Introduction : Motivations de l'intégrale de Lebesgue

À la fin du dix-neuvième siècle, les limitations de la théorie d'intégration de Riemann deviennent apparentes et plusieurs mathématiciens célèbres (Jordan, Borel, Young, ...) se mettent en devoir de la généraliser. C'est finalement la théorie de **Lebesgue**, exposée dans une note fondatrice de 1901, puis développée dans le Cours Peccot, qui sera adoptée par l'ensemble de la communauté mathématique. Elle se développe à partir du concept de **mesure**, introduit par **Borel** vers 1895.

La théorie de la mesure et l'intégration de Lebesgue seront ensuite perfectionnées et généralisées par de nombreux mathématiciens au cours du vingtième siècle, en particulier (par ordre chronologique approximatif) Carathéodory, Vitali, Radon, Riesz, Hausdorff, Kolmogorov et Besicovich. L'histoire de la théorie de la mesure est associée au développement de la théorie des probabilités, à celui de l'analyse harmonique moderne et même à celui de la logique axiomatique.

Avant d'entrer dans le vif du sujet, je vais passer en revue certaines des caractéristiques les plus notables de l'intégrale de Lebesgue, par comparaison avec d'autres théories de l'intégration.

Elargissement de la classe des fonctions intégrables

Le cadre classique le plus simple pour définir une intégrale est celui des fonctions en escalier sur un intervalle $[a, b]$, ou sa complétion pour la topologie de la convergence uniforme, l'espace des fonctions réglées (admettant une limite finie à droite et à gauche). La théorie de Riemann permet d'atteindre une plus grande généralité. Cependant, Riemann lui-même a conscience que l'intégrabilité "au sens de Riemann" impose encore des conditions relativement fortes : il démontre qu'une fonction $f : [a, b] \to \mathbb{R}$ est intégrable si et seulement si, pour tout $\alpha > 0$ donné, on peut choisir une décomposition de $[a, b]$ en sous-intervalles suffisamment fins pour que la somme des longueurs des sous-intervalles sur lesquels l'oscillation de la fonction dépasse α soit arbitrairement petite [Kahane, p. 64].

Plus tard, Lebesgue montrera qu'une fonction $f : [a, b] \to \mathbb{R}$ est Riemann-intégrable si et seulement si l'ensemble de ses points de discontinuité est *de mesure nulle*, au sens où on peut l'inclure dans une union d'intervalles ouverts dont la somme des longueurs est arbitrairement petite.

Ces conditions peuvent sembler assez faibles, puisqu'elles autorisent par exemple une fonction qui ne serait discontinue qu'en une quantité dénombrable de points. Mais il est facile de construire des fonctions bornées "naturelles" ne remplissant pas ces conditions : le contre-exemple qui vient tout de suite à l'esprit est la fonction indicatrice de $\mathbb{Q}$, ou sa restriction à un segment. Dans de nombreux problèmes d'analyse, on rencontre des fonctions qui ne sont pas forcément Riemann-intégrables.

Dans la théorie de Lebesgue, la classe des fonctions intégrables est beaucoup plus grande. Par exemple, toute fonction bornée "raisonnable" (en gros, que l'on peut décrire par un énoncé mathématique) est Lebesgue-intégrable. En outre, sa théorie généralise bien celle de Riemann.

Solution du problème des primitives

C'est par ce problème que Lebesgue motive sa construction dans sa note de 1901. L'intégrale de Riemann permet d'intégrer des fonctions discontinues, mais ne permet pas d'intégrer n'importe quelle fonction dérivée, même bornée ! Si donc f est une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$, il n'est pas garanti que l'identité

$$f(b) - f(a) \ = \ \int_a^b f'(t)\,dt \tag{1}$$

ait un sens. En fait, au début du siècle, divers auteurs (Volterra, Köpcke, Brodén, Schoenflies) construisent des classes de fonctions dérivables, dont la dérivée est bornée mais non Riemann-intégrable [Hawkins, pp. 57 et 108-110].

Au contraire, dans la théorie de Lebesgue, la dérivation et l'intégration deviennent des opérations inverses, sous des hypothèses simples. C'est ainsi que l'identité (1) est automatiquement vérifiée dès que f est continue sur $[a,b]$ et dérivable sur $]a,b[$, de dérivée bornée.

Bon comportement face aux limites

Peut-on échanger les opérations limite et intégrale ? C'est un problème classique, source de milliers d'exercices dans le cadre de l'intégrale de Riemann. On ne peut même pas y formuler le problème de manière suffisamment générale, car une limite de fonctions Riemann-intégrables n'est pas forcément Riemann-intégrable, même si ces fonctions sont uniformément bornées. Pour s'en convaincre, on peut noter que la fonction indicatrice de $[0,1]\cap\mathbb{Q}$, non intégrable au sens de Riemann, est limite de limites de fonctions continues puisque

$$1_{\mathbb{Q}}(x) = \lim_{n\to\infty}\lim_{m\to\infty}[\cos(2\pi n!x)]^m.$$

Au contraire, Lebesgue parvient à définir un concept de fonctions intégrables qui est stable par passage à la limite dans de nombreuses situations, et pour lequel l'échange intégrale-limite est presque automatique, sous des hypothèses simples et faciles à vérifier. Ce bon comportement par rapport aux limites trouve un intérêt même dans le cadre des fonctions Riemann-intégrables ! Pour s'en convaincre, on pourra méditer sur l'exercice suivant :

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues $[a,b]\to[0,1]$, convergeant (ponctuellement) vers 0. Alors $\int_a^b f_n \longrightarrow 0$.

Cet énoncé a bien sûr un sens dans le cadre de l'intégrale de Riemann, pourtant sa démonstration au moyen d'outils classiques est délicate (l'hypothèse naturelle dans cette théorie est la convergence uniforme et non la convergence simple) ; alors que la théorie de Lebesgue résout le problème sans douleur !

Mauvais traitement des compensations

Pour puissante qu'elle soit, la théorie de Lebesgue est impuissante à traiter la “semi-convergence” des intégrales, c'est-à-dire les situations où une fonction f se trouve être intégrable du fait de compensations entre valeurs positives et négatives, alors que sa valeur absolue $|f|$ n'est pas intégrable. D'autres théories sont plus habiles à tirer parti des compensations : ainsi les intégrales M (intégrale de Denjoy-Perron) ou M^2 présentées dans [Zygmund, T.2, pp.83–91].

Pourquoi alors n'enseigne-t-on pas ces théories alternatives plutôt que celle de Lebesgue ? D'une part, parce que dans l'immense majorité des applications, le mauvais traitement des intégrales semi-convergentes s'avère sans gravité (il s'agit en fait de situations relativement exceptionnelles) ; d'autre part, parce que la théorie de Lebesgue dispose d'une grande souplesse qui lui permet de se généraliser de manière abstraite, comme il est expliqué ci-après.

Insensibilité à la topologie

Dans l'intégrale de Riemann, un rôle particulier est joué par les propriétés de régularité, en un sens très lâche, des fonctions que l'on veut intégrer (variation importante au voisinage d'un point...) On a déjà mentionné, par exemple, des critères d'intégrabilité faisant intervenir l'ensemble l'ensemble des points de discontinuité. La topologie de l'espace de définition des fonctions (en l'occurrence la droite réelle) intervient donc. Cela se reflète sur les généralisations abstraites : pour adapter la construction de Riemann à des espaces plus généraux, on est tout de suite amené à faire des hypothèses de nature topologique assez fortes. (Ce constat vaut aussi pour les intégrales de Denjoy-Perron ou M^2.)

Au contraire, comme le comprit Radon vers 1913, l'intégrale construite par Lebesgue peut être adaptée à un cadre extrêmement général, sans qu'aucune hypothèse topologique ne soit faite sur l'espace d'intégration, et l'expérience montre que l'on peut construire ainsi des théories maniables. Le cas échéant, cet espace pourra même être un espace fonctionnel de dimension infinie : un exemple classique et très important dans les applications est l'**espace de Wiener**, qui n'est autre que l'espace $C([0,1],\mathbb{R}^d)$ des fonctions continues sur $[0,1]$ à valeurs dans $\mathbb{R}^d$. La **mesure de Wiener** est une mesure “gaussienne” sur cet espace, d'importance capitale en probabilités et en physique (mouvement brownien).

Idée de départ : intégration par tranches

L'idée de départ de Lebesgue semble très simple. Comme dans le cas de l'intégrale de Riemann, il s'agit d'approcher l'aire sous le graphe de la fonction par une union de rectangles. Mais ces rectangles sont définis de manière différente ! Dans le cas de Riemann, on s'intéresse aux variations de la fonction sur son domaine de définition : une base étant donnée, on définit le rectangle comme l'ensemble des points au-dessus de cette base, qui sont situés en-dessous de la courbe. Au contraire, dans le cas de Lebesgue, on définit le rectangle en fonction des *valeurs* de la fonction, sans jamais s'intéresser trop à l'espace de départ. Ce n'est donc pas la base du rectangle que l'on se donnera au départ, mais une variation dans les valeurs atteintes (côté “vertical”). Evidemment, on doit admettre que plusieurs rectangles partagent un même côté vertical. C'est ici que l'intégrale de Lebesgue va gagner toute sa complexité : alors que dans l'intégrale de Riemann, une brique élémentaire est un simple rectangle, dans celle de Lebesgue, il pourra s'agir de plusieurs ou même d'une infinité de rectangles.

L'exemple de la fonction indicatrice de $\mathbb{Q}$ est révélateur : bien que discontinue partout, cette fonction est très facile à décrire en fonction de ses valeurs. Dans la théorie de Riemann, on tenterait vainement de découper le segment $[0,1]$ en tout petits intervalles où cette fonction ne varie guère ; dans celle de Lebesgue, on partage $[0,1]$ en seulement deux morceaux qui sont assez complexes (totalement discontinus) mais sur chacun desquels la fonction est effectivement constante.

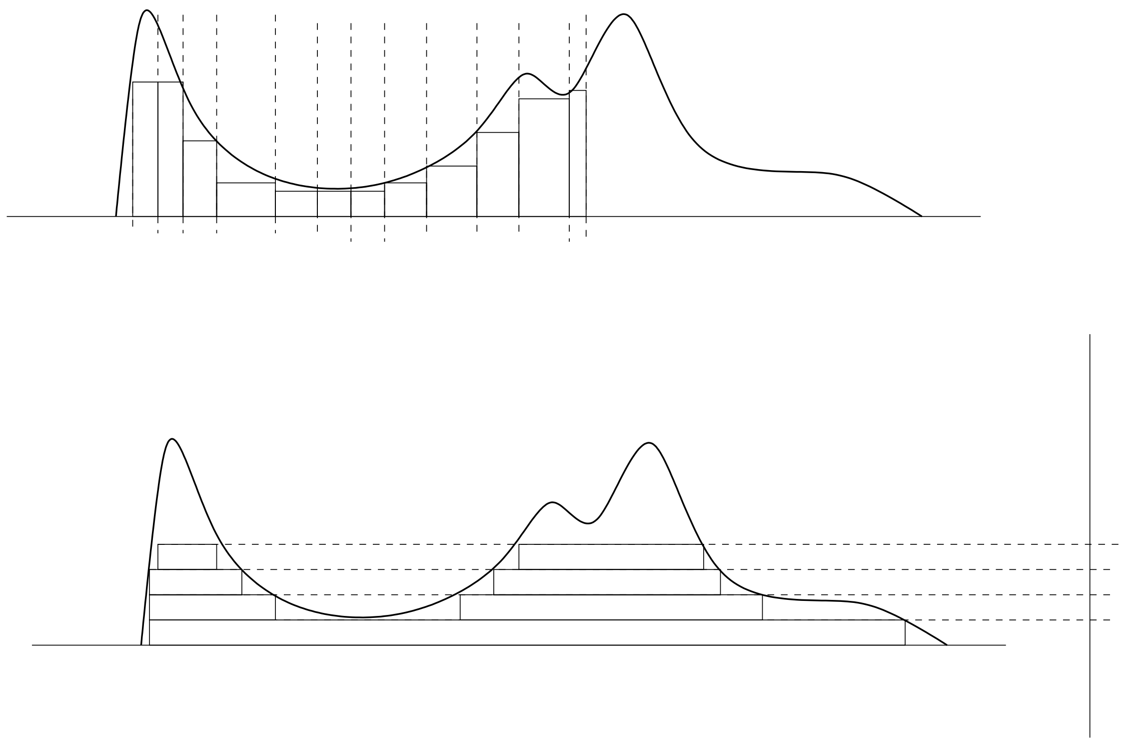

Fig. 1. Procédé de Riemann vs. procédé de Lebesgue