Cours de mathématiques

TOME VIII

THÉORIE DE LA MESURE

Mathématiques générales

France ∼ 2024

Écrit et réalisé par Louis Lascaud

Chapitre 1

Intégration de Riemann

Résumé

L'intégrale de Riemann a été introduite dans l'article de Bernard Riemann *Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe*, « Sur la représentabilité d'une fonction par une série trigonométrique ». Riemann a présenté ce travail à l'université de Göttingen en 1854 comme mémoire d'habilitation. Il a été ensuite publié en 1868 dans les Actes de la Société royale des sciences de Göttingen, vol. 13. Première théorie de l'intégration historiquement, elle n'est pas sans défaut ; aussi faudra-t-il attendre le vingtième siècle pour que Lebesgue généralise l'intégrale de Riemann dans une théorie complète de la mesure réelle. Cependant, l'intégrale de Riemann, d'abord pour des fonctions bornées sur un intervalle borné, ensuite généralisée grâce aux intégrales impropres, est largement suffisante pour l'analyse réelle. En particulier, nous fournissons la construction originale de l'intégrale, applicable aux fonctions quelconques, et non seulement aux fonctions continues par morceaux comme on en a souvent l'habitude.

1.1 Définition de l'intégration des fonctions réelles de la variable réelle sur un segment

DANS toute la suite, on fixe deux réels a,b afin de définir le segment $[a,b]$.

Exercice 1

Vérifier logiquement que le segment $[a,b]$ est non trivial si et seulement si $a < b$.

▷ **Éléments de réponse.**

Écrire la définition d'un segment avec la relation d'ordre de $\mathbb{R}$. Par définition, un segment est trivial s'il est vide ou réduit à un singleton.

1.1.1 Idée

Exemple. ***(Intégrale du demi-cercle)***

Considérons la courbe d'équation $y = \sqrt{1-x^2}$ définie pour $x \in [-1,1]$. C'est le graphe d'une fonction f définie sur $[-1,1]$, positive sur $[-1,1]$.
On pose $D_f = \{(x,y) \in \mathbb{R}^2, 0 \leqslant y \leqslant f(x)\}$.

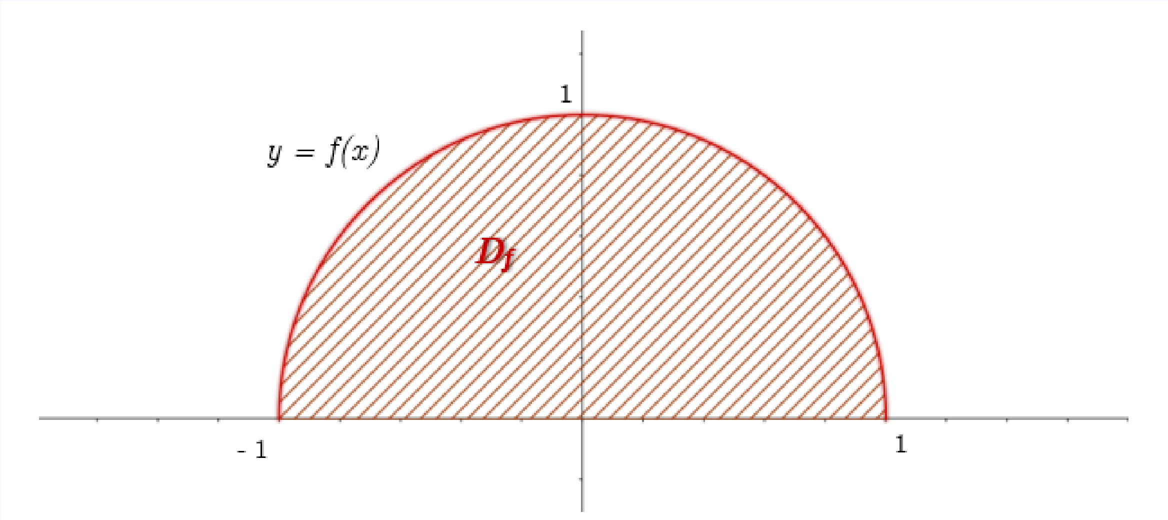

FIGURE A. *Aire sous la courbe du demi-cercle unité supérieure.* —

Nous allons définir, lorsque c'est possible, « l'aire » de D_f qui sera *l'intégrale* de f sur $[a,b] = [-1,1]$, notée $\displaystyle\int_{[a,b]} f$.
On trouve : $\displaystyle\int_{[a,b]} f = \frac{\pi}{2}$.

Considérons l'exemple général d'une fonction positive continue et bornée sur un segment $[a,b]$. En considérant le rectangle de longueur $[a,b]$ et de hauteur $[0, \sup(f)]$, son aire vaut $(b-a)\sup(f)$ et il est évident que cette quantité est un majorant de $\displaystyle\int_{[a,b]} f$; de même, en considérant le rectangle de longueur $[a,b]$ et de hauteur $[0, \inf(f)]$, son aire vaut $(b-a)\inf(f)$ et il est évident que cette quantité est un minorant de $\displaystyle\int_{[a,b]} f$. Considérons maintenant un troisième point $c \in]a,b[$. Alors on peut réitérer le procédé précédent en deux étapes, en encadrant chacun des deux morceaux de la courbe par les mêmes rectangles. On obtient alors, en sommant les aires de ces rectangles, que $(c-a)\sup(f_{|[a,c]}) + (b-c)\sup(f_{|[c,b]})$ est un majorant de $\displaystyle\int_{[a,b]} f$ intuitivement plus fin que le précédent et que $(c-a)\inf(f_{|[a,c]}) + (b-c)\inf(f_{|[c,b]})$ est un minorant de $\displaystyle\int_{[a,b]} f$ intuitivement plus fin que le précédent. En continuant, on voudra que $\displaystyle\int_{[a,b]} f$ soit le majorant le plus fin possible de la forme précédente (on introduira ainsi une borne inférieure) ou bien le minorant le plus fin possible de la forme précédente (on introduira ainsi une borne supérieure). **Pour que la construction soit viable, il faudra imposer que ces**

deux quantités soit égales.

Pour calculer cet aire, dans ce cas premier cas non dégénéré, l'idée naturelle est donc de l'encadrer par des polygones simples (et c'est possible, si l'on suppose la fonction bornée). Afin de raffiner progressivement ces encadrements, nous introduisons la notion de subdivision d'un segment.

1.1.2 Subdivisons d'un segment

1.1.2.1 Rappels sur les ensembles finis de réels

1.1.2.2 Treillis des subdivisions d'un segment

Remarque. Si B_1 représente] ou [et B_2 représente] ou [, alors pour tous réels $x \leqslant y$ et $z \leqslant t$, on a $B_1 x,y B_2 = B_1 z,t B_2$ si et seulement si $x = z$ et $y = t$.

Définition. ***(Subdivision d'un segment)***

On appelle *subdivision* du segment $[a,b]$ toute partie finie $\sigma \subseteq [a,b]$ telle que $a \in \sigma$ et $b \in \sigma$.
Autrement dit, l'ensemble des subdivisions de $[a,b]$ est défini comme :

$$\mathrm{Sub}([a,b]) = \mathrm{Sub}(a,b) := \{A \in \mathrm{Fin}([a,b]) \mid a,b \in A\}.$$

On définit d'ores et déjà une relation d'ordre sur $\mathrm{Sub}(a,b)$, permettant de *raffiner* une subdivision.

Proposition. ***(Relation de raffinement)***

Étant données $\sigma, \tau \in \mathrm{Sub}(a,b)$, on pose $\sigma \preceq \tau \iff \tau \subseteq \sigma$. Alors $\preceq$ est une relation d'ordre sur $\mathrm{Sub}(a,b)$.

▷ Conséquence directe du fait que $\subseteq$ est une relation d'ordre. ■

Attention à ne pas méprendre le sens de l'inclusion dans la relation de raffinement : une subdivision est *plus fine* qu'une autre si elle possède plus de points que celle-ci. Il est complètement bénin d'inverser le sens de cette relation d'ordre, mais alors les bornes inférieures et les bornes supérieures s'interchangent, ainsi que les sens de variation des fonctions dépendant de subdivisions *(voir dans la suite)*.

Propriété. ***(Structure de l'ensemble des subdivisions d'un segment)***

$(\mathrm{Sub}(a,b), \preceq)$ est un treillis.

▷ On rappelle qu'un treillis est un ensemble ordonné tel que toute paire possède une borne supérieure et une borne inférieure. Soient donc $\sigma, \tau \in \mathrm{Sub}(a,b)$. On remarque que $\inf(\sigma, \tau) = \sigma \cup \tau$ (oui !) et $\sup(\sigma, \tau) = \sigma \cap \tau$ (oui !) et que $\sigma \cup \tau, \sigma \cap \tau \in \mathrm{Sub}(a,b)$, car l'intersection et la réunion de parties finies sont finies et si σ et τ contiennent a et b, leur réunion et leur intersection également. ■

Définition. ***(Raffinement commun)***

Le *raffinement commun* à un nombre fini de subdivisions $\sigma_1, ..., \sigma_p \in \mathrm{Sub}(a,b)$ est la borne inférieure de l'ensemble $\{\sigma_1, ..., \sigma_p\}$ pour $\preceq$.

Exercice 2

Montrer que le raffinement commun à σ_1,...,σ_p est $\bigcup_{i=1}^{p} \sigma_i$.

Dans la suite, on fixe σ une subdivision du segment $[a,b]$.

Définition. ***(Hauteur d'une subdivision)***

On appelle *hauteur* de la subdivision σ l'application $h_\sigma : \sigma \longrightarrow \mathbb{N}$ définie pour tout $x \in \sigma$ par $h_\sigma(x) = \mathrm{card}(y \in \sigma, y < x)$.

Exercice 3

Montrer les propositions suivantes :

1. $h_\sigma(a) = 0$;

2. $h_\sigma(b) + 1 = \mathrm{card}(\sigma)$;

3. h_σ est strictement croissante.

▷ **Éléments de réponse.**

Pour le dernier point, prendre $t,x \in \sigma$ tels que $t < x$ et considérer $\{y \in \sigma, y < t\} \cup \{t\}$.

$\longrightarrow$ *Notation.* Dans toute la suite, on écrit $\sigma \setminus \{b\} = \sigma^\bullet$ la subdivision privée de la borne supérieure du segment sur lequel elle est définie.

Définition. ***(Fonction successeur)***

On définit une fonction dite *successeur* sur la subdivision pointée en b, c'est-à-dire $\sigma^\bullet$, à valeurs dans σ, qui à tout point x fait correspondre $\min\{y > x, y \in \sigma\}$. On note suc_σ ou s_σ cette fonction.

Exercice 4

Justifier la bonne définition de la fonction successeur.

▹ **Éléments de réponse.**

Se rappeler la propriété fondamentale de $\mathbb{N}$.

Lemme

$\forall x \in \sigma^\bullet,\ h_\sigma(\mathrm{suc}_\sigma(x)) = h_\sigma(x) + 1.$

▹ Soit $x \in \sigma^\bullet$. La stricte croissance de la fonction hauteur nous donne déjà $h_\sigma(\mathrm{suc}_\sigma(x)) \geqslant h_\sigma(x) + 1$, car $\mathrm{suc}_\sigma(x) \geqslant x$ par définition et h_σ est à valeurs discrètes.

Si maintenant $y < \mathrm{suc}_\sigma(x)$ et si $y \in \sigma$, $y \neq x$, alors $y < x \implies \{y \in \sigma, y < x\} \cup \{x\} \supseteq \{y \in \sigma, y < \mathrm{suc}_\sigma(x)\}$, donc $\{y \in \sigma, y < x\} \cup \{x\} = \{y \in \sigma, y < \mathrm{suc}_\sigma(x)\}$ et on conclut en passant au cardinal. ■

Définition. ***(Longueur d'une subdivision)***

On définit la *longueur* d'une subdivision par $\ell(\sigma) := h_\sigma(b)$.

Définition. ***(Vision alternative des subdivisions)***

Toute subdivision est en bijection avec l'intervalle d'entier $[\![0, \ell(\sigma)]\!] = [0..\ell(\sigma)]$ par l'application h_σ.

▹ En effet, $h_\sigma : \sigma \longrightarrow [\![0, \ell(\sigma)]\!]$ clairement par croissance de l'application. h_σ est même strictement croissante, en particulier elle est injective, car σ est totalement ordonné. Il ne reste plus qu'à montrer qu'elle est surjective. Supposons le contraire : supposons qu'il existe un entier m, $0 \leqslant m \leqslant \ell(\sigma)$ tel que pour tout $x \in \sigma$, $m \neq h_\sigma(x)$. Prenons le plus petit de ceux-là (toute partie non vide de $\mathbb{N}$ admet un plus petit élément). Alors $m \neq 0$, car $0 = h_\sigma(a)$. Ainsi $0 \leqslant m - 1 \leqslant \ell(\sigma)$ donc il existe un point $x \in \sigma$ tel que $h_\sigma(x) = m - 1$. Mais d'après le lemme, $h_\sigma(\mathrm{suc}_\sigma(x)) = m$, et $\mathrm{suc}_\sigma(x) \in \sigma$, ce qui est absurde. ■

Conséquence

Toute subdivision σ de longueur $\ell(\sigma) = n$ peut être décrite par la suite finie $(x_i)_{i \in [\![0,n]\!]}$ telle que $h_\sigma(x_i) = i$ pour tout i, où :

$$a = x_0 < x_1 < \ldots < x_n = b.$$

Ne pas confondre l'ordre naturellement défini sur une subdivision σ, qui est l'ordre induit de celui sur tous les réels de $[a,b]$, de l'ordre « être plus fin que » défini sur l'ensemble des subdivisions du segment $[a,b]$, soit $\mathrm{Sub}(a,b)$. Si le premier est évidemment total, **le second est partiel** (<u>le vérifier</u>).

Définition. ***(Pas d'une subdivision)***

Le *pas* de la subdivision σ est défini comme le nombre $\max\{s_\sigma(x)-x, x\in\sigma^\bullet\}$ (il est existe, car $\sigma^\bullet$ est fini). On note : $p(\sigma)$.

Exercice 5

(Un exemple fondamental) Montrer que, pour tout segment $[a,b]$, pour tout entier naturel n, il existe une unique subdivision régulière (c'est-à-dire : telle que $s_\sigma(x)-x$ soit constant pour $x\in\sigma$) du segment $[a,b]$ à n points.

▷ **Éléments de réponse.**

Commencer par exhiber cette subdivision. (C'est facile, car très intuitif.) Quel est son pas et que peut-on en dire ? Pour l'unicité, supposer qu'il y en a deux.

Définition. ***(Pic d'une subdivision))*** **HP**

On appelle *pic* de σ tout entier $i\in[\![1,\ell(\sigma)]\!]$ tel que $x_{i+1}-x_i=p(\sigma)$.

Exercice 6

Vérifier que toute subdivision non triviale (c'est-à-dire, non vide ni réduite à un singleton) admet un pic.

Astuce !

Il est possible de construire les sommes de Darboux, et donc, l'intégrale de Riemann, avec la vision ensembliste comme avec la vision ordonnée des subdivisions. L'avantage de la vision ensembliste, malgré la lourdeur des notations dans la somme (notamment la présence de la fonction successeur), est de nettement clarifier les concepts. L'exemple le plus frappant est de ne pas introduire de sommation jusqu'à un certain n qui, dans la vision ordonnée des subdivisions, dépend bel et bien de σ : $n=n(\sigma)=\ell(\sigma)$.

Définition. ***(Convergence au sens des subdivisions)*** **HP**

Soit g une application de la variable $\sigma\in\mathrm{Sub}(a,b)$. On dit que g converge vers u au sens des subdivisions si $\forall\varepsilon>0\ \exists\eta>0\ \forall\sigma\in\mathrm{Sub}(a,b)\quad p(\sigma)<\eta\implies|g(\sigma)-u|<\varepsilon$, ce qu'on écrit : $\lim\limits_{p(\sigma)\to+\infty} g(\sigma)=u$.

1.1.3 Sommes de Darboux

Dans toute la suite, on considère une fonction $f:[a,b]\longrightarrow\mathbb{R}$ **bornée**.

1.1.3.1 Définition

Définition. ***(Somme de Darboux associéesà une subdivision)***

Soit $\sigma \in \mathrm{Sub}(a,b)$.

On définit la *somme de Darboux inférieure de f associée à σ* par :

$$s(f,\sigma) = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) \inf \left(f_{|[x,s_\sigma(x)[}\right)$$

et la *sommes de Darboux supérieure de f associées à σ* par :

$$S(f,\sigma) = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) \sup \left(f_{|[x,s_\sigma(x)[}\right).$$

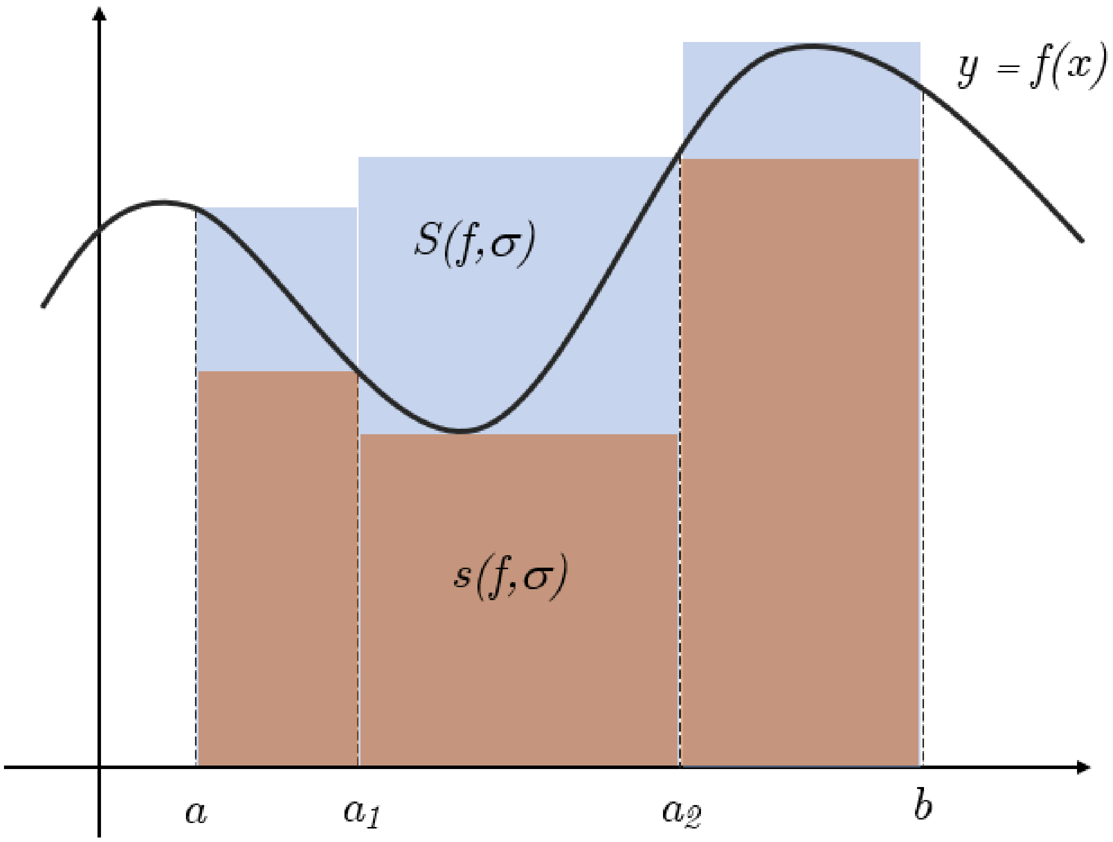

Figure 1.1.2 : *Sommes de Darboux inférieures (en rouge) et supérieures (en bleu) superposées pour une fonction visiblement continue sur un segment subdivisé en trois sous-intervalles.* — Les grandeurs considérées sont donc sommes d'aires de rectangles. Les largeurs de ces rectangles sont déterminées par l'écart entre deux points de la subdivision et leur hauteur est déterminée, disons, dans le cas des sommes de Darboux inférieures, par l'infimum de la fonction sur ce sous-intervalle. Lorsque la fonction n'est pas continue, il ne coïncide pas nécessairement avec la plus petite valeur prise par la fonction.

Exercice 7

Expliciter les sommes de Darboux inférieures et supérieures pour la subdivision régulière d'ordre n.

▷ **Éléments de réponse.**

Il suffit de prendre la définition.

Propriété **HP**

Lorsque f est continue, pour tout $x \in \sigma^\bullet$, $\sigma \in \mathrm{Sub}(a,b)$, $\inf f_{|[x,s_\sigma(x)[}$ est la plus petite valeur prise par f sur $[x, s_\sigma(x)]$.

▷ Soit $\sigma \in \mathrm{Sub}(a,b)$. Soit $x \in \sigma^\bullet$. On rappelle que la borne inférieure d'une fonction sur un ensemble A n'est autre que la borne inférieure de l'ensemble $f(A)$. Si f est continue sur $[a,b]$, elle l'est également sur $[x, s_\sigma(x)]$. Or cet intervalle est un segment, donc $f_{|[x,s_\sigma(x)]}$ est continue sur un segment ; d'après le théorème des bornes atteintes (théorème de compacité), f est bornée sur $[x, s_\sigma(x)]$ et atteint ses bornes. En particulier elle admet un minimum (et pas seulement une borne inférieure, le minimum étant par définition atteint) en $x_0 \in [x, s_\sigma(x)]$. Si $x_0 \neq s_\sigma(x)$, c'est terminé : en effet, $f(x_0)$ est un minimum dans $f([x, s_\sigma(x)[)$, donc en particulier sa borne inférieure. Sinon, on considère une suite (x_n) à valeurs dans $[x, s_\sigma(x)[$ tendant vers $s_\sigma(x)$, qui existe car le segment n'est autre que l'adhérence de cet intervalle. Par continuité de f, $f(x_n)$ tend vers $f(s_\sigma(x))$ lorsque n tend vers $+\infty$, les $f(x_n)$ étant tous dans $f([x, s_\sigma(x)[)$. De plus, par définition de la borne inférieure, $y \geqslant f(s_\sigma(x))$ pour tout $y \in f([x, s_\sigma(x)[)$. Par caractérisation séquentielle de la borne inférieure, $f(s_\sigma(x))$ est la borne inférieure de $f([x, s_\sigma(x)[)$, ce qu'il fallait montrer. (Remarquons que la séparation du cas $x_0 \neq s_\sigma(x)$ est inutile.) ■

Remarque. On a de façon identique le résultat analogue pour la borne supérieure : lorsque f est continue, $\sup f_{|[x,s_\sigma(x)[}$ est la plus grande valeur prise par f sur $[x, s_\sigma(x)]$.

Exercice 8

Donner un contre-exemple dans le cas discontinu.

▷ **Éléments de réponse.**

Exhiber une fonction continue par morceaux sur un segment, décroissante sur l'intérieur et dont la borne inférieure sur le segment entier n'est pas égal à la valeur de f en b.

Proposition

Pour toute subdivision σ de $[a,b]$, $\displaystyle\sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) = b - a$.

▷ D'après la vision ordonnée des subdivisions, cette somme est télescopique... ■

Propriété. ***(Sens de variations des sommes de Darboux)***

On a :

- $s(f,\,\cdot\,)$ est décroissante pour la relation de raffinement (et donc croissante pour la relation d'inclusion sur $\mathrm{Sub}(a,b)$) ;
- $S(f,\,\cdot\,)$ est croissante pour la relation de raffinement (et donc décroissante pour la relation d'inclusion sur $\mathrm{Sub}(a,b)$).

$\triangleright$ Soit f une fonction bornée. Soient σ, τ deux subdivisions telles que $\tau \preceq \sigma$, doit $\sigma \subseteq \tau$. Montrons que $s(f,\tau) \geqslant s(f,\sigma)$.

Si $x \in \sigma^\bullet$, on note $T_x = \{y \in \sigma, x \leqslant y < s_\sigma(x)\}$, ensemble contenant x. On voit aisément que la famille des T_x pour x parcourant $\sigma^\bullet$ partage $\tau^\bullet$ (c'en est une partition à parties éventuellement vides). Autrement dit $\tau^\bullet$ est l'union disjointe $\sqcup_{x\in\sigma^\bullet} T_x$: on dit qu'on partage la subdivision en paquets. Ainsi, on peut écrire

$$\begin{aligned} s(f,\tau) &= \sum_{y\in\tau^\bullet} (s_\tau(y) - y) \inf f_{|[y,s_\tau(y)[} \\ &= \sum_{x\in\sigma^\bullet} \sum_{y\in T_x} (s_\tau(y) - y) \inf f_{|[y,s_\tau(y)[}. \end{aligned}$$

Or si $y \in T_x$, alors $[y, s_\tau(y)[\subseteq [x, s_\sigma(x)[$ donc $\inf f_{|[y,s_\tau(y)[} \geqslant \inf f_{|[y,s_\sigma(y)[}$. Par conséquence,

$$s(f,\tau) \geqslant \sum_{x\in\sigma^\bullet} \inf f_{|[y,s_\tau(y)[} \sum_{y\in T_x} s_\tau(y) - y \geqslant s(f,\sigma)$$

puisque $\sum\limits_{y\in T_x} s_\tau(y) - y = s_\sigma(x) - x$.

On fait de même pour les sommes de Darboux supérieures. ■

Corollaires

1. Pour toute subdivision σ, $s(f,\sigma) \leqslant S(f,\sigma)$.
2. Pour toutes subdivisions σ, τ, $s(f,\sigma) \leqslant S(f,\tau)$.

$\triangleright$ Pour le premier point, il suffit de constater que pour tout $x \in \sigma^\bullet$, $\inf f_{|[x,s_\sigma(x)[} \leqslant \sup f_{|[x,s_\sigma(x)[}$ (c'est toujours le cas pour un même ensemble). Pour le second point, on introduit la borne inférieure pour la relation de raffinement $\sigma \cup \tau$. En effet, $s(f,\sigma) \leqslant s(f,\sigma \cup \tau) \leqslant S(f,\sigma \cup \tau) \leqslant S(f,\tau)$, la deuxième inégalité découlant du premier point. ■

Corollaires

On suppose $a \leqslant b$.

1. L'ensemble $\{s(f,\sigma), \sigma \in \mathrm{Sub}(a,b)\}$ est majoré.
2. L'ensemble $\{S(f,\sigma), \sigma \in \mathrm{Sub}(a,b)\}$ est minoré.
3. Les quantités $\sup\limits_{\sigma\in\mathrm{Sub}(a,b)} s(f,\sigma)$ et $\inf\limits_{\sigma\in\mathrm{Sub}(a,b)} S(f,\sigma)$ existent.

4. De plus $\sup_{\sigma\in\mathrm{Sub}(a,b)} s(f,\sigma) \leqslant \inf_{\sigma\in\mathrm{Sub}(a,b)} S(f,\sigma)$.

▷ Successivement :

1. Il existe au moins une subdivision du segment $[a,b]$, car $a \leqslant b$. On la note σ_0. D'après le second point du corollaire précédent, pour toute subdivision σ, on a $s(f,\sigma) \leqslant S(f,\sigma_0)$ majorant constant de l'ensemble considéré.
2. Même chose.
3. C'est une conséquence directe des deux points précédents et de la propriété de la borne supérieure dans $\mathbb{R}$.
4. C'est un double passage aux bornes : pour toutes subdivisions σ,τ, $s(f,\sigma) \leqslant S(f,\tau)$. En passant au supremum sur $\sigma \in \mathrm{Sub}(a,b)$, on obtient, comme il existe, $\sup_{\sigma\in\mathrm{Sub}(a,b)} s(f,\sigma) \leqslant S(f,\tau)$ pour toute subdivision τ. On fait de même en passant à l'infimum sur τ, d'où le résultat (les variables sont bien évidemment muettes !). ■

Définition. ***(Mesure d'inintégrabilité)*** **HP**

On appelle *mesure d'inintégrabilité* de f la quantité, positive d'après ce qui précède :

$$\overline{\mu}(f) = \inf_{\sigma\in\mathrm{Sub}(a,b)} S(f,\sigma) - \sup_{\sigma\in\mathrm{Sub}(a,b)} s(f,\sigma).$$

1.1.3.2 Notions d'oscillation

Définition. ***(Oscillation sur un ensemble)***

L'oscillation d'une fonction f bornée sur une partie $E \subseteq [a,b]$ est la quantité :

$$\mathrm{Osc}_E(f) = \sup f_{|E} - \inf f_{|E} \geqslant 0.$$

Cette quantité est bien définie, car la fonction f est supposée bornée sur $[a,b]$ donc en particulier sur E.

Propriété. ***(Différences de Darboux)***

Soit f une fonction bornée sur $[a,b]$. Soit σ une subdivision de $[a,b]$. Alors :

$$\delta(f,\sigma) = S(f,\sigma) - s(f,\sigma) = \sum_{x\in\sigma^\bullet} \big(s_\sigma(x) - x\big)\mathrm{Osc}_{[x,s_\sigma(x)[}(f).$$

▷ C'est une simple réécriture de la définition. ■

Propriété. ***(Définition alternative de la mesure d'inintégrabilité)*** **HP**

$\overline{\mu}(f) = \inf\limits_{\sigma \in \mathrm{Sub}(a,b)} \delta(f,\sigma)$.

▷ Le faire. ■

Définition. ***(Oscillation locale)***

L'oscillation d'une fonction f bornée en un point $x \in [a,b]$ est la quantité :

$$\mathrm{Osc}(f,x) = \inf_{\alpha \in \mathbb{R}_+^*} \{\mathrm{Osc}_{]x-\alpha,x+\alpha[\cap[a,b]}(f)\}$$

Nous nous permettons, dans toute la suite, de ne pas préciser que nous intersectons le voisinage de x avec l'intervalle $[a,b]$.

Exercice 9

Que faut-il supposer de plus faible sur la fonction pour définir l'oscillation locale de f en tout point de $[a,b]$?

1.1.4 Fonctions intégrables

On fixe $[a,b]$ un segment non vide de $\mathbb{R}$.

Définition. ***(Intégrabilité d'une fonction sur un segment)***

Une fonction $f : [a,b] \longrightarrow \mathbb{R}$ est dite *intégrable au sens de Riemann* sur $[a,b]$ si elle est bornée et sa mesure d'inintégrabilité est nulle. Autrement dit, il faut et il suffit que :

1. f soit bornée ;
2. $\sup\limits_{\sigma \in \mathrm{Sub}(a,b)} s(f,\sigma) = \inf\limits_{\sigma \in \mathrm{Sub}(a,b)} S(f,\sigma)$.

Cette terminologie, « intégrable » ne semble pas compatible avec la définition de l'intégrabilité sur les intervalles quelconques : on définira l'intégrabilité sur I par la convergence *absolue* de l'intégrale de f sur l'intervalle I, condition plus forte que la simple convergence de l'intégrale.

Définition. ***(Intégrale d'une fonction sur un segment)***

On appelle *intégrale* sur $[a,b]$ d'une fonction $f : [a,b] \longrightarrow \mathbb{R}$ intégrable au sens de Riemann sur $[a,b]$ la quantité $\sup_{\sigma \in \mathrm{Sub}(a,b)} s(f,\sigma) = \inf_{\sigma \in \mathrm{Sub}(a,b)} S(f,\sigma)$. On note

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

Propriété. ***(Caractérisation analytique de l'intégrabilité parmi les fonctions bornées)***

Soit f une fonction bornée. Alors f est intégrable si et seulement si : $\forall \varepsilon > 0\ \exists \sigma, \tau \in \mathrm{Sub}(a,b) \quad S(f,\sigma) \leqslant s(f,\tau) + \varepsilon$ (ou, ce qui est équivalent, $\forall \varepsilon > 0\ \exists \sigma \in \mathrm{Sub}(a,b) \quad S(f,\sigma) \leqslant s(f,\sigma) + \varepsilon$).

▷ On laisse au lecteur le soin de vérifier que les deux dernières formules sont équivalentes. Pour la première équivalence, le sens réciproque est évident ; pour l'autre implication, faire intervenir $\sigma \cup \tau$. ■

Propriété. ***(Caractérisation asymptotique de l'intégrabilité parmi les fonctions bornées)***

Soit f une fonction bornée. Alors f est intégrable si et seulement si $S(f, \cdot\,) - s(f, \cdot\,)$ converge vers 0 au sens des subdivisions.

▷ Supposons : $\forall \varepsilon > 0\ \exists \eta > 0\ \forall \sigma \in \mathrm{Sub}(a,b) \quad p(\sigma) < \eta \implies S(f,\sigma) - s(f,\sigma) < \varepsilon$. Soit $\varepsilon > 0$. On considère la subdivision régulière d'ordre N : $\sigma_N = \{a, a + \frac{b-a}{N}, ..., a + (N-1)\frac{b-a}{N}, b\}$. Le pas de σ_N est $\frac{b-a}{N}$ qui tend vers 0 lorsque N tend vers $+\infty$. Si $p(\sigma_N) < \eta_\varepsilon$, alors $S(f,\sigma_N) - s(f,\sigma_N) < \varepsilon$ par hypothèse. D'après la caractérisation précédente, f est intégrable.

Réciproquement, supposons f intégrable. Il existe donc une subdivision $\sigma \in \mathrm{Sub}(a,b)$ telle que $S(f,\sigma) - s(f,\sigma) < \frac{\varepsilon}{2}$ d'après la caractérisation précédente. Posons $M = \mathrm{Osc}_{[a,b]}(f)$. Si f est identiquement nulle, le résultat est évident. Sinon, $M > 0$. Posons également $\eta_0 = \frac{\varepsilon}{2\ell(\sigma)M}$, qui existe dans $\mathbb{R}$ car $\ell(\sigma) \geqslant 1$ par définition d'une subdivision. Posons $\eta_1 = \min\{s_\sigma(x) - x, x \in \sigma^\bullet\}$ et enfin $\eta = \min(\eta_0, \eta_1)$. Soit τ une subdivision telle que $p(\tau) < \eta$. Alors $S(f,\tau) - s(f,\tau) = \sum_{y \in \tau^\bullet} (s_\tau(y) - y)\mathrm{Osc}_{[y, s_\tau(y)[}(f)$. On définit $A = \{y \in \tau^\bullet, [y, s_\tau(y)[\cap \sigma \neq \emptyset\}$ et $B = t^\bullet \setminus A$. Dans ce cas, $S(f,\tau) - s(f,\tau) = \sum_{y \in A} (s_\tau(y) - y)\mathrm{Osc}_{[y, s_\tau(y)[}(f) + \sum_{y \in B} (s_\tau(y) - y)\mathrm{Osc}_{[y, s_\tau(y)[}(f)$. Définissons une application sur A qui à un élément $y \in A$ fait correspondre un point $x(y)$ de σ. On peut écrire : $\sum_{y \in A} (s_\tau(y) - y)\mathrm{Osc}_{[y, s_\tau(y)[}(f) \leqslant \sum_{y \in A} (s_\tau(x(y)) - x(y))\mathrm{Osc}_{[x(y), s_\tau(x(y))[}(f) \leqslant S(f,\sigma) - s(f,\sigma) \leqslant \frac{\varepsilon}{2}$. D'autre part, $\sum_{y \in B} (s_\tau(y) - y)\mathrm{Osc}_{[y, s_\tau(y)[}(f) \leqslant M \sum_{y \in B} (s_\tau(y) - y) \leqslant Mp(\tau) \sum_{y \in B} 1$. Or f de B dans $\sigma^\bullet$ qui à y fait correspondre $x \in [y, s_\tau(y)[\cap \sigma$ est une injection, donc le cardinal de B est majoré par $\ell(\sigma)$, d'où la majoration de la somme sur B par $Mp(\tau)\ell(\sigma) \leqslant M\eta_0(\sigma) = \frac{\varepsilon}{2}$. En sommant, on obtient bien $S(f,\tau) - s(f,\tau) < \varepsilon$. ■

1.1.5 Cas des fonctions en escalier

Théorème

Toutes les fonctions constantes sont intégrables.

▷ Trivial. Les infima et les suprema dans les formules des sommes de Darboux sont constants. Il ne reste plus que $\sum_{x\in\sigma^\bullet}(s_\sigma(x) - x) = b - a$; les sommes de Darboux sont toutes les mêmes. ■

Définition. ***(Fonction en escalier (fee))***

Une fonction $f : [a,b] \longrightarrow \mathbb{R}$ est dite *en escalier* s'il existe une subdivision $\sigma \in \mathrm{Sub}(a,b)$ telle que pour tout $x \in \sigma^\bullet$, $f_{|]x,s_\sigma(x)[}$ soit constante. Il revient à dire que pour tout $x \in \sigma^\bullet$, il existe $c_x \in \mathbb{R}$ tel que pour tout $t \in \mathbb{R}$, $x < t < s_\sigma(x) \implies t = c_x$. On dit que σ est *adaptée* à f.

Une fonction est en escalier, s'il existe une subdivision en sous-segments sur chacun desquels la fonction est constante *à l'intérieur*. A priori, les valeurs sur les points de la subdivision (en nombre fini, donc ça ne pose pas souvent problème) sont quelconques. Plus généralement, les classes de fonctions définies par morceaux sont telles que sur chaque sous-segment de la subdivision, la régularité n'est exigée que par prolongement par continuité sur l'adhérence de l'intérieur, les valeurs aux points extrémaux étant encore quelconques.

Proposition

Soit f une fonction en escalier, σ une subdivision adaptée à f. Toute subdivision plus fine que σ est encore adaptée à f.

▷ En exercice. ■

Remarque. Ce résultat est valable pour toutes les classes de fonctions définies par morceaux (continuité par morceaux, fonctions affines par morceaux, etc.).

Théorème

Toutes les fonctions en escalier sont intégrables.

▷ La fonction f est bornée, car elle ne prend qu'un nombre fini de valeurs. Soit σ une subdivision adaptée à f. Soit $n \in \mathbb{N}^*$. On pose $\tau_n = \sigma \cup \bigcup_{x\in\sigma^\bullet}\{x + \frac{\eta}{2n\ell(\sigma)}\}$, en posant $\eta = \min(s_\sigma(x) -$

$x, x \in \sigma^\bullet$). Pour tout $x \in \sigma^\bullet$, $s_{\tau_n}(x) = x + \frac{\eta}{2n\ell(\sigma)} < x + \eta \leqslant s_\sigma(x)$, puis $s_{\tau_n}(x + \frac{\eta}{2n\ell(\sigma)}) = s_\sigma(x)$. Nous écrivons astucieusement :

$$S(f,\tau_n) - s(f, tau_n) = \sum_{x\in\sigma^\bullet} \frac{\eta}{2n\ell(\sigma)} \mathrm{Osc}_{[x,x+\frac{\eta}{2n\ell(\sigma)}[}(f) + \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x - \frac{\eta}{2n\ell(\sigma)}) \mathrm{Osc}_{[x+\frac{\eta}{2n\ell(\sigma)}, s_\sigma(x)[}(f).$$

Cette dernière oscillation est nulle. Il ne reste plus qu'à majorer le premier terme, grossièrement : $S(f,\tau_n) - s(f,\tau_n) \leqslant \frac{\eta}{2n\ell(\sigma)} \mathrm{Osc}_{[a,b]}(f) \sum_{x\in\sigma^\bullet} 1$. Or $\sum_{x\in\sigma^\bullet} 1 = \ell(\sigma)$. Ainsi $S(f,\tau_n) - s(f,\tau_n) \leqslant \frac{\eta \mathrm{Osc}_{[a,b]}(f)}{2n} < \varepsilon$ pour n assez grand. ■

Exercice 10

Quel est l'intégrale d'une fonction en escalier ?

INDICATION Expliciter les sommes de Darboux inférieures.

▷ **Éléments de réponse.**

En reprenant les notations précédentes, on écrit $s(f, \tau_n) = \sum_{x\in\sigma^\bullet} \frac{\eta}{2n\ell(\sigma)} \inf f_{|[x,x+\frac{\eta}{2n\ell(\sigma)}[} + \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x - \frac{\eta}{2n\ell(\sigma)}) c_x = \frac{\eta}{2n\ell(\sigma)} \sum_{x\in\sigma^\bullet} (\min(f(x), c_x) - c_x) + \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) c_x$ d'où $\sup_{n\geqslant 1} s(f, \tau_n) = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) c_x \leqslant \int_a^b f \leqslant \inf_{n\geqslant 1} S(f,\tau_n) = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) c_x$, car $S(f,\tau_n) = \frac{\eta}{2n\ell(\sigma)} \sum_{x\in\sigma^\bullet} (\max(f(x), c_x) - c_x) + \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) c_x$. Ainsi $\int_a^b f = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) c_x$.

1.1.6 Cas des fonctions monotones

Théorème

Toute fonction monotone sur $[a,b]$ est intégrable.

▷ Si f est constante, le cas a déjà été traité. On suppose donc $f(a) \neq f(b)$. Sans perte de généralité, on suppose f croissante. (On peut voir aisément que l'intégrabilité de f équivaut à l'intégrabilité de $-f$). f est bornée par $f(b)$. Pour toute subdivision $\sigma \in \mathrm{Sub}(a,b)$, $S(f,\sigma) - s(f,\sigma) = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x) \mathrm{Osc}_{[x,s_\sigma(x)[}(f)$. Si $t \in [x, s_\sigma(x)[$, alors $f(x) \leqslant f(t) \leqslant f(s_\sigma(x))[$, donc $\mathrm{Osc}_{[x,s_\sigma(x)[}(f) \leqslant f(s_\sigma(x)) - f(x)$. Ainsi $\delta(f,\sigma) \leqslant \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x)(f(s_\sigma(x)) - f(x))$. Pour $p(\sigma) < \eta$, on a donc $\delta(f,\sigma) \leqslant \eta \sum_{x\in\sigma^\bullet} (f(s_\sigma(x)) - f(x))$. Or cette somme égale $\sum_{x\in\sigma^\bullet} f(s_\sigma(x)) - \sum_{x\in\sigma^\bullet} f(x) = f(b) - f(a)$ par télescopage. On n'a qu'à prendre $\eta \leqslant \frac{\varepsilon}{2[f(b)-f(a)]}$. ■

Exercice 11

1. Une fonction monotone peut-elle être discontinue en tout point ?
2. Une fonction monotone peut-elle être discontinue sur une partie dense ?

▷ **Éléments de réponse.**

La réponse à la première question est, de façon assez attendue, non : on montre plus fortement que l'ensemble des points de discontinuité d'une fonction monotone est fini ou dénombrable. Dans le cas f croissante, si f n'est continue ni en a, ni en b, les intervalles $]\lim_{x\to a^-} f(x), \lim_{x\to a^+} f(x)[$ et $]\lim_{x\to b^-} f(x), \lim_{x\to b^+} f(x)[$ sont disjoints. Puisque $\mathbb{Q}$ est dense, on peut définir une injection de l'ensemble des points de discontinuité de f dans $\mathbb{Q}$ et le tour est joué. Un intervalle non trivial n'étant jamais dénombrable, une fonction monotone sur $[a,b]$ ne peut être discontinue partout.

On peut imaginer alors qu'il existera des fonctions monotones dont le lieu de discontinuité serait un dénombrable dense, par exemple, $\mathbb{Q}$. Pour cela, il faut avoir un peu d'idée. Puisque $\mathbb{Q}$ est dénombrable, il suffit de prendre une énumération $(q_n)_{n\in\mathbb{N}}$ des rationnels, et alors la fonction $x \longmapsto \sum_{n=0} +\infty \frac{1}{2^n} \mathbb{1}_{[q_n+\infty[}(x)$ convient.

1.1.7 Intégrale des limites uniformes de fonctions

1.1.7.1 Théorème fondamental

Définition. ***(Limite uniforme)***

Si $f : [a,b] \longrightarrow \mathbb{R}$ est bornée, on définit sa norme uniforme par $\|f\|_u = \sup\{f(t), t \in [a,b]\} \in \mathbb{R}_+$. Une suite de fonctions $(f_n)_n$ converge uniformément vers f si la suite $\|f_n - f\|_u$ est définie à partir d'un certain rang et tend vers zéro.

Théorème. ***(Intégration d'une limite uniforme)***

Si (f_n) est une suite de fonctions définie sur $[a,b]$, intégrables, convergeant uniformément vers f, alors f est intégrable et $\int_a^b f = \lim_n \int_a^b f_n(t)\mathrm{d}t$.

▷ Soit $\varepsilon > 0$. Alors il existe un rang $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$, $\|f_n - f\| < \frac{\varepsilon}{4(b-a)}$. En particulier, $\|f_{n_0} - f\| < \frac{\varepsilon}{4(b-a)}$. Or f_{n_0} est intégrable, donc il existe une subdivision $\sigma \in \mathrm{Sub}(a,b)$ telle que $S(f_{n_0},\sigma) - s(f_{n_0},\sigma) < \frac{\varepsilon}{2}$. Or $S(f,\sigma) - s(f,\sigma) = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x)\mathrm{Osc}_{[x,s_\sigma(x)[}(f)$. De plus $\mathrm{Osc}_E(f) = \sup_E f - \inf_E f = \sup_{t\in E} f_{n_0}(t) + \underbrace{(f(t) - f_{n_0}(t))}_{<\frac{\varepsilon}{4(b-a)}} - \left(\inf_{t\in E} f_{n_0}(t) + \underbrace{(f(t) - f_{n_0}(t))}_{<\frac{\varepsilon}{4(b-a)}}\right)$. Par conséquent, $\mathrm{Osc}_E(f) \leqslant \sup_E f_{n_0} + \frac{\varepsilon}{4(b-a)} - \left(\inf_E f_{n_0} - \frac{\varepsilon}{4(b-a)}\right) \leqslant \mathrm{Osc}_E(f_{n_0}) + \frac{\varepsilon}{2(b-a)}$, d'où $S(f,\sigma) - s(f,\sigma) \leqslant \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x)\left(\mathrm{Osc}_{[x,s_\sigma(x)[}(f_{n_0}) + \frac{\varepsilon}{2(b-a)}\right) = \sum_{x\in\sigma^\bullet} (s_\sigma(x) - x)\mathrm{Osc}_{[x,s_\sigma(x)[}(f_{n_0}) + \frac{\varepsilon}{2(b-a)} \underbrace{\sum_{x\in\sigma^\bullet} (s_\sigma(x) - x)}_{b-a}$. Ainsi $S(f,\sigma) - s(f,\sigma) \leqslant S(f_{n_0},\sigma) - s(f_{n_0},\sigma) + \frac{\varepsilon}{2} \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leqslant \varepsilon$, donc f est intégrable.

Calculons l'intégrale de f. D'une part, $\int_a^b f = \sup_{\tau\in\mathrm{Sub}(a,b)} s(f,\tau)$. On a $s(f,\tau) - s(f_{n_0},\tau) = \sum_{x\in\tau^\bullet} (s_\tau(x) - x)\left(\inf_{[x,s_\tau(x)[} f - \inf_{[x,s_\tau(x)[} f_{n_0}\right)$. Mais pour tout $t \in [a,b]$, $f_{n_0}(t) - \frac{\varepsilon}{4(b-a)} \leqslant f(t) \leqslant f_{n_0}(t) + \frac{\varepsilon}{4(b-a)}$ donc $\inf_E f_{n_0} - \frac{\varepsilon}{4(b-a)} \leqslant \inf_E f \leqslant \inf_E f_{n_0} + \frac{\varepsilon}{4(b-a)}$. Ainsi $s(f,\tau) - s(f_{n_0},\tau) \leqslant \sum_{x\in\tau^\bullet} (s_\tau(x) - x)\frac{\varepsilon}{4(b-a)} \leqslant$

$\dfrac{\varepsilon}{4(b-a)}\underbrace{\sum_{x\in\tau^\bullet}(s_\tau(x)-x)}_{b-a}=\dfrac{\varepsilon}{4}$. Par suite, $s(f,\tau)\leqslant s(f_{n_0},\tau)+\frac{\varepsilon}{4}$ puis $s(f,\tau)\leqslant\int_a^b f_{n_0}(t)\mathrm{d}t+\frac{\varepsilon}{4}$, puis $\int_a^b f\leqslant\int_a^b f_{n_0}+\frac{\varepsilon}{4}$.

D'autre part, semblablement, $\int_a^b f=\inf\limits_{\tau\in\mathrm{Sub}(a,b)}S(f,\tau)$. On a $S(f,\tau)-S(f_{n_0},\tau)=\sum\limits_{x\in\tau^\bullet}(s_\tau(x)-x)\left(\sup\limits_{[x,s_\tau(x)[}f-\sup\limits_{[x,s_\tau(x)[}f_{n_0}\right)$. Mais pour tout $t\in[a,b]$, $f_{n_0}(t)-\frac{\varepsilon}{4(b-a)}\leqslant f(t)\leqslant f_{n_0}(t)+\frac{\varepsilon}{4(b-a)}$ donc $\sup\limits_E f_{n_0}-\dfrac{\varepsilon}{4(b-a)}\leqslant\sup\limits_E f\leqslant\sup\limits_E f_{n_0}+\dfrac{\varepsilon}{4(b-a)}$. Ainsi $S(f,\tau)-S(f_{n_0},\tau)\geqslant\sum\limits_{x\in\tau^\bullet}(s_\tau(x)-x)(\dfrac{-\varepsilon}{4(b-a)})\geqslant -\dfrac{\varepsilon}{4(b-a)}\underbrace{\sum_{x\in\tau^\bullet}(s_\tau(x)-x)}_{b-a}=\dfrac{\varepsilon}{4}$. Par suite, $S(f,\tau)\geqslant s(f_{n_0},\tau)-\frac{\varepsilon}{4}$ d'où $\int_a^b f\geqslant\int_a^b f_{n_0}-\frac{\varepsilon}{4}$.

Ceci étant vrai pour tout $n\geqslant n_0$, on a $\left|\int_a^b f-\int_a^b f_n\right|\leqslant\frac{\varepsilon}{4}$, ce qui permet de conclure. ■

1.1.7.2 Fonctions réglées

Définition. ***(Fonction réglée)***

Une fonction $f:[a,b]\longrightarrow\mathbb{R}$ est dite *réglée* si elle est limite uniforme d'une suite de fonctions en escalier.

Corollaire

Toute fonction réglée est intégrable.

1.1.7.3 Cas des fonctions continues par morceaux

Théorème. ***(Heine)***

Toute fonction continue sur un compact est uniformément continue.

▷ Fastoche. ■

Théorème. ***(Approximation uniforme)***

Toute fonction continue par morceaux sur un segment est réglée.

▷ On utilise le théorème de Heine. Puisque $\lim\frac{b-a}{n}=0$, on choisit n tel que $\frac{b-a}{n}<\eta$ et on pose σ la subdivision régulière de $[a,b]$ d'ordre n, $\sigma_n=\{a+k\frac{b-a}{n}\mid k\in[\![0,n]\!]\}$. On construit g telle que $\|f-g\|_u\leqslant\varepsilon$. On pose $g(t)=f(t)$ si $t\in\sigma$ et $g(t)=\frac{f(x)+f(s_\sigma(x)}{2}$ si $t\in[x,s_\sigma(x)[$. g est constante sur ces intervalles donc g est en escalier. De plus, $|g(t)-f(t)|=\begin{cases}0 \text{ si } t\in\sigma\\ \left|\frac{f(x)-f(t)}{2}+\frac{f(s_\sigma(x))-f(t)}{2}\right| \text{ si } t\in]x,s_\sigma(x)[\end{cases}$.
Or si $x<t<s_\sigma(x)$, $0<t-x<s_\sigma(x)-x=\frac{b-a}{n}<\eta$, donc par la propriété d'uniforme continuité, dès que $|f(x)-f(t)|<\varepsilon$, $0<s_\sigma(x)-t<s_\sigma(x)-x=\frac{b-a}{n}<\eta$ donc encore une fois $|f(s_\sigma(x)-f(t)|<\varepsilon$. Ainsi pour tout $t\in[a,b]$, si $|g(t)-f(t)|<\varepsilon$, alors $\|g-f\|_u\leqslant\varepsilon$. ■

Corollaire

Toute fonction continue par morceaux est intégrable.

Corollaire

Toute fonction continue est intégrable.

1.1.8 Sommes de Riemann

Définition. ***(Subdivision pointée)***

On appelle *subdivision pointée* d'un segment $[a,b]$ de la droite réelle tout couple (σ, ξ) constitué d'une subdivision σ de $[a,b]$ et d'une application $\xi : \sigma \setminus \{b\} \longrightarrow \mathbb{R}$ telle que pour tout $x \in \sigma^\bullet$, $x \leqslant \xi(x) \leqslant \mathrm{suc}_\sigma(x)$.

Définition. ***(Somme de Riemann)***

Étant donnée $f : [a,b] \longrightarrow \mathbb{R}$, on appelle *somme de Riemann de la fonction f attachée à la subdivision pointée* (σ,ξ) la somme

$$SR(f,\sigma,\xi) = \sum_{x\in\sigma^\bullet} f(\xi(x))(\mathrm{suc}_\sigma(x) - x).$$

Heuristique

Dans le cas où f est positive, c'est l'aire d'une réunion de rectangles contigus.

Proposition

Si $\forall x \in \sigma^\bullet,\ \xi(x) < \mathrm{suc}_\sigma(x)$ et si f est bornée, alors $s(f,\sigma) \leqslant SR(f,\sigma,\xi) \leqslant S(f,\sigma)$.

▷ Pour tout $x \in \sigma^\bullet$, $x \leqslant \xi(x) < \mathrm{suc}_\sigma(x)$. Ainsi $\inf\limits_{t\in[x,\mathrm{suc}_\sigma(x)[} f(t) \leqslant f(\xi(x)) \leqslant \sup\limits_{t\in[x,\mathrm{suc}_\sigma(x)[} f(t)$, d'où $(\mathrm{suc}_\sigma(x) - x) \inf\limits_{t\in[x,\mathrm{suc}_\sigma(x)[} f(t) \leqslant (\mathrm{suc}_\sigma(x) - x) f(\xi(x)) \leqslant (\mathrm{suc}_\sigma(x) - x) \sup\limits_{t\in[x,\mathrm{suc}_\sigma(x)[} f(t)$. En sommant pour x décrivant $\sigma^\bullet$, $s(f,\sigma) \leqslant SR(f,\sigma,\xi) \leqslant S(f,\sigma)$. ■

Proposition

Si f est bornée, alors $\overline{s}(f,\sigma) \leqslant SR(f,\sigma,\xi) \leqslant \overline{S}(f,\sigma)$.

▷ Pour tout $x \in \sigma^\bullet$, $x \leqslant \xi(x) \leqslant \mathrm{suc}_\sigma(x)$. Ainsi $\inf\limits_{t\in[x,\mathrm{suc}_\sigma(x)]} f(t) \leqslant f(\xi(x)) \leqslant \sup\limits_{t\in[x,\mathrm{suc}_\sigma(x)]} f(t)$ c'est-à-dire $\overline{m}_\sigma(x) \leqslant f(\xi(x)) \leqslant \overline{M}_\sigma(x)$, d'où $\overline{s}(f,\sigma) \leqslant SR(f,\sigma,\xi) \leqslant \overline{S}(f,\sigma)$. ■

Proposition

Si f est intégrable, alors pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour toute $\sigma \in Sub(a,b)$ telle que $p(\sigma) < \eta$, $|I - SR(f,\sigma,\xi)| < \varepsilon$ en posant $I = \int_a^b f$.

$\triangleright$ On utilise l'une des propositions précédentes. De la question précédente, $-\varepsilon \leqslant \overline{s}(f,\sigma) - \overline{S}(f,\sigma) < I - \overline{S}(f,\sigma) \leqslant I - SR(f,\sigma,\xi) \leqslant I - \overline{s}(f,\sigma) \leqslant \overline{S}(f,\sigma) - \overline{s}(f,\sigma) < \varepsilon$ d'où le résultat. ■

Théorème. ***(Réciproque au théorème des sommes de Riemann)***

On suppose que les sommes de Riemann tendent vers I au sens des subdivisions. Alors f est intégrable d'intégrale I.

$\triangleright$ Réécrire l'hypothèse. Montrons que f est bornée. Il existe $\eta > 0$ tel que $p(\sigma) < \eta$ implique $I - 1 \leqslant SR(f,\sigma,\xi) \leqslant I + 1$. Posons $\sigma_n = \{a + k\frac{b-a}{n}, 0 \leqslant k \leqslant n\}$. Il existe $n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$, $n \geqslant n_0 \implies p(\sigma_n) < \eta$. Fixons $n \geqslant n_0$: pour tout $t \in [a,b]$, il existe t_0 tel que $a + k_0\frac{b-a}{n} \leqslant t \leqslant a + (k_0+1)\frac{b-a}{n}$. Posons $\xi_k = a + k\frac{b-a}{n}$ pour $k \neq k_0$, $k < n$. $\xi_{k_0} = t$. On a $I - 1 < SR(f,\sigma,\xi)$, $\displaystyle\sum_{k \neq k_0, 0 \leqslant k < n} f(a + k\frac{b-a}{n})\frac{b-a}{n} + f(t)\frac{b-a}{n} < I + 1$ donc $\displaystyle I - 1 - \sum_{k \neq k_0, 0 \leqslant k < n} f(a + k\frac{b-a}{n}) < f(t) < I + 1 - \sum_{k \neq k_0, 0 \leqslant k < n} f(a + k\frac{b-a}{n})$ donc f est borné sur chaque sous intervalle de la subdivision donc f est bornée. Montrons maintenant qu'elle est intégrable. Nous étudions $S(f,\sigma) - s(f,\sigma)$. On introduit l'oscillation. On introduit m_x, M_x les bornes de l'intervalle entre x et son successeur. $M - x - \frac{\varepsilon}{4(b-a)} < M_x$. Prenons ξ_x dans cet écart, et de même un ξ'_x pour m_x. Alors $M_x - m_x - \frac{\varepsilon}{4(b-a)} < f(\xi_x) - f(\xi'_x) < M_x - m_x = Osc$. Il est facile de conclure alors. ■

1.1.9 Mesurabilité au sens de Riemann

Définition. ***(Mesurabilité au sens de Riemann)***

Une partie $E \subseteq [a,b]$ est dite *mesurable (au sens de Riemann)* si $\mathbb{1}_E$ est intégrable au sens de Riemann.

Définition. ***(Mesure de Riemann)***

Soit E une partie du segment $[a,b]$ mesurable. Sa mesure est par définition :

$$\mu(E) = \int_{[a,b]} \mathbb{1}_E.$$

Le résultat, démontré en exercice, donnant la valeur de l'intégrale d'une fonction en escalier (c'est-à-dire, constante par morceaux) donnent les résultats suivants :

Propriété. ***(Mesure d'une partie finie)***

Toute partie finie de $[a,b]$ est mesurable, de mesure nulle.

Propriété. ***(Mesure du segment entier)***

Le segment $[a,b]$ est mesurable, de mesure $\mu([a,b]) = b - a$.

Heuristique

Pour les parties d'un segment $[a,b]$, la mesurabilité n'est pas une question de taille, puisque le segment entier est lui-même mesurable. La question de la mesurabilité se ramenant à l'intégrabilité d'une fonction ne prenant que deux valeurs, la question porte véritablement sur la « régularité » de la répartition des points de E dans le segment.

Exercice 12

L'ensemble des rationnels compris entre 0 et 1 est-il mesurable ?

▷ **Éléments de réponse.**

Non. On va montrer que la fonction $\mathbb{1}_{\mathbb{Q}}$ restreinte à $[0,1]$ n'est pas intégrable. Soit $\sigma \in \mathrm{Sub}(0,1)$. Soit $x \in \sigma \setminus \{1\}$. Puisque $\mathbb{Q}$ est dense, $\sup\limits_{x \leqslant t \leqslant s_\sigma(x)} \mathbb{1}_E(t) = 1$ et puisque $\mathbb{R} \setminus \mathbb{Q}$ est dense, $\inf\limits_{x \leqslant t \leqslant s_\sigma(x)} \mathbb{1}_E(t) = 0$. Ainsi $s(f,\sigma) = 0$ et $S(f,\sigma) = \sum\limits_{x \in \sigma \setminus \{1\}} \mathrm{suc}_\sigma(x) - x = 1$. En passant aux bornes, la mesure d'inintégrabilité vaut $1 > 0$.

Propriété. ***(Mesure d'un intervalle)***

Soit I un sous-intervalle de $[a,b]$. Alors I est mesurable, de mesure $\sup(I) - \inf(I)$.

▷ La fonction indicatrice de I est en escalier. ■

Exercice 13

On considère l'ensemble $E = \{\frac{1}{n}, n \in \mathbb{N}^*\}$. Est-il mesurable ? Si oui, déterminer sa mesure.

▷ **Éléments de réponse.**

On considère la subdivision dépendant de N définie par $\sigma_N = \{1, \frac{1}{2}, \dots, \frac{1}{N}, 0\}$. (Ce n'est pas la subdivision régulière.) Dans ce cas, $s(\mathbb{1}_E,\sigma_N) = 0$ et $S(\mathbb{1}_E,\sigma_N) = 1$. On définit $\tau_N = \sigma_N \cup \{\frac{1}{2N^2}, \frac{1}{N} + \frac{1}{2N^2}, \frac{1}{N-1} + \frac{1}{2N^2}, \dots, \frac{1}{2} + \frac{1}{2N^2}, 1\}$. Toujours $s(\mathbb{1}_E, \tau_N) = 0$. De plus $0 < \frac{1}{2N^2} \leqslant \frac{1}{N} \leqslant \frac{1}{N} + \frac{1}{2N^2} \leqslant \frac{1}{N-1} \dots$ Ainsi $S(\mathbb{1}_E,\tau_N) = \frac{1}{2N^2} \times 1 + (\frac{1}{N} - \frac{1}{2N^2}) \times 1 + \frac{1}{2N^2} \times 1 + (\frac{1}{N-1} - \frac{1}{N} - \frac{1}{2N^2}) \times 0 + \frac{1}{2N^2} + (\dots) \times 0$ d'où $S(\mathbb{1}_E,\tau_N) = \frac{1}{2N^2}(\underbrace{1 + 1 + \dots + 1}_{n \text{ fois}}) + 0 + \frac{1}{N} - \frac{1}{2N^2}$. Ceci tend vers 0 lorsque N tend vers 0. Ainsi E est mesurable, de mesure $\mu(E) = 0$, bien qu'il soit infini.

Méthode. ***(Ensembles de mesure nulle)***

Pour montrer qu'un ensemble E est de mesure nulle, il suffit d'exhiber une suite de subdivisions (σ_N) telle que $S(\mathbb{1}_E,\sigma_N)$ tend vers zéro.

▷ En effet, puisque les indicatrices sont positives, le théorème des gendarmes permet de conclure grâce à l'encadrement de $\mu(E)$ en tant qu'infimum des sommes de Darboux supérieures. ■

1.2 Théorème de Lebesgue-Vitali

1.2.1 Parties négligeables de $\mathbb{R}$

Définition. ***(Partie négligeable des réels)***

Soit $A \subseteq \mathbb{R}$. L'ensemble A est négligeable si pour tout $\varepsilon > 0$, il existe une suite d'intervalles ouverts $(I_n)_{n\in\mathbb{N}^*} = (]a_n, b_n[)_{n\in\mathbb{N}^*}$, $a_n \leqslant b_n$ $\forall n$, qui soit un recouvrement de A tel que $\sum_{n=1}^{+\infty} \ell_n \leqslant \varepsilon$ en posant $\ell_n = b_n - a_n$.

Remarque. Cette dernière condition équivaut, par positivité des longueurs des intervalles, à $\forall N \geqslant 1 \quad \sum_{n=1}^{N} \ell_n \leqslant \varepsilon$.

Exercice 14

Montrer que, dans cette dernière condition, l'inégalité large peut être remplacée par une inégalité stricte.

Propriété. ***(Sous-partie d'une négligeable)***

Si A est négligeable et $B \subseteq A$, alors B est négligeable.

▷ Easy peasy. ■

Propriété. ***(Négligeabilité de l'ensemble vide)***

$\emptyset$ est négligeable.

▷ On écrit pour tout $\varepsilon > 0$, $I_n =]0,0[= \emptyset$. ■

Propriété. ***(Négligeabilité des singletons)***

Tout singleton est négligeable.

▷ On écrit pour tout $\varepsilon > 0$, $I_n =]x - \frac{\varepsilon}{2^{n+1}}, x + \frac{\varepsilon}{2^{n+1}}[\ni x$. Ainsi $\{x\} \in \bigcup\limits_{n \geqslant 1} I_n$ et pour tout $N \in \mathbb{N}^*$, $\sum\limits_{n=1}^{N} \frac{\varepsilon}{2^n} = \frac{\varepsilon}{2} \frac{1 - (\frac{1}{2^N})}{1 - \frac{1}{2}} < \varepsilon$. ■

Propriété. ***(Non-négligeabilité des intervalles)***

Un intervalle non réduit à un point ni vide n'est jamais négligeable.

▷ Il suffit de vérifier $[a,b]$, $a < b$, non négligeable et l'on peut conclure pour tous les types d'intervalles grâce à la propriété des sous-parties. Si $[a,b]$ était négligeable, soit $\varepsilon < b - a$, par exemple $\varepsilon = \frac{b-a}{2}$. Soit (I_n) un recouvrement de $[a,b]$ par des intervalles bornés ouverts dont la somme de série des longueurs soit inférieure à ε. D'après la propriété de Borel-Lebesgue, il existe $N \in \mathbb{N}$ tel que $[a,b] \subseteq \bigcup\limits_{n=1}^{N} I_n$. On écrit astucieusement :

$$\varepsilon < b - a = b - a_n + (a_n - a_{n-1}) + ... + (a_2 - a) < b_n - a_n + (b_{n-1} - a_{n-1}) + ... + (b_1 - a_1) \leqslant \varepsilon,$$

car $\sum\limits_{n=1}^{N} (b_n - a_n) \leqslant \varepsilon$. En particulier, $\varepsilon < \varepsilon$, ce qui est absurde. ■

Propriété. ***(Réunion dénombrable de négligeables)***

Toute réunion dénombrable de parties négligeables est négligeable.

▷ Soit D un ensemble dénombrable. Soit $(A_x)_{x \in D}$ une famille dénombrable d'ensembles négligeables. Soit donc $\varphi : D \hookrightarrow \mathbb{N}^*$. Pour tout $x \in D$, A_x est négligeable. Avec les notations évidentes, on a $A_x \subseteq \bigcup\limits_{n=1}^{\infty} I_{n,x}$. Or pour tout $N \geqslant 1$, $\sum\limits_{n=1}^{N} b_{n,x} - a_{n,x} \leqslant \frac{\varepsilon}{\varphi(x)}$ avec $I_{n,x} =]a_{n,x}, b_{n,x}[$. De plus, par produit fini de dénombrables, $\mathbb{N}^* \times D$ est dénombrable. Soit donc $\varphi : \mathbb{N}^* \twoheadrightarrow \mathbb{N}^* \times D$. On pose pour tout $n \in \mathbb{N}^*$, $J_n = I_{\varphi(n)}$. Écrivons également $\bigcup\limits_{x \in D} A_x \subseteq \bigcup\limits_{x \in D} \bigcup\limits_{n \in \mathbb{N}^*} I_{n,x} = \bigcup\limits_{n \geqslant 1} J_n$. De plus, pour tout $N \geqslant 1$,
$\sum\limits_{n=1}^{N} b_n - a_n \leqslant \sum\limits_{(n,x) \in \mathbb{N}^* \times D} b_{n,x} - a_{n,x} \leqslant \sum\limits_{x \in D} \sum\limits_{n=1}^{+\infty} b_{n,x} - a_{n,x} \leqslant \sum\limits_{x \in D} \frac{\varepsilon}{2^{\varphi(n)}} \leqslant \sum\limits_{n \geqslant 1} \frac{\varepsilon}{2^n} \leqslant \varepsilon$. ■

Corollaire. ***(Réunion de deux parties négligeables)***

Si A et B sont deux parties négligeables des réels, alors $A \cup B$ l'est également.

▷ Les ensembles finis sont dénombrables. ■

Corollaire. ***(Négligeabilité d'un dénombrable)***

Toute partie dénombrable de $\mathbb{R}$ est négligeable au sens de l'intégration.

▷ On écrit A grâce à sa partition en événements élémentaires, ici à support dénombrable. ■

Deux conséquences directes du corollaire précédent :

Exemples

1. $\mathbb{Q}$ est négligeable.
2. L'image de toute suite est négligeable.

1.2.2 Le théorème de Lebesgue-Vitali : caractérisation des fonctions intégrables au sens de Riemann

Soit $f : [a,b] \longrightarrow \mathbb{R}$ une fonction bornée.

$\longrightarrow$ *Notation.* On note $Disc(f)$ l'ensemble des points de discontinuité de f, c'est-à-dire l'ensemble des points de $[a,b]$ en lequel la fonction f n'est pas continue.

Lemme. ***(Caractérisation des points de discontinuité par l'oscillation)***

On a $x \in Disc(f) \iff \mathrm{Osc}(f,x) > 0$.

Remarque. Cette condition est équivalente à $\mathrm{Osc}(f,x) \neq 0$, car une oscillation locale, de même qu'une oscillation sur un ensemble, est toujours positive.

$\triangleright$ On sait que f est continue en $x \in [a,b]$ si et seulement si : $\forall \varepsilon > 0\ \exists \eta > 0\ \forall t \in [a,b] \quad |t-x| < \eta \implies |f(t) - f(x)| < \varepsilon$. Donc si f est continue en x,

$$\forall \varepsilon > 0\ \exists \eta > 0 \quad \mathrm{Osc}_{]x-\eta,x+\eta[}(f) \leqslant 2\varepsilon$$

d'où $\forall \varepsilon > 0 \quad \mathrm{Osc}(f,x) \leqslant 2\varepsilon$. Or un résultat d'analyse plus ou moins célèbre nous dit $\forall x\ (\forall \varepsilon > 0 \quad 0 \leqslant x \leqslant \varepsilon) \implies x = 0$. Or $\mathrm{Osc}(f,x)$ est positive par définition donc $\mathrm{Osc}(f,x) = 0$.

Réciproquement, si $\mathrm{Osc}(f,x) = 0$, alors pour $\varepsilon > 0$ quelconque, $\inf(\mathrm{Osc}_{]x-\alpha,x+\alpha[}(f), \alpha \in \mathbb{R}_+^*) < \varepsilon$ donc il existe $\alpha > 0$ tel que $\mathrm{Osc}_{]x-\alpha,x+\alpha[}(f) < \varepsilon$. Si $t \in]x-\alpha, x+\alpha[$, $\inf\{f(t), t \in]x-\alpha, x+\alpha[\} \leqslant f(t) \leqslant \sup\{f(t), t \in]x-\alpha, x+\alpha[\}$ d'où $|f(t) - f(x)| < 2\varepsilon$. ■

Lemme

Pour tout $h > 0$, $E_h = \{x \in [a,b], \mathrm{Osc}(f,x) \geqslant h\}$ est fermé.

$\triangleright$ Soit $x \in \overline{E_h}$. Soit $\beta > 0$. Alors $]x-\beta, x+\beta[\cap E_h \neq \emptyset$. ■

Théorème. ***(Lebesgue-Vitali)***

Si $f : [a,b] \longrightarrow \mathbb{R}$ est une fonction bornée, alors f est intégrable au sens de Riemann sur $[a,b]$ si et seulement si $Disc(f)$ est négligeable.

1.2.3 Corollaires

1.2.3.1 Somme de fonctions intégrables

Corollaire. *(Somme de deux fonctions intégrables)*

Si f,g sont deux fonctions intégrables sur $[a,b]$, alors $f+g$ l'est aussi.

▷ $Disc(f+g) \subseteq Disc(f) \cup Disc(g)$. Or la réunion de deux parties négligeables est négligeable, donc $Disc(f+g)$ l'est aussi. Il faut aussi justifier que $f+g$ est bornée ! On aurait pu le voir aussi avec les sommes de Darboux, car l'infimum des $f(t)+g(t)$ sur $[x,y[$ est plus grand que la somme des infima, donc $s(f+g,\sigma) \geqslant s(f,\sigma)+s(f,\sigma)$; de même, dans l'autre sens, pour les suprema. L'inégalité pour les sommes de Darboux inférieures entraîne que l'intégrale de $f+g$ est supérieure pour toute subdivision à $s(f,\sigma)+s(g,\sigma)$. Ainsi l'intégrale sur $-s(f,s)$ est supérieure à $s(f,g)$, donc supérieure à l'intégrale de g. On échange l'intégrale de g et s de place, puis rebelote, et donc l'intégrale de la somme est supérieure à la somme des intégrales. De même, avec les sommes supérieures, d'où l'égalité.

On a montré que l'intégrale est additive. ■

Corollaire. *(Combinaison linéaires de fonctions intégrables)*

La combinaison linéaire de fonctions intégrables est intégrable. Autrement dit, l'ensemble des fonctions intégrables sur un segment est un espace vectoriel.

▷ On montre de même que l'intégrale est linéaire : $\int \alpha f + \beta g = \alpha \int f + \beta \int g$. ■

1.2.3.2 Produit de fonctions intégrables

Corollaire. *(Produit de deux fonctions intégrables)*

Si f,g sont deux fonctions intégrables sur $[a,b]$, alors fg l'est aussi.

▷ La justification est la même que celle faisant intervenir le théorème de Lebesgue-Vitali. Mais attention ! L'intégrale n'est pas multiplicative. (Regarder où ça coince dans la preuve.) ■

1.2.3.3 Quotient de fonctions intégrables

Corollaire. *(Quotient de deux fonctions intégrables)*

Si f,g sont deux fonctions intégrables sur $[a,b]$, et g ne s'annule pas sur $[a,b]$, alors **si le quotient est borné sur le segment**, $\frac{f}{g}$ est également intégrable.

Remarque. C'est faux sans l'hypothèse de bornitude du quotient !

1.2.3.4 Valeur absolue d'une fonction intégrable

Corollaire. ***(Valeur absolue d'une fonction intégrable)***

Si f est intégrable sur le segment $[a,b]$, alors $|f|$ est intégrable.

▷ On a, mais ça ne marche que parce que $a < b$, $S(f,s) \leqslant S(|f|,s)$, donc $-\int |f| \leqslant \int f \int |f|$. On obtient du même coup la formule suivante. ■

Formule. ***(Intégrale de la valeur absolue)***

Soit f intégrable sur $[a,b]$. Alors $|f|$ est intégrable et alors

$$\left|\int f\right| \leqslant \int |f|.$$

La réciproque est fausse ! On a vu que $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ est non intégrable. Ainsi la fonction $f = -1 + 2\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ est non intégrable mais sa valeur absolue 1 l'est. C'est donc, pour les lecteurs ayant déjà travaillé cette notion, le contraire de ce qu'il se passe généralement pour les fonctions continues par morceaux sur un intervalle quelconque, prendre garde !

1.2.3.5 Invariances de l'intégrale par modification en « quelques » points

Corollaire

Soit f,g deux fonctions bornées sur $[a,b]$. On note $E = \{x \in [a,b]\ f(x) \neq g(x)\}$. Si E est *fini*, alors f est intégrable si et seulement si g l'est, et le cas échéant, $\int f = \int g$.

▷ En effet E est fermé donc $Disc(g) \subseteq Disc(f) \cup E$ et $Disc(f) \cup Disc(g) \cup E$. ■

Corollaire

Soit f,g deux fonctions bornées sur $[a,b]$. On note $E = \{x \in [a,b]\ f(x) \neq g(x)\}$. Si E est *fermé et négligeable*, même conclusion.

Corollaire

Soit f,g deux fonctions bornées sur $[a,b]$. On note $E = \{x \in [a,b]\ f(x) \neq g(x)\}$. Si *l'adhérence de E est négligeable*, même conclusion.

Exemple. ***(Un exemple garde-fou)***

$\mathbb{1}_{\mathbb{Q}\cap[0,1]} = 0$ en dehors de $\mathbb{Q} \cap [0,1]$. Mais l'adhérence de cet ensemble est $[0,1]$ qui n'est pas négligeable.

De même, si $f \neq g$ sur un ensemble au plus dénombrable, on ne peut conclure a priori que $\int f = \int g$, en considérant un contre-exemple sur $\mathbb{Q}$, dénombrable, négligeable et d'adhérence non négligeable.

1.2.3.6 Relation de Chasles

Corollaire. *(Chasles)*

Soient $a < b < c$ trois points réels. On considère $f : [a,c] \longrightarrow \mathbb{R}$. f est intégrable sur $[a,c]$ si et seulement si $f_{|[a,b]}$ et $f_{|[b,c]}$ sont intégrables.

▷ En effet $Disc(f) \setminus \{b\}$ = l'union des discontinuités de f pour chaque restriction, privées de l'élément b. ■

Corollaire. *(Relation de Chasles)*

De plus l'intégrale vérifie une relation de Chasles.

▷ On déduit le précédent de cette égalité. Proprement $s(f,\sigma) = s(f_{|[a,b]},\sigma \cap [a,b]) + s(f_{|[b,c]},\sigma \cap [b,c]$. Pareil avec S. ■

Le théorème est vrai pour les points placés n'importe comment en fixant la convention de l'intégrale nulle pour deux bornes égales et d'échange des signes par échange des bornes. La propriété de la valeur absolue subsiste aussi.

1.3 Calcul intégral

1.3.1 Primitives

Définition. *(Primitive)*

Soit $f : [a,b] \longrightarrow \mathbb{R}$. On appelle *primitive* de f toute fonction $F : [a,b] \longrightarrow \mathbb{R}$ dérivable en tout point de $[a,b]$ et telle que $\forall t \in [a,b] \quad F'(t) = f(t)$.

Rappel. $F'(a) = \lim\limits_{t \to a^+} \frac{F(t)-F(a)}{t-a}$ dérivée à droite, de même pour b à gauche.

Définition. *(Primitivabilité)*

Une fonction $f : [a,b] \longrightarrow \mathbb{R}$ est primitivable si elle admet une primitive. On dit aussi que f est *dérivée*. (Mais ça, on le sait depuis le lycée...)

Théorème. *(Darboux)*

Toute fonction primitivable vérifie la propriété des valeurs intermédiaires.

Exemple

La fonction indicatrice de]1,2] sur [0,1]. Ainsi toute fonction intégrable (ici en escalier) n'est pas primitivable.

1.3.2 Lien entre primitives et intégrales

Théorème. *(Théorème fondamental de l'analyse)*

Soit $f : [a,b] \longrightarrow \mathbb{R}$ intégrable, si $x \in [a,b]$, alors f est intégrable sur $[a,x]$ d'après le théorème de Chasles. On définit :

$$\begin{array}{rccl} \mathrm{F}: & [a,b] & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \int_a^b f(t)\mathrm{d}t. \end{array}$$

Alors

1. F est continue ;
2. Si f est continue au point $x \in [a,b]$, alors F est dérivable en x, avec $F'(x) = f(x)$.

▷ Successivement :

1. On montre que F est lipschitzienne pour la constante C borné f, f étant intégrable ;
2. Traduisons pour ε, α la continuité de f en x. Les calculs nous montrent que la valeur absolue de $F(y) - F(x) - f(x)(y-x)$ est inférieur à $|\int_x^y |f(t) - f(y)|\mathrm{d}t|$. Donc cette quantité est inférieur à $\varepsilon|y-x|$. On divise par $|y-x|$ est c'est bon. ■

Corollaire

Toute fonction continue sur $[a,b]$ est primitivable.

Définition. *(Continuité par morceaux)*

Soit $f : [a,b] \longrightarrow \mathbb{R}$. f est *continue par morceaux* sur $[a,b]$ s'il existe $\sigma \in \mathrm{Sub}([a,b])$ telle que $\forall x \in \sigma\{b\}$ $\exists g_x : [x, suc_\sigma(x)] \longrightarrow \mathbb{R}$ continue telle que $f_{|]x,suc_\sigma(x)[} = g_{x]x,suc_\sigma(x)}$.

Propriété

Toute fonction en escalier est continue par morceaux.

Propriété. *(T)*

oute fonction continue par morceaux est intégrable.

▷ En effet, en posant la même intégrable que précédemment, F est dérivable sauf en un nombre fini de points et en excluant ces points, $F' = f$ (puis théorème d'invariance). ■

Théorème. ***(Newton)***

Soit $f : [a,b] \longrightarrow \mathbb{R}$ une fonction intégrable et primitivable, de primitive $F : [a,b] \longrightarrow \mathbb{R}$. Alors $\int_a^b f(t)\mathrm{d}t = F(b) - F(a)$.

$\triangleright$ Pour toute $\sigma \in \mathrm{Sub}(a,b)$. ■

Remarque. Certains auteurs échangent les noms du théorème fondamental et du corollaire de Newton.

Théorème. ***(Cas d'une fonction continue par morceaux)***

Soit $f : [a,b] \longrightarrow \mathbb{R}$ une fonction continue par morceaux. Il existe F telle que et dérivable sauf en un nombre fini de points... et aussi dans ce cas l'égalité de Newton.

1.3.3 Quelques formules de calcul intégral

1.3.3.1 Formule d'intégration par parties

Formule. ***(Intégration par parties)***

Soient $f,g : [a,b] \longrightarrow \mathbb{R}$ intégrables et primitivables, F,G leurs primitives respectives. Alors Fg est intégrable, fG est intégrable et on a la formule :

$$\int_a^b F(t)g(t)\mathrm{d}t = F(b)G(b) - F(a)G(a) - \int_a^b f(t)G(t)\mathrm{d}t.$$

$\triangleright$ G, puis F est continue donc intégrable. f,g est intégrable donc ces deux produits sont intégrables. On écrit :

$$\begin{aligned}\int_a^b (F(t)g(t) + f(t)G(t))\mathrm{d}t &= \int_a^b (F(t)G'(t) + F'(t)G(t))\mathrm{d}t \\ &= \int_a^b (FG)'(t)\mathrm{d}t\end{aligned}$$

et l'on conclut par linéarité grâce au théorème fondamental de l'analyse (ici appelé théorème de Newton). ■

On a le même résultat quand f et g sont intégrables et seulement continues par morceaux. Dans ce cas, on pose une certaine $F(x) = \displaystyle\int_a^x f(t)\mathrm{d}t + C$ et idem pour G avec une constante C' et la formule d'intégration par parties tient constantes à près.

1.3.3.2 Formules de la moyenne

Formule. ***(Deuxième formule de la moyenne)***

Soit F dérivable décroissante positive, F' intégrable et g intégrable. Alors $\int_a^b F(t)g(t)\mathrm{d}t = F(a)\int_a^x g(t)\mathrm{d}t$ pour un $x \in [a,b]$.

$\triangleright$ *(Preuve dans le cas g positif)* On pose $G(x) = \int_a^x g(t)\mathrm{d}t$. Fg est intégrable par produit de fonctions intégrables. $\int_a^b F(t)g(t)\mathrm{d}t$.

On a $F(a) \geqslant F(t) \geqslant F(b)$, d'où $F(a)g(t) \geqslant F(t)g(t) \geqslant F(b)g(t)$. En intégrant, $F(a)\int_a^b g(t) \geqslant \int_a^b F(t)g(t)\mathrm{d}t \int_a^b F(b)\mathrm{d}t \int_a^b g(t)\mathrm{d}t$ puis en divisant $G(b) \geqslant \frac{1}{F(a)}\int_a^b F(t)g(t)\mathrm{d}t \geqslant G(a) = 0$. Comme G est continue, il existe $x \in [a,b]$ tel que $\frac{1}{F(a)}\int_a^b F(t)g(t)\mathrm{d}t = G(x)$. ■

$\triangleright$ *(Preuve dans le cas général)* On fait une intégration par parties en prenant la primitive de Newton de G. Alors G est continue donc bornée sur $[a,b]$. On l'encadre donc par ses bornes. On multiplie ensuite par $-f$ qui est positive, puis on intègre en a et b ; par le théorème fondamental, on peut remplacer les intégrables par les termes limites $F(b) - F(a)$. En simplifiant, on vérifie qu'il reste $\inf G \cdot F(b) \leqslant F(b)G(b) \leqslant \sup G \cdot F(b)$, d'où $\inf G \cdot F(a) \leqslant \int_a^b F(t)g(t)\mathrm{d}t \leqslant \sup G \cdot F(a)$. Pour toute valeur de $[\inf G, \sup G]$ est un certain $G(x)$ pour $x \in [a,b]$, car G est continue, on a $\frac{1}{F(a)}\int_a^b F(t)g(t)\mathrm{d}t = G(x)$ pour un certain $x \in [a,b]$, ce qu'il fallait montrer. ■

1.3.3.3 Formule de changement de variable

Formule. ***(Changement de variable)***

Soit $f : [a,b] \longrightarrow \mathbb{R}$ primitivable et intégrable de primitive $\varphi : [c,d] \longrightarrow [a,b]$. φ est dérivable, $\varphi(c) = a$, $\varphi(d) = b$. On suppose que φ' et $f \circ \varphi$ sont intégrables sur $[c,d]$. Alors $\int_a^b f(t)\mathrm{d}t = \int_c^d (f \circ \varphi)(t)\varphi'(t)\mathrm{d}t$.

$\triangleright$ C'est un cas particulier du théorème fondamental de l'analyse. ■

Formule. ***(Changement de variable, autre version)***

Le théorème vrai en supposant f continue par morceaux et en supposant φ continue et de classe $\mathcal{C}^1$ par morceaux.

Définition. ***(Classe de régularité par morceaux)***

Si $r \in \mathbb{N}$, on dit que $\varphi : [c,d] \longrightarrow \mathbb{R}$ est de classe $\mathcal{C}^r$ par morceaux s'il existe une subdivision σ de $[c,d]$ tel que pour tout $x \in \sigma^\bullet$, il existe une fonction $g_x : [x, s_\sigma(x)] \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ telle que pour tout $t \in]x, s_\sigma(x)[$, $g_x(t) = \varphi(t)$.

Formule. ***(Troisième version du changement de variable)***

Supposons $f : [a,b] \longrightarrow \mathbb{R}$, $\varphi : [c,d] \longrightarrow [a,b]$ dérivable. Alors f primitivable, f intégrable sur l'intervalle entre $\varphi(c)$ et $\varphi(d)$. $f \circ \varphi$ et φ' intégrables sur $[c,d]$. Alors $\displaystyle\int_{\varphi(c)}^{\varphi(d)} f(t)\mathrm{d}t = \int_c^d (f\circ\varphi)'(t)\varphi'(t)\mathrm{d}t$.

1.4 Intégrales paramétriques

Définition. ***(Intégrale dépendant d'un paramètre)***

C'est une intégrale de la forme suivante : étant donnée $f : [a,b] \times D \longrightarrow \mathbb{R}$ où D est une partie de $\mathbb{R}$, $F : D \longrightarrow \mathbb{R}$, $y \longmapsto \displaystyle\int_a^b f(t,y)\mathrm{d}t$.

1.4.1 Continuité

Théorème. ***(Continuité des intégrables à paramètre)***

Si f est continue, alors F est bien définie et continue.

$\triangleright$ Pour y fixé, $t \longmapsto f(t,y)$ est comprise de $t \longmapsto (t,y)$ et de f qui sont continues. Elle est donc continue et donc intégrable, de sorte que $F(y)$ existe. ■

Lemme

Si la suite (y_n) converge vers y, alors la suite de fonction f_n définies sur $[a,b]$ par $f_n(t) = f(t,y_n)$ converge uniformément vers la fonction $t \longmapsto f(t,y)$.

$\triangleright$ Montrons que $\forall \varepsilon > 0\ \exists \theta > 0\ \forall y \in D, \forall z \in D \quad \|y - z\| < \theta \implies \forall t \in [a,b]\ |f(t,y) - f(t,z)| < \varepsilon$, car si $y_n \longrightarrow y$ alors il existe $n_0 \in \mathbb{N}\ \forall n \in \mathbb{N}, n \geqslant n_0 \implies \|y_n - y\| < \theta \implies \|f(t,y_n) - f(t,y)\|_u \leqslant \varepsilon$.

On fixe $\varepsilon > 0$ arbitraire. Soit $A = \{(t,z) \in [a,b] \times D \mid |f(t,y) - f(t,z) \geqslant \varepsilon\}$. C'est l'image réciproque du fermé $[\varepsilon, +\infty[$ par l'application continue $(t,z) \longmapsto |f(t,y) - f(t,z)|$, donc A est fermé dans $[a,b] \times D$. Définissons $g : [a,b] \longrightarrow \mathbb{R}$ qui à t fait correspondre $\inf\limits_{m\in A}\{\|(t,y) - m\|\} = d((t,z),A)$. Pour tout $t \in [a,b]$, $g(t) > 0$. g est continue sur le compact $[a,b]$ donc atteint ses bornes. Donc il existe $\theta > 0$ tel que pour tout t, $g(t) \geqslant 2\theta$. Ainsi $\|y - z\| < \theta \implies \|y - z\| > \frac{1}{2} \inf\limits_{m\in A} \|(t,y) - m\|$ la norme produit. Si

$(t,z) \in A$, $\|(t,y)-(t,z)\| = \|y-z\| \geqslant g(t) \geqslant g(t) \geqslant 2\theta$. Donc $\|y-z\| < \theta \implies \forall t \in [a,b] \quad (t,z) \in A$ donc $\forall t \in [a,b] \quad |f(t,y) - f(t,z)| < \varepsilon$. On a montré la convergence uniforme. ■

On conclut le théorème de continuité en disant que l'intégrale des $f(t,y_n)$ égalant celle des $f_n(t)$ elle tend vers l'intégrale de f. On a montré que pour toute suite y_n convergeant vers y, $F(y_n)$ tend vers $F(y)$. Par caractérisation séquentielle de la continuité, F est continue.

1.4.2 Dérivabilité

On suppose que D est un ouvert de $\mathbb{R}$.

Théorème. ***(Dérivation)***

Si $f : [a,b] \longrightarrow \mathbb{R}$ continue, $\frac{\partial f}{\partial y}(t,y)$ existe en tout point $(t,y) \in [a,b] \times D$. De plus on suppose que $(t,y) \longmapsto \frac{\partial f}{\partial y}(t,y)$ est continue sur $[a,b] \times D$. Alors $F : D \longrightarrow \mathbb{R}$, $y \longmapsto \int_a^b f(t,y)\mathrm{d}t$ est définie, de classe $\mathbb{C}^1$ sur D avec $F'(y) = \int_a^b \frac{\partial f}{\partial y}(t,y)\mathrm{d}t$.

Lemme

$\forall \varepsilon > 0 \ \forall y_0 \in D \ \exists \theta > 0 \ \forall y \in D \quad |y - y_0| < \theta \implies \forall t \in [a,b] \quad |f(t,y) - f(t,y_0) - \frac{\partial f}{\partial y}(t,y)(y - y_0)| \leqslant \varepsilon |y - y_0|$.

▷ On fixe $\varepsilon > 0$ et $y_0 \in D$. On considère $h : [a,b] \times D \longrightarrow \mathbb{R}$ qui à (t,y) fait correspondre $f(t,y) - f(t,y_0) - \frac{\partial f}{\partial y}(t,y)(y-y_0)$. h est évidemment continue sur $[a,b] \times D$. De plus $\frac{\partial h}{\partial y}(t,y)$ existe pour tout $(t,y) \in [a,b] \times D$, avec $\frac{\partial h}{\partial y}(t,y) = \frac{\partial f}{\partial y}(t,y) - \frac{\partial f}{\partial y}(t,y_0)$ fonction continue de (t,y). Comme dans la partie continuité, on peut montrer $\forall \varepsilon > 0$, $\exists \theta > 0 \ \forall (t,y) \in [a,b] \times D, \quad |y - y_0| \leqslant \theta \implies |\frac{\partial h}{\partial y}(t,y) \leqslant \varepsilon$ (comme D est ouverte, on peut choisir θ tel que $|y - y_0| \leqslant \theta \implies y \in D$). Le théorème des accroissements finis donne que $h(t,y) - h(t,y_0) = \frac{\partial h}{\partial y}(t,\xi)(y-y_0)$, donc en valeur absolue cette quantité est inférieure à $\varepsilon|y - y_0|$. ■

On peut conclure la preuve :

Preuve.
▷ $\left|F(y) - F(y_0) - \left(\int_a^b \frac{\partial f}{\partial y}(t,y_0)(y-y_0)\mathrm{d}t\right)\right| = \left|\int_a^b (f(t,y) - f(t,y_0)) \frac{\partial f}{\partial y}(t,y_0)(y-y_0)\mathrm{d}t\right| \leqslant$ à une intégrale donnee par l'inégalité de la valeur absolue, et à condition que $y - y_0 < \theta$, elle-même inférieure à $\varepsilon(b-a)(y-y_0)$. Donc, pourvu que cette condition soit vérifiée, le taux d'accroissement entre y et y_0 auquel on soutrait l'intégrale de la dérivée partielle de f en t,y_0 est inférieur à $\varepsilon(b-a)$, donc la limite du taux d'accroissement tend bien vers l'intégrale et voilà. ■

1.5 Intégrales impropres

1.5.1 Définition

Soit $b \in \mathbb{R} \cup \{+\infty\}$.

Définition. ***(Intégrabilité locale)***

Soit $f : [a,b[\longrightarrow \mathbb{R}$. On dit que f est *localement intégrable* lorsque $f_{|[a,x]}$ est intégrable pour tout $x \in [a,b[$.

Définition. ***(Intégrale impropre)***

Soit $f : [a,b[\longrightarrow \mathbb{R}$ une fonction localement intégrable. L'intégrale impropre de f, si elle existe dans $\mathbb{R}$, est $\lim\limits_{x \to b} \displaystyle\int_a^b f(t)\mathrm{d}t$. Cette limite est appelée la *valeur* de $\displaystyle\int_a^b f(t)\mathrm{d}t$ et l'on note en conséquence (on écrit $I = \int_a^b f$ pour signifier $I = \lim\limits_{x \to b} \displaystyle\int_a^b f(t)\mathrm{d}t$).

Sa *nature* est soit convergente, soit divergente. On dit que l'intégrale de a à b est *convergente* si l'intégrale impropre existe dans $\mathbb{R}$.

Définition. ***(Intégrabilité généralisée)***

Soit $f : [a,b[\longrightarrow \mathbb{R}$ une fonction. On dit que l'intégrale de f est *existante* sur $[a,b[$ si elle est localement intégrable et son intégrale impropre existe.

On dit que l'intégrale de a à b est divergente lorsqu'elle n'est pas convergente, c'est-à-dire si $\int_a^b f = \pm\infty$ ou n'existe pas.

Propriété. ***(Caractère asymptotique de l'intégrabilité généralisée)***

Soit f localement intégrable sur $[a,b[$. Soit $c \in [a,b[$. Alors les intégrales impropres $\displaystyle\int_a^b f$ et $\displaystyle\int_c^b f$ sont de même nature.

▷ C'est une relation de Chasles (l'écrire). ■

Théorème. ***(Condition de Cauchy)***

Étant donnée une fonction $f : [a,b] \longrightarrow \mathbb{R}$, $\displaystyle\int_a^b f$ converge si et seulement si $\lim\limits_{\substack{c \to b \\ d \to b}} \displaystyle\int_c^d f = 0$, c'est-à-dire si pour tout $\varepsilon > 0$, il existe A_ε tel que pour tous $(c,d) \in [a,b[^2$, $c > A$ et $d > A$ impliquent que l'intégrale de c à d est majorée par ε.

▷ La preuve de ce fait évident n'est pas du tout immédiate. Nous montrons le sens réciproque, puis le sens direct. ■

1.5.2 Intégrales de référence

Dans tous ces cas, il faut, maintenant que l'on travaille dans le cadre de la théorie de l'intégration générale de Riemann, vérifier l'intégrabilité locale avant de parler d'intégrabilité généralisée.

1.5.2.1 Intégrales de Riemann

1.5.3 Théorèmes de comparaison

1.5.3.1 Convergence absolue

Théorème. ***(Premier théorème de comparaison)***

Toute intégrale absolument convergente est convergente.

▷ Par hypothèse, la condition de Cauchy est vérifiée pour $|f|$. Or inégalité de la valeur absolue donc condition de Cauchy pour f et voilà. ■

1.5.3.2 Domination

Théorème. ***(Deuxième théorème de comparaison)***

Si f,g localement intégrables sur $[a,b[$ eti g domine f, soit $|f| \leqslant Mg$, alors la convergence de l'intégrale de g entraine celle de celle f.

▷ C'est encore une fois la condition de Cauchy (l'écrire). ■

Heuristique

Conjuguée aux intégrales de Riemann, les théorèmes de comparaison donnent une idée du « paysage » des intégrales impropres.
Cependant, il ne faut qu'il vienne à l'idée de comparer toutes les intégrales impropres à des intégrales de Riemann ! Des considérations annexes à ce sujet sont l'objet des exercices qui suivent.

1.5.3.3 L'équivalence

Théorème. ***(Troisième théorème de comparaison)***

Si f,g localement intégrables sur $[a,b[$ et g positive sur V(b) et $f \sim g$ en b, alors leurs intégrales sont de même nature.

▷ Le coup du $\frac{1}{2}$ et l'on applique le TC2. ■

1.5.4 Exemples

1.5.4.1 Comparaison à une intégrale de référence

Exemple

$\displaystyle\int_1^{+\infty} e^{-t^u} \mathrm{d}t$; comparons-le à t^s.

1.5.4.2 Intégrale de Dirichlet

Exemple

$\displaystyle\int_1^{+\infty} \frac{\sin(t)}{x} \mathrm{d}t$. La fonction se prolonge par continuité en 0 ; il n'y a de problème qu'en l'infini.
On étudie les intégrales partielles en intégrant par parties, ce qui est courant pour les intégrales semi-convergentes (c'est un point méthode classique). Prouver le reste est plus délicat. Une façon de faire :
L'intégrande est négatif si et seulement si il existe k tel que $2kn < t < (2k+1)n$. Même chose pour les négatif après un impair pour -1.
Alors $\displaystyle I(x) = \sum_{k=1}^{\lfloor \frac{x}{n} \rfloor} \int_{(k-1)n}^{kn} \frac{\sin(t)}{t} \mathrm{d}t + \int_{\lfloor \frac{x}{n} \rfloor n}^{+\infty} \frac{\sin(t)}{t} \mathrm{d}t$ et l'on conclut facilement.
On donne une autre façon de montrer la non-convergence absolue, plus *cool* : on utilise la deuxième formule de la moyenne avec la condition de Cauchy. On majore alors l'intégrale par $2/c$ avec un certain x donné par celle-là.

1.5.5 Intégrales impropres dépendant d'une paramétrique

Définition. ***(Convergence uniforme d'une famille d'intégrales impropres)***

Soit $(f_i)_{i \in I}$ une famille de fonctions de $[a,b[$ dans $\mathbb{R}$. On suppose que chaque fonction f_i est localement intégrable sur $[a,b[$. On dit que les intégrales $\displaystyle\int_a^b f_i(t)\mathrm{d}t$ convergent uniformément sur I lorsqu'il existe une fonction $\Phi : I \longrightarrow \mathbb{R}$ telle que

$$\forall \varepsilon > 0\ \exists A \in]a,b[\ \forall i \in I\ \forall x \in \mathbb{R} \qquad A < x < b \implies \left| \int_a^x f_i(t)\mathrm{d}t - \Phi(i) \right| < \varepsilon.$$

Remarque. Ceci entraîne en particulier que pour tout $i \in I$, $\Phi(i) = \int_a^b f_i(t)\mathrm{d}t$.

Propriété. ***(Condition de Cauchy uniforme)***

Les intégrales $\int_a^b f_i(t)\mathrm{d}t$ convergent uniformément sur I si et seulement si pour tout $\varepsilon > 0$, il existe $A \in]a,b[$ tel que pour tous $x,y \in \mathbb{R}^2$, dès que $A < \min(x,y) \leqslant \max(x,y) < b$, pour tout $i \in I$, $\left| \int_x^y f_i(t)\mathrm{d}t \right| < \varepsilon$.

Théorème. ***(Condition suffisante de domination)***

S'il existe une fonction $g : [a,b] \longrightarrow \mathbb{R}$ localement intégrable telle que $\int_a^b g(t)\mathrm{d}t$ converge et telle que pour tout $i \in I$, pour tout $t \in [a,b[$, $|f_i(t)| \leqslant g(t)$, alors les intégrales I_i convergent uniformément sur I.

Théorème. ***(Théorème d'interversion)***

Soit $(f_n)_n$ une suite de fonctions localement intégrables sur $[a,b[$. On suppose que cette suite converge uniformément vers f sur tout sous-segment de $[a,b[$ et que les I_n convergent uniformément sur $\mathbb{N}$, alors l'intégrale de f converge et c'est cette limite.

$\triangleright$ Soit (x_p) une suite de points de $[a,b[$ telle que x_p tende vers b. Posons $I_{n,p} == \int_a^{x_p} f_n$. Alors comme ce sont des intégrales propres, $\lim_{n\to\infty} I_{n,p} = \int_a^{x_p} f$. La deuxième hypothèse nous donne qu'il existe $\varphi : \mathbb{N} \longrightarrow \mathbb{R}$ telle que pour tout $\varepsilon > 0$, il existe $A \in [a,b[$ tel que pour tout $x \in \mathbb{R}$, si $A < x < b$, pour tout n, $\left| \int_x^y f_n(t)\mathrm{d}t - \Phi(n) \right| < \varepsilon$. Puisque x_p tend vers b, il existe $p_0 \in \mathbb{N}$ tel que pour tout $p \in \mathbb{N}$ tel que dès que $p \geqslant p_0$, $b - x_p < b - A$ puis $x_p > A$, puis pour tout $p \geqslant p_0$, pour tout n, $|I_{n,p} - \Phi(n)| < \varepsilon$ donc pour tout n, $\lim_{p\to\infty} \int_a^{x_p} f_n = \Phi(n) = \int_a^b f_n$. Montrons la convergence de $(\Phi(n))_n$. Pour cela on montre qu'elle est de Cauchy. On a :

$$|\Phi(n) - \Phi(m)| \leqslant \left| \Phi(n) - \int_a^{x_p} f_n \right| + \int_a^{x_p} |f_n - f_m| + \left| \Phi(m) - \int_a^{x_p} f_m \right|.$$

Or les deux termes à gauche et à droite sont majorés par ε dès que $p \geqslant p_0$, et pour n et m assez grand, le terme central également.

Posons L la limite. Montrons que $\int_a^{x_p} f$ tend vers L. La suite est laissée en exercice. ■

Heuristique

Par conséquent, les théorèmes d'échange de l'intégrale de Riemann ne se généralisent que pour des fonctions à valeurs dans des espaces de Banach.

Chapitre 2

Intégrale de Lebesgue

Résumé

L'intégrale de Borel-Lebesgue est une théorie développée il y a plus d'un siècle pour essayer de contourner les problèmes inhérents à l'intégrale de Riemann. Pour obtenir alors une classe de fonctions bien plus large que dans cette théorie, le prix à payer est la théorie de la mesure ; une fois ce cadre établi, l'intégration des fonctions est facilitée.

2.1 Idée

On prendra garde de distinguer la théorie de la mesure de Borel-Lebesgue (ou, par abus, de Lebesgue) : c'est la théorie qui est à Borel-Lebesgue, et la mesure de Lebesgue (sur $\mathbb{R}^n$) qui est un exemple de mesure de la théorie de Borel-Lebesgue.

Voyons intuitivement la différence de point de vue entre les intégrales de Riemann et de Lebesgue.

Heuristique

$\star$ **Intégrale de Riemann**.
Soit $f : I \longrightarrow \mathbb{R}$, I un intervalle de $\mathbb{R}$.
On pose pour définition $\displaystyle\int_I f \approx \sum_{n=0}^{N} f(x_n)(x_{n+1} - x_n)$.

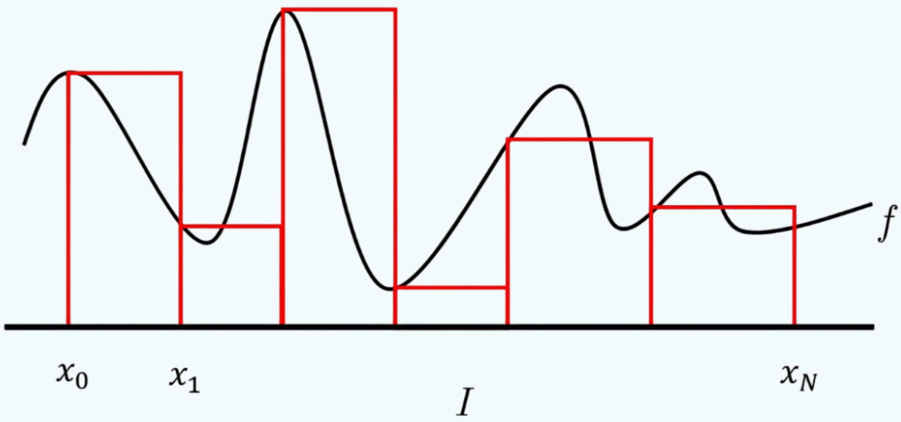

Figure A. *Principe de l'intégrale de Riemann.* —

On passe ensuite à la limite dans l'approximation précédente.
MAIS il faut que $f(x_n)$ « représente » bien l'intervalle $[x_n, x_{n+1}]$, autrement dit, que la quantité précédente ne dépende pas des points choisis.
La contrainte induite par cette propriété, porte sur la régularité de f : sa continuité, ou continuité par morceaux... mais l'intégrale de Riemann ne pourra guère intégrer des fonctions moins régulières que cela. ⋆ **Intégrale de Borel et Lebesgue**.

On change diamétralement de point de vue : on ne découpe pas l'intervalle I, c'est-à-dire l'axe des abscisses, mais l'axe des ordonnées, c'est-à-dire les valeurs prises par f.

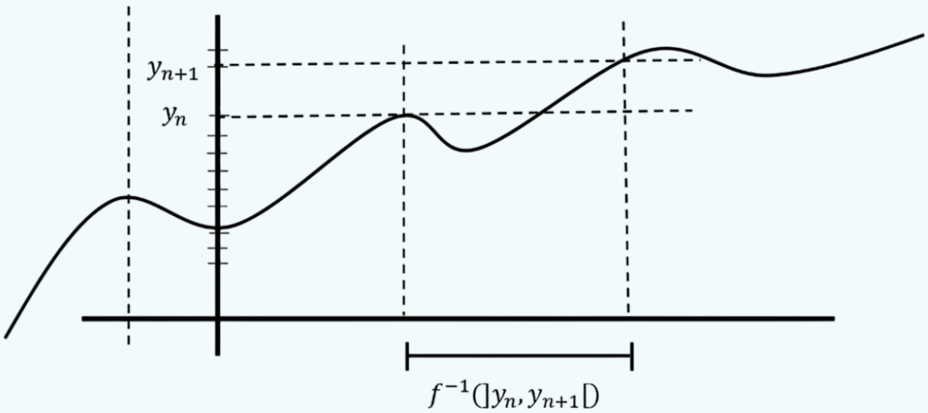

Figure A. *Principe de l'intégrale de Borel et Lebesgue.* —

On pose : $\int_I f \approx \sum_n y_n \mu(f^{-1}([y_n, y_{n+1}[))$.
Sous cette définition, la contrainte de régularité de f est remplacée par une contrainte de mesurabilité. Celle-ci est nettement moindre mais elle est sous-tendue par la construction d'une mesure sur I, ce qui est conceptuellement difficile.

De plus, notons que l'intégrale de Lebesgue généralise I à un ensemble quelconque, pourvu qu'il soit muni d'une tribu.

L'intégrale de Jordan

Théorie intermédiaire entre l'intégrale de Riemann et l'intégrale de Lebesgue, elle voit les choses dans le « mauvais sens »... Par exemple, elle approxime un mesurable par des ouverts par au-dessus et par des compacts par dessous, ce qui est curieux.

2.2 Tribus (ou σ-algèbres)

L'intégrale de Borel-Lebesgue se fonde sur la notion de fonctions mesurables, qui, par définition, préservent l'appartenance à une certaine tribu de parties d'un ensemble (quelconque). Nous commençons donc par étudier la notion de tribu.

2.2.1 Théorème d'unicité des mesures

Contre-exemple. ***(Très instructif)***

Soient μ, ν les deux mesures σ-finies : μ la mesure de Lebesgue sur $\mathbb{R}_+$ et ν la mesure de comptage sur $\mathbb{R}+$. On considère l'exhaustion de compact $K_n = [0,n]$, *i.e.* $\bigcup_{k\in\mathbb{N}} K_n = \mathbb{R}_+$.
Alors $\mu([0,n]) = \nu([0,n]) = n$.
D'autre part, pour le π-système $\mathcal{C} = \{[x,\infty[, x > 0\}$ engendrant les boréliens, on a également $\mu([x,\infty[) = \nu([x,\infty[) = +\infty$. Mais μ, ν ne coïncident pas ! Le π-système engendrant les boréliens ne contient pas l'exhaustion de compact de σ-finitude, le théorème ne s'applique pas ! $\square$

2.2.2 Tribu borélienne

Propriété. ***(Préservation des boréliens par homéomorphisme)***

Soit $(E,\mathcal{O})$ un espace topologique et h un automorphisme de E. L'image d'un borélien par h est un borélien de E.

▷ Soit $h : E \longrightarrow E$ un homéomorphisme. Soit $A \in \mathcal{B}(E)$. Alors $h(A) = g^{-1}(A)$ où $g = h^{-1}$ est continue. Or par le lemme de transport (section suivante), il suffit de montrer que si A est un ouvert (les ouverts engendrent les boréliens), alors $g^{-1}(A)$ est un borélien. C'est immédiat, car g étant continue, cet ensemble est un ouvert donc a fortiori un borélien. ■

On a même le lemme suivant, intuitif et qui pourra se révéler très utile.

Propriété. ***(Préservation de la tribu des boréliens par homéomorphisme)***

Soit $(E,\mathcal{O})$ un espace topologique et h un automorphisme de E. Alors h induit une permutation de $\mathcal{B}(E)$. En particulier, $h(\mathcal{B}(E)) = \mathcal{B}(E)$.

▷ On applique la propriété précédente et les propriétés élémentaires sur les images et préimages dans le cas d'une bijection. ■

2.2.2.1 Tribu borélienne de $\mathbb{R}$

2.2.2.1.1 Négligeabilité au sens de la mesure réelle de Lebesgue

On vérifie que la notion de négligeabilité au sens de Riemann coïncide bien avec les ensembles inclus dans des mesurables de mesure nulle au sens de la mesure de Lebesgue sur l'intervalle $\mathbb{R}$.

Propriété. ***(Caractérisation des négligeables de $\mathbb{R}$)***

Une partie $A \in \mathcal{P}(\mathbb{R})$ est λ-négligeable, c'est-à-dire qu'il existe $B \in \mathcal{B}(\mathbb{R})$ tels que $A \subseteq B$ et $\lambda(B) = 0$, si et seulement si, pour tout $\varepsilon > 0$, il existe une famille dénombrable d'intervalles ouverts $(I_j)_{j\in\mathbb{N}}$ tels que $E \subseteq \bigcup_{j\in\mathbb{N}} I_j$ et $\sum_{n=0}^{+\infty} \lambda(I_j) \leqslant \varepsilon$.

▷ Il s'agit d'exhiber les mesurables de λ nulle pour l'extension de Kolmogorov construisant la mesure de Lebesgue. ■

L'adhérence d'un négligeable n'est pas nécessairement négligeable ! Il suffit de penser à $\mathbb{Q}$, dénombrable, donc λ-négligeable, mais dense dans $\mathbb{R}$, d'où $\overline{\mathbb{Q}} = \mathbb{R}$ qui n'est certainement pas négligeable, car la mesure de Lebesgue n'est pas nulle.

2.2.2.1.2 Parties non boréliennes des réels

Il n'est pas évident d'exhiber un élément de $\mathcal{P}(\mathbb{R})$ qui ne soit pas un élément de $\mathcal{B}(\mathbb{R})$. En particulier, on aura recours à l'axiome du choix qui permet dans un premier temps d'en construire un.

Dans toute la suite, on note λ la mesure de Lebesgue sur $\mathbb{R}$.

Définition-propriété. ***(Ensembles de Vitali)***

On appelle *ensemble de Vitali* tout ensemble de représentants de $\mathbb{R}/\mathbb{Q}$ choisis dans $[0,1]$.

$\triangleright$ En effet, tout élément de $\mathbb{R}$ est associé modulo la relation d'équivalence $x \sim y \iff x-y \in \mathbb{Q}$ à un élément de $[0,1]$ par sa partie fractionnaire. L'existence d'un ensemble de représentants est assurée par l'axiome du choix **quelconque**, car on est en présence d'un ensemble indénombrable ($\mathbb{R}$). ■

Propriété. ***(Existence de non-boréliens)***

Tout ensemble de Vitali est non mesurable.

$\triangleright$ Soit V un ensemble de Vitali. On suppose que V est mesurable. On considère $Q = \mathbb{Q} \cap]-1,1[$ et on définit l'ensemble :

$$A = \bigcup_{r \in Q} (V + r).$$

On observe que par translation $A \subseteq [-1,2]$. De plus, A est mesurable par réunion mesurable de translatés de mesurables (donc qui leur sont homéomorphes...). On remarque que cette réunion est formée d'ensembles deux à deux disjoints puisque V ne contient qu'un réel par classe d'équivalence modulo $\mathbb{Q}$. Par conséquent :

$$\lambda(A) = \sum_{r \in Q} \lambda(V + r) = \sum_{r \in Q} \lambda(V).$$

Puisque la mesure de A est bornée par $3 = \lambda([-1,2]$, cela force cette famille constante à être sommable, d'où $\lambda(V) = 0$. En particulier, $\lambda(A) = 0$.

Cependant, pour tout $x \in [0,1]$, il existe $y \in V$ tel que $x - y \in \mathbb{Q}$, et donc, par hypothèse sur x et sur V, $x - y \in Q$. Ainsi $x \in V + (x - y)$ donc $x \in A$. Par suite, l'ensemble $[0,1]$ est inclus dans A, donc $\lambda(A) \geqslant 1$, ce qui fournit une contradiction. ■

Nécessité de l'axiome du choix

Pour construire les ensembles de Vitali, on utilise la forme forte (à valeurs dans une puissance du continu) de l'axiome du choix.

La théorie des ensembles peut démontrer qu'en présence de l'axiome du choix, grâce à une récurrence transfinie (théorie de la hiérarchie de Borel), on a $\mathrm{card}(\mathcal{B}(\mathbb{R})) = \mathrm{card}(\mathbb{R})$. Ceci redonne immédiatement l'existence de parties de $\mathbb{R}$ qui ne sont pas boréliennes, et même, montre que l'ensemble des non-boréliens de $\mathbb{R}$ a pour cardinal $2^{\mathbb{R}}$.

On peut se demander si, si l'axiome du choix est infirmé, ou en présence seule de l'axiome du choix dépendant, toutes les parties de $\mathbb{R}$ sont boréliennes, *i.e.* si $\mathcal{B}(\mathbb{R}) = \mathcal{P}(\mathbb{R})$. Ce n'est pas incohérent en considérant le fait suivant : si l'on admet que $\mathbb{R}$ est une réunion dénombrable d'ensembles dénombrables (puisque, pour montrer qu'une réunion dénombrable d'ensembles dénombrables est dénombrable, il faut l'axiome du choix), alors on peut montrer que toute partie de réelle est borélienne.

Plus précisément :

Proposition

Si $\mathbb{R}$ est réunion dénombrable d'ensembles dénombrables, alors $\mathcal{B}(\mathbb{R}) = \mathcal{P}(\mathbb{R})$.

▷ En effet, si $\mathbb{R} = \bigcup\limits_{n\in\mathbb{N}} A_n$ où les A_n sont sont des parties dénombrables de $\mathbb{R}$, si B est une partie de $\mathbb{R}$, alors $B = \bigcup\limits_{n\in\mathbb{N}} B \cap A_n$, mais pour tout $n \in \mathbb{N}$, $B \cap A_n$ est dénombrable, donc borélien comme réunion dénombrable de singletons. Par réunion dénombrable de boréliens, B est borélien, et c'est terminé. ■

Si l'on dispose de l'axiome du choix, on peut montrer bien plus fort, comme annoncé dans la note historique.

Théorème. ***(Cardinal des boréliens))***

Sous AC, $\mathrm{card}(\mathcal{B}(\mathbb{R})) = \mathrm{card}(\mathbb{R})$.

▷ On va montrer qu'il existe un ensemble fini X à cinq éléments tel que $X^{\mathbb{N}}$ se surjecte sur l'ensemble des boréliens. On pose $X = \{e, c, \cup, \cap, 0, 1\}$. Soient également $\phi : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}^*$ une injection, par exemple, qui à la décomposition en facteurs premiers d'un nombre associe ce nombre. Soit aussi $\theta : \{0,1\}^{\mathbb{N}} \longrightarrow \mathcal{P}(\mathbb{R})$ un injection d'image l'ensemble des intervalles de $\mathbb{R}$. On définit les éléments *bien formés* de $X^{\mathbb{N}}$ par récurrence transfinie, de la façon suivante : un élément u est dit bien formé s'il commence par e et s'il ne contient après e, que des 0 ou des 1. On définit alors $f(u) := \theta(n \mapsto u_{n+1})$. Soit $u \in X^{\mathbb{N}}$. On suppose que u commence par $\cup$ et que pour tout k, $n \mapsto u_{\phi(k,n)}$ est un élément bien formé. Alors on dit aussi que u est bien formé et l'on pose $f(u) := \bigcup_k f(n \mapsto u_{\phi(k,n)})$. Soit $u \in X^{\mathbb{N}}$. On suppose que u commence par $\cap$ et que pour tout k, $n \mapsto u_{\phi(k,n)}$ est un élément bien formé. Alors on dit aussi que u est bien formé et l'on pose $f(u) := \bigcap_k f(n \mapsto u_{\phi(k,n)})$. Soit enfin $u \in X^{\mathbb{N}}$. On suppose que u commence par c et que $n \mapsto u_{n+1}$ est bien formé. Alors on dit aussi que u est bien formé et l'on pose $f(u) := f(n \mapsto u^c_{n+1}$. Il ne reste plus qu'à montrer que f est une fonction sur le sous-ensemble de $X^{\mathbb{N}}$ contenant les suites bien formées, et que l'image de f est la tribu des boréliens. ■

On retrouve en particulier :

Corollaire. ***(Existence de non-boréliens)***

Sous AC, il existe au moins une partie non borélienne de $\mathbb{R}$. Autrement dit, $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$.

$\triangleright$ Conséquence de ce qui suit. $\blacksquare$

On a cependant le résultat beaucoup plus fort :

Corollaire. ***(Cardinal des non-boréliens)***

Sous AC, le cardinal de l'ensemble des parties non boréliennes de $\mathbb{R}$ est de $2^{\mathbb{R}}$.

$\triangleright$ En effet, un ensemble de cardinal $2^{\mathbb{R}}$ privé d'un ensemble de cardinal $\mathbb{R}$ est de cardinal $2^{\mathbb{R}}$. $\blacksquare$

2.3 Mesurabilité

2.3.1 Application mesurable

Définition. ***(Application mesurable)***

Soient (S, Σ) et $(F, \mathcal{F})$ deux espaces mesurables. On dit que $h : S \longrightarrow F$ est $(\Sigma, \mathcal{F})$-mesurable si $\forall A \in \mathcal{F}, h^{-1}(A) \in \Sigma$.
Naturellement, on dira aussi que h est un morphisme d'espaces mesurables.

Remarques.

1. Dans la suite du cours, on prendra fréquemment, mais pas toujours, $F = \mathbb{R}$ et $\mathcal{F} = \mathcal{B}(\mathbb{R})$.
2. On dira qu'une application est mesurable (sans préciser les tribus de départ et d'arrivée) lorsque les tribus mises en jeu le sont sans ambiguïté.
3. Parfois on se contentera de ne préciser que la tribu de départ : $h : S \longrightarrow \mathbb{R}$ est Σ-mesurable veut dire (la plupart du temps) que $(\Sigma, \mathcal{B}(\mathbb{R}))$-mesurable.
4. On est parfois amené à considérer à la place $F = \overline{\mathbb{R}}$ muni de sa topologie usuelle et de la tribu borélienne.

Pour conclure cette section, on mentionne un résultat en pratique inutile mais d'ordre épistémologique important.

Théorème. ***(Lien entre tribu et fonction mesurable)***

La donnée d'une tribu est équivalente à celle d'une fonction mesurable.

⟶ *Notation.* On se permettra donc d'écrire qu'une fonction mesurable appartient à une tribu, pour dire qu'elle est mesurable par rapport à cette tribu (que la tribu qu'elle engendre y est contenue).

2.3.2 Exemples de fonctions mesurables

Propriété. ***(Mesurabilité des applications constantes)***

Une application constante entre deux espaces mesurables est mesurable *quelles que soient* les tribus considérées.

▷ On reprend les notations précédentes. Soit $c \in \mathcal{F}$ et $h : S \longrightarrow F$, $x \longmapsto c$. Alors pour tout $A \in \mathcal{F}$, etc. ■

Propriété. ***(Mesurabilité des fonctions continues)***

Toute application continue entre deux espaces topologiques est mesurable par rapport aux tribus boréliennes associées.

▷ On utilise la définition topologique de la continuité. ■

Propriété. ***(Mesurabilité des fonctions continues par morceaux)***

Toute fonction continue par morceaux de $\mathbb{R}$ dans $\mathbb{K}$ est mesurable par rapport aux tribus boréliennes.

▷ Le nombre de discontinuité d'une fonction continue par morceaux est au plus dénombrable. ■

Propriété. ***(Mesurabilité des fonctions monotones)***

Toute fonction monotone de $\mathbb{R}$ dans $\mathbb{R}$ est mesurable.

▷ De même, on sait que le nombre de discontinuité d'une telle fonction est au plus dénombrable, ce qui permet de tracter au niveau des mesurables. ■

2.4 Mesure générale

2.4.1 Atomes, mesure diffuse

Définition. ***(Atome d'une mesure)***

Soit $(E, \mathcal{T}, \mu)$ mesuré. On dit que $x \in E$ est un *atome* de μ si $\{x\}$ est mesurable et $\mu(\{x\}) > 0$.

Propriété

L'ensemble des atomes d'une mesure quelconque est au plus dénombrable.

Définition. ***(Mesure continue, mesure diffuse)***

Soit $(E, \mathcal{T}, \mu)$ mesuré. On dit que μ est *diffuse*, ou parfois *continue*, si elle est continue.

Propriété

Tout mesure à densité contre la mesure de Lebesgue est continue.

2.4.2 Théorème de décomposition des mesures

Définition. ***(Mesure purement atomique)***

Une mesure sur un espace $(E, \mathcal{T})$ mesurable est dite *purement atomique*, si on peut l'écrire sous la forme

$$\sum_{x \in \mathcal{A}} \mu(\{x\})$$

où $\mathcal{A}$ est un ensemble dénombrable.

VOC Les anglais disent *pure point*, d'où la notation.

Définition. ***(Partie atomique)***

Soit $(E, \mathcal{T}, \mu)$ mesuré. On appelle *partie atomique* de μ la mesure atomique (le vérifier) $\mu_{pp} = \sum_{x \in \mathcal{A}} \mu(\{x\})$ où $\mathcal{A}$ est l'ensemble des atomes de μ.

Propriété

Soit μ une mesure. Alors $\mu_c = \mu - \mu_{pp}$ est diffuse.

Théorème. ***(Décomposition de Lebesgue)***

Soit μ une mesure borélienne finie. Alors :

$$\mu = \mu_{pp} + \mu_c = \mu_{pp} + \mu_{ac} + \mu_{sc}$$

où μ_c est continue, μ_{ac} est absolument continu par rapport à la mesure de Lebesgue, et μ_{sc} est un reste appelé *mesure singulière.*

2.5 Intégration

$\longrightarrow$ *Notation.* Soit $f : S \longrightarrow \mathbb{R}$. Soit $c \in \mathbb{R}$. On note $\{f \leqslant c\} = f^{-1}(]-\infty, c])$.

2.5.1 Intégration de fonctions positives

2.5.2 Bestiaire de suites de fonctions bizarrement intégrables... ou pas

Remarque importante. Pour obtenir un exemple de la forme d'une fonction à partir des suites de fonctions suivantes, il suffit parfois de sommer : $f = \sum_{n=0}^{\infty} f_n$. La justification de la sommation infinie est sans problème puisque dans la plupart des cas les termes des suites suivantes sont continus à support compacts, donc uniformément continus.

2.5.2.1 Bosses et vagues

Définition. ***(Bosse glissante)***

On appelle *bosse glissante*, la suite de fonctions affines par morceaux donnée pour tout $n \in \mathbb{N}$ par f_n, affine sur les intervalles $[n-1, n]$ et $[n, n+1]$, telle que $f(n-1) = f(n+1) = 0$ et $f(n) = 1$, et nulle en dehors des intervalles $[n-1, n+1]$.

Remarque. De toute bosse glissante de masse 1 comme exposée précédemment on peut déduire une bosse glissant de masse m pour tout $m \in \mathbb{R}$. Il suffit en effet de considérer la suite $(mf_n)_{n\in\mathbb{N}}$.

Définition. ***(Double bosse)***

On appelle *double bosse glissante*, la suite de fonctions donnée pour tout $n \in \mathbb{N}$ par $f_n(x) = \mathbb{1}_{n+1>|x|>n}$ pour tout $x \in \mathbb{R}$.

Propriété

La double bosse glissante est composée de fonctions L^1 mais n'est pas dominée par une fonction L^1. De plus, sa somme n'est pas L^1.

Définition. ***(Escalier belge)***

On appelle *escalier belge*, une suite de fonctions du type suivant : $(ng_n)_{n\in\mathbb{N}}$, où (g_n) est une bosse glissante.

Définition. ***(La montagne intégrable)***

On appelle *montagne intégrable*, la **série** de terme général l'escalier belge suivant : les fonctions affines par morceaux données pour tout $n \in \mathbb{N}$ par f_n, affine sur les intervalles $[n - \frac{1}{4^n}, n]$ et $[n, n + \frac{1}{4^n}]$, telle que $f(n - \frac{1}{4^n}) = f(n + \frac{1}{4^n}) = 0$ et $f(n) = 2^n$, et nulle en dehors des intervalles $[n - \frac{1}{4^n}, n + \frac{1}{4^n}]$.

Propriété

La montagne intégrable est intégrable et non bornée.

Définition. ***(Vague infinie)***

On appelle *vague infinie*, la suite de fonctions de $\mathbb{R}$ dans $\overline{\mathbb{R}}$ donnée par $f_n(x) = +\infty$ si $x = n$ et $f_n(x) = 0$ sinon, pour tout $n \in \mathbb{N}$.

Définition. ***(Chameau élastique)***

On appelle *chameau élastique*, la suite de fonctions données par $f_0 = g_0$ et $f_n = g_0 + g_n$ pour tout $n \geqslant 1$, où (g_n) est une bosse glissante.

2.5.2.2 Puits infinis de la théorie de la mesure

Définition. ***(Puits infini droit)***

On appelle *puits infini droit*, la suite de fonctions donnée pour tout $n \in \mathbb{N}$ par f_n la bosse affine par deux morceaux de largeur $\frac{1}{2n}$, centrée en $\frac{1}{n}$, de hauteur n.

Propriété

Le puits infini droit est non borné, a la suite de ses supports strictement décroissante, est à support uniforme compact strictement décroissant, mais de somme non intégrable.

Définition. ***(Puits infini droit rétréci)***

On appelle *puits infini droit rétréci*, la suite de fonctions donnée pour tout $n \in \mathbb{N}$ par f_n la bosse affine par deux morceaux de largeur $\frac{1}{2n^2}$, centrée en $\frac{1}{n}$, de hauteur n^2.

Propriété

Le puits infini droit rétréci est non borné, a la suite de ses supports strictement décroissante, est à support uniforme compact **terme général d'une série convergente**, mais de somme non intégrable.

Définition. ***(Puits infini inversé)***

On appelle *puits infini inversé*, l'opposé du puits infini droit.

2.5.2.3 Stroboscopes

Définition. ***(Les stroboscopes infernaux)***

On appelle *stroboscope*, toute suite de fonctions donnée par un éclairage (a_n) qui se déplace le long de $[0,1]$ modulo 1, tel que la série $\sum a_n$ soit divergente, par exemple $a_n = 1/n$.

Définition. ***(Stroboscope classique)***

On appelle *stroboscope classique*, la suite (g_n) donnée par la construction suivante. Pour tous $n \geqslant 1$ et pour tout $j \in [\![1,\ldots,n]\!]$, on pose $Y_{n,j}(\omega) = \mathbb{1}_{\left[\frac{j-1}{n},\frac{j}{n}\right]}(\omega)$ pour tout $\omega \in [0,1]$ et 0 ailleurs. On pose $X_1 = Y_{1,1}$, $X_2 = Y_{2,1}$, $X_3 = Y_{2,2}$, puis $X_4 = Y_{3,1}$, etc. Ainsi $X_m = Y_{n,j}$ pour un unique couple (n,j) avec $n \in \mathbb{N}$ et $j \in [\![1,n]\!]$. Il est clair que $n(m) \longmapsto +\infty$ quand $m \to \infty$. On pose $g_n = X_n$ pour tout $n \in \mathbb{N}$.

Propriété

Le stroboscope tend en mesure vers zéro mais ne converge nulle part.

2.5.3 Comparaison de l'intégrale de Riemann et de l'intégrale de Lebesgue

Le fait général est le suivant : l'intégrale de Lebesgue étend celle de Riemann. Ce qui est vrai sur un segment ne l'est plus nécessairement sur un intervalle impropre, d'où quelques précisions.

Exemple. ***(Fonction intégrable à discontinuité de deuxième espèce)***

La fonction $x \mapsto \sin(\frac{1}{x})$ n'a qu'une discontinuité en 0. Par Lebesgue-Vitali, elle est Riemann-intégrable sur $[-1,1]$ et donc aussi Lebesgue-intégrable.

Contre-exemple. ***(Fonction non réglée Riemann-intégrable)***

L'exemple précédent convient. □

Contre-exemple. ***(Fonction qui n'est pas limite simple d'applications continues par morceaux)***

L'application $\mathbb{1}_{C_3}$ n'est pas limite simple d'applications continues par morceaux. Pourtant, elle est Lebesgue-intégrable. □

Contre-exemple. ***(Fonction non continue par morceaux Lebesgue-intégrable)***

$\mathbb{1}_{\mathbb{Q}}$ convient. Notons que cette fonction n'est continue nulle part donc ne satisfait pas aux hypothèses du théorème de caractérisation de Lebesgue-Vitali, d'où le résultat suivant. □

Contre-exemple. ***(Fonction non Riemann-intégrable mais Lebesgue-intégrable)***

Même exemple. □

Sur un intervalle quelconque, les choses se compliquent.

Contre-exemple. ***(Fonction Riemann-intégrable non Lebesgue-intégrable)***

Toute intégrable semi-convergente convient, le sinus cardinal par exemple.
Remarquons la malhonnêteté de ce exemple. La fonction $x \mapsto \frac{\sin(x)}{x}$ est d'intégrale convergente pour l'intégrale de Riemann généralisée, mais par compensation ! En effet, sa partie positive *n'est pas intégrable* : autrement, on aurait par quelques manipulations techniques l'intégrabilité de sa partie négative, et donc son intégrabilité en valeur absolue, qui est infirmée ! □

2.5.4 Théorèmes fondamentaux de l'intégrale de Lebesgue

Conclusion de l'intégrale de Lebesgue

Les théories de l'intégration ont originellement pour but de résoudre le problème bien connu en langue anglaise sous le nom d'*antiderivative* : étant donné une fonction, trouver la fonction dont elle est dérivée. Le problème est que les fonctions dérivées peuvent écoper de propriétés très mauvaises, ces mêmes propriétés que l'intégrale de Riemann, la seule connue jusqu'au dix-neuvième siècle (avec ses sommes de Riemann), ne permet pas d'intégrer ou de primitiver.

L'intégrale de Lebesgue, comme l'intégrale de Jordan, souhaite remédier à ce problème. Cependant **elle échoue** à résoudre le problème de l'antidérivée. D'autres théories de l'intégration, plus récentes, comme l'intégrale d'Arnaud Denjoy, un français, ou l'intégrale de Kurzweil-Henstock, y parviennent.

Et pourtant... on s'en fout. L'intégrale de Lebesgue fournit à la place deux merveilles : d'abord, le théorème de convergence dominée qui fournit un outil extrêmement puissant pour intégrer des limites de fonctions relativement quelconques. D'autre part, elle sert d'assise théorique globale à la théorie des probabilités qui devait naître une quarante d'années après sa formalisation par Lebesgue.

2.5.5 Mesures à densité

Définition-propriété. ***(Mesure à densité)***

Soit $(E, \mathcal{A}, \mu)$ un espace mesuré et f un fonction mesurable positive sur E. Pour tout $A \in \mathcal{A}$, on pose : $\nu(A) = \int_E \mathbb{1}_A f \mathrm{d}\mu$. Alors ν est une mesure, dite *mesure à densité*, sur $(E, \mathcal{A})$.

$\triangleright$ On a, par définition, $\nu(A) = \int_A f(x)\mathrm{d}\mu$. On a $\nu(\emptyset) = 0$, car nulle partout. Soit (A_n) mesurables, disjoints. Par définition $\mu = \bigcup_n A_n = ...$ Or $\mathbb{1}_{\bigcup_n A_n} = \sum_{n=1}^{\infty} \mathbb{1}_{A_n}$ car les A_i sont disjoints. ■

Lemme

Tout événement μ-presque certain est ν-presque certain.

Propriété. ***(Intégration d'une fonction positive par rapport à une mesure à densité)***

Soit h un fonction $\mathcal{A}$-mesurable de E dans $\mathbb{R}$, positive. On a :

$$\int_E h \mathrm{d}\nu = \int_E hf \mathrm{d}\mu.$$

$\triangleright$ Si $h = \mathbb{1}_A$, alors $\int_E h\mathrm{d}\nu = \nu(A) = \int_E \mathbb{1}_A f \mathrm{d}\mu = \int_E hf\mathrm{d}\mu$ par définition de ν, donc la propriété est vraie pour les indicatrices. Si h est maintenant étagée, soit $h = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$, alors par linéarité,

$$\int_E h\mathrm{d}\nu = iintgE \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} \mathrm{d}\nu = \sum_{i=1}^{n} a_i \int_E \mathbb{1}_{A_i} \mathrm{d}\nu = \sum_{i=1}^{n} a_i \int_E \mathbb{1}_{A_i} f \mathrm{d}\mu = \int_E \left(\sum_{i=1}^{n} a_i \mathbb{1}_{A_i}\right) f \mathrm{d}\mu = \int_E hf\mathrm{d}\mu.$$

Enfin, si h est positive mesurable, alors il existe une suite de fonctions étagées (h_n) telle que $h_n \uparrow h$ μ-presque sûrement. Remarquons que si $\mu(A) = 0$ alors $\nu(A) = 0$ d'après le lemme précédent. En particulier, $h_n \uparrow h$ ν-presque sûrement. Par le théorème de convergence monotone, $\int_E h\mathrm{d}\nu =$

$\lim_{n\to\infty}\int_E h_n \mathrm{d}\nu = \lim_{n\to\infty}\int_E h_n f \mathrm{d}\mu = \int_E \lim_{n\to\infty} h_n f \mathrm{d}\mu = \int_E hf\mathrm{d}\mu$, car clairement, $h_n f \uparrow hf$ μ-presque sûrement. ■

Propriété. ***(Intégration par rapport à une mesure à densité)***

Soit h une fonction $\mathcal{A}$-mesurable de E dans $\mathbb{R}$. Alors h est ν-intégrable si et seulement si hf est μ-intégrable, et dans ce cas, on a encore :

$$\int_E h\mathrm{d}\nu = \int_E hf\mathrm{d}\mu.$$

▷ Posons $h = h_+ - h_-$ ses parties positive et négative à supports disjoints. On a $|h| = h_+ + h_-$. On a $h \in L^1(\nu)$ si et seulement si $\int_E |h|\mathrm{d}\nu < \infty$, c'est-à-dire d'après l'expression de $|h|$ si et seulement si $h_+ \in L^1(\nu)$ et $h_- \in L^1(\nu)$. En outre, h_+ est positive, donc $h_+ \in L^1(\nu)$ si et seulement si $\int_E h_+\mathrm{d}\nu < \infty$ si et seulement si $\int_E h_+ f\mathrm{d}\mu < \infty$ si et seulement si $h_+ f \in L^1(\mu)$. De même, h_- est positive, donc $h_- \in L^1(\nu)$ si et seulement si $h_- f \in L^1(\mu)$. Or f est positive, donc $(hf)_+ = h_+ f$ et $(hf)_- = h_- f$ puisque f est positive, donc $h \in L^1(\nu)$ si et seulement si $hf \in L^1(\mu)$. Pour conclure à la relation précédente, il suffit d'utiliser la linéarité sur parties positive et négative. ■

Chapitre 3

Exercices

Difficulté des exercices :

- $\bullet\circ\circ\circ\circ$ Question de cours, application directe, exercice purement calculatoire sans réelle difficulté technique
- $\bullet\bullet\circ\circ\circ$ Exercice faisable, soit intuitivement, soit en employant des moyens rudimentaires ou des techniques déjà vues
- $\bullet\bullet\bullet\circ\circ$ Exercice relativement difficile et dont la résolution appelle à une réflexion plus importante à cause d'obstacles techniques ou conceptuels, qui cependant devraient être à la portée de la plupart des étudiants bien entraînés
- $\bullet\bullet\bullet\bullet\circ$ Exercice très exigeant, destiné aux élèves prétendant aux concours les plus difficiles, exercice « classique ».
- $\bullet\bullet\bullet\bullet\bullet$ La résolution de l'exercice requiert un raisonnement et des connaissances extrêmement avancés, dépassant les attentes du prérequis. Il est presque impossible de le mener à terme sans indication. Bien qu'exigibles à très peu d'endroits, ces exercices sont très intéressants et présentent souvent des résultats forts.

Appendice

Table des matières

Bibliographie

[1] *Titre du livre*, Auteur du livre, date, maison d'édition

Table des figures

Liste des tableaux