Chapitre 1
Introduction

The beginner should not be discouraged if he finds that he does not have the prerequisites for reading the prerequisites. *Paul Halmos*

1.1 Rappels sur l'intégrale de Riemann

Le concept d'intégrale repose sur la notion d'aire : soit f une fonction continue sur l'intervalle $[a,b] \subset \mathbb{R}$, à valeurs positives, alors la quantité

$$\int_a^b f(x)\,dx$$

s'interprète comme l'aire du domaine

$$D = \{(x,y) \in \mathbb{R}^2,\ a \leq x \leq b,\ 0 \leq y \leq f(x)\}.$$

Plus généralement, si les variables x, y sont liées à des grandeurs physiques, l'intégrale a une signification physique induite par ces grandeurs (énergie, travail, aire,...).

En 1853, B. Riemann a introduit une définition rigoureuse de l'intégrale de la manière suivante. On considère une fonction $f : [a,b] \longrightarrow \mathbb{R}$. On introduit une subdivision de l'intervalle $[a,b]$:

$$a = x_0 < x_1 < ... < x_{n-1} < x_n = b$$

de pas

$$h = \sup\big\{|x_i - x_{i-1}|,\ i \in \{1,...,n\}\big\}.$$

(On notera que le pas h et le nombre n de points de la subdivision sont intimement liés : pour une subdivision régulière $h = (b-a)/n$ et $x_i = a + ih$). On pose alors

$$I_h = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$$

où ξ_i est un point choisi dans $[x_{i-1}, x_i]$. L'intégrale $\int_a^b f(t)\,dt$ est définie comme étant la limite, si elle existe indépendamment du choix des ξ_i, de I_h quand h tend vers 0.

Par exemple, considérons une fonction f continue sur $[a,b]$. On définit alors les quantités I_h^M et I_h^m, obtenues en prenant respectivement $f(\xi_i^M) = \max\big\{f(\xi),\ x_{i-1} \leq \xi \leq x_i\big\}$ et $f(\xi_i^m) = \min\big\{f(\xi),\ x_{i-1} \leq \xi \leq x_i\big\}$. Alors, les suites $\big(I_h^M\big)_{h>0}$ et $\big(I_h^m\big)_{h>0}$ sont bornées et quelque soit le choix de ξ_i on a toujours $I_h^m \leq I_h \leq I_h^M$. Mais, pour une telle fonction continue f, le théorème de Heine[1] assure que f est en fait uniformément continue sur $[a,b]$ et pour tout $\varepsilon > 0$, il existe $h(\varepsilon) > 0$ tel que si $0 < h < h(\varepsilon)$ alors $0 \leq I_h^M - I_h^m \leq \varepsilon$. On en déduit que $\liminf_{h\to 0} I_h^M = \limsup_{h\to 0} I_h^m = I$, puis que I_h admet une même limite I quand $h \to 0$ pour tout choix des ξ_i. En particulier, cette définition permet donc de considérer l'intégrale d'une fonction continue sur $[a,b]$. L'intégrale est alors liée à la notion de primitive car on a

$$\int_a^b f(t)\,dt = F(b) - F(a), \qquad F'(x) = f(x)$$

où on notera que F n'est définie qu'à une constante près. Sans difficulté, on peut aussi définir l'intégrale d'une fonction f continue par morceaux sur $[a,b]$, c'est-à-dire admettant un nombre fini de points de discontinuité (on remarquera qu'une telle fonction est bornée) : si $a_0, a_1, ..., a_n$ sont les points de discontinuité de f alors

$$\int_a^b f(t)\,dt = \int_a^{a_0} f(t)\,dt + \int_{a_0}^{a_1} f(t)\,dt + ... + \int_{a_n}^b f(t)\,dt.$$

Plus généralement encore, on peut définir l'intégrale (de Riemann) de *certaines* fonctions bornées ayant une infinité de points de discontinuité (ce sont les fonctions réglées qui sont limites uniformes sur $[a,b]$ de fonctions en escaliers, c'est-à-dire constantes par morceaux sur $[a,b]$)... la restriction, grosso modo, est que la fonction f "ne doit pas trop varier sur $[a,b]$". Enfin, l'intégrale de Riemann s'étend en une *intégrale généralisée* (ou *impropre*) pour des fonctions non bornées au voisinage d'un point $c \in [a,b]$ ou en considérant des intervalles d'intégration infinis. Cette théorie, et les règles de base du *calcul intégral* qui y sont liées, seront considérées comme des prérequis. On rappelle en particulier les exemples fondamentaux (dits de Riemann) :

$$\begin{cases} x \mapsto 1/x^\alpha \text{ admet une intégrale impropre sur }]0,1] \text{ si et seulement si } \alpha < 1, \\ x \mapsto 1/x^\alpha \text{ admet une intégrale impropre sur } [1,+\infty[\text{ si et seulement si } \alpha > 1. \end{cases}$$

(Et on rappelle aussi que $x \mapsto 1/x$ n'admet pas d'intégrale impropre ni en 0, ni en l'infini.) Cependant, la définition de Riemann soulève quelques difficultés qui la rendent peu appropriée pour poursuivre les développements de l'analyse mathématique.

1.2 Insuffisances de l'intégrale de Riemann

1.2.1 L'espace n'est pas complet

Tout comme on peut mesurer la distance entre deux points de l'axe réel, par exemple avec $|x-y|$, ou entre deux points de $\mathbb{R}^N$ avec les normes

[1] Voir par exemple [21, Théorème IV.5].

$$\begin{cases} \|x-y\|_1 = \displaystyle\sum_{i=1}^{N} |x_i - y_i|, \\ \|x-y\|_2 = \sqrt{\displaystyle\sum_{i=1}^{N} |x_i - y_i|^2}, \\ \|x-y\|_p = \left(\displaystyle\sum_{i=1}^{N} |x_i - y_i|^p\right)^{1/p}, \\ \|x-y\|_\infty = \max\{|x_i - y_i|,\, i \in \{1, ..., N\}\}, \end{cases}$$

on peut chercher à évaluer la distance qui sépare des fonctions définies sur $[a,b]$. On dispose par exemple de la norme de la convergence uniforme

$$\|f-g\|_\infty = \sup\{|f(x) - g(x)|,\, x \in [a,b]\}.$$

On désigne par $C^0([a,b])$ l'ensemble des fonctions continues sur l'intervalle $[a,b]$, à valeurs réelles (ou complexes, auquel cas $|\cdot|$ designe le module). Il est pertinent d'équiper $C^0([a,b])$ de cette norme. On dispose alors des propriétés suivantes :

- si $\big(f_n\big)_{n\in\mathbb{N}}$ est une suite de fonctions continues sur $[a,b]$ et $\|f_n - f\|_\infty$ tend vers 0 quand n tend vers l'infini alors f est continue sur $[a,b]$,
- $(C^0([a,b]), \|\cdot\|_\infty)$ est un un espace vectoriel normé *complet* : si $\big(f_n\big)_{n\in\mathbb{N}}$ est une suite composée de fonctions continues sur $[a,b]$ qui est de Cauchy pour la norme $\|\cdot\|_\infty$, alors il existe une fonction f continue sur $[a,b]$ telle que f_n converge uniformément vers f sur $[a,b]$.

Toutefois, on peut envisager d'autres manières de mesurer les distances entre fonctions, par exemple avec les quantités :

$$\begin{cases} N_1(f-g) = \displaystyle\int_a^b |f(t) - g(t)|\, dt, \\ N_2(f-g) = \sqrt{\displaystyle\int_a^b |f(t) - g(t)|^2\, dt}. \end{cases}$$

Ces quantités interviennent de manière naturelle dans de nombreux problèmes. Plus généralement, on peut considérer pour $p \geq 1$,

$$N_p(f-g) = \left(\int_a^b |f(t) - g(t)|^p\, dt\right)^{1/p}.$$

La difficulté provient du fait que $(C^0([a,b]), N_1(\cdot))$ n'est pas complet : il existe des suites de fonctions continues qui sont de Cauchy pour N_1 mais qui n'admettent pas de limite continue pour N_1.

Exemple 1.1. On pose

$$f_n(x) = \begin{cases} -1 & \text{sur }]-1, -1/n], \\ +1 & \text{sur } [1/n, +1[, \\ nx & \text{sur }]-1/n, 1/n[. \end{cases}$$

Il est instructif de dessiner le graphe de f_n. On montre aisément que $\big(f_n\big)_{n\in\mathbb{N}}$ est de Cauchy pour N_1. Par ailleurs, f_n converge simplement sur $[-1,+1]$ vers

$$f(x) = \begin{cases} +1 & \text{si } \ 0 < x \leq 1, \\ 0 & \text{si } \ x = 0, \\ -1 & \text{si } \ -1 \leq x < 0. \end{cases}$$

Comme la limite simple f n'est pas continue sur $[-1,+1]$, la convergence ne peut pas être uniforme sur $[-1,+1]$, et la suite $\big(f_n\big)_{n\in\mathbb{N}}$ n'est pas de Cauchy dans $(C^0([a,b]), \|\cdot\|_\infty)$. Cependant, un calcul direct, guidé par le graphe de f_n, montre que $\lim_{n\to\infty} N_1(f_n - f) = 0$. On pourrait aussi remarquer que $|f_n(x) - f(x)| \leq 1$, et appliquer le théorème de Lebesgue (voir plus loin) pour aboutir à cette conclusion. Enfin, on peut aussi montrer qu'on ne peut pas avoir $\lim_{n\to\infty} N_1(f_n - g) = 0$ avec g continue sur $[-1,+1]$ (en utilisant l'Exercice 1.14). En effet, supposons qu'une telle fonction existe. Alors, pour tout $\varepsilon > 0$, on a

$$\int_{-1}^{+1} |f_n(x) - g(x)|\, dx \geq \int_{\varepsilon}^{+1} |f_n(x) - g(x)|\, dx = \int_{\varepsilon}^{+1} |1 - g(x)|\, dx,$$

la dernière relation s'appliquant pour $n \geq 1/\varepsilon$. Comme on suppose que le terme de gauche tend vers 0 lorsque $n \to \infty$, on en déduit que $g(x) = 1$ pour $x \in [\varepsilon, 1]$, $\varepsilon > 0$ quelconque, donc $g(x) = 1$ sur $]0,1]$. Un raisonnement identique montre que $g(x) = -1$ sur $[-1,0[$. Ce raisonnement ne définit pas la valeur de g en $x = 0$, néanmoins on ne pourra pas rendre g continue sur $[-1,+1]$... On en conclut que $C^0([-1,+1]), N_1(\cdot))$ n'est pas complet.

Exemple 1.2. En fait, on peut même trouver des suites de Cauchy (pour N_1) de fonctions Riemann-intégrables qui convergent simplement vers des fonctions qui ne sont pas Riemann-intégrables. On note $\chi_{\mathbb{Q}}$ la fonction caractéristique des rationnels sur $[0,1]$:

$$\chi_{\mathbb{Q}}(x) = 1 \text{ si } x \in \mathbb{Q} \cap [0,1], \qquad \chi_{\mathbb{Q}}(x) = 0 \text{ si } x \in [0,1] \backslash \mathbb{Q}.$$

On obtient alors

$$I_h = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) = \begin{cases} \displaystyle\sum_{i=1}^{n} 1(x_i - x_{i-1}) = 1 \text{ si les } \xi_i \in \mathbb{Q}, \\ \displaystyle\sum_{i=1}^{n} 0(x_i - x_{i-1}) = 0 \text{ si les } \xi_i \in [0,1]\backslash\mathbb{Q}. \end{cases}$$

Donc I_h n'a pas de limite quand $h \to 0$, indépendante des ξ_i, et $\chi_{\mathbb{Q}}$ n'est pas Riemann-intégrable. Or, $\mathbb{Q}$ est dénombrable, donc on peut représenter $\mathbb{Q} \cap [0,1]$ par une suite $\{q_0, q_1, ..., q_n, ...\}$. On pose

$$f_n(x) = 1 \text{ si } x \in \{q_0, q_1, ..., q_n\}, \qquad f_n(x) = 0 \text{ sinon.}$$

Les fonctions f_n sont continues par morceaux donc Riemann-intégrables et on a

$$\int_0^1 f_n(x)\, dx = \int_0^{q_0} f_n(x)\, dx + \int_{q_0}^{q_1} f_n(x)\, dx + ... + \int_{q_n}^{1} f_n(x)\, dx = 0.$$

De plus, $\big(f_n\big)_{n\in\mathbb{N}}$ converge simplement vers $\chi_{\mathbb{Q}}$ et est une suite de Cauchy pour N_1.

On se trouve donc dans une situation analogue à celle rencontrée lorsqu'on construit l'ensemble des nombres réels $\mathbb{R}$ à partir de l'ensemble des nombres rationnels $\mathbb{Q}$: $\mathbb{Q}$ n'est pas complet et $\mathbb{R}$ complète $\mathbb{Q}$. L'intégrale de Lebesgue est un procédé permettant de construire le complété de C^0 pour N_1.

1.2.2 Passage à la limite

Dans le cadre de l'intégrale de Riemann, les théorèmes de "passage à la limite sous le signe somme" (ou de "dérivation sous le signe somme") sont assez difficiles. La convergence simple de f_n vers f ne suffit pas à assurer que

$$\lim_{n\to\infty}\int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

L'un des seuls résultats ("facile" à obtenir[2]) dont on dispose est le

Théorème 1.3. *Si $\big(f_n\big)_{n\in\mathbb{N}}$ est une suite de fonctions continues qui converge uniformément vers f sur $[a,b]$ alors*

$$\lim_{n\to\infty}\int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

En pratique, la convergence uniforme est un concept difficile et une propriété qui n'est parfois pas évidente à vérifier. Des exemples simples montrent que ce théorème, qui de surcroît ne s'applique qu'avec a et b finis, est loin d'être optimal.

Exemple 1.4. Les suites $f_n(x) = x^n$ ou e^{-nx} convergent simplement sur $]0,1[$ vers 0, mais pas uniformément, alors que les intégrales passent à la limite car on a $\lim_{n\to\infty}\int_0^1 f_n(x)\,dx = 0$.

L'intégrale de Lebesgue permet d'obtenir des énoncés d'un usage beaucoup plus souple et plus général. Le théorème central de ce manuscrit est le :

Théorème 1.5 (Théorème de Convergence Dominée de Lebesgue). *Soit $\big(f_n\big)_{n\in\mathbb{N}}$ une suite de fonctions (Lebesgue-)intégrables sur I (intervalle borné ou non). On suppose que*

i) Hypothèse de convergence : $\lim_{n\to\infty} f_n(x) = f(x)$, *pp* $x\in I$,
ii) Hypothèse de domination : il existe une fonction g, (Lebesgue-)intégrable sur I, telle que pour tout $n\in\mathbb{N}$ et pp $x\in I$ on a $|f_n(x)|\le g(x)$.

Alors la limite f est (Lebesgue-)intégrable sur I et on a

$$\lim_{n\to\infty}\int_I f_n(x)\,dx = \int_I f(x)\,dx$$

(et même $\lim_{n\to\infty}\int_I |f_n(x)-f(x)|\,dx = 0$.*)*

[2] Pour des résultats proches du théorème de Lebesgue, voir [9]... mais bien sûr les démonstrations demeurent lourdes et délicates.

Cet énoncé appelle quelques commentaires :

1. il utilise la notion d'intégrale de Lebesgue, mais sans préciser pour le moment davantage cette notion, on peut appliquer le résultat en disant que :
   - les fonctions Riemann-intégrables sont Lebesgue-intégrables,
   - les fonctions admettant une intégrale impropre absolument convergente sont Lebesgue-intégrables.
2. il utilise un nouveau quantificateur *pp*, qui signifie *presque partout*, mais sans préciser pour le moment davantage cette notion, on peut appliquer le résultat en disant que :
   - si les propriétés ont lieu pour tout x alors elles ont lieu *pp*,
   - si les propriétés ont lieu pour tout x, sauf pour un nombre fini de points, alors elles ont lieu *pp*,
   - si les propriétés ont lieu pour tout x, sauf pour un ensemble dénombrable de points, alors elles ont lieu *pp*.
3. Ce théorème assure aussi que la limite "ponctuelle" (ou *pp*) de fonctions Lebesgue-intégrables dominées est elle-même une fonction Lebesgue-intégrable (voir aussi sur ce point le Lemme de Fatou, Lemme 3.27).
4. L'hypothèse de domination est en pratique beaucoup plus facile à vérifier que les propriétés de convergence uniforme.

Exemple 1.6. En reprenant l'exercice 1.2, et en appliquant le théorème de Lebesgue, on obtient que $\chi_{\mathbb{Q}}$ est (Lebesgue-)intégrable et que son intégrale est nulle.

Exemple 1.7. L'hypothèse de domination ne peut pas être levée. Lorsque f_n converge simplement (ou *pp*) vers f, les obstacles à la convergence de $\int_I f_n\,dx$ vers $\int_I f\,dx$ sont illustrés par les exemples fondamentaux

- du créneau $f_n(x) = n\chi_{[-1/n,1/n]}(x)$,
- de la bosse glissante $g_n(x) = \chi_{[n,n+1]}(x)$.

Ces deux fonctions convergent *pp* vers 0 : pour f_n, le seul problème est en 0 où $f_n(0)$ tend vers $+\infty$, mais que pour $x \neq 0$, il existe $n(x) \in \mathbb{N}$ tel que $|x| > 1/n$ lorsque $n > n(x)$ donc $f_n(x) = 0$ pour n assez grand ; pour g_n on note que pour x fixé il existe $n(x) \in \mathbb{N}$ tel que $x < n(x)$ donc $g_n(x) = 0$ pour $n > n(x)$. Toutefois, leurs intégrales ne tendent pas vers 0 car on a

$$\int_{\mathbb{R}} f_n(x)\,dx = 2, \qquad \int_{\mathbb{R}} g_n(x)\,dx = 1.$$

Le premier exemple illustre un phénomène de concentration, le second un phénomène où "la masse s'échappe à l'infini".

1.2.3 Construction de l'intégrale de Lebesgue

La définition d'intégrale proposée par Henri Lebesgue dans sa thèse *Intégrale, longueur, aire* en 1902 généralise celle de Riemann[3] :

- les fonctions Riemann-intégrables sont Lebesgue-intégrables,
- il existe des fonctions Lebesgue-intégrables qui ne sont pas Riemann-intégrables (par exemple $\chi_{\mathbb{Q}}$),
- les fonctions qui admettent une intégrale impropre *absolument* convergente sont Lebesgue-intégrables,
- les théorèmes de passage à la limite deviennent plus faciles à manier et plus puissants,
- les fonctions Lebesgue-intégrables forment un espace complet pour N_1.

Dans la pratique, on utilise la théorie de l'intégrale de Lebesgue, mais on ne calcule que des intégrales de Riemann (de fonctions continues), tout comme on utilise les réels en ne faisant des calculs qu'avec des rationnels (et même des décimaux le plus souvent...).

La construction de l'intégrale de Lebesgue repose sur la théorie de la mesure (initialisée par E. Borel). Au lieu d'utiliser une subdivision "verticale" comme Riemann, Lebesgue introduit une subdivision "horizontale" de l'intervalle $m \le f(x) \le M$:

$$m = y_0 < y_1 < ... < y_n = M$$

de pas

$$h = \max\{y_i - y_{i-1},\, i \in \{1, ..., n\}\}.$$

Il attribue ensuite une mesure (un "poids") m_i à l'ensemble

$$J_i = \{x \in [a, b],\, y_{i-1} \le f(x) < y_i\}$$

et pose

$$J_h = \sum_{i=1}^{n} m_i \eta_i$$

où η_i est une valeur choisie dans $[y_{i-1}, y_i[$. L'intégrale est définie comme la limite, si elle existe indépendamment du choix des η_i, de J_h quand h tend vers 0. La première étape de cette construction consiste à définir la mesure d'un ensemble (c'est-à-dire comment obtenir les m_i) ; c'est ce que nous verrons au chapitre suivant. En fait, la stratégie est parfaitement illustrée par H. Lebesgue lui-même :

---

[3] On pourra consulter la Note de H. Lebesgue reproduite et commentée dans [1], ainsi que la présentation [26].

Je dois payer une certaine somme ; je fouille dans mes poches et j'en sors des pièces et des billets de différentes valeurs. Je les verse à mon créancier dans l'ordre où elles apparaissent jusqu'à atteindre le total de ma dette. C'est l'intégrale de Riemann.
Mais je peux procéder autrement...
Ayant sorti tout mon argent, je réunis les billets de même valeur, les pièces semblables et j'effectue le paiement en donnant ensemble les signes monétaires de même valeur. C'est mon intégrale.

Remarque 1.8. Outre son intérêt propre pour l'intégration et l'analyse, la théorie de la mesure a de nombreuses applications en théorie des probabilités et de nombreux résultats d'intégration se traduisent en langage probabiliste.

Remarque 1.9. La définition de Lebesgue conduit aussi à une géométrie fonctionnelle : on travaille dans des espaces vectoriels où les points sont des fonctions sur lesquelles on peut faire agir des opérateurs. C'est le domaine de l'analyse fonctionnelle.

Remarque 1.10. On a comparé N_1, N_2, ..., N_∞ sur des fonctions avec les normes $\|\cdot\|_1$, $\|\cdot\|_2$, ..., $\|\cdot\|_\infty$ dans $\mathbb{R}^N$. Dans $\mathbb{R}^N$ ces normes sont équivalentes, ce n'est plus vrai dans des epaces de fonctions qui sont de dimension infinie. On rappelle aussi que de toute suite bornée dans $\mathbb{R}^N$, on peut extraire une sous-suite convergente (c'est le théorème de Bolzano-Weierstrass). Là encore ce résultat est restreint à la dimension finie. On méditera avec profit l'exemple de $f_n(x) = \sin(nx)$ sur $[0, 2\pi]$. On vérifie que, pour $n \geq 1$, $N_2(f_n)^2 = \pi$, donc $\left(f_n\right)_{n\in\mathbb{N}}$ est bornée pour N_2. Mais on a aussi

$$\int_0^{2\pi} f_n(x) f_m(x)\, dx = 0$$

et donc, pour tous n, m distincts,

$$N_2(f_n - f_m)^2 = \int_0^{2\pi} |f_n(x) - f_m(x)|^2\, dx = 2\pi.$$

Ceci prouve qu'il n'est pas possible d'extraire de f_n une sous-suite convergente (puisqu'aucune suite extraite ne peut être de Cauchy). En revanche, on peut démontrer le théorème de Riemann-Lebesgue, à savoir que

$$\lim_{n\to\infty} \int_0^{2\pi} f_n(x)\varphi(x)\, dx = 0$$

pour toute fonction φ continue sur $[0, 2\pi]$. Le schéma de la preuve est le suivant. A l'aide d'intégrations par parties, on commence par calculer, pour $k \in \mathbb{N}$ fixé, $\int_0^{2\pi} f_n(x) x^k\, dx$. On en déduit que l'énoncé est vrai lorsque φ est une fonction polynômiale. On conclut par le théorème de Stone-Weierstrass qui assure que les polynômes sont denses dans les fonctions continues, c'est-à-dire que pour tout $\varepsilon > 0$ et toute fonction φ continue sur un intervalle borné I, il existe un polynôme P, dépendant de ε et de φ tel que $\|\varphi - P\|_\infty = \sup_{x\in I} |\varphi(x) - P(x)| \leq \varepsilon$.