty[) \cap g^{-1}(]a - r, +\infty[) \right).$$

En effet, il est clair que l'union est contenue dans $(f + g)^{-1}(]a, +\infty[)$. Réciproquement, si on a un x qui vérifie $f(x) + g(x) > a$, on choisit $r \in \mathbf{Q}$ qui vérifie $a - g(x) < r < f(x)$ et x est dans l'ensemble de l'union correspondant à r . Dans le cas complexe on utilise la décomposition $f = \text{Re } f + i \text{Im } f$.

Pour le produit, le raisonnement est analogue à celui utilisé pour la somme lorsque les fonctions sont positives et on se ramène à ce cas en employant les décompositions du type $f^+ - f^-$ et $\text{Re } f + i \text{Im } f$.

Enfin le point 6) vient du fait que si f est égale presque partout à g on a $f^{-1}(U) \sim g^{-1}(U)$.

4.1.6 Fonctions étagées

Les fonctions étagées vont jouer pour l'intégrale de Lebesgue le rôle que jouaient les fonctions en escalier pour celle de Riemann :

4.1.27 Proposition-Définition. *Soit $f : \mathbf{R}^d \rightarrow \mathbf{R}$ une fonction. Les propriétés suivantes sont équivalentes :*

- 1) *La fonction f est mesurable et ne prend qu'un nombre fini de valeurs.*
- 2) *La fonction f est de la forme $f = \sum_{i=1}^m \alpha_i \chi_{A_i}$ où les A_i sont mesurables et disjoints et où les α_i sont des réels (que l'on peut supposer distincts).*
- 3) *La fonction f est de la forme $f = \sum_{j=1}^n \beta_j \chi_{B_j}$ où les B_j sont mesurables et où les β_j sont des réels.*

Une fonction $f : \mathbf{R}^d \rightarrow \mathbf{R}$ qui vérifie les conditions précédentes est dite étagée.

Démonstration. L'équivalence de 1) et 2) est claire en prenant pour α_i les valeurs non nulles prises par la fonction. L'écriture de 2) avec les α_i distincts est donc uniquement déterminée par f (on parle de l'écriture **canonique** de f). Il est clair aussi que 2) implique 3). Enfin si f est de la forme évoquée en 3), elle ne prend qu'un nombre fini de valeurs (qui sont parmi les $\beta_{j_1} + \dots + \beta_{j_r}$).

4.1.28 Remarque. On notera que l'ensemble des fonctions étagées est un \mathbf{R} -espace vectoriel (utiliser les caractérisations 1) ou 3)).

Le théorème suivant relie les fonctions mesurables positives et les fonctions étagées :

4.1.29 Théorème. *Soit $f : \mathbf{R}^d \rightarrow \mathbf{R}^+$ une fonction mesurable positive. Il existe une suite croissante (u_n) de fonctions étagées positives qui converge simplement vers f .*

Démonstration. Pour $r \in \mathbf{Q}^+$ on définit une fonction w_r en posant $w_r(x) = r$ si $f(x) \geq r$ et $w_r(x) = 0$ sinon. La fonction w_r est étagée car elle prend les valeurs 0 et r sur $f^{-1}([0, r])$ et $f^{-1}([r, +\infty[)$ qui sont mesurables. Posons $g = \sup_{r \in \mathbf{Q}^+} w_r$. On a $f = g$. En effet, on a $w_r \leq f$ pour tout r , donc $g \leq f$. Réciproquement, si r est un rationnel vérifiant $0 \leq r < f(x)$, on a $w_r(x) = r$, d'où $g(x) \geq r$. Comme cela vaut pour tous les rationnels $< f(x)$, on a $g(x) \geq f(x)$. On numérote alors les rationnels en une suite r_n et on pose $v_n = w_{r_n}$, puis $u_n = \text{Max}_{k \leq n} v_k$. Il est clair que la suite (u_n) convient.

4.2 L'intégrale de Lebesgue

La définition de l'intégrale de Lebesgue se fait par étapes.

4.2.1 Définition de l'intégrale : le cas des fonctions mesurables positives

1) On commence par les fonctions caractéristiques. Si A est mesurable, on pose $\int_{\mathbf{R}^d} \chi_A = \mu(A)$, comme dans le cas de \mathbf{R} , mais, à l'envers : cette fois on connaît la mesure et on veut l'intégrale !

2) On passe ensuite aux fonctions étagées positives.

Soit f une fonction étagée positive écrite sous la forme canonique $\sum_{i=1}^n \alpha_i \chi_{A_i}$ (avec les A_i disjoints et les α_i distincts). On a vu que cette écriture est bien définie. Puisqu'on veut que l'intégrale soit linéaire, on doit définir $\int_{\mathbf{R}^d} f = \sum_{i=1}^n \alpha_i \mu(A_i) \in [0, +\infty]$. Pour les autres écritures on a le lemme suivant :

4.2.1 Lemme. *Soit $f = \sum_{j=1}^n \beta_j \chi_{B_j}$ une fonction étagée positive. On a $\int_{\mathbf{R}^d} f = \sum_{j=1}^n \beta_j \mu(B_j)$ (même si les B_j ne sont pas disjoints et si certains des β_j peuvent être égaux).*

Démonstration. Pour le cas où les B_j sont disjoints, mais où certains des β_j sont égaux il suffit de regrouper les B_j correspondants et d'appliquer l'additivité de la mesure. Pour le cas général, posons $I = \{1, 2, \dots, n\}$. On vérifie qu'on a $f = \sum_{J \subset I} \gamma_J \chi_{C_J}$, où l'on a posé $C_J = \bigcap_{j \in J} B_j - \bigcup_{k \in I-J} B_k$ (les points qui sont dans les B_j pour $j \in J$ et pas dans les autres) et $\gamma_J = \sum_{j \in J} \beta_j$. Comme les C_J sont disjoints on a $\int f = \sum_J \gamma_J \mu(C_J)$. Par ailleurs, on a $B_j = \bigcup_{j \in J} C_J$ et cette union est disjointe. On conclut en utilisant l'additivité de la mesure : $\mu(B_j) = \sum_{j \in J} \mu(C_J)$.

4.2.2 Corollaire. *Si f et g sont des fonctions étagées positives, on a $\int(f+g) = \int f + \int g$.*

Démonstration. On écrit $f = \sum_{i=1}^m \alpha_i \chi_{A_i}$, $g = \sum_{j=1}^n \beta_j \chi_{B_j}$ et on calcule $\int(f+g)$ avec l'écriture $f+g = \sum_{i=1}^m \alpha_i \chi_{A_i} + \sum_{j=1}^n \beta_j \chi_{B_j}$.

4.2.3 Corollaire. *Si f et g sont des fonctions étagées positives et si on a $f \leq g$ on a $\int f \leq \int g$.*

Démonstration. Il suffit d'écrire $g = f + (g - f)$ et d'appliquer le résultat précédent.

3) On passe ensuite aux fonctions mesurables positives.

4.2.4 Définition. *Soit f une fonction mesurable positive, on pose $\int_{\mathbf{R}^d} f = \sup \int_{\mathbf{R}^d} u$ (le sup peut être fini ou infini) où u parcourt l'ensemble des fonctions étagées $u \leq f$. On dit que f est **intégrable** si et seulement si $\int_{\mathbf{R}^d} f$ est finie.*

4.2.5 Remarques.

1) Le lecteur notera que cette définition est compatible avec celle donnée pour les fonctions étagées en vertu de 4.2.3.

2) On note aussitôt l'implication $f \leq g \implies \int f \leq \int g$.

4.2.2 Le théorème de convergence monotone pour les fonctions positives

On commence par deux lemmes sur la convergence des suites de fonctions étagées.

4.2.6 Lemme. *Soit (u_n) une suite décroissante de fonctions étagées positives qui converge simplement vers 0. On suppose que $\int_{\mathbf{R}^d} u_0$ est finie. Alors $I_n = \int_{\mathbf{R}^d} u_n$ converge vers 0.*

4.2.7 Remarque. L'exemple de la suite constante $u_n = 1/n$ montre la nécessité de l'hypothèse sur la finitude de $\int_{\mathbf{R}^d} u_0$ ou de l'une des $\int_{\mathbf{R}^d} u_n$.

Démonstration. Le cas où u_0 est nulle presque partout est trivial (toutes les intégrales sont nulles) et nous l'écartérons désormais. Soient m et M les minimum et maximum des valeurs non nulles de u_0 et soit $S = \{x \in \mathbf{R}^d \mid u_0(x) > 0\}$. On a $m\mu(S) \leq \int u_0 \leq M\mu(S)$, ce qui montre que $\mu(S)$ est finie. De plus, comme f n'est pas nulle presque partout, on a $\mu(S) > 0$. Soit $\epsilon > 0$. Nous allons montrer que $\int u_n$ est $\leq 2\epsilon$ pour n assez grand. Pour cela, appelons A_n l'ensemble des x tels que $u_n(x) > \epsilon/\mu(S)$. Comme u_n est nulle en dehors de S , on peut écrire $\int_{\mathbf{R}^d} u_n = \int_{A_n} u_n + \int_{S-A_n} u_n$. Comme (u_n) tend vers 0 en décroissant, la suite (A_n) est une suite décroissante de parties mesurables, d'intersection vide, et contenues dans S (donc de mesure finie). En vertu de 4.1.19.4, on en déduit que $\mu(A_n)$ tend vers 0. Comme u_n est $\leq M$ on a $\int_{A_n} u_n \leq M\mu(A_n)$ et, pour n grand on peut rendre ce terme $< \epsilon$. Par ailleurs, on a $\int_{S-A_n} u_n \leq \frac{\epsilon}{\mu(S)}\mu(S - A_n) \leq \epsilon$ et, en définitive, on a bien $\int u_n \leq 2\epsilon$.

4.2.8 Lemme. *Soit (u_n) une suite croissante de fonctions étagées positives qui converge simplement vers une fonction étagée u . On suppose $I = \int_{\mathbf{R}^d} u$ finie. Alors $I_n = \int_{\mathbf{R}^d} u_n$ converge vers $I = \int_{\mathbf{R}^d} u$.*

Démonstration. C'est évident en appliquant le lemme précédent à la suite $v_n = u - u_n$.

On peut maintenant prouver le théorème de convergence monotone (pour les fonctions mesurables positives) :

4.2.9 Théorème. Soit (f_n) une suite croissante de fonctions mesurables positives qui converge simplement vers une fonction f . Alors on a $I := \int_{\mathbf{R}^d} f = \lim I_n$ où l'on a posé $I_n = \int_{\mathbf{R}^d} f_n$. Si de plus on suppose que les f_n sont intégrables et que les intégrales $I_n = \int_{\mathbf{R}^d} f_n$ sont bornées, la fonction f est intégrable.

Démonstration. Comme on a $f_n \leq f$, on a $I_n \leq I$. Il en résulte déjà que le théorème est évident si l'une des fonctions f_n n'est pas intégrable (car on a alors $I_n = +\infty = I$) ou si les I_n ne sont pas bornées (car on a $I \geq \sup I_n = +\infty$). Supposons donc les f_n intégrables et les I_n bornées. La suite (I_n) est croissante et majorée, donc converge et on a $l = \lim I_n \leq I$ et il s'agit de voir qu'on a l'égalité. Sinon, il existe une fonction étagée $u \leq f$, d'intégrale finie¹, qui vérifie $l < \int u < \int f$. De plus, on peut supposer qu'on a, pour tout $x \in \mathbf{R}^d$, $u(x) < f(x)$. En effet, si on a $u = \sum_{i=1}^r \alpha_i \chi_{A_i}$, on peut toujours diminuer les α_i de $\epsilon > 0$ assez petit en conservant² l'inégalité $\int u > l$.

On définit alors une suite de fonctions étagées positives u_n comme suit : $u_n(x) = u(x)$ si $u(x) \leq f_n(x)$ et $u_n(x) = 0$ sinon. On vérifie que la suite (u_n) est croissante. Montrons qu'elle converge simplement vers u . Comme on a $u(x) < f(x)$ et que $f_n(x)$ converge vers $f(x)$, il existe p avec $u(x) < f_p(x) < f(x)$. Mais alors on a $u_n(x) = u(x)$ pour $n \geq p$. En vertu de 4.2.8, la suite $\int u_n$ converge vers $\int u$. Comme on a $u_n \leq f_n$, on en déduit à la limite $\int u \leq l$ et c'est absurde.

4.2.10 Corollaire. Si f et g sont des fonctions intégrables positives, $f + g$ est intégrable et on a $\int(f + g) = \int f + \int g$.

Démonstration. On peut écrire f et g comme limites croissantes de fonctions étagées positives f_n et g_n (cf. 4.1.29). La suite $f_n + g_n$ converge vers $f + g$ en croissant. On a $\int(f_n + g_n) = \int f_n + \int g_n$ en vertu de 4.2.2 et ces intégrales sont majorées par $\int f + \int g$. On peut appliquer le théorème de convergence monotone 4.2.9 aux suites (f_n) , (g_n) et $(f_n + g_n)$. On voit que $f + g$ est intégrable et que son intégrale est bien $\int f + \int g$.

4.2.3 Définition de l'intégrale : le cas général

On peut maintenant achever la définition de l'intégrale.

1) Pour une fonction mesurable réelle quelconque f , on décompose f en $f^+ - f^-$. La fonction est dite intégrable si et seulement si f^+ et f^- le sont,

¹Si on a une fonction étagée $u = \sum_{i=1}^r \alpha_i \chi_{A_i} \leq f$ d'intégrale infinie, on en obtient une d'intégrale finie plus grande que l en tronquant les A_i par les pavés $[-n, n]^d$ pour n assez grand.

²Il suffit de prendre $\epsilon < (\int u - l) / \sum \mu(A_i)$ et $\epsilon < \text{Min } \alpha_i$.

ou encore si et seulement si $|f|$ l'est (c'est la même chose en vertu de 4.2.10). On pose alors $\int f = \int f^+ - \int f^-$.

2) Enfin, pour les fonctions mesurables à valeurs complexes, on décompose f en partie réelle et partie imaginaire : f est dite intégrable si ces deux fonctions le sont (ou encore si $|f|$ l'est) et on pose $\int f = \int \operatorname{Re} f + i \int \operatorname{Im} f$.

4.2.11 Notation. Si on note $x = (x_1, \dots, x_d)$ un point de \mathbf{R}^d on notera aussi l'intégrale de f sous la forme :

$$\int_{\mathbf{R}^d} f = \int_{\mathbf{R}^d} f(x) dx = \int_{\mathbf{R}^d} f(x_1, \dots, x_d) dx_1 \dots dx_d.$$

Le critère suivant est immédiat, mais fondamental :

4.2.12 Proposition. *Soit f une fonction de \mathbf{R}^d dans \mathbf{C} . Alors f est intégrable si et seulement si elle est mesurable et si $|f|$ est intégrable. Cette dernière condition est réalisée si $|f|$ est majorée par une fonction intégrable g .*

Intégrale sur une partie mesurable

Il est très commode de pouvoir parler d'intégrale non seulement sur \mathbf{R}^d , mais aussi sur une partie mesurable de \mathbf{R}^d (par exemple un pavé, ou une boule, etc.).

4.2.13 Définition. *Soit A une partie mesurable de \mathbf{R}^d , soit f une fonction de A dans \mathbf{C} et soit $\tilde{f} = f$ la fonction obtenue en prolongeant f par 0 en dehors de A . On dit que f est mesurable (resp. intégrable) sur A si \tilde{f} est mesurable (resp. intégrable) sur \mathbf{R}^d et on pose $\int_A f = \int_{\mathbf{R}^d} \tilde{f}$.*

4.2.14 Remarques.

1) Une fonction $f : A \rightarrow \mathbf{C}$ continue sur A est mesurable. En effet, si U est un ouvert de \mathbf{C} , $f^{-1}(U)$ est un ouvert de A , donc de la forme $\Omega \cap A$ avec Ω ouvert de \mathbf{R}^d . C'est donc une partie mesurable. De plus, $\tilde{f}^{-1}(U)$ est égal à $f^{-1}(U)$ si U ne contient pas 0 ou à $f^{-1}(U) \cup A^c$ si U contient 0. Il est donc mesurable. La propriété est encore vraie si f est continue sauf éventuellement sur un ensemble négligeable $B \subset A$ (cf. 4.1.26.1).

2) Si f est mesurable sur \mathbf{R}^d , sa restriction à une partie mesurable A quelconque est mesurable sur A . En effet, on a $\tilde{f}|_A = f\chi_A$.

Le lecteur se convaincra que la plupart des théorèmes énoncés dans ce chapitre se généralisent au cas des intégrales sur une partie mesurable. En particulier, une fonction est intégrable sur A si elle est mesurable sur A et si sa valeur absolue est majorée par une fonction intégrable sur A . On a une sorte de relation de Chasles :

4.2.15 Proposition. Soient A, B deux parties mesurables disjointes et soit f une fonction intégrable sur $A \cup B$. Alors f est intégrable sur A et B et on a $\int_{A \cup B} f = \int_A f + \int_B f$.

Démonstration. Si on note respectivement \widetilde{f} , \widetilde{f}_A et \widetilde{f}_B les prolongements par 0 de f , $f|_A$ et $f|_B$, la proposition vient de la formule $\widetilde{f} = \widetilde{f}_A + \widetilde{f}_B$.

4.2.4 Propriétés de l'intégrale

Le théorème suivant montre que l'intégrale a les propriétés usuelles : linéarité, croissance, etc.

4.2.16 Théorème.

- 1) L'ensemble $\mathcal{L}(\mathbf{R}^d)$ des fonctions Lebesgue-intégrables $f : \mathbf{R}^d \rightarrow \mathbf{R}$ est un sous-espace vectoriel de l'espace vectoriel de toutes les fonctions de \mathbf{R}^d dans \mathbf{R} . On a une assertion analogue pour les fonctions à valeurs dans \mathbf{C} .
- 2) L'application $I : f \mapsto \int_{\mathbf{R}^d} f(t) dt$ est une forme linéaire sur $\mathcal{L}(\mathbf{R}^d)$.
- 3) L'application I est croissante : si on a $f \leq g$ on a $I(f) \leq I(g)$. En particulier, on a l'inégalité $|\int_{\mathbf{R}^d} f(t) dt| \leq \int_{\mathbf{R}^d} |f(t)| dt$.
- 4) Si f et g sont intégrables, $\text{Max}(f, g)$ et $\text{Min}(f, g)$ le sont aussi.
- 5) Si f est presque partout nulle elle est intégrable et son intégrale est nulle. La réciproque est vraie si l'on suppose de plus $f \geq 0$. Si f est intégrable et si g est égale à f presque partout, alors g est intégrable et on a $\int_{\mathbf{R}^d} f = \int_{\mathbf{R}^d} g$.

Démonstration. 1) et 2). Le cas des fonctions positives a été traité en 4.2.10.

Pour les fonctions de signe quelconque, on écrit $f = f^+ - f^-$, $g = g^+ - g^-$ et, en posant $h = f + g$, $h = h^+ - h^-$. D'abord, on a $|h| \leq |f| + |g|$, de sorte que h est intégrable si f et g le sont. Ensuite, on a $h^+ - h^- = f^+ - f^- + g^+ - g^-$, ou encore $h^+ + f^- + g^- = h^- + f^+ + g^+$. Par l'additivité de l'intégrale des fonctions positives, on en déduit $\int h^+ + \int f^- + \int g^- = \int h^- + \int f^+ + \int g^+$, soit encore $\int h = \int f + \int g$.

Montrons la croissance. Si on a $f \leq g$, cela signifie $f^+ - f^- \leq g^+ - g^-$, donc $f^+ + g^- \leq f^- + g^+$. Mais, comme la croissance est vraie pour les fonctions positives, on en déduit : $\int f^+ + \int g^- \leq \int f^- + \int g^+$, ce qui donne l'inégalité cherchée.

Le point 4) résulte de formules du type : $\text{Max}(f, g) = \frac{1}{2}(f + g + |f - g|)$ et de 1) et 3).

Si f est une fonction nulle presque partout sa valeur absolue est presque partout nulle et il suffit de montrer que $|f|$ est intégrable et d'intégrale nulle. On est ainsi ramené au cas $f \geq 0$. Dans ce cas, si $u = \sum_{i=1}^n \alpha_i \chi_{A_i}$ (avec

$\alpha_i > 0$) est une fonction étagée positive et $\leq f$, les A_i sont contenus dans $\{x \mid f(x) > 0\}$. Comme cet ensemble est de mesure nulle, les A_i aussi et on a $\int u = 0$. Pour la réciproque, si A (resp. A_n) est l'ensemble sur lequel f est > 0 (resp. $\geq 1/n$), on a $\int f \geq \mu(A_n)/n$ (c'est la croissance appliquée à f et à la fonction étagée qui vaut $1/n$ sur A_n et 0 ailleurs). On en déduit $\mu(A_n) = 0$, de sorte que A_n est négligeable, donc aussi A qui est l'union des A_n .

Enfin, si $f = g$ presque partout, il suffit d'écrire $g = f + (g - f)$. Comme $g - f$ est nulle presque partout, le premier point de 5) conclut.

4.2.17 Définition. Pour $f \in \mathcal{L}(\mathbf{R}^d)$ on définit la "norme" L^1 : $\|f\|_1 = \int_{\mathbf{R}^d} |f|$.

4.3 Les théorèmes de convergence

Comme dans le cas de la dimension 1, c'est dans ce domaine que l'intégrale de Lebesgue montre toute sa puissance.

4.3.1 Le théorème de convergence monotone

4.3.1 Théorème. Soit $f_n : \mathbf{R}^d \rightarrow \mathbf{R}$ une suite monotone de fonctions intégrables. On suppose que la suite des intégrales $I_n = \int_{\mathbf{R}^d} f_n$ est bornée. Alors, la suite (f_n) converge vers une fonction intégrable f , à la fois au sens de la norme L^1 et au sens de la convergence simple presque partout et la suite $\int f_n$ converge vers $\int f$.

Démonstration. Supposons par exemple la suite croissante et soit M un nombre > 0 qui majore les I_n . On considère la fonction \widehat{f} définie par $\widehat{f}(x) = \sup f_n(x)$. Montrons d'abord que l'ensemble A des points où \widehat{f} est infinie est négligeable. En effet, A est l'intersection des A_p , avec $A_p = \{x \in \mathbf{R}^d \mid \widehat{f}(x) > p\}$. Si $\mu(A)$ est > 0 , soit m fini tel que $0 < m \leq \mu(A)$. On a $\mu(A_p) \geq m$ pour tout p . Fixons un p qui vérifie $pm > 2M$. Posons $B_n = \{x \mid f_n(x) > p\}$. La suite B_n est croissante et sa réunion est A_p . On a donc $\mu(A_p) = \lim \mu(B_n)$ en vertu de 4.1.19.3 et il s'ensuit qu'on a, pour n assez grand, $\mu(B_n) \geq m/2$. Mais alors on a $I_n \geq \mu(B_n)p \geq mp/2 > M$, contrairement à l'hypothèse.

On considère alors les fonctions f'_n définies par $f'_n(x) = f_n(x)$ pour $x \notin A$ et $f'_n(x) = 0$ pour $x \in A$. Soit $f = \sup f'_n$. On a $f(x) = \widehat{f}(x)$ pour $x \notin A$ et $f(x) = 0$ pour $x \in A$. Les fonctions f'_n étant égales aux f_n presque partout

sont intégrables et on a $\int f'_n = \int f_n$. La suite (f'_n) est croissante et converge simplement vers f (de sorte que la suite initiale (f_n) converge vers f presque partout).

On considère ensuite les fonctions $g_n = f'_n - f'_0$. La suite (g_n) est une suite croissante de fonctions intégrables positives, qui converge vers $f - f'_0$ et dont les intégrales $I_n - I_0$ sont bornées. En vertu de 4.2.9, $f - f'_0$ est intégrable et on a $\int(f - f'_0) = \lim \int(f'_n - f'_0)$. Il en résulte que f est intégrable et que $\int f'_n = \int f_n$ tend vers $\int f$. L'assertion sur la convergence L^1 vient de l'hypothèse de monotonie (on a $\int |f - f_n| = \int |f - f'_n| = \int(f - f'_n)$).

4.3.2 Remarque. Si dans 4.3.1 on suppose que la suite (f_n) converge simplement vers f , la démonstration se simplifie en se réduisant au dernier paragraphe.

Comme dans le cas de \mathbf{R} on a un corollaire qui concerne les séries à termes positifs :

4.3.3 Corollaire. Soit (u_n) une suite de fonctions intégrables positives. On suppose $\sum_{n=0}^{+\infty} \int_{\mathbf{R}^d} u_n < +\infty$. Alors, la série $\sum_{n=0}^{+\infty} u_n(x)$ converge presque partout et on a la formule :

$$\int_{\mathbf{R}^d} \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_{\mathbf{R}^d} u_n.$$

4.3.2 Le théorème de convergence dominée

4.3.4 Théorème. Soit (f_n) une suite de fonctions intégrables de \mathbf{R}^d dans \mathbf{C} . On suppose que f_n converge presque partout vers une fonction f et qu'il existe une fonction intégrable $g \geq 0$ qui majore toutes les fonctions $|f_n|$. Alors, la fonction f est intégrable, la suite (f_n) converge vers f au sens de la norme L^1 et on a $\lim \int f_n = \int f$.

Démonstration. La fonction f est intégrable. En effet, elle est mesurable (comme limite simple presque partout d'une suite de fonctions mesurables, cf. 4.1.26) et majorée en module par la fonction intégrable g (cf. 4.2.12). On considère les fonctions $g_n = |f - f_n|$ qui convergent simplement vers 0 presque partout et sont dominées par $2g$, puis les fonctions $h_n = \sup_{k \geq n} g_k$. Ces fonctions sont finies (majorées par $2g$) et intégrables (elles sont mesurables par 4.1.26 et majorées par $2g$). La suite (h_n) est décroissante et tend vers 0 presque partout. Le théorème de convergence monotone montre que l'intégrale $\int h_n$ tend vers 0, donc, *a fortiori*, $\int g_n$, soit $\int |f - f_n|$ et on a gagné.

4.3.5 Corollaire. Soit (g_n) une suite de fonctions intégrables. On suppose que la série $\sum_n \|g_n\|_1$ converge. Alors, la série de terme général g_n converge, à la fois au sens de la norme L^1 et de la convergence presque partout, vers une fonction intégrable g et on a $\sum_n \int g_n = \int g$.

Démonstration. On considère la série de terme général $|g_n|$. On a $\sum_{n=0}^{+\infty} \int_{\mathbf{R}} |g_n| = \sum_{n=0}^{+\infty} \|g_n\|_1 < +\infty$. On peut donc appliquer 4.3.3 qui montre que la série $|g_n|$ converge presque partout vers une fonction intégrable positive h et qu'on a $\int h = \sum_{n=0}^{+\infty} \int_{\mathbf{R}} |g_n|$. Il en résulte que la série $g_n(x)$ est (presque partout) absolument convergente, donc convergente. De plus, si on pose $h_n = \sum_{k=0}^n g_k$, les fonctions h_n sont dominées par h . On conclut par le théorème de convergence dominée.

4.3.3 Complétude de $\mathcal{L}(\mathbf{R}^d)$

4.3.6 Théorème. L'espace $\mathcal{L}(\mathbf{R}^d)$ est complet pour la norme L^1 .

Démonstration. Soit (f_n) une suite de Cauchy. Il suffit de montrer qu'elle contient une sous-suite qui converge au sens de L^1 . En raisonnant comme en 2.3.2 ou en 3.2.1, on se ramène au cas où l'on a, pour tout n , $\|f_{n+1} - f_n\|_1 \leq 1/2^n$. Mais alors, par le corollaire précédent, la série $f_{n+1} - f_n$ converge pour la norme L^1 , donc aussi la suite (f_n) .

4.3.4 Intégrales dépendant d'un paramètre

Les résultats et les démonstrations sont à peu près identiques à ceux obtenus sur \mathbf{R} .

Limite et continuité

4.3.7 Théorème. Soit Λ un espace métrique quelconque (par exemple une partie de \mathbf{R}^n) et soit $f : \mathbf{R}^d \times \Lambda \rightarrow \mathbf{C}$ une fonction vérifiant les conditions suivantes :

- 1) $\forall \lambda \in \Lambda$, la fonction $x \mapsto f(x, \lambda)$ est intégrable sur \mathbf{R}^d ,
- 2) Il existe un ensemble négligeable $A \subset \mathbf{R}^d$ tel que :
 - a) pour $x \notin A$, $f(x, \lambda)$ tend vers $g(x)$ quand λ tend vers λ_0 ,
 - b) il existe une fonction intégrable $h \geq 0$ telle que l'on ait, pour tout $\lambda \in \Lambda$ et tout $x \notin A$, $|f(x, \lambda)| \leq h(x)$.

Alors, la fonction g est intégrable sur \mathbf{R}^d et $F(\lambda) = \int_{\mathbf{R}^d} f(x, \lambda) dx$ tend vers

$\int_{\mathbf{R}^d} g(x) dx$ quand λ tend vers λ_0 .
 En particulier, si, pour tout $x \notin A$, la fonction $\lambda \mapsto f(x, \lambda)$ est continue en λ_0 (toujours avec les hypothèses 1 et 2.a) la fonction $F(\lambda) = \int_{\mathbf{R}^d} f(x, \lambda) dx$ est continue en λ_0 .

Démonstration. Il suffit de montrer que si (λ_n) est une suite convergeant vers λ_0 , la suite $(F(\lambda_n))$ converge vers $\int_{\mathbf{R}^d} g(x) dx$. C'est exactement le théorème de convergence dominée appliqué aux fonctions $f_n(x) = f(x, \lambda_n)$. Le cas de la continuité s'obtient en appliquant le résultat avec $g(x) = f(x, \lambda_0)$.

Dérivabilité

4.3.8 Théorème. Dérivation sous le signe somme

Soit I un intervalle ouvert de \mathbf{R} et soit $f : \mathbf{R}^d \times I \rightarrow \mathbf{R}$ une fonction vérifiant les conditions suivantes :

- 1) $\forall \lambda \in I, x \mapsto f(x, \lambda)$ est intégrable.
- 2) Il existe un ensemble négligeable $A \subset \mathbf{R}^d$ tel que :
 - a) pour $x \notin A$, la fonction $\lambda \mapsto f(x, \lambda)$ est dérivable sur I ,
 - b) il existe une fonction intégrable $h \geq 0$ telle que l'on ait, pour tout $\lambda \in I$

et tout $x \notin A, \left| \frac{\partial f}{\partial \lambda}(x, \lambda) \right| \leq h(x)$.

Alors, la fonction $\lambda \mapsto F(\lambda) = \int_{\mathbf{R}^d} f(x, \lambda) dx$ est dérivable sur I et on a :

$$F'(\lambda) = \int_{\mathbf{R}^d} \frac{\partial f}{\partial \lambda}(x, \lambda) dx.$$

Démonstration. Fixons $\lambda \in I$. Il s'agit de calculer la limite, quand h_n tend vers 0, de $\frac{F(\lambda + h_n) - F(\lambda)}{h_n}$. Cette quantité est l'intégrale de la fonction $\varphi_n(x) = \frac{f(x, \lambda + h_n) - f(x, \lambda)}{h_n}$. Par définition de la dérivée, pour $x \notin A$, cette fonction tend vers $\frac{\partial f}{\partial \lambda}(x, \lambda)$ quand n tend vers $+\infty$. Par ailleurs, le théorème des accroissements finis montre qu'on a $\varphi_n(x) = \frac{\partial f}{\partial \lambda}(x, \theta_n)$ avec $\theta_n \in]\lambda, \lambda + h_n[$. On a donc $|\varphi_n(x)| \leq h(x)$ pour $x \notin A$. On conclut avec le théorème de convergence dominée.

4.4 Les théorèmes de Fubini

Soit d un entier ≥ 2 écrit sous la forme $d = p + q$ avec $p, q \in \mathbf{N}^*$. Les théorèmes de Fubini permettent de ramener l'étude des intégrales sur \mathbf{R}^d à

celle d'intégrales sur \mathbf{R}^p et \mathbf{R}^q et, en itérant le procédé, à des intégrales sur \mathbf{R} . On considère donc des fonctions $f : \mathbf{R}^d = \mathbf{R}^p \times \mathbf{R}^q \rightarrow \mathbf{C}$. On note $f(x, y)$ l'image par f du point (x, y) avec $x \in \mathbf{R}^p$ et $y \in \mathbf{R}^q$. La mesure de Lebesgue sur \mathbf{R}^d (resp. \mathbf{R}^p , resp. \mathbf{R}^q) est notée μ_d (resp. μ_p , resp. μ_q). Les intégrales sur \mathbf{R}^d (resp. \mathbf{R}^p , resp. \mathbf{R}^q) seront notées avec le symbole $dx dy$ (resp. dx , resp. dy).

4.4.1 Énoncés des deux théorèmes

Fubini-Lebesgue

4.4.1 Théorème. *Soit $f : \mathbf{R}^p \times \mathbf{R}^q \rightarrow \mathbf{C}$ une fonction intégrable.*

- 1) *Pour presque tout $y \in \mathbf{R}^q$, la fonction f_y qui à x associe $f(x, y)$ est intégrable sur \mathbf{R}^p .*
- 2) *La fonction qui à y associe $\int_{\mathbf{R}^p} f_y(x) dx$ est définie presque partout, elle est intégrable sur \mathbf{R}^q et on a la formule :*

$$\int_{\mathbf{R}^q} \left(\int_{\mathbf{R}^p} f_y(x) dx \right) dy = \int_{\mathbf{R}^p \times \mathbf{R}^q} f(x, y) dx dy.$$

Fubini-Tonelli

4.4.2 Théorème. *Soit $f : \mathbf{R}^p \times \mathbf{R}^q \rightarrow \mathbf{R}^+$ une fonction mesurable positive.*

- 1) *Pour presque tout $y \in \mathbf{R}^q$, la fonction f_y qui à x associe $f(x, y)$ est mesurable sur \mathbf{R}^p .*
- 2) *Si, de plus, on suppose que f_y est intégrable pour presque tout $y \in \mathbf{R}^q$, la fonction définie presque partout qui à y associe $\int_{\mathbf{R}^p} f(x, y) dx$ est mesurable sur \mathbf{R}^q et on a la formule :*

$$\int_{\mathbf{R}^q} \left(\int_{\mathbf{R}^p} f_y(x) dx \right) dy = \int_{\mathbf{R}^p \times \mathbf{R}^q} f(x, y) dx dy.$$

Bien entendu ces intégrales sont à valeurs dans $[0, +\infty]$.

On notera les deux différences essentielles avec Fubini-Lebesgue : f est seulement supposée mesurable, mais on suppose f_y intégrable.

4.4.2 Démonstration du théorème de Fubini

Le cas des fonctions caractéristiques

On démontre les deux variantes du théorème en même temps. On commence par traiter le cas des fonctions caractéristiques.

Il est commode ici d'utiliser des fonctions pouvant prendre la valeur $+\infty$. On dira qu'une telle fonction est mesurable si l'ensemble A des points où elle est infinie est mesurable et si f est mesurable en restriction à $\mathbf{R}^d - A$. Avec cette convention, si (f_n) est une suite de fonctions mesurables positives, $\sup f_n$ est mesurable.

4.4.3 Lemme. (*Fubini pour une fonction caractéristique*) Soit $A \in \mathbf{R}^{p+q}$ une partie mesurable. On pose, pour $y \in \mathbf{R}^q$, $A_y = \{x \in \mathbf{R}^p \mid (x, y) \in A\}$ (A_y est la **coupe** de A selon y).

1) Pour presque tout y , l'ensemble A_y est mesurable et la fonction $\varphi_A : \mathbf{R}^q \rightarrow \mathbf{R} \cup \{+\infty\}$ définie par $\varphi_A(y) = \mu_p(A_y)$ est mesurable (avec la convention précédente).

2) Si la fonction φ_A est presque partout finie (ce qui est le cas, en particulier, si $\mu(A)$ est fini), on a $\mu_d(A) = \int_{\mathbf{R}^q} \mu_p(A_y) dy$.

4.4.4 Remarque. Comme on a l'égalité $(\chi_A)_y(x) = \chi_{A_y}(x)$, 4.4.3 est bien le cas particulier de 4.4.1 dans lequel on a $f = \chi_A$.

Démonstration. 1) On note d'abord que le lemme est évident si A est un ensemble élémentaire. En effet, on peut écrire $A = B_1 \times \dots \times B_p \times C_1 \times \dots \times C_q$ et A est de la forme $B \times C$ avec $B \subset \mathbf{R}^p$ et $C \subset \mathbf{R}^q$. Par définition, on a $\mu(A) = \prod_{i=1}^p \mu(B_i) \times \prod_{j=1}^q \mu(C_j) = \mu_p(B) \mu_q(C)$. On a $A_y = B$ pour tout $y \in C$ et $A_y = \emptyset$ sinon, ce qui montre que A_y est mesurable pour tout y . De plus, la fonction φ_A n'est autre que $\mu(B) \chi_C$, qui est bien mesurable. L'intégrale $\int_{\mathbf{R}^q} \mu_p(A_y)$ est donc l'intégrale de la fonction constante et égale à $\mu(B)$ sur C : c'est bien $\mu(B) \mu(C) = \mu(A)$.

Notons que si $\mu(A)$ est fini, φ_A est fini presque partout (si cette fonction prend la valeur $+\infty$ c'est qu'on a $\mu(B) = +\infty$, mais alors, on a $\mu(C) = 0$).

2) Montrons que le lemme est vrai si A est dans \mathcal{D} , i.e. réunion dénombrable disjointe d'élémentaires. Si on a $A = \bigcup_{n \in \mathbf{N}} A_n$, avec des A_n élémentaires, on note que l'on a $A_y = \bigcup_{n \in \mathbf{N}} A_{n,y}$ (réunion disjointe). Comme les $A_{n,y}$ sont mesurables, il en est de même de A_y et la fonction φ_A est mesurable comme somme des φ_{A_n} . On a $\mu_d(A) = \sum_{n \in \mathbf{N}} \mu_d(A_n)$, $\mu_p(A_y) = \sum_{n \in \mathbf{N}} \mu_p(A_{n,y})$ et $\mu_d(A_n) = \int_{\mathbf{R}^p} \mu_p(A_{n,y})$. On conclut par le théorème de convergence monotone (y compris pour la finitude de φ_A).

3) Il s'agit maintenant de montrer 4.4.3 pour une partie mesurable quelconque. On commence par le point 1). Appelons (P) la propriété requise dans le 1) du lemme (A_y est une partie mesurable de \mathbf{R}^p et φ_A est mesurable sur

\mathbf{R}^q) et notons \mathcal{T} l'ensemble des parties mesurables de \mathbf{R}^d qui vérifient (P). Montrons que l'ensemble \mathcal{T} est une tribu.

Soit (A_n) une suite de parties vérifiant (P), et soit A la réunion des A_n . Quitte à remplacer A_n par $A_0 \cup \dots \cup A_n$, on peut supposer la suite (A_n) croissante. On a $A_y = \bigcup_n A_{n,y}$. Soit X_n l'ensemble des $y \in \mathbf{R}^q$ tels que $A_{n,y}$ ne soit pas mesurable. Par hypothèse, on a $\mu_q(A_{n,y}) = 0$. Si y n'est pas dans $X = \bigcup_n X_n$, les parties $A_{n,y}$ sont donc toutes mesurables, donc aussi leur réunion A_y . Comme X est encore de mesure nulle, on a montré que A_y est mesurable pour presque tout y . Par ailleurs, on a $\mu_p(A_y) = \sup \mu_p(A_{n,y})$ (car la suite $A_{n,y}$ est croissante) de sorte que la fonction $y \mapsto \mu_p(A_y)$ est mesurable comme sup de fonctions mesurables.

Soit $A \in \mathcal{T}$. On a $(A^c)_y = (A_y)^c$. Comme A_y est mesurable pour presque tout y , il en est de même de $(A^c)_y$. Pour voir que φ_{A^c} est mesurable, on considère le pavé $P_n = [-n, n]^q$. La mesure $\mu(A_y^c)$ est la borne supérieure de $\mu(P_n \cap A_y^c) = \mu(P_n) - \mu(P_n \cap A_y) = \mu(P_n) - \mu(A_y \chi_{P_n})$. Comme $\mu(A_y)$ est mesurable, il en résulte que $\mu(A_y^c)$ l'est aussi (par produit, somme et passage à la borne supérieure).

Pour voir que \mathcal{T} contient \mathcal{M} , comme on a vu que \mathcal{T} contient les ensembles élémentaires, il reste à voir qu'elle contient les ensembles négligeables. Cela résulte du lemme suivant :

4.4.5 Lemme. *Soit $A \subset \mathbf{R}^d$ un ensemble négligeable. Alors, pour presque tout $y \in \mathbf{R}^q$, la coupe A_y est négligeable (et donc la fonction φ_A est presque partout nulle, donc mesurable).*

Démonstration. On sait (cf. 4.1.22) qu'on peut trouver des ensembles A_n , réunion dénombrables d'élémentaires, qui vérifient $A \subset A_n$ et $\lim \mu(A_n) = 0$. Les ensembles A_n sont dans la tribu \mathcal{T} et il en est de même de $A' = \bigcap_n A_n$. On a $A \subset A' \subset A_n$, donc $\mu(A') \leq \mu(A_n)$ et comme $\mu(A_n)$ tend vers 0 on a $\mu(A') = 0$.

Comme A' est dans \mathcal{T} , les A'_y sont mesurables pour presque tout y et la fonction $y \mapsto \mu(A'_y)$ est mesurable. Comme les A_n sont dans \mathcal{D} , ils vérifient Fubini comme on l'a vu ci-dessus. On a donc $\int_{\mathbf{R}^q} \mu(A_{n,y}) = \mu(A_n)$. Comme A'_y est contenu dans $A_{n,y}$, on a $\int_{\mathbf{R}^q} \mu(A'_y) \leq \int_{\mathbf{R}^q} \mu(A_{n,y}) = \mu(A_n)$ et il en résulte qu'on a $\int_{\mathbf{R}^q} \mu(A'_y) = 0$. Cela montre que la fonction $\mu(A'_y)$ est nulle presque partout, autrement dit que A'_y est négligeable pour presque tout $y \in \mathbf{R}^q$. Comme on a $A_y \subset A'_y$, la propriété est encore vraie pour A_y .

4) Il reste à montrer le point 2) de 4.4.3 pour une partie mesurable quelconque A . En vertu de 4.1.22, il existe des parties $A_n \in \mathcal{D}$, contenant A , telles que l'on ait $\lim \mu(A_n) = \mu(A)$ et, plus précisément, $\lim \mu(A_n - A) = 0$. En vertu du point 2), on a $\mu(A_n) = \int_{\mathbf{R}^q} \mu_p(A_{n,y})$.

Supposons que $\varphi_A(y) = \mu_p(A_y)$ est fini presque partout³. Comme A_y est contenu dans $A_{n,y}$, on en déduit $\int_{\mathbf{R}^q} \mu_p(A_y) \leq \mu(A_n)$ et, par passage à la limite, $\int_{\mathbf{R}^q} \mu_p(A_y) \leq \mu(A)$. On va appliquer ce résultat à la partie $A_n - A$. Cette partie est de mesure finie, de sorte que $\mu_p((A_n - A)_y)$ est fini pour presque tout y . Comme $\mu_p(A_y)$ est fini aussi, on en déduit que $\mu_p(A_{n,y})$ est fini. Le résultat précédent montre qu'on a $\int_{\mathbf{R}^q} \mu_p((A_n - A)_y) \leq \mu(A_n - A)$ et cette quantité tend vers 0 quand n tend vers l'infini. Mais on a $(A_n - A)_y = A_{n,y} - A_y$, et, comme les mesures sont finies, le premier membre est égal à $\int_{\mathbf{R}^q} \mu_p(A_{n,y}) - \int_{\mathbf{R}^q} \mu_p(A_y)$, ce qui montre que $\mu(A_n) = \int_{\mathbf{R}^q} \mu_p(A_{n,y})$ tend vers $\int_{\mathbf{R}^q} \mu_p(A_y)$. Comme on sait que $\mu(A_n)$ tend vers $\mu(A)$, on a terminé.

Le cas général

1) Comme le théorème de Fubini est vrai pour les fonctions caractéristiques, par linéarité, il est vrai aussi pour les fonctions étagées positives. En effet, une telle fonction est de la forme $f = \sum_{i=1}^n \alpha_i \chi_{A_i}$. Si les A_i sont disjoints et les $\alpha_i > 0$, f est intégrable si et seulement si les A_i sont de mesure finie.

2) Si maintenant f est une fonction mesurable positive quelconque, on écrit f comme limite croissante d'une suite de fonctions u_n étagées positives (cf. 4.1.29). En vertu de 1) on a $I_n = \int u_n(x, y) dx dy = \int \left(\int u_n(x, y) dx \right) dy$ et ces intégrales sont majorées par $I = \int f(x, y) dx dy$. Posons $g_n(y) = \int u_n(x, y) dx$. En vertu de 4.2.9, $g_n(y)$ tend vers $g(y) = \int f(x, y) dx$. Supposons f intégrable donc I finie (variante Fubini-Lebesgue⁴) et montrons que $g(y)$ est fini pour presque tout y , c'est-à-dire que l'ensemble $A = \{y \in \mathbf{R}^q \mid g(y) = +\infty\}$ est négligeable. Pour cela, on introduit, pour $M \in \mathbf{N}^*$, les ensembles $A_M = \{y \in \mathbf{R}^q \mid g(y) > M\}$ et $A_{M,n} = \{y \in \mathbf{R}^q \mid g_n(y) > M\}$. L'ensemble A est intersection décroissante des A_M et, comme g_n tend vers g en croissant, A_M est réunion croissante des $A_{M,n}$ et on a donc $\mu(A_M) = \lim \mu(A_{M,n})$ (cf. 4.1.19.3). Mais, on a $M\mu(A_{M,n}) \leq I_n = \int g_n(y) dy \leq I$. On en déduit $\mu(A_{M,n}) \leq I/M$ et, à la limite, $\mu(A_M) \leq I/M$. Comme $\mu(A)$ est la limite des $\mu(A_M)$ quand M tend vers l'infini (cf. 4.1.19.4), on a bien $\mu(A) = 0$.

Le théorème de convergence monotone appliqué à u_n sur \mathbf{R}^d et à g_n sur \mathbf{R}^q permet de conclure à l'égalité $\int f = \int g(y) dy$ dans les deux variantes.

3) Pour une fonction intégrable quelconque, à valeurs réelles, on écrit $f = f^+ - f^-$ et on est ramené au cas précédent.

³On notera que, si $\mu(A)$ est finie, il en est de même de $\mu(A_n)$ pour n assez grand, de sorte que $\varphi_{A_n}(y) = \mu_p(A_{n,y})$ est finie presque partout (par le cas numéro 2). Comme on a $\varphi_A \leq \varphi_{A_n}$, φ_A est finie elle aussi et l'hypothèse est réalisée.

⁴Dans la variante Fubini-Tonelli, on fait l'hypothèse supplémentaire que $g(y)$ est fini pour presque tout y .

4) Enfin, pour une fonction f à valeurs complexes, on applique le théorème aux parties réelle et imaginaire de f .

Le lemme 4.4.3 permet de prouver le résultat suivant, bien utile pour montrer que certaines parties sont négligeables (par exemple le cercle unité de \mathbf{R}^2).

4.4.6 Corollaire. *Les notations sont celles de 4.4.1. Soit A un ensemble mesurable de \mathbf{R}^d . On suppose que pour presque tout $y \in \mathbf{R}^q$, la coupe A_y est négligeable. Alors A est négligeable.*

Démonstration. On applique 4.4.3 et on a $\mu(A) = \int_{\mathbf{R}^p} \mu_p(A_y)$. Mais, comme la fonction $\mu_p(A_y)$ est presque partout nulle, cette intégrale est nulle.

4.4.3 Utilisation pratique des théorèmes de Fubini

On considère une fonction $f : \mathbf{R}^p \times \mathbf{R}^q \rightarrow \mathbf{C}$. Pour appliquer Fubini-Lebesgue, il faut prouver que f est intégrable. Pour cela, on montre d'abord que f est mesurable, ce qui est facile (le plus souvent f sera continue, au moins presque partout). Pour voir que f est intégrable il suffit alors de montrer que $|f|$ l'est. Pour cela on va utiliser Fubini-Tonelli appliqué à $|f|$ (ou à une fonction g qui majore $|f|$). Cela signifie qu'on regarde la fonction $G(y) = \int_{\mathbf{R}^p} |f(x, y)| dx$, à valeurs dans $[0, +\infty]$. De deux choses l'une :

- Soit l'ensemble $Y = \{y \mid G(y) = \infty\}$ n'est pas négligeable. Alors, par Fubini-Lebesgue, on en déduit que $|f|$ n'est pas intégrable, donc f non plus.
- Soit Y est négligeable et on peut appliquer Fubini-Tonelli, on a donc :

$$\int_{\mathbf{R}^q} G(y) dy = \int_{\mathbf{R}^q \times \mathbf{R}^p} |f(x, y)| dx dy$$

et on est ramené à voir si cette intégrale est finie, c'est-à-dire si G est intégrable sur \mathbf{R}^q .

4.4.7 Remarques.

- 1) Bien entendu, dans Fubini, les variables x et y jouent des rôles symétriques.
- 2) Très souvent, ce qui est utile dans les applications, c'est d'invertir l'ordre des intégrations.
- 3) En itérant Fubini on peut théoriquement ramener toute intégrale multiple à des intégrales simples.
- 4) Le théorème de Fubini est valable pour des intégrales sur des parties mesurables E autres que \mathbf{R}^{p+q} .

4.4.8 Exemple. Soit a un réel > 0 . On considère la fonction définie sur $D := [0, a] \times \mathbf{R}^+$ par $f(x, y) = e^{-xy^2} \sin x$. Il s'agit de montrer que cette fonction est intégrable sur D . On considère la fonction \tilde{f} obtenue en prolongeant f par 0 à l'extérieur de D est intégrable. La fonction \tilde{f} est continue sauf peut-être sur la frontière F de D . Comme F est de mesure nulle (elle est contenue dans la réunion des droites $x = 0$, $x = a$ et $y = 0$), \tilde{f} est mesurable (cf. 4.1.26.1), donc aussi f (cf. 4.2.13). On majore alors $|f|$ par $g(x, y) = e^{-xy^2}$ (mesurable elle aussi) et il s'agit de voir que g est intégrable sur D .

- *Méthode 1.* On intègre d'abord en x . On a $G(y) = \int_0^a e^{-xy^2} dx$. On sait calculer cette intégrale : on a $G(y) = \frac{1 - e^{-ay^2}}{y^2}$. Il reste à voir que G est intégrable sur $[0, +\infty[$, mais G est continue en 0 (un développement limité montre qu'elle tend vers a) et elle est équivalente à $1/y^2$ à l'infini. L'intégrale généralisée est donc convergente et on a gagné.

- *Méthode 2.* On intègre d'abord en y . On a $F(x) = \int_0^{+\infty} e^{-xy^2} dy$. On calcule cette intégrale en faisant le changement de variables $u = \sqrt{x}y$. On se ramène à l'intégrale de Gauss $\int_0^{+\infty} e^{-u^2} du$, dont on sait qu'elle vaut $\sqrt{\pi}/2$. On a donc $F(x) = \frac{\sqrt{\pi}}{2\sqrt{x}}$ et cette fonction est bien intégrable sur $[0, a]$ (c'est une intégrale absolument convergente).

4.4.9 Exemple. L'exemple suivant est à méditer. On considère la fonction f définie sur $D = [0, 1]^2$ par :

$$f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2} \text{ pour } (x, y) \neq (0, 0) \text{ et } f(0, 0) = 0.$$

Appliquons Fubini-Lebesgue sans précautions pour calculer $I = \int_D f$. On calcule facilement (surtout avec une calculatrice!) :

$$F(x) = \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy = \frac{1}{1 + x^2},$$

puis $I = \int_0^1 F(x) dx = \pi/4$. De même, on a :

$$G(y) = \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx = -\frac{1}{1 + y^2},$$

puis $I = \int_0^1 G(y) dy = -\pi/4!$

Le problème c'est que la fonction f n'est pas intégrable sur D . En effet, si elle l'était, $|f|$ le serait aussi (tout est dans la valeur absolue!), donc elle serait aussi intégrable sur D' qui est la partie de D sur laquelle on a $x \geq y$. Calculons cette intégrale par Fubini-Tonelli. On doit considérer

$$H(x) = \int_0^x \frac{x^2 - y^2}{(x^2 + y^2)^2} dy = \left[\frac{y}{x^2 + y^2} \right]_0^x = \frac{1}{2x}$$

et on voit que $H(x)$ n'est pas intégrable sur $[0, 1]$.

4.4.4 Applications

Un graphe est négligeable

La proposition suivante donne de nombreux exemples d'ensembles négligeables :

4.4.10 Proposition. Soit $f : \mathbf{R}^d \rightarrow \mathbf{R}$ une fonction mesurable et soit $G(f)$ son graphe dans \mathbf{R}^{d+1} :

$$G(f) = \{(x, y) \in \mathbf{R}^d \times \mathbf{R} \mid y = f(x)\}.$$

Alors $G(f)$ est une partie négligeable de \mathbf{R}^{d+1} .

Démonstration. On considère la fonction $F : \mathbf{R}^{d+1} \rightarrow \mathbf{R}$ qui à (x, y) associe $f(x) - y$. Comme f est mesurable, on vérifie qu'il en est de même de F et on a $G(f) = F^{-1}(0)$, de sorte que $G(f)$ est mesurable.

Soit G_x la coupe de $G(f)$ pour $x \in \mathbf{R}^d$: $G_x = \{y \in \mathbf{R} \mid (x, y) \in G(f)\} = \{y \in \mathbf{R} \mid y = f(x)\}$. Comme f est une fonction, G_x est réduit à un point, donc négligeable. On conclut par le corollaire 4.4.6.

L'intégrale de Gauss, première méthode

Il s'agit de calculer $I = \int_{-\infty}^{+\infty} e^{-x^2} dx = 2 \int_0^{+\infty} e^{-x^2} dx$. Le changement de variables $x = \sqrt{t}$ donne $I = \int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}} dt$.

L'astuce est d'introduire, pour $t > 0$, $J = 2 \int_0^{+\infty} e^{-tx^2} dx$ et d'y faire $u = \sqrt{t}x$, ce qui donne $J = \frac{I}{\sqrt{t}}$, puis de calculer $K = 2 \int_{\mathbf{R}^+ \times \mathbf{R}^+} e^{-t(1+x^2)} dt dx$. Le théorème de Fubini-Tonelli s'applique : on intègre en t , on trouve $2/(1+x^2)$

que l'on intègre de 0 à l'infini, ce qui donne $K = \pi$. Si on intègre en x d'abord, on trouve $K = 2 \int_0^{+\infty} e^{-t} \left(\int_0^{+\infty} e^{-tx^2} dx \right) dt = \int_0^{+\infty} I \frac{e^{-t}}{\sqrt{t}} dt = I^2$. En définitive, on a $I = \sqrt{\pi}$.

Volume de la boule unité de \mathbf{R}^d

Soit $B_d(R)$ la boule de centre l'origine et de rayon R de \mathbf{R}^d pour la norme euclidienne :

$$B_d(R) = \{x = (x_1, \dots, x_d) \in \mathbf{R}^d \mid \sum_{i=1}^d x_i^2 \leq R^2\}.$$

Il s'agit de calculer $\mu_d(B_d(R))$. On considère, pour $y \in [-R, R]$ la coupe $B_y = \{x \in \mathbf{R}^{d-1} \mid (x, y) \in B_d(R)\}$. La condition donne $x_1^2 + \dots + x_{d-1}^2 + y^2 \leq R^2$, soit $x_1^2 + \dots + x_{d-1}^2 \leq R^2 - y^2$. On a donc $B_y = B_{d-1}(\sqrt{R^2 - y^2})$. On en déduit,

par Fubini : $\mu_d(B_d(R)) = \int_{-R}^R \mu_{d-1}(B_{d-1}(\sqrt{R^2 - y^2})) dy$. Mais, l'examen des

premiers cas : $\mu_1(B_1(R)) = 2R$, $\mu_2(B_2(R)) = \pi R^2$, etc. nous fait subodorer une formule du genre $\mu_d(B_d(R)) = a_d R^d$. On en déduit alors $\mu_{d+1}(B_{d+1}(R)) = \int_{-R}^R a_d (R^2 - y^2)^{d/2} dy$. L'intégrale se calcule avec le changement de variables

$y = R \sin t$ et on obtient la relation de récurrence : $a_{d+1} = 2a_d \int_0^{\pi/2} \cos^{d+1} t dt$.

On note la présence des intégrales de Wallis. Voilà les premières valeurs : $a_1 = 2$, $a_2 = \pi$, $a_3 = 4\pi/3$, $a_4 = \pi^2/2$, $a_5 = 8\pi^2/15$, d'où les volumes $2R$, πR^2 , $(4/3)\pi R^3$, $\pi^2 R^4/2$, $8\pi^2 R^5/15$.

L'intégrale de Fresnel

Il s'agit de calculer l'intégrale $F = \int_0^{+\infty} e^{ix^2} dx$ (c'est-à-dire la limite,

quand R tend vers $+\infty$, de $F_R = \int_0^R e^{ix^2} dx$). On commence par faire le

changement de variables $t = x^2$, qui montre qu'on a $F_R = \int_0^{R^2} \frac{e^{it}}{2\sqrt{t}} dt$. Par

ailleurs, on a, pour $t > 0$, $\int_0^{+\infty} e^{-tx^2} dx = \frac{\sqrt{\pi}}{2\sqrt{t}}$ (on se ramène à l'intégrale de Gauss par le changement de variables $u = x\sqrt{t}$). On en déduit qu'on a :

$$F_R = \frac{1}{\sqrt{\pi}} \int_0^{R^2} e^{it} dt \int_0^{+\infty} e^{-tx^2} dx.$$

Considérons l'intégrale $\int_D e^{it} e^{-tx^2} dt dx$, où D désigne l'ensemble $[0, R^2] \times [0, +\infty[$. En appliquant Fubini-Tonelli à $e^{-tx^2} = |e^{it} e^{-tx^2}|$, on voit que la fonction est intégrable, ce qui légitime l'application du théorème de Fubini-Lebesgue et permet d'inverser l'ordre des sommations. On obtient :

$$F_R = \frac{1}{\sqrt{\pi}} \int_0^{+\infty} dx \int_0^{R^2} e^{t(i-x^2)} dt = \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \frac{1 - e^{-R^2(x^2-i)}}{x^2 - i} dx.$$

Quand R tend vers $+\infty$, la fonction à intégrer $g_R(x)$ tend simplement vers $1/(x^2 - i)$ sauf en $x = 0$. De plus, on a $|g_R(x)| \leq \frac{2(x^2 + 1)}{x^4 + 1}$. Comme cette fonction est intégrable sur $[0, +\infty[$, on peut appliquer le théorème de convergence dominée qui montre que F_R tend vers $\frac{1}{\sqrt{\pi}} \int_0^{+\infty} \frac{1}{x^2 - i} dx$. L'intégrale $\int_0^{+\infty} \frac{1}{x^2 - i} dx$ se calcule en décomposant $\frac{1}{x^2 - i} = \frac{x^2 + i}{x^4 + 1}$ en éléments simples et on trouve $\frac{\pi}{2} e^{i\pi/4}$. On en déduit $F = \frac{\sqrt{\pi}}{2} e^{i\pi/4} = \frac{\sqrt{\pi}}{2\sqrt{2}} (1 + i)$.

4.5 Changement de variables

4.5.1 L'énoncé

4.5.1 Théorème. Soient U et V deux ouverts non vides de \mathbf{R}^d et $\varphi : U \rightarrow V$ un C^1 -difféomorphisme⁵. On note J_φ la matrice jacobienne de φ . Soit $f : V \rightarrow \mathbf{R}$ une fonction mesurable et positive (resp. Lebesgue-intégrable). Alors, la fonction $x \mapsto f(\varphi(x)) | \det J_\varphi(x) |$ est mesurable positive (resp. Lebesgue-intégrable) sur U et on a la formule :

$$\int_V f(y) dy = \int_U f(\varphi(x)) | \det J_\varphi(x) | dx.$$

Nous **admettrons** ce théorème. La démonstration utilise le théorème d'inversion locale. Nous démontrons seulement la variante linéaire, cf. ci-dessous.

4.5.2 Le cas d'un changement de variables linéaire

C'est le théorème précédent dans le cas où l'on a $U = V = \mathbf{R}^d$ et où φ est linéaire.

⁵C'est cela que l'on nomme changement de variables, au moins en dimension $d \geq 2$, car en dimension 1 on a un énoncé plus fort, voir 1.3.3.

4.5.2 Théorème. Soit $\varphi : \mathbf{R}^d \rightarrow \mathbf{R}^d$ une application linéaire bijective. Soit $f : \mathbf{R}^d \rightarrow \mathbf{R}$ une fonction mesurable et positive (resp. Lebesgue-intégrable). Alors, la fonction $x \mapsto f(\varphi(x)) |\det \varphi|$ est mesurable positive (resp. Lebesgue-intégrable) sur \mathbf{R}^d et on a la formule :

$$\int_{\mathbf{R}^d} f(y) dy = |\det \varphi| \int_{\mathbf{R}^d} f(\varphi(x)) dx.$$

4.5.3 Remarque. Le résultat est encore vrai si φ n'est pas inversible. Dans ce cas l'image de φ est contenue dans un hyperplan, donc est de mesure nulle, et on a $\det \varphi = 0$, de sorte que les deux membres sont nuls.

Démonstration. Elle va découler de la proposition suivante :

4.5.4 Proposition. Soit E une partie mesurable de \mathbf{R}^d . On a $\mu(\varphi(E)) = |\det \varphi| \mu(E)$.

Réduction de 4.5.2 à 4.5.4

Il suffit de montrer le théorème pour une fonction positive. En effet, on utilise ensuite la décomposition $f = f^+ - f^-$. La proposition n'est autre que le cas particulier $f = \chi_{\varphi(E)}$ du théorème (car on a $\chi_{\varphi(E)} \circ \varphi = \chi_E$). Si elle est vraie, on en déduit par linéarité que le théorème vaut pour une fonction étagée positive $f = \sum_{i=1}^n a_i \chi_{E_i}$. Mais, si f vérifie l'une des hypothèses du théorème, l'intégrale de f est la borne supérieure des intégrales des fonctions étagées $g \leq f$. Or, si g est une telle fonction, $g \circ \varphi$ est encore étagée et on a $g \circ \varphi \leq f \circ \varphi$. Inversement, si h est étagée et $h \leq f \circ \varphi$, $h \circ \varphi^{-1}$ est étagée et $h \circ \varphi^{-1} \leq f$. Le résultat s'obtient donc à partir du cas étagé par passage à la borne supérieure.

Preuve de 4.5.4

Il suffit de montrer la proposition lorsque E est un ensemble élémentaire. En effet, il est clair que les applications μ et $|\det \varphi| \mu \circ \varphi$ sont toutes deux σ -additives. Si elles coïncident sur les ensembles élémentaires, on en déduit aussitôt qu'elles coïncident aussi sur les ensembles de \mathcal{D} , donc sur les ensembles mesurables en vertu de 4.1.21 (et plus précisément de la condition $(*)$ vue dans la preuve de cet énoncé).

Nous aurons besoin de deux résultats d'invariance de la mesure de Lebesgue sur \mathbf{R} :

4.5.5 Lemme. Soit $A \subset \mathbf{R}$ une partie mesurable et soient α, β des réels non nuls. On a $\lambda(\alpha A) = |\alpha|\lambda(A)$ et $\lambda(\beta + A) = \lambda(A)$.

Démonstration. Les formules sont un cas particulier des suivantes : $\int_{\mathbf{R}} f(\alpha x)dx = \frac{1}{\alpha} \int_{\mathbf{R}} f(x)dx$ et $\int_{\mathbf{R}} f(\beta + x)dx = \int_{\mathbf{R}} f(x)dx$. Ces formules se montrent en se ramenant au cas des fonctions en escalier et en faisant un changement de variables.

Rappelons aussi le lemme d'algèbre suivant :

4.5.6 Lemme. Toute application linéaire inversible $\varphi : \mathbf{R}^d \rightarrow \mathbf{R}^d$ est composée d'une dilatation et de transvections, c'est-à-dire d'applications linéaires de matrices :

$$D(\alpha) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \alpha \end{pmatrix} \quad \text{et} \quad E_{i,j}(\beta) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & \beta & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 1 \end{pmatrix}$$

avec α et β non nuls, β étant en position (i, j) avec $i \neq j$.

Soit $A = A_1 \times \cdots \times A_{d-1} \times A_d$ un ensemble élémentaire. Pour montrer 4.5.4 il suffit de le montrer pour une dilatation et une transvection.

Il faut montrer d'abord que si φ est une dilatation $D(\alpha)$, on a $\mu(\varphi(A)) = |\alpha|\mu(A)$. Comme $\varphi(A)$ est égal à $A_1 \times \cdots \times A_{d-1} \times (\alpha A_d)$, cela résulte de 4.5.5.

Il faut montrer ensuite que le volume de A ne change pas si on lui applique une transvection. Quitte à permuter les coordonnées on peut supposer que la transvection est $E_{n-1,n}(\beta)$, autrement dit qu'on a :

$$\varphi(x_1, \dots, x_{d-1}, x_d) = (x_1, \dots, x_{d-1} + \beta x_d, x_d).$$

On calcule le volume de l'image $\varphi(A)$ par Fubini. On fixe la dernière coordonnée y . La coupe $\varphi(A)_y$ est l'ensemble des $(x_1, \dots, x_{d-1} + \beta y)$, avec $(x_1, \dots, x_{d-1}, y) \in A$. C'est donc exactement l'ensemble :

$$A_1 \times \cdots \times (\beta y + A_{d-1}),$$

dont la mesure (cf. 4.5.5) est égale à $\lambda(A_1) \cdots \lambda(A_{d-1})$. La fonction $y \mapsto \mu(\varphi(A)_y)$ est donc constante et on a :

$$\mu(\varphi(A)) = \int_{A_d} \mu(\varphi(A)_y) dy = \lambda(A_1) \cdots \lambda(A_{d-1})\lambda(A_d).$$

4.5.7 Corollaire. *Tout hyperplan de \mathbf{R}^d est négligeable (et donc toute partie contenue dans un hyperplan est négligeable).*

Démonstration. On sait (cf. 4.1.8) que la propriété est vraie pour un hyperplan de coordonnées, par exemple pour H_0 défini par $x_d = 0$. Si H est un hyperplan quelconque, il existe une application φ , composée d'une application linéaire et d'une translation, telle que l'on ait $\varphi(H_0) = H$. En vertu de 4.5.4 on a $\mu(H) = |\det \varphi| \mu(H_0) = 0$.

4.5.3 Application : l'intégrale de Gauss, deuxième méthode

Nous calculons l'intégrale de Gauss par une autre méthode. On considère l'intégrale $K = \int_{\mathbf{R}^2} e^{-x^2-y^2} dx dy$. Une application de Fubini montre que l'on a $K = I^2$, où $I = \int_{-\infty}^{+\infty} e^{-x^2}$ est l'intégrale de Gauss. Considérons l'application de \mathbf{R}^2 dans \mathbf{R}^2 définie par $\Phi(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta)$ (passage en coordonnées polaires). On vérifie que Φ induit un difféomorphisme de $D =]0, +\infty[\times]0, 2\pi[$ sur l'ouvert $\mathbf{R}^2 - A$ où A est l'ensemble $\{(x, y) \in \mathbf{R}^2 \mid y = 0 \text{ et } x \geq 0\}$. Le jacobien de Φ vaut ρ . On peut appliquer la formule de changement de variables qui donne (puisque A est de mesure nulle) $K = \int_D e^{-\rho^2} \rho d\rho d\theta$. Avec Fubini on a $K = \int_0^{2\pi} d\theta \int_0^{+\infty} e^{-\rho^2} \rho d\rho = \frac{1}{2} \int_0^{2\pi} d\theta = \pi$. On retrouve bien la formule $I = \sqrt{\pi}$.