LES THÉORIES DE L'INTÉGRATION
DE CAUCHY À HENSTOCK
(ou la quête de l'intégrale) [1]

par Éric Charpentier
Institut de Mathématiques, Université Bordeaux I,
351 cours de la Libération,
33405 Talence Cedex, France

Table des matières

1. Version résumée et provisoire.

1 Introduction

Cet exposé est un rapide survol des principales théories rigoureuses de l'intégration, de Cauchy à aujourd'hui, ma ligne directrice étant d'essayer de rendre apparente la continuité (et parfois la discontinuité) des idées, et de dégager quelques points de vue unificateurs (Daniell, Berberian, Foran ...).

2 L'intégrale de Cauchy

2.1 L'intégrale des fonctions continues

C'est Cauchy qui fournit la première définition (constructive[2]) rigoureuse de l'intégrale (*Résumé des leçons données à l'École Royale Polytechnique sur le calcul infinitésimal*, 1823).

Soit $[a,b]$ $(b > a)$ un intervalle compact de $\mathbb{R}$. Cauchy commence par considérer le cas des fonctions f *continues* sur $[a,b]$. À toute subdivision de $[a,b]$: $\sigma = \{a_0, \ldots, a_n\}$, $a = a_0 \leq a_1 \leq a_2 \leq \cdots \leq a_{n-1} \leq a_n = b$, il associe le nombre $|\sigma| = \sup_{1\leq i\leq n+1} |a_i - a_{i-1}|$ (le "pas" de la subdivision) et la somme

$$S(f,\sigma) = \sum_{i=1}^{n} f(a_{i-1})(a_i - a_{i-1}).$$

f étant uniformément continue sur $[a,b]$[3], Cauchy prouve facilement que pour toute suite de subdivisions $(\sigma_n)_n$ dont le pas tend vers 0, la suite $(S(f,\sigma_n))_n$ est de Cauchy, donc converge, et que la limite ne dépend pas de la suite $(\sigma_n)_n$ choisie : cette limite, dit Cauchy, est *par définition* l'intégrale $\int_a^b f(t)\,dt$.

Cauchy démontre ensuite que, si f est continue sur $[a,b]$, $F(x) = \int_a^x f(t)dt$ est une fonction dérivable de x, et que $F'(x) = f(x)$: c'est le premier énoncé rigoureux du Théorème Fondamental (TF) de Newton.

2. Avant Cauchy, on utilisait la définition *descriptive* de Newton (*l'intégrale, c'est la primitive*), et la définition *constructive* (approximation par des sommes finies, dont l'idée fondamentale remonte au moins à Démocrite). Ces définitions étaient encore assez vagues (surtout la constructive), et la question de l'existence n'était pas clairement posée.

3. Cauchy le supposait implicitement, sans se rendre compte qu'il y a là une difficulté ; c'est Heine (1870) qui explicitera la notion de continuité uniforme et prouvera (1872) qu'une fonction continue sur un compact y est uniformément continue.

2.2 L'intégrale des fonctions discontinues

Cauchy savait donc intégrer les fonctions continues sur un segment $[a,b]$. Maintenant, se demande-t-il, comment intégrer une fonction discontinue?
Cauchy ne considère que des fonctions qui ont un nombre *fini* de discontinuités car, dit-il, c'est suffisant en général pour les applications ... (Comme on va le voir, c'est Dirichlet qui va essayer de poursuivre le processus inauguré par Cauchy, pour atteindre des fonctions discontinues un peu plus compliquées, et Lebesgue poussera la méthode jusqu'à ses limites – N.B. ce *n'est pas* l'intégrale de Lebesgue).

Soit donc f une fonction discontinue en un nombre *fini* de points $c_1, \dots, c_N$ de $[a,b]$, continue partout ailleurs. (Notez bien que Cauchy *ne suppose pas* que f a une limite à droite ou à gauche en ces points, ni même qu'elle reste bornée.) f est intégrable sur $[a,b]$ au sens de Cauchy ("C-intégrable") si, et seulement si, $\forall i \in \{1, \dots, N\}$,

$$\lim_{h \to 0^+} \int_{c_{i-1}}^{c_i - h} f(t)dt + \lim_{k \to 0^+} \int_{c_i + k}^{c_{i+1}} f(t)dt$$

existe (en notant $c_0 = a$ et $c_{N+1} = b$). On note alors ce nombre $\int_a^b f(t)\,dt$.

Avec cela, on peut intégrer les fonctions en escalier, et on a les intégrales impropres (sur $]a,b[$). L'hypothèse que a et b sont finis ne joue aucun rôle ci-dessus, et on a donc obtenu du même coup les intégrales impropres de la forme $\int_{-\infty}^{b} f(t)dt$, $\int_a^{+\infty} f(t)dt$, $\int_{-\infty}^{+\infty} f(t)dt$.
Il y a là, déjà, de quoi couvrir la plupart des besoins de l'époque, pour les calculs d'aires, y compris entre une courbe et son asymptote, et pour les calculs de longueurs, de figures d'équilibre, etc.
De plus, les surprises commencent :

1. Il existe des fonctions discontinues $f : [a,b] \to \mathbb{R}$, intégrables au sens de Cauchy, et qui ne sont pas des dérivées (ex. la fonction caractéristique de $\{0\}$: elle n'a pas la propriété de la valeur intermédiaire).

2. Le fait que des fonctions discontinues puissent être intégrables, alors que $F' = f$ n'a lieu *a priori* qu'aux points de continuité de f, indique qu'il y a une différence essentielle entre "intégrale" et "primitive" : l'énoncé naïf (newtonien) du "Théorème Fondamental" est inexact, les choses sont plus compliquées que prévu.

On pourrait penser que ce n'est pas bien grave, car les points exceptionnels sont en nombre fini. Mais les choses ne font que commencer. Par exemple, la définition de Cauchy ne suffit pas pour l'étude des séries de Fourier. Car, selon Fourier (1810), on devrait pouvoir représenter des fonctions même *très discontinues* par des séries trigonométriques, et calculer les coefficients de ces séries par ses formules intégrales – ce qui suppose qu'on puisse intégrer ces fonctions ...
Dirichlet (1829) avait démontré cette assertion de Fourier dans le cas de fonctions continues et monotones *par morceaux* et il avait essayé de passer, de là, aux fonctions ayant une infinité d'extréma et/ou de discontinuités. Cela l'avait amené à la fin de son article, à s'interroger sur la définition de l'intégrale (exprimant les coefficients de Fourier) pour de telles fonctions.

C'est très probablement pour cette raison que Dirichlet[4] dans ses cours essaye de pousser le processus de Cauchy un cran plus loin : supposons, dit-il, que f ait des points de discontinuité dans $[a,b]$, dont l'ensemble E n'a qu'un nombre *fini* de points d'accumulation $c_1, \dots, c_N$; dans chaque intervalle $]c_{i-1}+h, c_i-k[$ $(h,k>0$, $h+k<c_i-c_{i-1})$, f n'a qu'un nombre fini de points de discontinuité, on peut donc définir pour h,k assez petits l'intégrale $\int_{c_{i-1}+h}^{c_i-k} f(t)dt$ à la façon de Cauchy (en supposant qu'elle existe), puis $\int_{c_{i-1}}^{c_i} f(t)dt$ comme la limite $\lim_{h,k\to 0}\int_{c_{i-1}+h}^{c_i-k} f(t)dt$ si cette limite existe, et enfin $\int_a^b f(t)dt$ comme la somme des $\int_{c_{i-1}}^{c_i} f(t)dt$ (si toutes les limites nécessaires existent).

Dirichlet ne le dit pas (du moins, pour autant que je le sache), mais on voit tout de suite qu'on peut réitérer (c'est ce que fait Lebesgue [17]) : si E est *l'adhérence* de l'ensemble des points de discontinuité de f (le fait de le fermer est commode et ne change pas grand chose dans l'argument de Dirichlet), si E' désigne l'ensemble des points d'accumulation de E (E' est un fermé de E, que depuis Cantor on appelle le *dérivé* de E), et si on suppose que l'ensemble E'' des points d'accumulation de E' est fini, $E''=\{c_1,\dots,c_N\}$, alors sur chaque compact de $]c_{i-1},c_i[$ on est ramené au cas de Dirichlet, comme Dirichlet se ramenait au cas de Cauchy.
Si E'' n'est pas fini mais que $E^{(3)}$ l'est, on se ramène encore au cas précédent, etc. On peut donc définir $\int_a^b f(t)dt$ si l'un des ensembles dérivés $E^{(n)}$ est fini, à condition que toutes les limites requises existent.
Si tous les $E^{(n)}$ sont infinis, mais que leur intersection, notée $E^{(\omega)}$, est finie, $E^{(\omega)} = \{c_1,\dots,c_N\}$, dans tout compact de chaque $]c_{i-1},c_i[$ l'un des $E^{(n)}$ est fini[5] (et tous les suivants vides, donc), et on peut appliquer ce qui précède, et donc définir

4. Dans un travail non publié, que nous ne connaissons que par des notes de cours prises par Lipschitz. Je suis ici l'exposé qu'en donne Lebesgue [17].

5. Par Bolzano-Weierstrass.

encore l'intégrale $\int_{c_{i-1}}^{c_i} f(t)dt$ comme la limite $\lim_{h\to 0^+}\lim_{k\to 0^+}\int_{c_{i-1}+h}^{c_i-k} f(t)dt$ (si elle existe), puis de là, $\int_a^b f(t)dt$ comme la somme des $\int_{c_{i-1}}^{c_i} f(t)dt$ si elles existent toutes.
Si $E^{(\omega)}$ est infini mais que son dérivé $E^{(\omega+1)}$ est fini, on se ramène au cas précédent sur tout compact ne contenant pas de point de $E^{(\omega+1)}$, puis on passe à la limite, etc.
On sait donc définir $\int_a^b f(t)dt$ si l'adhérence E de l'ensemble des points de discontinuité de f dans $[a,b]$ vérifie la propriété suivante : l'un des dérivés $E^{(n)}$ est fini, ou bien leur intersection $E^{(\omega)}$ l'est, ou bien l'un des dérivés $E^{(\omega+n)}$ de $E^{(\omega)}$ l'est, ou bien leur intersection, notée $E^{(2\omega)}$ l'est, ou bien l'un des dérivés $E^{(2\omega+n)}$ de $E^{(2\omega)}$ l'est, ou bien ... [6, 7] et si toutes les limites nécessaires existent.
On voit facilement [17] que cette "intégrale transfinie" de Cauchy (et Dirichlet, et Lebesgue), ou "TC-intégrale", peut être définie de manière *descriptive* (Lebesgue) :

Proposition 2.1 *f est intégrable sur $[a,b]$ au sens de l'intégrale transfinie de Cauchy ("TC-intégrable")* ssi *existe dans $[a,b]$ une fonction continue F,* unique, à une constante additive près*, telle que*[8]

$$\int_\alpha^\beta f(t)dt = F(\beta) - F(\alpha) \qquad (1)$$

dans tout intervalle où f est continue (i.e. telle que $F' = f$ dans tout intervalle où f est continue), et alors

$$\int_a^b f(t)dt = F(a) - F(b).$$

La classe des fonctions TC-intégrables englobe, cette fois, beaucoup de fonctions discontinues (mais pas toutes les fonctions réglées ; même pas toutes les fonctions

6. Je n'insiste pas trop sur ce que signifient ces points de suspension. Pour une présentation élémentaire des ordinaux transfinis dénombrables (car c'est de cela qu'il s'agit), cf. par exemple la Note qui termine [17], ou [7], ou [2], ou [8] ...
7. On dit alors que E est *réductible*. Un fermé E est réductible *ssi* il est dénombrable (cf. [17]).
8. Si E n'est pas réductible, ou bien il contient un intervalle, et il est clair que F, si elle existe, n'y est *pas* déterminée qu'à une constante près, ou bien l'un de ses dérivés est un ensemble de Cantor, et F, si elle existe, n'est définie à une constante près que sur chaque composante connexe du complémentaire de ce dérivé (Lebesgue [17]).

monotones). Comme le fait remarquer Lebesgue [17], partant d'un fermé dénombrable E *quelconque*, et d'une énumération arbitraire $x_1, x_2, \ldots$ de ses éléments, on peut former sur $\mathbb{R} \setminus E$ la fonction (bornée) :

$$f(x) = \sum_{n \geq 1} 2^{-n} \sin \frac{1}{x - x_n}$$

(la série est uniformément convergente sur toute composante connexe du complémentaire de E) ; E est exactement l'ensemble des points de discontinuité de f ; notez que f n'est pas réglée ; mais elle est intégrable au sens de l'intégrale transfinie de Cauchy et on peut l'intégrer terme à terme !

On peut estimer que toute bonne théorie de l'intégration doit être capable de donner un sens (et le bon !) à des intégrales comme $\int_{-\infty}^{+\infty} \frac{\sin x}{x} dx$ ou $\int_{-\infty}^{+\infty} \sin(x^2) dx$. Mais un seul passage à la limite peut ne pas suffire : considérez une somme (ou même une série) d'intégrales impropres du type $\int_a^b (x-a)^{-1}(b-x)^{-1}(a+b-2x) \cos \frac{1}{(x-a)^2(b-x)^2}$. Pourquoi s'arrêter aux intégrales impropres qui s'obtiennent par un seul passage à la limite ? Ou deux, ou trois, ou n ? Ou à la réunion pour n décrivant $\mathbb{N}$ des classes de fonctions intégrables ainsi construites ? Pourquoi exclure alors les intégrales impropres construites par un nouveau passage à la limite à partir de cette classe, etc. ? On est ainsi amené naturellement à une itération transfinie. Et cela signifie que, si on y regarde bien (et si on exagère un petit peu, peut-être), toute bonne théorie de l'intégration doit contenir l'intégrale transfinie de Cauchy. Nous verrons que ce n'est pas le cas de l'intégrale de Riemann, ni de celle de Lebesgue – mais ce sera le cas de l'intégrale de Denjoy (et, de façon plus cachée, de l'intégrale de Riemann complète).

2.3 Propriétés

Il existe un procédé "régulier" permettant de savoir si une fonction donnée arbitrairement est intégrable : il suffit de voir si elle est continue, et sinon, si ses points de discontinuité forment une partie réductible (i.e. d'adhérence dénombrable) et si toutes les limites nécessaires (intégrales impropres) existent – mais ce processus "régulier" peut nécessiter une succession transfinie de vérifications, c'est-à-dire ce qu'on appelle un "esprit infini" : nous y reviendrons.

Les deux intégrales ont les propriétés basiques : normalisation ($\int_0^1 dx = 1$), positivité, linéarité, relation de Chasles, continuité de l'intégrale indéfinie, restriction (pour les intervalles), i.e. f intégrable sur I et $J \subset I$ (I, J intervalles) $\Longrightarrow f$ intégrable sur J ; le lien avec les primitives est démontré pour les dérivées continues.

On en déduit les formules d'intégration par parties (pour des fonctions C^1) et de changement de variable.
Le lien avec les sommes de Cauchy permet de faire du calcul numérique.

On a un théorème d'interversion dû à Cauchy [9]

Théorème 2.1 *(Cauchy, 1823 et 1853) Si une suite* (f_n) *de fonctions* continues *sur* $[a,b]$ *converge* uniformément *vers une fonction* f *sur* $[a,b]$*, alors* f *est continue sur* $[a,b]$*, et la* C*-intégrale de* f *est la limite des* C*-intégrales des* f_n*.*

On a un théorème de "convergence monotone" :

Théorème 2.2 *(Dini, vers 1880) Si une suite* (f_n) *de fonctions* continues *sur* $[a,b]$ *converge* en croissant *vers une fonction* f continue *sur* $[a,b]$*, alors la convergence est uniforme sur* $[a,b]$*, et la* C*-intégrale de* f *est la limite des* C*-intégrales des* f_n*.*

On a aussi un théorème de "convergence dominée" :

Théorème 2.3 *(Osgood, 1897) Si une suite* (f_n) *de fonctions* continues *sur* $[a,b]$ *converge vers une fonction* f continue *sur* $[a,b]$*, et s'il existe une constante* M *t.q.* $\forall n \in \mathbb{N}$*,* $\forall x \in [a,b]$*,* $f_n(x) \leq M$*, alors la* C*-intégrale de* f *est la limite des* C*-intégrales des* f_n*.*

Le théorème de Fourier (décomposition des fonctions périodiques en séries trigonométriques) est vrai pour les fonctions continues et monotones par morceaux (*théorème de Dirichlet*, 1829), et les coefficients de Fourier sont donnés par l'intégrale de Cauchy des fonctions discontinues.

Le passage aux intégrales multiples se fait sans problème pour les fonctions continues. C'est dans ce cadre, historiquement, que le théorème d'intégrations successives fut démontré par Cauchy, et que le théorème de changement de variables C^1 fut démontré par Catalan (1840) et par Jacobi (1841).
Le passage aux fonctions à valeurs vectorielles ne pose pas de réel problème.
Mais les fonctions continues (ou même TC-intégrables) sont trop particulières pour qu'on puisse construire sur le même moule une intégration abstraite.

9. Énoncé d'abord de façon incorrecte (1823), puis de façon correcte en 1853. La notion de convergence uniforme apparaît chez Gudermann en 1838, chez son élève Weierstrass aussitôt après, chez von Seidel, Stokes et Björling vers 1850, enfin chez Cauchy en 1853, qui en donne la première définition claire et est le premier à prouver le théorème d'interversion.

3 L'intégrale des fonctions réglées (“r-intégrale”)

3.1 Définition

Elle est attribuée aussi à Cauchy [10].
Le théorème de convergence de Cauchy, valable pour l'intégrale C, ne s'étend pas à l'intégrale TC : par exemple, (f_n) continues par morceaux et $f_n \to f$ uniformément n'impliquent pas que f soit TC-intégrable. L'intégrale des fonctions réglées est la complétion correspondante pour les fonctions continues par morceaux.

3.2 Comparaison avec l'intégrale TC

L'ensemble des points de discontinuité d'une fonction réglée est dénombrable, mais son adhérence n'a aucune raison de l'être (cf. l'exemple de Riemann ci-après). Donc il y a des fonctions qui sont r-intégrables mais pas TC-intégrables. Inversement, il y a des fonctions TC- intégrables qui ne sont pas réglées. (Exemple : la fonction définie sur $[0,1]$ par $f(x) = \cos\frac{1}{x}$, et $f(0) = 0$, disons.)

3.3 Propriétés

Les propriétés sont essentiellement les mêmes que pour l'intégrale C. On gagne beaucoup de fonctions intégrables. Il y a un procédé régulier pour savoir si une fonction donnée est intégrable ou non : il suffit de regarder si elle a en tout point une limite à gauche et une limite à droite, ou si elle est limite uniforme de fonctions en escalier.

4 L'intégrale de Newton

4.1 La définition

Dès que Cauchy (vers 1820) eut donné une définition rigoureuse des notions de limite et de dérivée (définitions qui n'étaient d'ailleurs pas si différentes de celles qu'avaient proposées d'Alembert et, avant lui, Newton), on pouvait définir l'intégrale exactement comme Newton l'avait fait, i.e. comme l'opération inverse de la dérivation :

Définition 4.1 *Une fonction f est intégrable au sens de Newton (N-intégrable) sur un intervalle $[a,b]$ si, et seulement si elle est sur cet intervalle la dérivée d'une fonction F, et alors son intégrale (au sens de Newton) est $\int_a^b f(t)dt = F(b)-F(a)$.*

10. Je ne l'ai pas vue dans ses oeuvres complètes, mais je n'ai pas cherché très attentivement.

Ce n'est pas ce qu'a fait Cauchy, mais c'est ce que font par exemple Duhamel (dans ses cours à l'École Polytechnique, vers 1850), et Serret (dans ses cours à la Sorbonne, 1863).

4.2 Propriétés

L'intégrale de Newton permet d'intégrer toutes les fonctions continues, car, comme l'a montré Cauchy, toute fonction continue est une dérivée (mais pour le prouver, Cauchy a dû utiliser une définition constructive de l'intégrale).
De plus, elle a certaines propriétés intéressantes :

1. L'intégrale de Newton a toutes les propriétés basiques déjà énumérées pour les intégrales C et TC.

2. Le TF est une définition.

3. On peut intégrer par parties en toute généralité : si fg' est une dérivée, alors $f'g = (fg)' - fg'$ en est aussi une, et $\int_a^b f'g = fg|_a^b - \int_a^b fg'$.

4. La formule de changement de variable $\int_a^b f\left(g(t)\right) g'(t)dt = \int_{g(a)}^{g(b)} f(s)ds$ est vraie dès que g est dérivable et que f est une dérivée sur $g\left([a,b]\right)$.

5. Un théorème de convergence monotone : si une suite (f_n) de fonctions N-intégrables converge *en croissant* vers une fonction f N-*intégrable* (il faut le *supposer* !), la N-intégrale de f est la limite des N-intégrales des f_n.

6. Le théorème de convergence de Cauchy (convergence uniforme) est valable pour l'intégrale de Newton (et la limite *est* N-intégrable, il n'y a pas à le supposer).

7. La longueur d'une courbe X dans $\mathbb{R}^n$ est donnée par l'intégrale de Newton de $\sqrt{\sum_{i=1}^n \left(X_i'\right)^2}$ à chaque fois que cette intégrale existe.

8. Si une fonction dérivée *bornée* admet un développement trigonométrique, les coefficients sont donnés par les formules de Fourier où les intégrales sont des intégrales de Newton.

9. Contrairement au cas des intégrales C et TC (et des intégrales de Riemann et de Lebesgue que nous verrons plus loin), f N-intégrable n'implique pas $|f|$ N-intégrable (on dit que l'intégrale de Newton n'est pas "absolue" ; ce sera le cas aussi de l'intégrale de Denjoy). C'est le premier exemple d'une telle intégrale.

Dans certains des énoncés ci-dessus figure la clause *"si l'intégrale de Newton de f existe"*. Or, c'est là un problème délicat : la définition de Newton n'est pas "constructive", en ce sens qu'elle ne répond pas à la question suivante :
Partant d'une fonction f, on peut dire si elle est dérivable, et calculer sa dérivée en tout point par un simple passage à la limite : $f'(x) = \lim_{y\to x} \frac{f(y)-f(x)}{y-x}$. Mais, *si on se donne les valeurs $f'(x)$, peut-on remonter de même à f, uniquement par des passages à la limite*? Et, d'ailleurs, *si on se donne une fonction g arbitrairement, peut-on déterminer, simplement par des passages à la limite, si c'est une dérivée*? Autrement dit, existe-t-il un procédé un tant soit peu régulier permettant de savoir si une fonction donnée arbitrairement est ou non intégrable au sens de Newton ?

L'intégrale de Newton laisse également de côté une autre question : existe-t-il un moyen d'intégrer des fonctions qui ne *sont pas* des dérivées, et si oui, quelle est la signification de leur intégrale, et est-ce qu'elle présente un intérêt quelconque [11] ? L'intégrale de Newton ne permet pas d'intégrer la fonction caractéristique de $\{0\}$, qui n'est pas une dérivée (car elle n'a pas la propriété de la valeur intermédiaire), ce qui est un peu gênant ! Les fonctions discontinues en un nombre fini de points ne sont en général N-intégrables que sur les compacts ne rencontrant pas ces points. On peut éventuellement utiliser le procédé de Cauchy (intégrale impropre) pour prolonger l'intégrale jusqu'en ces points de discontinuité, et puis la relation de Chasles (prise comme définition) pour intégrer de telles fonctions sur tout intervalle. Mais que faire pour une fonction dont les points de discontinuité ont des points d'accumulation ? Il faut procéder comme Dirichlet, puis réitérer transfiniment, comme Lebesgue ! On retrouve la même difficulté que chez Cauchy : la construction, si on la mène à son terme, est trop "théorique" (elle nécessite un "esprit infini"), et pourtant il serait artificiel de s'arrêter à tel ou tel ordinal : la nature doit aller jusqu'au bout [12] !

Remarque (intégrales multiples) : On peut définir les intégrales multiples selon les idées newtoniennes, comme anti-dérivées multiples, l'intégrale indéfinie $\int\int f(x,y)dxdy$ étant considérée comme l'une des fonctions $F(x,y)$ telles que

11. Les définitions constructives de l'intégrale proposées par Cauchy (pour les fonctions discontinues), puis Riemann, permettent en effet d'intégrer des fonctions qui ne sont pas des dérivées. Mais la première réponse vraiment satisfaisante sera apportée par Lebesgue : on peut intégrer des fonctions qui ne sont pas *partout* des dérivées, mais elles sont *presque partout* des dérivées !

12. Cette difficulté – la *fuite transfinie* – va entacher toutes les théories de l'intégration jusqu'à la découverte de l'intégrale de Riemann complète (∼1960), cf. plus loin.

$\partial^2 F/\partial x\partial y = f$. C'est ce qu'on faisait au XVIII$^{\text{ième}}$ siècle, mais ce n'est pas très commode sur le plan théorique ...

5 L'intégrale de Riemann ("R-intégrale")

5.1 Riemann

Dirichlet et Riemann admettent, avec Cauchy, que l'intégrale de Cauchy permet d'intégrer toutes les fonctions qui interviennent usuellement dans les problèmes de physique (*c'était vrai à leur époque, plus maintenant !*).
Mais, ajoutent-ils, il pourrait être instructif, pour mieux comprendre les principes mêmes du calcul différentiel et intégral, de regarder ce qu'il en est des formules de Fourier (par exemple) pour des fonctions plus irrégulières. D'ailleurs, ajoute Riemann (1854), les séries trigonométriques ne s'appliquent pas qu'à la physique : on commence, à cette époque, à entrevoir des applications dans d'autres domaines, comme la théorie des nombres, où les fonctions à développer (et donc à intégrer, pour obtenir les coefficients de Fourier) risquent fort d'être très irrégulières.
Par exemple, dit Riemann, on peut rencontrer des fonctions périodiques dont les discontinuités sont partout denses, fonctions que l'intégrale de Cauchy (même sa version "transfinie") ne peut donc pas intégrer.
L'exemple de Riemann est

$$f(x) = \sum_{n\geq 1} \frac{(nx)}{n^2},$$

où (x) désigne la différence entre x et l'entier le plus proche si $x \notin \mathbb{N} + \frac{1}{2}$, 0 sinon. C'est une fonction périodique, de période 1 : est-elle la somme de sa série de Fourier ?
f a des discontinuités de première espèce aux points $\frac{2k+1}{2m}$ (m premier à $2k+1$), et on a :

$$f\left(\frac{2k+1}{2m} \pm 0\right) = f\left(\frac{2k+1}{2m}\right) \mp \frac{\pi^2}{16m^2}.$$

La théorie de l'intégration de Cauchy (même TC) n'est même pas capable de donner un sens aux coefficients de Fourier (mais la r-intégrale – que Riemann semble n'avoir pas connue – le peut). Il est donc justifié, conclut Riemann, d'étendre la notion d'intégrale.

Riemann commence par généraliser un peu les sommes de Cauchy, en piochant un élément x_i dans chaque intervalle $I_i = [a_{i-1}, a_i]$: $S(f,P) = \sum_{i=1}^{n} f(x_i)|I_i|$

et donne la définition que l'on sait pour l'intégrale de Riemann. Il énonce explicitement ce que l'on appelle aujourd'hui le critère de Darboux, sous la forme suivante : pour $f : [a,b] \to \mathbb{R}$, notons $\omega(f,I) = \sup_I f - \inf_I f$ l'oscillation de f sur un sous-intervalle I de $[a,b]$; f est R-intégrable *ssi* $\forall\varepsilon > 0$, $\exists\delta_0 > 0$ t.q. pour toute subdivision de pas $\leq \delta_0$, $\sum_{i=1}^n \omega(f,I_i)|I_i| < \varepsilon$.
Riemann en déduit cette caractérisation des fonctions intégrables :

Théorème 5.1 *(Riemann) Une fonction $f : [a,b] \to \mathbb{R}$ bornée est Riemann-intégrable sur $[a,b]$ ssi $\forall\sigma > 0$, $\forall\varepsilon > 0$ $\exists\delta_0 > 0$ t.q. pour toute subdivision de pas $\leq \delta_0$, la somme ℓ_σ des longueurs des intervalles I_i tels que $\omega(f,I_i) > \sigma$ est $< \varepsilon$.*

Avec la définition (due essentiellement à Borel) :

Définition 5.1 *On dit qu'une partie de $\mathbb{R}$ est de mesure nulle (ou négligeable) si on peut l'inclure dans une union au plus dénombrable d'intervalles dont la somme des longueurs est aussi petite qu'on le veut,*

on voit facilement qu'une union dénombrable de parties de mesure nulle est aussi de mesure nulle et, de là, que :

Théorème 5.2 *f est Riemann-intégrable sur $[a,b]$ ssi l'ensemble de ses points de discontinuité est de mesure nulle.*

Cette dernière caractérisation rend particulièrement évidentes tout un tas de propriétés de l'intégrale de Riemann (dont le fait que les fonctions réglées de $[a,b]$ dans $\mathbb{R}$ sont R-intégrables ...)

5.2 Les intégrales de Darboux

Si une fonction $f : [a,b] \to \mathbb{R}$, même non Riemann-intégrable, est bornée, on peut minorer ses sommes de Riemann par ses "sommes de Darboux" inférieures $\sum_{i=1}^n \inf_{I_i} f|I_i|$, et les majorer par ses "sommes de Darboux" supérieures $\sum_{i=1}^n \sup_{I_i} f|I_i|$. Lorsque le pas de la subdivision tend vers 0, les "sommes de Darboux" inférieures croissent vers une limite qui s'appelle "l'intégrale de Darboux inférieure", et les "sommes de Darboux" supérieures décroissent vers une limite qui s'appelle "l'intégrale de Darboux supérieure". Comme on l'a dit, il y a un critère de Darboux (1875) : f est Riemann intégrable *ssi* ses intégrales de Darboux coïncident, et leur valeur commune est alors l'intégrale de Riemann de f.

5.3 Propriétés

L'intégrale de Riemann a toutes les propriétés basiques (normalisation, positivité, linéarité, relation de Chasles, continuité, restriction). Les sommes de Riemann permettent le calcul approché des intégrales : il suffit de prendre des subdivisions suffisamment fines pour avoir des approximations suffisamment précises de l'intégrale.

5.3.1 Théorèmes de convergence

L'intégrale de Riemann possède d'importantes propriétés d'interversion limite - intégrale. La première est analogue au théorème de Cauchy :

Théorème 5.3 *(Weierstrass, 1861) Soient* $f, f_1, ..., f_n... : [a, b] \to \mathbb{R}$ *; si pour tout* n*,* f_n *est R-intégrable et si* f *est limite uniforme de la suite* (f_n)*, alors* f *est R-intégrable et*

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

Corollaire 5.1 *On retrouve ainsi que l'intégrale de Riemann permet d'intégrer toutes les fonctions réglées de* $[a, b]$ *dans* $\mathbb{R}$.

Arzelà [1] (1885) donne un autre théorème d'interversion limite - intégrale pour l'intégrale de Riemann, analogue à celui d'Osgood[13], et qui est l'ancêtre du théorème de *convergence dominée* de Lebesgue :

Théorème 5.4 *(Arzelà) Soit* $f_1, ..., f_n \ldots : [a, b] \to \mathbb{R}$ *une suite de fonctions convergeant simplement sur* $[a, b]$ *vers une fonction* f *; si toutes les* f_n *sont Riemann-intégrables, s'il existe* $g : [a, b] \to \mathbb{R}$ *Riemann-intégrable telle que* $\forall n$, $|f_n| \leq g$ *sur* $[a, b]$ *(il revient au même de supposer qu'existe une constante* M *telle que* $|f_n| \leq M$ *sur* $[a, b]$*), et enfin* si f est Riemann-intégrable*, alors :*

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

L'intégrabilité de f doit être *supposée* : rien ne la garantit *a priori*.
Il y a aussi un “théorème de convergence monotone”, qui généralise le théorème de Dini :

Théorème 5.5 *(cf. Lebesgue [17]) Si une suite* (f_n) *de fonctions R-intégrables converge en croissant vers une fonction* f *R-intégrable, la R-intégrale de* f *est la limite des R-intégrales des* f_n.

13. Avant Osgood, en fait, qui le redémontrera indépendamment, en 1897, dans le cas particulier où f et les f_n sont *continues*.

La convergence n'a pas de raison d'être uniforme, contrairement au cas particulier des fonctions continues (cf. le théorème de Dini).
Ici aussi, comme dans le théorème d'Arzelà, l'intégrabilité de f doit être supposée.

5.3.2 L'intégrale de Riemann et le TF

Le TF de Cauchy est généralisé par Darboux (1875) à l'intégrale de Riemann :

Théorème 5.6 *Soit $f : [a,b] \to \mathbb{R}$ R-intégrable. L' "intégrale indéfinie" $F(x) = \int_a^x f(t)\,dt$ est une fonction continue. En tout point où f est* continue*, F est dérivable et $F'(x) = f(x)$. En particulier, F est dérivable* presque partout *(i.e.* sauf peut-être *sur un ensemble de* mesure nulle*).*

Remarque 5.1 *Si f est bornée mais non intégrable sur $[a,b]$, ses intégrales indéfinies de Darboux, inférieure et supérieure, sont encore continues, et admettent encore f pour dérivée en tous les points où f est continue.*

Corollaire 5.2 *Soit g continue sur $[a,b]$ et dérivable sur $[a,b]$ (dérivable à gauche en a et dérivable à droite en b).* Si *g' est R-intégrable sur $[a,b]$, alors $\int_a^b g'(t)\,dt = g(b) - g(a)$.*

Remarque 5.2

1. *Il existe des fonctions g, dérivables, dont la dérivée g' n'est pas R-intégrable, par exemple : $a \leq 0 \leq b$, $g(x) = x^2 \sin(x^{-2})$, prolongée par 0 en 0 (la dérivée n'est bornée dans aucun voisinage de 0). Volterra a donné un exemple où la dérivée est bornée mais non R-intégrable.*
2. *Si g' n'est pas R-intégrable mais est cependant bornée, $g(b) - g(a)$ est compris entre les intégrales de Darboux inférieure et supérieure de g'.*

Le corollaire ci-dessus apporte une réponse partielle au problème de savoir si une fonction donnée f est ou non une dérivée :

Corollaire 5.3 *Une fonction $f : [a,b] \to \mathbb{R}$ R-*intégrable *est une dérivée si et seulement si elle est la dérivée de son intégrale indéfinie.*

Comme les dérivées ne sont généralement pas R-intégrables, ce critère ne permet pas de décider, en toute généralité, si une fonction donnée est ou non une dérivée.

Remarque 5.3 *L'intégrale de Riemann ne contient pas l'intégrale TC. On peut construire une intégrale de Riemann transfinie analogue à l'intégrale TC. Hölder (1884) en donne une définition descriptive.*

5.4 Jordan : la définition géométrique de l'intégrale

Avec la présentation de l'intégrale par Cauchy, Dirichlet et Riemann, on s'était éloigné du point de vue géométrique au profit d'un point de vue de plus en plus "analytique". Le retour au point de vue géométrique va être motivé par trois raisons :

1. L'apparition, notamment chez Cantor (dans ses études sur les séries trigono métriques) d'ensembles de points beaucoup plus compliqués que tout ce qu'on avait rencontré jusque-là. Comment les "mesurer", i.e. comment définir et calculer leurs longueurs, aires, etc., alors que les définitions usuelles par une intégrale risquaient de ne pas s'appliquer[14] ?

2. Si on parvenait à définir une notion générale de mesure des aires, on pouvait définir l'intégrale d'une fonction f comme étant l'aire comprise sous la courbe (*"définition géométrique de l'intégrale"*). Allait-on retrouver ainsi l'intégrale de Riemann, ou bien obtenir une notion différente?

3. Il fallait aussi étendre l'intégrale en dimension ≥ 2 : quelles fonctions pouvait-on intégrer, et sur quels types de domaines? Là aussi, un point de vue plus géométrique pouvait faciliter les choses.

5.4.1 Définition géométrique de l'aire

Stolz, Harnack et du Bois-Reymond (1882) définissent la mesure d'une union finie d'intervalles de $\mathbb{R}$ de la façon évidente, puis la mesure d'une partie bornée E quelconque de $\mathbb{R}$ comme la *borne inférieure* des mesures de leurs recouvrements par des unions finies d'intervalles. C'est ce que Jordan appellera la *mesure extérieure* de E, $\mu_e(E)$. Cantor adoptera la même définition dans $\mathbb{R}^n$, en remplaçant les intervalles par des boules ; Peano (1887) prendra des prismes, Jordan (1892) des pavés de côtés parallèles aux axes.
Avec cette définition, toute partie de $\mathbb{R}^n$ est mesurable, mais il y a un défaut majeur : la mesure de deux parties disjointes peut être strictement plus petite que la somme de leurs mesures. Par exemple, pour $A = \mathbb{Q} \cap [0,1]$ et $B = [0,1] \setminus \mathbb{Q}$, on a $\mu_e(A) = \mu_e(B) = 1$ mais $\mu_e(A \cup B) = 1$.
Peano et Jordan remarquent qu'il aurait été tout aussi naturel de définir une *mesure intérieure* de E, $\mu_i(E)$, *borne supérieure* des mesures des unions finies d'intervalles (ou de pavés aux côtés parallèles aux axes, si $n \geq 2$) contenues

14. L'ambition de *mesurer* des ensembles de points si généraux fut une source de progrès, accomplis notamment par Cantor, Peano, Jordan, mais ce fut bien vite aussi un obstacle. C'est en se *restreignant* à des ensembles "constructibles" que Borel pourra enfin aller plus loin.

dans E. Là encore, toute partie de $\mathbb{R}^n$ a une mesure intérieure (nulle *ssi* E est d'intérieur vide), avec le même défaut : reprenant $A = \mathbb{Q} \cap [0,1]$ et $B = [0,1] \setminus \mathbb{Q}$, on a cette fois $\mu_i(A) = \mu_i(B) = 0$ mais $\mu_i(A \cup B) = 1$.
L'idée de Peano et de Jordan est de dire qu'une partie E de $\mathbb{R}^n$ est "quarrable" si $\mu_i(E) = \mu_e(E)$, et de définir alors sa mesure par $\mu(E) = \mu_i(E) = \mu_e(E)$. Alors, pour toutes parties *quarrables* disjointes A et B, $A \cup B$ est quarrable et $\mu(A \cup B) = \mu(A) + \mu(B)$.

N.B. Un ouvert borné peut ne pas être quarrable. Exemple : soit $(r_n)_{n \geq 1}$ une énumération des rationnels de $]0,1[$, $\Omega =]0,1[\cap \cup_{n \geq 1}]r_n - \frac{1}{2^{n+2}}, r_n + \frac{1}{2^{n+2}}[$; Ω est un ouvert de $]0,1[$, $\overline{\Omega} = [0,1]$, $\mu_e(\Omega) = 1$, $\mu_i(\Omega) \leq \sum_{n \geq 1} \frac{1}{2^{n+1}} = \frac{1}{2}$.

5.4.2 Définition géométrique de l'intégrale

Cette notion de mesurabilité (quarrabilité) est cependant assez bonne pour qu'on puisse tenter une définition géométrique de l'intégrale :

Définition 5.2 *Soit f une fonction d'un intervalle $[a,b] \subset \mathbb{R}$ dans $\mathbb{R}$, et pour tout $x \in [a,b]$ soit $E(x)$ la région du plan comprise sous la courbe, i.e. l'ensemble des $(s,t) \in \mathbb{R}^2$ tels que $a \leq s \leq x$ et que $0 \leq t \leq f(x)$ si $0 \leq f(x)$, $f(x) \leq t \leq 0$ si $f(x) \leq 0$. Notons $E^+(x) = E(x) \cap \{t > 0\}$, $E^-(x) = E(x) \cap \{t < 0\}$. f est intégrable au sens de Jordan sur $[a,b]$ si $\forall x \in [a,b]$, $E^+(x)$ et $E^-(x)$ sont quarrables, et alors on pose :*

$$\int_a^x f(s)ds = \mu\left(E^+(x)\right) - \mu\left(E^-(x)\right). \tag{2}$$

Si f n'est pas intégrable au sens de Jordan, on peut toujours poser :

$\underline{J} = \mu_i\left(E^+(x)\right) - \mu_e\left(E^-(x)\right)$, $\overline{J} = \mu_e\left(E^+(x)\right) - \mu_i\left(E^-(x)\right)$.

Il est facile de vérifier que ce sont les intégrales supérieure et inférieure de Darboux. Grâce au critère de Darboux, il en résulte donc ceci :

Proposition 5.1 *L'intégrale de Jordan est exactement l'intégrale de Riemann ; i.e. f est R-intégrable sur $[a,x]$ ssi $E(x)$ est quarrable, et alors son intégrale de Riemann est donnée par* (2).

5.4.3 Application à l'intégrale multiple

Dès la parution du mémoire de Riemann (posthume, 1867), on a essayé de construire les intégrales multiples par un procédé analogue de sommes de Riemann. Mais on remarqua très vite que $\int\int f(x,y)dxdy$ peut exister sans que les intégrales simples successives existent au sens de Riemann : la théorie de Riemann n'est pas très bien adaptée pour justifier le calcul des intégrales multiples par intégrations simples successives. (Ce n'était pas trop gênant cependant pour les applications à la physique, car le théorème d'intégrations successives pouvait être démontré pour des intégrands continus sur des domaines compacts : Cauchy l'avait déjà fait ...)

6 L'intégrale de Lebesgue

6.1 La construction de Lebesgue

Nous avons vu que la quarrabilité était une notion bien faible : même un ouvert borné peut n'être pas quarrable. Borel (1898) change complètement de point de vue : il définit la mesure d'un ouvert de $\mathbb{R}$ comme la somme des longueurs (finie ou infinie) de ses composantes connexes, et par passages au complémentaire et unions dénombrables, obtient la classe des boréliens[15] et la mesure de Borel. Cette dernière est “complètement additive” (ou “σ-additive”), i.e. la mesure d'une union dénombrable d'ensembles deux à deux disjoints est la somme de leurs mesures.
Lebesgue construit (très simplement, et *sans utiliser la construction de Borel*, comme on va le voir) une mesure qui *complète* la mesure de Borel, et puis il étend cette construction à $\mathbb{R}^n$.
Pour obtenir l'intégrale de Lebesgue, il suffit de remplacer dans la définition géométrique de l'intégrale, la quarrabilité par la mesurabilité au sens de Lebesgue. Lebesgue expose cette construction “géométrique” de son intégrale, mais ce n'est pas par celle-ci qu'il commence, et il est très intéressant de suivre pas à pas sa démarche (qui répond, au passage, à beaucoup de questions fondamentales).

Voici comment il présente son idée centrale [16] : “Pour définir l'intégrale d'une fonction continue croissante $\gamma(x)$ $(a \leq x \leq b)$, on divise l'intervalle $[a,b]$ en intervalles partiels et l'on fait la somme des quantités obtenues en multipliant la longueur de chaque intervalle partiel par l'une des valeurs de y quand x est dans cet intervalle. Si x est dans l'intervalle $[a_i, a_{i+1}]$, y varie entre certaines limites

15. La plus petite classe de parties de $\mathbb{R}$ contenant les ouverts et stable par union dénombrable et par passage au complémentaire.

μ_i,m_{i+1}, et réciproquement si y est entre μ_i et m_{i+1}, x est entre a_i et a_{i+1}. De sorte qu'au lieu de se donner la division de la variation de x, c'est-à-dire de se donner les nombres a_i, on aurait pu se donner la division de la variation de y, c'est-à-dire les nombres μ_i. De là deux manières de généraliser la notion d'intégrale. On sait que la première (se donner les a_i) conduit à la définition donnée par Riemann et aux définitions des intégrales par excès et par défaut données par M. Darboux. Voyons la seconde".

Pour bien comprendre l'origine de cette idée, suivons l'exposé qu'en donne son auteur. Lebesgue [17] commence par énumérer les conditions qui lui paraissent indispensables pour une bonne théorie de l'intégration. Il suppose que l'intégrale à définir doit vérifier :
l'invariance par translation (T)[16] ; la relation de Chasles (C) ; l'additivité (A) ; la positivité (P) ; la normalisation (N) ; et enfin, la propriété de monotonie (M) : Si f_n tend en croissant vers f, et si toutes les f_n sont intégrables sur $[a,b]$, alors f est intégrable sur $[a,b]$ et $\int_a^b f_n(t)dt \longrightarrow \int_a^b f(t)dt$ quand $n \longrightarrow 0$.
Lebesgue remarque que la condition (M) n'a pas le même caractère de simplicité et de nécessité que les cinq premières, mais nous allons voir quel rôle elle joue exactement et d'où elle vient[17].
Lebesgue établit d'abord un résultat très simple et très général :

Proposition 6.1 *Une intégrale qui vérifie les cinq premières conditions ci-dessus, (T) à (N), vérifie aussi celles-ci (pour les fonctions bornées) :*

1. *elle est linéaire (L)*
2. *elle vérifie* $\forall c,d \in [a,b]$, $\int_c^d dt = d-c$
3. *elle vérifie l'inégalité de la moyenne ;*
4. $\int_a^b f$ *est comprise entre l'intégrale de Darboux inférieure et l'intégrale de Darboux supérieure ; donc, si* f *est R-intégrable, son intégrale ci-dessus coïncide avec son intégrale de Riemann (autrement dit, l'intégrale de Riemann est la seule intégrale raisonnable, quand elle existe) ;*
5. *si* f *et* $|f|$ *sont intégrables,* $|\int_a^b f| \leq \int_a^b |f|$ *;*

16. i.e. $\forall c \in \mathbb{R}$ et pour toute fonction intégrable f sur un intervalle $[a,b]$, la fonction $x \mapsto f(x-c)$ est intégrable sur $[a+c,b+c]$, et $\int_a^b f(t)dt = \int_{a+c}^{b+c} f(s-c)ds$. (C'est un cas très particulier de la formule de changement de variable).

17. La propriété (M) est vraie pour les intégrales C et N. Lebesgue en inféra que la bonne notion de mesure devait être complètement additive, comme celle de Borel, et contrairement à celle de Jordan. Cf. plus loin.

6. *Si la somme d'une série uniformément convergente de fonctions intégrables est intégrable, on peut intégrer terme à terme ;*

7. *l'intégrale (du moins sur les fonctions bornées) est entièrement définie par les intégrales des fonctions caractéristiques des ensembles de la forme* $\{c \leq f < d\}$ *(*$a \leq c < d \leq b$*,* f *intégrable).*

On est donc ramené à intégrer les fonctions caractéristiques des ensembles de la forme $E = \{c \leq f < d\}$ ou $E = \{c < f \leq d\}$ ($a \leq c < d \leq b$, f intégrable). Posons $\int_a^b 1|_E = \ell(E)$ et appelons cela la “mesure de E”. Il résulte de ce qui précède que la “mesure” d'un intervalle [18] est sa longueur, de (T) que la “mesure” est invariante par translation, de (P) que la mesure est positive, et de (A) que la “mesure” d'une réunion *finie* d'ensembles disjoints deux à deux est la somme de leurs mesures. C'est “l'additivité finie”.
Maintenant, Lebesgue estime que la “bonne” notion de mesure doit posséder la propriété (de Borel) “d'additivité complète”, i.e. la “mesure” d'une réunion *dénombrable* d'ensembles disjoints deux à deux est la somme de leurs mesures [19]. C'est ce que traduit la propriété (M) quand on l'applique aux fonctions caractéristiques d'ensembles disjoints deux à deux, et il est facile de voir que si on l'a pour ces fonctions particulières on doit l'avoir pour toutes les fonctions intégrables : c'est de là que vient la propriété (M).

Faisons le point : Lebesgue ramène le problème de l'intégration à celui de la mesure des ensembles de la forme $E = \{c \leq f < d\}$ ou $E = \{c < f \leq d\}$ ($a \leq c < d \leq b$) avec f intégrable. Mais jusqu'ici il n'a dit ni quelles sont les fonctions intégrables f, ni quels sont les ensembles mesurables – sauf que les ensembles $\{c \leq f < d\}$, $\{c < f \leq d\}$ doivent l'être pour f intégrable.
Lebesgue va d'abord définir les ensembles mesurables, puis il dira qu'une fonction f est mesurable (intégrable) si les ensembles $\{c \leq f < d\}$ ($a \leq c < d \leq b$) (ou $\{c < f \leq d\}$: en fait ce sera équivalent) sont mesurables.

Comment donc définir la mesurabilité et la mesure ? Vu ce qui précède, il est immédiat que si on note μ_i et μ_e les mesures intérieure et extérieure de Jordan, on aura $\mu_i(E) \leq \ell(E) \leq \mu_e(E)$, donc *pour un ensemble quarrable la mesure de Lebesgue coïncidera avec la mesure de Jordan.*

18. Les intervalles font évidemment partie des ensembles E qui nous intéressent, puisqu'ils sont de la forme $\{c \leq f < d\}$ avec f affine par morceaux, et que toute bonne théorie de l'intégration doit être capable d'intégrer les fonctions affines par morceaux.

19. C'est très naturel : un ouvert étant toujours une réunion dénombrable d'intervalles ouverts disjoints deux à deux, il est très tentant de définir sa mesure comme la somme des longueurs de ces intervalles !

Pour construire sa mesure, Lebesgue s'inspire des idées de Peano et Jordan sur les mesures extérieures, mais en gardant l'idée d'additivité complète de Borel : il définit la *mesure extérieure* $\ell_e(E)$ d'une partie $E \subset \mathbb{R}$ comme la borne inférieure des mesures (de Borel) des ouverts qui contiennent E (i.e. il fait la somme des longueurs d'une infinité dénombrable d'intervalles ouverts disjoints recouvrant E, alors que Peano et Jordan ne considéraient qu'un nombre fini d'intervalles, et il prend la borne inférieure des sommes obtenues). Évidemment, $\ell_e(E) \leq \mu_e(E)$. Si on peut donner un sens précis à la mesure $\ell(E)$, vu (P) on devra avoir $\ell(E) \leq \ell_e(E)$.

Supposons que E soit contenu dans un segment I. Quand on aura réussi à donner un sens précis à la notion de mesure, on pourra écrire, pour E mesurable : $\ell(E) + \ell(I \setminus E) = \ell(I)$, donc $\ell(E) = \ell(I) - \ell(I \setminus E) \geq \ell(I) - \ell_e(I \setminus E)$. Cela conduit Lebesgue à définir la *mesure intérieure* de E par : $\ell_i(E) = \ell(I) - \ell_e(I \setminus E)$, où ℓ_e est la mesure extérieure (et $\ell(I)$ la longueur de I). $\ell_i(E)$ ne dépend pas de I. Évidemment elle est $\geq \mu_i(E)$[20].

Supposons E borné et I compact. On a toujours $\ell_i(E) \leq \ell_e(E)$. En effet, si on recouvre E par des intervalles I_i et $I \setminus E$ par des intervalles J_j, les I_i, J_j recouvrent I donc, extrayant un sous-recouvrement fini de I,

$$\sum_i \ell(I_i) + \sum_j \ell(J_j) \geq \sum_{\text{finie}} \ell(I_i) + \sum_{\text{finie}} \ell(J_j) \geq \ell(I)$$

ce qui montre que $\ell_e(E) + \ell_e(I \setminus E) \geq \ell(I)$, i.e. $\ell_i(E) \leq \ell_e(E)$ vu la définition de $\ell_i(E)$.

Lebesgue dit alors que E est *mesurable* si $\ell_i(E) = \ell_e(E) \equiv \ell(E)$: c'est la mesure de Lebesgue.

La mesure ainsi définie associe à un intervalle sa longueur, et elle est invariante par translation, positive et additive. Reste à vérifier qu'elle est *complètement additive*, ce que Lebesgue fait en une page ([17], p.115 de la seconde édition).

Il est clair que le complémentaire d'un ensemble mesurable l'est aussi, et que l'intersection de deux ensembles mesurables l'est encore, et que donc une union dénombrable d'ensembles mesurables (même non disjoints deux à deux) est mesurable. Bref, la classe des ensembles mesurables est une tribu, comme on dit aujourd'hui.

Elle contient la tribu de Borel, et pour tout borélien, la mesure de Lebesgue coïncide avec la mesure de Borel.

20. Avec les mesures de Jordan on avait aussi $\mu_i(E) = \ell(I) - \mu_e(I \setminus E)$: les deux caractérisations (celle-ci, et comme sup des longueurs des unions finies d'intervalles inclus dans E) coïncident. Ce n'est plus vrai pour l_i. Exemple : $E = [0,1] \setminus \mathbb{Q}$: la première définition (celle de Lebesgue) donne 1, la seconde donnerait 0 ...

De plus, la mesure de Lebesgue *complète* la mesure de Borel, en ce sens qu'un ensemble inclus dans un borélien de mesure nulle est lebesguien (et sa mesure de Lebesgue est nulle).

Lebesgue définit ensuite les fonctions mesurables :

Une fonction f (bornée ou non) est mesurable si les ensembles $\{c \leq f < d\}$ $(a \leq c < d \leq b)$ sont mesurables. La classe des fonctions mesurables est stable par combinaison linéaire, produit, passage à la limite. Une fonction caractéristique d'un ensemble mesurable est mesurable.

Les fonctions constantes et id sont mesurables, donc les polynômes aussi, donc les fonctions continues aussi. Une fonction continue p.p. est mesurable aussi, et donc toute fonction R-intégrable est mesurable.

Ce qui précède permet de définir l'intégrale (de Lebesgue) d'une fonction mesurable bornée. Évidemment, si E est une partie mesurable de $]a,b[$, et f une fonction mesurable bornée sur $]a,b[$, on pose $\int_E f(t)dt = \int_a^b 1_E(t)f(t)dt$. Si (E_n) est une suite de parties mesurables de $]a,b[$ disjointes deux à deux, on a $\int_E f(t)dt = \sum_n \int_{E_n} f(t)dt$.

En particulier, il suffit de parler de l'intégrale $\int_{\mathbb{R}} f(t)dt$ car l'intégrale sur toute partie mesurable de $\mathbb{R}$ s'y ramène.

On peut étendre l'intégrale à des fonctions non bornées : une fonction mesurable positive $f \ : \mathbb{R} \to \mathbb{R}_+$ est Lebesgue intégrable (L-intégrable) *ssi* pour toute subdivision $(x_k)_{k\in\mathbb{Z}}$ de $\mathbb{R}$, $\sum_{k\in\mathbb{Z}} x_k \ell(x_k \leq f < x_{k+1})$ tend vers une limite finie quand $\sup_{k\in\mathbb{Z}}(x_{k+1} - x_k) \longrightarrow 0$ (l'intégrale étant alors cette limite). Pour une fonction non positive, f est intégrable *ssi* $f^+ = \max\{f,0\}$ et $f^- = -\min\{f,0\}$ le sont, et alors $\int f = \int f^+ - \int f^-$. Par exemple, si (x_n) est une suite de points distincts de $[a,b]$, si $f : [a,b] \to \mathbb{R}$ est définie par $f(x_n) = n$ et $f(x) = 0$ si $x \in [a,b] \backslash \{x_n; n \in \mathbb{N}\}$, alors f est L-intégrable et non bornée (par conséquent, f n'est pas R-intégrable).

Lebesgue appelle *fonction étagée* une fonction mesurable qui ne prend qu'un nombre fini de valeurs. L'intégrale d'une fonction étagée est définie de manière évidente : $\sum_i y_i \ell\left(f^{-1}\left(\{y_i\}\right)\right)$, où les y_i sont les valeurs que prend f.

On voit sur ce qui précède que l'intégrale de Lebesgue est entièrement définie à partir de l'intégrale des fonctions étagées, qui jouent dans sa théorie un rôle analogue à celui que jouaient les fonctions en escalier dans les intégrales C, r et R.

6.2 Propriétés

L'intégrale de Lebesgue a toutes les propriétés basiques. La propriété de restriction est vraie non seulement pour les intervalles, mais pour toutes les parties mesurables : si f est intégrable sur E mesurable, elle l'est sur toute $F \subset E$ mesurable.
Étant donné une fonction *bornée*, savoir si elle est intégrable sur un segment équivaut à savoir si elle est mesurable, et la question générale est indécidable dans la théorie des ensembles de Zermelo-Fraenkel.

6.2.1 Interversion intégrale-limite :

Lebesgue démontre d'abord, un analogue du théorème d'Arzelà :

Théorème 6.1 *(Lebesgue, 1902) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions L-intégrables sur $[a,b]$ convergeant simplement vers une fonction f, et telle qu'existe une constante M telle que $\forall n$, $|f-f_n| \leq M$. Alors f est L-intégrable sur $[a,b]$ et $\int f = \lim_n \int f_n$.*

La preuve est immédiate :
$\forall \varepsilon > 0$, $|\int(f-f_n)| \leq \varepsilon\ell(\{|f-f_n|<\varepsilon\}) + M\ell([a,b]\setminus\{|f-f_n|<\varepsilon\})$,
et $\ell([a,b]\setminus\{|f-f_n|<\varepsilon\}) \longrightarrow 0$ quand $n \longrightarrow 0$.
Lebesgue en déduit le théorème de convergence monotone (propriété (M)), sous l'hypothèse restrictive que les f_n sont uniformément bornées. C'est Beppo Levi qui parvient à s'affranchir de cette hypothèse :

Théorème 6.2 *(B. Levi, 1906) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions L-intégrables (sur une partie mesurable E de $\mathbb{R}$) t.q. $sup_{n\to\infty}\int f_n < +\infty$. Alors la suite (f_n) converge p.p. vers une fonction f, celle-ci est L-intégrable, et $\int f = \lim_n \int f_n$.*

Puis Lebesgue généralise son théorème de type Arzelà, et c'est le fameux théorème de convergence dominée :

Théorème 6.3 *(Lebesgue, 1909) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions L-intégrables sur une partie mesurable E de $\mathbb{R}$ convergeant simplement vers une fonction f. Supposons qu'existe une fonction g L-intégrable sur E telle que $\forall n$, $|f_n| \leq g$. Alors f est L-intégrable sur E et $\int f = \lim_n \int f_n$.*

Preuve. Montrons-le seulement pour E de mesure finie. Pour $\varepsilon > 0$, on considère $E_n = \cap_{p,q\geq 0}\{|f_{n+p}-f_{n+q}|<\varepsilon\}$. On a : $|\int_E f - \int_E f_n| \leq \int_{E_n}|f-f_n| + \int_{E\setminus E_n}(|f|+|f_n|) \leq \varepsilon\ell(E_n) + 2\int_{E\setminus E_n} g$, et, par additivité complète de la mesure (appliquée aux $E_{n+1}\setminus E_n$), on a $\ell(E_n)\uparrow\ell(E)$ et $\ell(E\setminus E_n)\downarrow 0$ quand $n \longrightarrow 0$.

6.2.2 L'intégrale de Lebesgue et le TF

Si f est L-intégrable sur $[a,b]$, la fonction $F(x) = \int_a^x f$ a p.p. une dérivée, égale à f.
La fonction $f(x) = \left(x^2 \sin \frac{1}{x^2}\right)' = 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}$ (complétée par 0 en 0) n'est pas L-intégrable : l'intégrale de Lebesgue ne contient donc pas l'intégrale de Newton. Cependant, puisque les fonctions dérivées sont des limites de fonctions continues, elles sont mesurables, et donc toute fonction dérivée *bornée* est L-intégrable, et par convergence dominée on a :

Théorème 6.4 *(Lebesgue, 1902) Si f' est une dérivée bornée sur $[a,b]$, elle y est L-intégrable, et $\forall x \in [a,b]$, $\int_a^x f'(t)dt = f(x) - f(a)$.*

En fait, le résultat reste vrai sous la seule hypothèse que f' soit L-intégrable, même si elle n'est pas bornée.

6.2.3 Définition descriptive de l'intégrale de Lebesgue.

Théorème 6.5 *Une fonction f (à valeurs réelles) est L-intégrable ssi existe une fonction F absolument continue*[21] *telle que $F' = f$ p.p.*

L'intégrale de Lebesgue est donc une variante de l'intégrale de Newton, à la fois plus restrictive (à cause de la clause d'absolue continuité[22]) et plus vaste (grâce au p.p.).

6.2.4 Définition géométrique de l'intégrale

La mesure de Lebesgue peut se définir dans le plan exactement comme sur la droite, en remplaçant seulement les intervalles par les carrés (avec la normalisation $\ell_2([0,1] \times [0,1]) = 1$).
Par analogie avec l'intégrale de Jordan, il est naturel de dire que f est intégrable au sens géométrique sur $[a,b]$ si $\forall x \in [a,b]$, $E^+(x)$ et $E^-(x)$ (définis plus haut) sont mesurables (pour la mesure de Lebesgue dans $\mathbb{R}^2$), et de poser alors :

$$\int_a^x f(s)ds = \ell_2\left(E^+(x)\right) - \ell_2\left(E^-(x)\right). \tag{3}$$

On voit facilement que l'intégrale géométrique de Lebesgue coïncide (pour les fonctions bornées) avec l'intégrale de Lebesgue définie plus haut.

21. i.e. t.q. $\forall \varepsilon > 0$, existe $\eta > 0$ t.q. pour toute subdivision $a \leq a_1 < b_1 \ldots < a_N < b_N \leq b$, $\sum_{i=1}^N (b_i - a_i) < \eta$ implique $\sum_{i=1}^N |f(b_i) - f(a_i)| < \varepsilon$.
22. Par exemple, $F(x) = x^2 \sin \frac{1}{x^2}$ (complétée par 0 en 0) est dérivable mais non absolument continue. Elle est donc N-intégrable mais pas L-intégrable.

6.2.5 Intégrales multiples

Avec ℓ_2 on peut définir une intégrale de Lebesgue dans $\mathbb{R}^2$ comme on l'a fait dans $\mathbb{R}$ avec $\ell = \ell_1$. Lebesgue, dans sa thèse, montre que si $(x, y) \mapsto f(x, y)$ est une fonction borélienne bornée alors c'est aussi le cas de $y \mapsto f(x, y)$ et de $x \mapsto \int f(x, y)dy$, et on a $\int\int f(x, y)dxdy = \int dx \int dy f(x, y)$ sur tout rectangle. Fubini (1907) montre que, plus généralement, si f est L-intégrable, l'ensemble $\{x|\ y \mapsto f(x, y)$ non intégrable $\}$ est de mesure nulle, et il en déduit que $\int\int_{\mathbb{R}^2} f(x, y)dxdy = \int_{\mathbb{R}} dx \int_{\mathbb{R}^2} dy f(x, y)$. (Si on veut écrire le résultat analogue avec $\int\int_E$ pour une partie E ℓ_2-mesurable quelconque de $\mathbb{R}^2$, il suffit de remplacer f par $f\mathbf{1}|_E$).

La définition descriptive de l'intégrale de Lebesgue s'étend au cas de plusieurs variables.

6.3 Young, Daniell

Daniell (1917), généralisant une méthode de Young (1903), donne une construction générale des intégrales de type Lebesgue : pour construire une intégrale sur un espace abstrait, il part d'un ensemble de "fonctions élémentaires" qu'on sait intégrer, et il définit une intégrale inférieure (resp. supérieure) d'une fonction f comme la borne sup. (resp. inf) des intégrales des fonctions élémentaires $g \leq f$ (resp. $g \geq f$), puis il dit que f est intégrable *ssi* ses intégrales inférieure et supérieure coïncident.

Par exemple, dans $\mathbb{R}$, si on prend comme fonctions élementaires les fonctions étagées, on retrouve exactement la construction de Lebesgue ; si on prend les fonctions s.c.i. et s.c.s (avec pour intégrales les intégrales de Darboux sup. et inf.), on retrouve la construction de Young (qui donne la même intégrale) ; on peut aussi partir des fonctions continues à support compact avec pour intégrale l'intégrale de Cauchy, comme le faisait Daniell lui-même ; si on prend comme fonctions élementaires les fonctions en escalier, on obtient la construction que présente Deville dans son cours ...

7 Bilan de la "révolution lebesguienne"

En forçant légèrement le trait : la propriété de convergence monotone équivaut à l'additivité complète de la mesure. L'autre grand théorème de convergence – le théorème de convergence dominée – sort aussi de là. L'interprétation de l'additivité complète en termes de théorèmes de convergence pour l'intégrale est

l'une des deux innovations fondamentales de Lebesgue, l'autre étant la notion de "presque partout" (*p.p.*).
L'apparition de ces deux concepts a notablement changé (en le simplifiant) le paysage de la théorie de l'intégration.

7.1 Développements

Si on remplace l'espace d'arrivée $\mathbb{R}$ par un espace de Banach, on obtient l'intégrale de Bochner (1933).
La notion générale de mesure (et d'intégrale "de Lebesgue" par rapport à cette mesure) dans $\mathbb{R}^n$ (fonction d'ensembles complètement additive quelconque, définie sur les parties L-mesurables de $\mathbb{R}^n$) est due à Radon (1913). Les notions de mesure et d'intégrale se généralisent facilement aux espaces abstraits munis d'une tribu (Fréchet, 1915).
Une bonne partie de l'analyse fonctionnelle est sortie tout droit de l'étude des espaces de Lebesgue L^p (espaces des classes modulo p.p. des fonctions f t.q. $|f|^p$ soit intégrable), introduits par Fréchet (1907, pour L^2) et Riesz (1910).
Enfin, l'intégrale de Lebesgue est particulièrement adaptée au calcul des probabilités : axiomatique de Kolmogorov (1933).

7.2 Le TF depuis l'intégrale de Lebesgue

Pour l'intégrale de Newton, le TF donnait une condition nécessaire et suffisante d'intégrabilité : f est intégrable *ssi* c'est une dérivée. Cela restait vrai pour l'intégrale de Cauchy des fonctions continues : une fonction f est C-intégrable ssi elle est continue, i.e. ssi elle est la dérivée d'une fonction C^1. Mais ce *ssi* est obtenu au prix d'une restriction un peu artificielle : pourquoi ne définir l'intégrale que pour les fonctions continues ? Cauchy s'était limité à ce cas parce qu'il pouvait alors démontrer la convergence des sommes de Cauchy et un bon énoncé du TF.
Pour toutes les intégrales qui ont suivi, le TF avait perdu sa forme d'une CNS : il ne l'a retrouvée qu'avec l'intégrale de Lebesgue. Depuis la théorie de Lebesgue, presque toutes (!) les théories de l'intégration ont conservé la forme que Lebesgue a donnée au TF, i.e. la forme suivante (dite de Lebesgue-Berberian) :
Une fonction f est "intégrable" *ssi* existe une "bonne" fonction F (et une seule, à une constante près), telle que F' existe et coïncide avec f partout en dehors d'un ensemble N "négligeable" en un certain sens. On a alors $\int_a^b f(t)dt = F(b) - F(a)$.
C'est le sens précis des mots "bonne" et "négligeable" qui distinguera les diverses théories de l'intégration pour lesquelles il existe un bon TF. Pour avoir l'unicité de F (à une constante près), et les propriétés indispensables, il faut que ces notions satisfassent certaines exigences. Berberian demande : que l'ensemble des

"bonnes" fonctions soit un sous-espace B de $C^0([a,b])$ contenant la fonction id ; que l'ensemble des parties "négligeables" de $\mathbb{R}$ soit stable par réunions finies, et que ses éléments soient des parties de $\mathbb{R}$ d'intérieur vide ; enfin, que l'image par une "bonne" fonction d'une partie "négligeable" soit "négligeable" (propriété "de Lusin").

Berberian [3] fait remarquer qu'il *n'existe pas* de sous-espace B de $C^0([a,b])$ et d'ensemble $\mathcal{N}$ de parties de $\mathbb{R}$ tels qu'on ait un énoncé du TF comme ci-dessus pour l'intégrale des fonctions réglées, ni pour l'intégrale de Riemann[23]. Le TF pour les intégrales de Denjoy dont nous allons parler maintenant sera bien, par contre, de la forme de Lebesgue-Berberian.

8 Les intégrales de Denjoy

8.1 La construction de Denjoy

8.1.1 Intégrer les dérivées

Le but de Denjoy (1912) était de prolonger l'intégrale de Lebesgue en une intégrale qui contienne celle de Newton (i.e. qui résolve le problème de la recherche des primitives). On a vu que l'intégrale de Lebesgue permettait de remonter aux primitives des fonctions dérivées *bornées*, mais que par exemple la fonction $f(x) = 2x\sin(x^{-2}) - 2x^{-1}\cos(x^{-2})$, prolongée par 0 en 0, qui n'est pas bornée, mais qui est la dérivée de la fonction $F(x) = x^2\sin(x^{-2})$ (prolongée par 0 en 0), n'est pas L-intégrable (car F n'est pas absolument continue au voisinage de 0).

Voici comment procède Denjoy. Définissons une "intégrale" comme une forme linéaire sur un espace de fonctions numériques (lesquelles sont alors dites "intégrables"), possédant les propriétés suivantes :

1. Elle est invariante quand on change les valeurs d'une fonction en un nombre fini de points : propriété (F). En particulier, l'intégrale d'une fonction f sur un intervalle borné I (i.e. par définition l'intégrale de $f\mathbf{1}_I$) est la même, que I contienne ou non ses extrémités.

2. Elle possède la propriété de restriction (R) pour les intervalles.

23. On voit facilement que $\mathcal{N}$ = l'ensemble des parties finies de $[0,1]$ pour toute intégrale qui ne sait intégrer que des fonctions bornées ; on ne peut donc avoir un bon TF avec une intégrale capable d'intégrer des fonctions ayant une infinité de discontinuités, que si celle-ci est assez puissante pour intégrer des fonctions non bornées.

3. Elle possède la propriété de Chasles (C).

Par (R), on peut définir la notion d'intégrabilité en un point (f est intégrable au point x_0 *ssi* elle est intégrable sur un intervalle ouvert contenant ce point). L'ensemble E_1 des points où f n'est pas intégrable est un fermé.
Par (R) et (C), si I désigne une composante connexe du complémentaire de E_1, f est intégrable sur tout sous-intervalle compact K de I. En effet, on peut recouvrir un tel intervalle par des intervalles ouverts sur l'adhérence desquels f est intégrable, en extraire un sous-recouvrement fini ; les extrémités de K et celles des intervalles de ce sous-recouvrement qui sont dans K définissent une partition de K ; par (R), f est intégrable sur chaque intervalle de cette subdivision, et par (C) elle l'est sur K.
Maintenant, soit f une fonction dérivable sur un intervalle ouvert I de $\mathbb{R}$, et supposons qu'on veuille intégrer f' sur un segment $[a, b] \subset I$. Si f' est L-intégrable sur $[a, b]$, on est satisfait. Sinon commence une récurrence. Soit E_1 l'ensemble (fermé) des points de $[a, b]$ en lesquels f' n'est pas Lebesgue-intégrable. Il faut noter qu'en vertu du théorème de Baire, f' est continue en tout point d'une partie dense de $[a, b]$, donc bornée (et donc Lebesgue-intégrable) au voisinage de tout point de cette partie dense, et donc E_1 est d'intérieur vide (i.e. nulle part dense : c'est la même chose, pour un fermé). Comme on vient de le voir, f' est L-intégrable sur tout segment $[c, d]$ du complémentaire de E_1, et on a $\int_c^d f'(t)dt = f(d) - f(c)$. (C'est "l'initialisation" : on commence par intégrer f au sens de Lebesgue sur tout segment où c'est possible).
Comme f est continue, si on note $]\alpha, \beta[$ une composante connexe de $]a, b[\setminus E_1$ et si on prend une suite de segments $[c_n, d_n] \subset]\alpha, \beta[$ t.q. $c_n \downarrow \alpha$ et $d_n \uparrow \beta$, on a $\int_{c_n}^{d_n} f'(t)dt = f(d_n) - f(c_n) \to f(\beta) - f(\alpha)$ quand $n \to \infty$. Procédant comme Cauchy, on peut donc *définir* $\int_\alpha^\beta f'(t)dt = \lim_n \int_{c_n}^{d_n} f'(t)dt$. (C'est ce que Denjoy appelle la "deuxième opération".)
De cette manière, f' est intégrable (au moins au sens des intégrales généralisées) sur toute composante connexe de $[a, b] \setminus E_1$. Notez que si E_1 a un point isolé x_0 (disons $E_1 \cap [x_1, x_2] = \{x_0\}$, $x_1 < x_0 < x_2$), la seconde opération permet de définir les intégrales $\int_{x_1}^{x_0} f'(t)dt$ et $\int_{x_0}^{x_2} f'(t)dt$, et donc $\int_{x_1}^{x_2} f'(t)dt$ par la *relation de Chasles*, qui est la "première opération".
On voit qu'on a déjà éliminé tout point isolé de E_1, et qu'on est donc ramené à considérer son dérivé E_1'. En particulier, on a d'ores et déjà intégré la dérivée de la fonction $f(x) = x^2 \sin \frac{1}{x^2}$ (complétée par 0 en 0), pour laquelle $E_1 = \{0\}$ – comme l'aurait fait Cauchy.
Si on n'utilisait que les deux "opérations" ci-dessus, on obtiendrait, en réitérant transfiniment, une intégrale qui serait l'exacte analogue de l'intégrale transfinie

de Cauchy, mais plus puissante car on serait parti de l'intégrale de Lebesgue au lieu de l'intégrale de Cauchy des fonctions continues.

Supposons qu'on ait prolongé l'intégrale de f' aux composantes connexes du complémentaire de E_1 (par la "deuxième opération"), puis éliminé les points isolés de E_1 (par la "première opération"), puis prolongé l'intégrale aux composantes connexes du complémentaire de E_1' (par la "deuxième opération"), puis éliminé les points isolés de E_1' (par la "première opération"), etc., et qu'on arrive en itérant ainsi à définir l'intégrale de f' sur toute composante connexe du complémentaire de tout $E_1^{(n)}$, $n \in \mathbb{N}$. Soit alors σ un segment inclus dans le complémentaire de $E_1^{(\omega)} = \cap_n E_1^{(n)}$. Il existe N tel que σ soit inclus dans le complémentaire de $E_1^{(N)}$, donc l'intégrale est déjà définie sur σ. Maintenant, la deuxième opération permet d'étendre l'intégrale à toute composante connexe du complémentaire de $E_1^{(\omega)}$, puis par la première opération on évacue les points isolés de $E_1^{(\omega)}$, puis par la deuxième opération on prolonge l'intégrale à toute composante connexe du complémentaire de $E_1^{(\omega+1)}$, etc. On peut donc réitérer ainsi transfiniment, s'il le faut.
D'après un résultat connu sous le nom de "théorème de Cantor-Baire", on aboutit alors à un ordinal α *fini ou dénombrable* tel que $E_1^{(\alpha)}$ soit *parfait*, i.e. tel que $E_1^{(\alpha+1)} = E_1^{(\alpha)}$. $E_1^{(\alpha)}$ s'appelle le noyau parfait de E_1. Notons-le K_1. S'il n'est pas vide, il a la puissance du continu, cf. [17], i.e. K_1 est vide *ssi* E_1 est dénombrable (et dans ce cas on peut intégrer f' uniquement avec les deux premières opérations de Denjoy). Supposons K_1 non vide.
Après avoir appliqué la "deuxième opération", f' est intégrable sur toute composante connexe du complémentaire de K_1 (et dès lors la "deuxième opération" ne s'applique plus), et comme K_1 n'a pas de point isolé, la "première opération" ne s'applique plus non plus. Ce cas peut effectivement se produire : Denjoy construit un exemple explicite, cf. [9]. Les deux premières opérations ne permettent donc pas d'intégrer toutes les dérivées.
Dans ce cas, que faire ?
Notons I_n les composantes connexes du complémentaire de K_1.
Disons que la série $\sum_n \int_{I_n} f'(t)dt$ converge absolument *au voisinage* d'un point $x \in K_1$ s'il existe un intervalle ouvert J de $[a, b]$, avec $x \in J$, tel que $\sum_n \int_{J\cap I_n} f'(t)dt$ converge absolument.
Une remarque cruciale de Lebesgue[24], est que l'ensemble (évidemment ouvert) des points x de K_1 tels que

- la série $\sum_n \int_{I_n} f'(t)dt$ converge absolument au voisinage de x ;

24. Lebesgue écrit dans [17] (seconde édition) que cet énoncé "marque le point extrême" qu'il avait atteint dans la recherche des fonctions primitives.

– $f'\mathbf{1}_{K_1}$ est bornée (donc Lebesgue intégrable) au voisinage de x,

est *non vide* – et même est dense dans K_1. (Donc son complémentaire, noté E_2, est un fermé d'intérieur vide de K_1). Appelons cela le *lemme de Lebesgue-Denjoy*. On pose alors, pour J comme ci-dessus :

$$\int_J f'(t)dt = \int_{J\cap K_1} f'(t)dt + \sum_n \int_{J\cap I_n} f'(t)dt,$$

ce qui constitue la “troisième opération”, qui généralise en fait une procédure déjà utilisée par Harnack, et qui permet de prolonger la définition de l'intégrale à un intervalle J sur lequel elle n'était pas définie jusque là (et donc sur tous ses sous-intervalles, et sur tout ce qui s'en déduit par les deux premières opérations).

Par les deux premières opérations, on arrive au noyau parfait K_2 de E_2. S'il est vide, on a terminé. Sinon, on a comme ci-dessus que l'ensemble (évidemment ouvert) des points x de K_2 tels que

– la série $\sum_n \int_{I_n} f'(t)dt$ converge absolument au voisinage de x ;

– $f'\mathbf{1}_{K_2}$ est bornée (donc Lebesgue intégrable) au voisinage de x,

est *non vide* – et même est dense dans K_2. Donc son complémentaire E_3, est un fermé d'intérieur vide de K_2. Et on recommence. Au besoin, on construit ainsi E_n, K_n ($K_n =$ le noyau parfait de E_n) pour tout $n \in \mathbb{N}^*$, puis on pose $E_\omega = \cap_n E_n$, $K_\omega = \cap_n K_n =$ le noyau parfait de E_ω, et si $K_\omega \neq \emptyset$, on considère $E_{\omega+1}$, qui est un fermé d'intérieur vide de K_ω, etc. Ayant construit E_α, K_α on construit comme ci-dessus $E_{\alpha+1}$, $K_{\alpha+1}$, et ayant construit les E_α, K_α pour $\alpha < \beta$ (β ordinal limite) on pose $E_\beta = \cap_{\alpha<\beta} E_\alpha$, $K_\beta = \cap_{\alpha<\beta} K_\alpha$.

On obtient ainsi une famille transfinie décroissante de fermés E_α dont chacun est d'intérieur vide dans le précédent. Par le théorème de Cantor-Baire, il existe un α au plus dénombrable tel que $E_\alpha = \emptyset$. À ce moment là, la “totalisation” est terminée : l'intégrale de f' est définie sur $[a,b]$ et $\int_a^b f'(t)dt = f(b) - f(a)$.

8.1.2 Récapitulation : conditions d'intégrabilité, intégrales restreinte et large

Une fonction g est Denjoy intégrable au sens large (D-intégrable) si l'itération transfinie des trois opérations ci-dessus (ce que Denjoy appelle la “totalisation” de g) ne s'arrête que sur un E_α vide, i.e. si, tant que $E_\alpha \neq \emptyset$, l'une des trois opérations s'applique.

Comme les équations de Maxwell, chacune des opérations de Denjoy porte le nom de quelqu'un d'antérieur (c'est commode pour se les remémorer) :

1. 1ère opération (Chasles) : intégrer sur les segments, et additionner les intégrales correspondant aux segments contigus ;

2. 2ème opération (Cauchy) : prolonger par continuité les intégrales aux intervalles maximaux (intégrales impropres) ;

3. 3ème opération (Harnack-Lebesgue) : dans un intervalle où c'est possible, sommer la série des intégrales sur les intervalles maximaux (si elle converge absolument – donc commutativement – car il n'y a pas d'ordre naturel) et ajouter l'intégrale sur le complémentaire (si elle existe) ;

4. (boucle) : réitérer, en essayant les opérations dans cet ordre.

Cette itération transfinie des trois opérations n'est pas un luxe, en ce sens que pour tout ordinal α fini ou transfini dénombrable, il existe une dérivée f' dont la "totalisation" se termine exactement à l'ordinal α. Denjoy [9] donne explicitement une telle dérivée, par une construction transfinie à partir de la dérivée de fonctions du type $(x-a)^2(b-x)^2 \sin \frac{1}{(x-a)^2(b-x)^2}$.
Donc, si on veut pouvoir intégrer toutes les dérivées, il faut pouvoir poursuivre l'itération transfinie jusqu'à n'importe quel ordinal α au plus dénombrable.
D'ailleurs, dire "on s'interdira de dépasser tel ordinal α_0" serait arbitraire et antinaturel : pourquoi s'arrêter à α_0 et non à $\alpha_0 + 1$? C'est toujours comme cela avec les ordinaux transfinis, il y a toujours un suivant, on ne peut pas dire "tiens je vais m'arrêter là", et c'est pour cette raison que Cantor a été bien obligé de les considérer. La seule borne naturelle pour la totalisation est Ω, le premier ordinal non dénombrable, parce qu'elle correspond à un phénomène naturel, à savoir le théorème de Cantor-Baire, qui prouve que Ω ne peut jamais être atteint dans une totalisation : on ne le *décrète* pas, on le *constate*.

On remarquera que les opérations s'appliquent dans un ordre parfaitement déterminé, il n'y a pas d'arbitraire : à chaque étape, on applique la première opération si elle est possible, sinon c'est la seconde – si elle est possible – et sinon c'est la troisième. Il s'agit donc d'un véritable algorithme (transfini), dont le théorème de Cantor-Baire garantit qu'il s'achève avant Ω. La fonction qu'on entre est intégrable au sens de Denjoy si l'algorithme s'arrête sur l'ensemble vide : il crache alors la valeur de l'intégrale. (Si la fonction entrée n'est pas intégrable au sens de Denjoy, la machine crache un ensemble E_α non vide sur lequel aucune des

trois opérations n'est applicable, et la valeur de l'intégrale sur les composantes connexes du complémentaire de cet ensemble.)
Bien sûr, aucune machine physique n'est capable d'accomplir une infinité de tâches en un temps fini (ni de survivre un temps infini, d'ailleurs), et donc "l'algorithme" de Denjoy ne donnera jamais lieu à un logiciel ... Mais le simple fait qu'un algorithme transfini soit possible est en soit assez remarquable pour que l'intégrale de Denjoy (et la définition que Denjoy en donne) soit la préférée des logiciens (cf. [13] et ses références, notamment aux travaux de Ajtai et de Dougherty et Kechris [11] qui ont démontré que, en un sens très précis, l'algorithme transfini de Denjoy – ou plutôt une variante, qui en fait un vrai algorithme, ne parlant que de nombres entiers – est la façon la plus simple et la plus constructive de remonter aux primitives).
Comme le dit Freiling [13], le but du logicien "n'est pas de présenter un algorithme qui puisse être accompli physiquement" mais "d'utiliser des programmes pour illustrer et expliquer la nature constructive de l'anti différentiation". Car celle-ci, dans l'approche de Denjoy, possède en effet "un haut degré de constructivité". "Supposons, ajoute Freiling, qu'on vous donne une suite infinie de 0 et de 1 et que vous vouliez savoir si la suite contient un 1. Tout le monde est d'accord pour dire qu'il y a un algorithme simple pour le savoir, même s'il peut être impossible à accomplir physiquement. Il semble qu'il faudrait un ordinateur ayant une aptitude spéciale (appelée quelquefois un 'esprit infini') pour accomplir cette tâche. L' 'esprit infini' de l'ordinateur lui permettrait de parcourir un nombre infini d'étapes dont on sait déjà qu'elles sont calculables, et de rapporter si un certain événement a été rencontré ou non. [...] Le résultat de Ajtai, Dougherty et Kechris (combiné avec le travail de Kleene) dit, de façon surprenante, que cet 'esprit infini' est le seul ingrédient nouveau dont on a besoin pour calculer les primitives !"

On remarque que dans la totalisation d'une dérivée f', le lemme de Lebesgue-Denjoy montrait en fait un résultat plus précis : l'ensemble des points x de K_1 tels que

- la série $\sum_n \omega(f, I_n)$ converge au voisinage de x :
- $f'\mathbf{1}_{K_1}$ est bornée (donc L-intégrable) au voisinage de x,

(où $\omega(f, I_n)$ désigne l'oscillation de f sur I_n) est dense dans K_1.
On peut donc remplacer la "troisième condition" énoncée plus haut par celle-ci :

"Troisième condition restreinte" : sur toute composante connexe I_n du complémentaire d'un parfait P, l'ensemble des points de P au voisinage desquels la série $\sum_n \sup_{J \subset I_n} |\int_J g(t)dt|$ ne converge pas, n'est nulle part dense sur P.

g est alors dite Denjoy intégrable au sens *restreint* ("D_*-intégrable")[25]. Les dérivées sont D_*-intégrables.

8.2 Définition descriptive de l'intégrale de Denjoy

Elle fait appel – comme pour l'intégrale de Lebesgue – à la notion d'absolue continuité, mais ici sous une forme un peu plus générale :

Définition 8.1 *Soit $E \subset \mathbb{R}$. Soit $f : \mathbb{R} \to \mathbb{R}$. On dit que f est absolument continue (au sens restreint) sur E ssi $\forall \varepsilon > 0$, existe $\eta > 0$ t.q. pour toute subdivision $a_0 < a_1 < \cdots < a_N$ de $\mathbb{R}$ par des points de E, $\sum_{i=1}^{N}(b_i - a_i) < \eta$ implique $\sum_{i=1}^{N} \omega(f, [a_i, b_i]) < \varepsilon$. On note $f \in AC_*(E)$.*

L'intégrale de Denjoy étant obtenue à partir de l'intégrale de Lebesgue par – entre autres – des passages à la limite (des intégrales généralisées), il faut refléter cette opération dans la notion de fonction absolument continue :

Définition 8.2 *Soit $E \subset \mathbb{R}$. Soit $f : \mathbb{R} \to \mathbb{R}$. On dit que f est absolument continue généralisée (au sens restreint) sur E ssi f est continue sur E et que E est une réunion dénombrable de parties E_n bornées sur chacune desquelles f est AC_*. On note $f \in ACG_*(E)$.*

Propriétés.

1. $AC_*(E)$ et $ACG_*(E)$ sont des $\mathbb{R}$-algèbres.

2. Une fonction ACG_* sur $E \subset \mathbb{R}$ est dérivable p.p. sur E.

3. Si f est continue sur $E \subset \mathbb{R}$, et dérivable en tout point de E (sauf peut-être ceux d'une partie dénombrable de E), f est ACG_*.

4. Si f est ACG_* sur un fermé E de $\mathbb{R}$, l'image par f de toute partie négligeable de E est négligeable ("propriété de Lusin").

25. L'intégrale D_* est due à Denjoy (1912) ; l'intégrale D à Denjoy et Khintchine (1916).

Un exemple. Considérons la fonction $f(x) = x^2 \sin \frac{1}{x^2}$ (complétée par 0 en 0), et $E =]-1, 1[$. f est dérivable en tout point de E. Mais elle n'est pas AC_* sur $]-1, 1[$. En effet, l'oscillation de f sur l'intervalle $I_n = [\frac{1}{\sqrt{2n\pi+\frac{\pi}{2}}}, \frac{1}{\sqrt{2n\pi}}]$ est $\omega_n = \frac{1}{2n\pi+\frac{\pi}{2}}$, donc la série des ω_n diverge, et donc il existe $\varepsilon > 0$ t.q. $\forall \eta > 0$, on puisse toujours trouver des intervalles I_n dans $[0, \eta]$ tels que la somme des ω_n correspondants soit $> \varepsilon$. Par contre, elle est AC_* sur tout $]-1, -\frac{1}{n}] \cup \{0\} \cup [\frac{1}{n}, 1[$ $(n \geq 1)$, donc elle est bien ACG_* sur $]-1, 1[$.

On peut énoncer maintenant la *définition descriptive de l'intégrale de Denjoy au sens restreint*:

Théorème 8.1 *Une fonction f (à valeurs réelles) est Denjoy- intégrable au sens restreint (D_*-intégrable) sur un intervalle $[a, b] \subset \mathbb{R}$ ssi existe une fonction $F \in ACG_*([a, b])$ telle que $F' = f$ p.p. L'intégrale de Denjoy de f sur $[a, b]$ est alors $\int_a^b f(t)dt = F(b) - F(a)$.*

Puisqu'une fonction dérivable en tout point de $[a, b]$ est ACG_*, sa dérivée est D_*-intégrable : on retrouve ainsi le fait que l'intégrale de Denjoy contient l'intégrale de Newton. Si F est dérivable en tout point de $[a, b]$ sauf peut-être à l'exception d'une quantité dénombrable de points, mais qu'elle est continue sur $[a, b]$, sa dérivée est encore (D_*-intégrable) – et toute fonction qui coïncide p.p. avec sa dérivée l'est aussi.

Il existe une définition descriptive de l'intégrale de Denjoy au sens large, qui est parfaitement analogue sauf que ACG_* y est remplacé par ACG (c'est la même chose mais avec les variations au lieu des oscillations), et que la dérivée est remplacée par une notion nouvelle, celle de "dérivée approximative" (dans le taux de variation, h tend vers 0 *le long d'un ensemble de densité 1 ...*)

8.3 Quelques propriétés de l'intégrale de Denjoy

Proposition 8.1 *Une fonction f D_*-intégrable et positive p.p. est L-intégrable. Par conséquent, une fonction f mesurable est L-intégrable ssi $|f|$ est D_*-intégrable.*

En effet, il existe F ACG_* t.q. $f = F'$ p.p., donc $F' \geq 0$ p.p. ; cela implique que F est croissante, donc absolument continue au sens classique (F est continue), et donc f est L-intégrable.
Ainsi, la théorie de Denjoy n'apporte, par rapport à celle de Lebesgue, que des intégrales semi-convergentes.

L'intégrale de Denjoy (aussi bien la D que la D_*) a toutes les propriétés basiques ; l'intégration par parties est valable : si f est intégrable, et F son intégrale indéfinie, et si g est C^1, alors $\int_a^b fg = F(b)g(b) - F(a)g(a) - \int_a^b Fg'$. La formule de changement de variable $\int_a^b f(g(t))\, g'(t)dt = \int_{g(a)}^{g(b)} f(s)ds$ est vraie dès que g est dérivable et que f est une dérivée sur $g([a,b])$.
Les théorèmes de convergence monotone et dominée sont valables pour les intégrales D et D_*.
Le théorème de convergence dominée peut s'exprimer ainsi ("théorème de convergence encadrée") :

Théorème 8.2 *Soient I un intervalle de $\mathbb{R}$ (borné ou non), et $f_n : I \to \mathbb{R}$ $(n \geq 0)$, $f : I \to \mathbb{R}$ des applications t.q. toutes les f_n soient D_*- intégrables sur I. Disons que la suite $(f_n)_n$ a la propriété de* convergence encadrée *vers f sur I si :*

1. *la suite (f_n) converge p.p. vers f sur I.*

2. *il existe des fonctions $g : I \to \mathbb{R}$ et $h : I \to \mathbb{R}$ D_*- intégrables sur I qui "encadrent" toutes les f_n : $\forall n,\ g \leq f_n \leq h$.*

Dans ce cas, f est D_*- intégrable sur I et $\int_I f = \lim_{n\to\infty} \int_I f_n$.

8.4 Remarques et généralisations

La notion de fonction ACG_* passe au cas de plusieurs variables, en remplaçant les intervalles par des pavés. Par exemple, pour deux variables, la variation de F sur un rectangle de sommets (x,y), $(x+h,y)$, $(x+h,y+k)$, $(x,y+k)$, est $V(x,y,h,k) = [F(x+h,y+k) - F(x+h,y)] - [F(x,y+k) - F(x,y)]$ (et le lecteur devine la généralisation en dimension supérieure) ; l'oscillation dans un rectangle (de côtés parallèlles aux axes, toujours) est le sup. des variations dans les sous-rectangles. La "totalisation" consiste alors à remonter de $\frac{\partial^2 F}{\partial x \partial y}$ ou, plus généralement de $\lim_{h,k\to 0} V(x,y,h,k)/hk$, à F. On peut donner une définition descriptive de l'intégrale double (ou n-uple) indéfinie de Denjoy. On peut même en donner un algorithme constructif qui se termine à un ordinal dénombrable, exactement comme en dimension 1 : cf. Looman 1923 [18] et Krzyżański 1934 [15]. Le théorème de Fubini est alors valable.
Si on remplace l'espace d'arrivée $\mathbb{R}$ par un espace de Banach, on obtient l'intégrale de Denjoy-Bochner (R. A. Gordon, 1989).

8.5 Perron

Perron (1914) donne une construction complètement différente de l'intégrale de Denjoy restreinte.
Soit f' une dérivée. Soient φ, ψ deux fonctions continues sur $[a,b]$ t.q. $\varphi(a) = \psi(a) = 0$. Perron (et déjà, avant lui, de la Vallée Poussin) dit que φ est une "majorante" (forte) de f' sur $[a,b]$ si

$$\underline{D}\varphi(x) = \underline{\lim}_{h\to 0}\frac{\varphi(x+h)-\varphi(x)}{h}$$

est $\geq f'(x)$, $\forall x \in [a,b]$, et que ψ est une "minorante" (forte) de f' sur $[a,b]$ si

$$\overline{D}\psi(x) = \overline{\lim}_{h\to 0}\frac{\psi(x+h)-\psi(x)}{h}$$

est $\leq f'(x)$, $\forall x \in [a,b]$.
Pour $x \in [a,b]$, on a alors $\varphi(x) \geq \psi(x)$. En effet, $\underline{D}(\varphi-\psi) \geq \underline{D}(\varphi) - \overline{D}\psi \geq 0$, donc $\varphi-\psi$ est croissante, et puisqu'elle est nulle en a, elle est ≥ 0 pour $x \in [a,b]$.
Mais $f - f(a)$ est elle-même à la fois une majorante et une minorante de f' sur $[a,b]$, donc $\forall x \in [a,b]$, et pour toutes φ, ψ resp. minorante et majorante, on a $\varphi(x) \geq f(x) \geq \psi(x)$.
Donc $f - f(a)$ peut être caractérisée comme étant *la borne inférieure de l'ensemble des majorantes*, et *la borne supérieure de l'ensemble des minorantes.*

L'idée de Perron est de considérer cela comme une *définition* de l'intégrale de f' et, plus généralement, de dire qu'une fonction f est intégrable sur $[a,x]$ *ssi* la borne inférieure de l'ensemble de ses majorantes coïncide avec la borne supérieure de l'ensemble des minorantes – la valeur commune étant alors par définition, l'intégrale de Perron de f sur $[a,x]$.
L'intégrale de Perron englobe l'intégrale de Lebesgue (c'est, essentiellement, un théorème de C. de la Vallée Poussin – antérieur en fait à la définition de Perron), et également l'intégrale de Newton. Mais, théorème (Hake, 1921 ; Alexandroff, 1924 ; Looman, 1925) : l'intégrale de Perron est exactement équivalente à l'intégrale de Denjoy restreinte (c'est pourquoi Saks [24] l'a appelée l'intégrale de Denjoy-Perron, ce qui n'a guère plu à Denjoy ...)

Denjoy n'aimait pas du tout la définition de Perron, qu'il trouvait totalement ineffective. Voici ce qu'il en disait [10] :
"Perron a dit : *"Donnez-moi une fonction finie quelconque f. Je considère TOUTES les majorantes fortes, TOUTES les minorantes fortes relatives à $f(x)$. Ne me demandez pas d'en calculer une seule. Ne me demandez pas si oui ou non il en existe*

aucune. Je ne saurais vous satisfaire. Mais si ces fonctions me tombent toutes *du ciel, et aussi bien dans l'une des espèces que dans l'autre, si je constate encore que l'écart minimum entre majorantes et minorantes est nul, je déclare que leur borne intermédiaire commune est l'intégrale de la fonction $f(x)$."* Et le public mathématique a convenu sans balancer : *"Perron a intégré la fonction $f(x)$"* ".
Il ajoute :
"On voudrait construire un avion dont la vitesse en régime permanent fût exactement 1500 km à l'heure. Il y a deux façons de concevoir la solution du problème : Les uns se livreront à des études patientes, laborieuses, d'aérodynamique expérimentale et théorique. Ils dessineront d'innombrables plans de pièces, de machines, d'appareils.
À l'opposé, un disciple de Perron se contentera de dire : *"Je considère* tous *les modèles d'avion dont la vitesse dépasse 1500 km à l'heure. Je considère* tous *les modèles d'avion dont la vitesse est inférieure à 1500 km à l'heure. Il n'y a pas de raison pour que les deux classes soient écartées l'une de l'autre. J'ai donc défini l'avion cherché".*"

9 Au-delà de l'intégrale de Denjoy : Foran

Foran [12] propose un cadre général[26] qui donne un assez bon point de vue unificateur :
Disons qu'une fonction F est de Lusin si elle satisfait à la "condition N de Lusin", à savoir qu'elle envoie toute partie négligeable (au sens de Lebesgue) de $\mathbb{R}$ sur une partie négligeable de $\mathbb{R}$.
Si F est continue, le fait qu'elle soit de Lusin est une condition nécessaire et suffisante pour que F envoie toute partie mesurable sur une partie mesurable.
Les fonctions AC_* sont continues et de Lusin, et donc les fonctions ACG_* aussi. (Idem pour les AC et les ACG).
Soit $\mathcal{F}$ une classe additive de fonctions continues de Lusin, et D une application de $\mathcal{F}$ dans l'espace des fonctions $\mathbb{R} \to \mathbb{R}$ (en fait on demande seulement que pour $F \in \mathcal{F}$, DF soit définie p.p.), telle que $\forall F, G \in \mathcal{F}$, $DF - DG = D(F - G)$ p.p., et qui prolonge la dérivation, en ce sens que $DF = F'$ p.p. sur l'ensemble des points où F' existe. Foran dit qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ est $\mathcal{F}D$-intégrable sur $[a, b] \subset \mathbb{R}$ s'il existe $F \in \mathcal{F}$ telle que $f = DF$ p.p. sur $[a, b] \subset \mathbb{R}$.
F est alors unique (à une constante près) car si on avait $f = DF = DG$ p.p., on aurait $D(G - F) = DG - DF = 0$ p.p. Or un théorème de Bary et Banach (cf. par ex. [24], théor. 7.7., p. 285-286) dit en particulier que si H est une fonction

26. Il en existe d'autres. Je n'en parlerai pas ici.

continue de Lusin t.q. $H' \geq 0$ en presque tout point où H' existe, H est croissante. On a donc ici $F - G =$cte.
On peut donc poser, avec Foran : $\int_a^b f(t)dt = F(b) - F(a)$, ce qui définit l'intégrale $\mathcal{F}D$ de Foran.
$\mathcal{F}$ est la classe des intégrales indéfinies pour l'intégrale $\mathcal{F}D$. Les classes d'intégrales indéfinies des intégrales usuelles sont multiplicatives, i.e. si F et G sont des intégrales indéfinies, FG en est une aussi (c'est vrai pour les intégrales de Newton, Cauchy, Riemann, Lebesgue, Denjoy ...) Mais Foran ne l'impose pas *a priori*, cette condition ne semblant pas nécessaire.
L'intégrale de Lebesgue correspond à $\mathcal{F}$ = l'espace des fonctions absolument continues (au sens classique) et D= la dérivation usuelle ; celle de Denjoy restreinte, à $\mathcal{F} = ACG_*$ (et D= la dérivation usuelle aussi) ; celle de Denjoy large, à $\mathcal{F} = ACG$ et D= la dérivation approximative. Il en existe d'autres.

10 L'intégrale de Riemann complète

10.1 L'idée

Dans les définitions de Cauchy et de Riemann, on se représentait le graphe de la fonction f (du moins, dans les cas les plus simples), et on l'approximait par celui d'une fonction en escalier. On voit tout de suite ce qui ne va pas avec une telle définition : prenez le graphe de la fonction $x \to x^{-1/2}$ sur $]0, 1]$. L'aire sous la courbe est finie. On voudrait donc que la fonction soit intégrable sur $]0, 1]$. Mais si vous essayez d'approximer son graphe par celui d'une fonction en escalier, il y a un problème avec la première marche : vous la choisissez à la hauteur $f(x_1)$ (avec x_1 dans l'intervalle I_1 de la subdivision), et en prenant x_1 arbitrairement voisin de 0, cette hauteur va être arbitrairement grande. Alors, même si vous avez pris soin de prendre toutes vos marches moins larges qu'un certain $\delta > 0$, aussi petit que soit ce δ, vous pourrez rendre vos sommes de Riemann arbitrairement grandes, et la définition de Cauchy et Riemann ne s'applique pas. Bien sûr, Cauchy s'en était aperçu et avait proposé cette parade : *imposer que la première marche soit à la hauteur zéro.* Autrement dit, considérer l'intégrale de f sur $]0, 1]$ comme la limite des intégrales sur $[a, 1]$ pour $a \downarrow 0$ – bref, une intégrale impropre. Bien sûr, au lieu d'imposer que la première marche soit à la hauteur 0, on pourrait tout aussi bien lui permettre de se placer à une hauteur quelconque *inférieure à un plafond donné indépendant de* δ. Dans tous les cas, cela revient à faire abstraction de la *vraie* forme de la courbe dans la première marche – en la tronquant !
La raison en est bien simple : une fonction en escalier (à un nombre *fini* de marches) ne peut épouser partout la forme d'une courbe non bornée – aussi

petites qu'on prenne les marches. Elle ne peut non plus épouser partout la forme d'une courbe qui oscille trop, même si elle est bornée : par exemple la caractéristique des rationnels. Là, même le procédé de Cauchy ne marche plus (même réitéré transfiniment !).

Une idée serait d'admettre un nombre infini de marches. Dans le cas de la courbe $x \to x^{-1/2}$ sur $]0,1]$, on voit bien ce que cela donnerait : un escalier qui monte le long de l'axe des y, avec des marches de plus en plus étroites, de largeur tendant vers 0. Mais l'aire comprise sous l'escalier est la somme d'une série, et on calcule la somme d'une série comme limite des sommes partielles – ce qui fait qu'on est ramené au procédé (d'intégrale impropre) de Cauchy.

Toutefois, cela suggère une idée : ce qui ne va pas dans la définition de Riemann, c'est qu'elle suppose qu'on prenne une même largeur δ tout le long du graphe de la fonction. Or on voit bien que là où la fonction varie rapidement, on doit diminuer le pas des marches.

Euler, dans ses *institutiones calculi integralis* (1768), parlant du calcul numérique des intégrales par ce qu'on appelle aujourd'hui les sommes de Riemann, écrivait : *"Nous avons déjà noté que les distances $a_j - a_{j-1}$, par lesquelles x est supposé croître successivement, doivent être prises très petites pour que les valeurs correspondantes $f(a_{j-1})$, $f(a_j)$, ne diffèrent à leur tour guère l'une de l'autre ; à partir de cela, il faut juger si les intervalles $a_1 - a$, $a_2 - a_1$, $\ldots$ doivent être pris égaux ou inégaux. En fait, là où la valeur de $f(x)$ ne change guère lorsque x varie, l'intervalle par lequel x croît peut être pris grand sans danger. D'autre part, là où des changements peu importants de x conduisent à des variations violentes de $f(x)$, on devra prendre l'intervalle très petit".* (Cité dans Mawhin [20].)

Or, c'est vrai non seulement pour le calcul pratique, mais aussi pour les questions théoriques, quand on a affaire à des fonctions qui présentent des irrégularités vraiment très violentes. Comme le dit Borel [5][27] : *"Un caractère important de cette définition [de Riemann] est le suivant : la division en intervalles est entièrement indépendante des propriétés de la fonction ; si l'on considère deux fonctions différentes, on prendra pour ces fonctions les mêmes intervalles, c'est-à-dire qu'on leur appliquera* un procédé de calcul uniforme. *C'est évidemment là un grand avantage pour le calcul. Mais c'est en même temps un inconvénient : un tel procédé, qui ne tient pas compte des propriétés particulières de la fonction à laquelle il s'applique, peut être comparé à ces vêtements confectionnés qui ne sauraient être exactement ajustés, surtout s'il s'agit d'habiller un individu difforme : certaines fonctions singulières ont pu être justement comparées aux types monstrueux de la biologie".*

La solution est là : *remplacer dans la définition de Cauchy et Riemann la constante*

27. Reproduit dans [6], pp. 1277-1307. La citation est p. 1301. Cité aussi dans Mawhin [20].

δ *par une* fonction *dont on peut adapter les propriétés à celles de la fonction* f *à intégrer* – ne garder que les subdivisions qui adhèrent au plus près au graphe de la fonction, en suivant ses irrégularités. D'où les définitions que voici.

10.2 Définition(s)*

Définition 10.1 *1. Une* subdivision pointée *(s.p.) d'un intervalle* $I = [a, b]$ *($a < b$) de* $\mathbb{R}$ *est un ensemble fini* $P = \{(x_1, I_1), \ldots, (x_n, I_n)\}$*, où* $I_1, \ldots, I_n$ *sont des intervalles de* I*, d'intérieurs disjoints deux à deux, dont la réunion est* I*, et où* $\forall i$*,* x_i *est un point de* I_i*.*

2. *On note* $\ell(I)$ *la longueur d'un intervalle* I

3. *On appelle* jauge *sur un segment* I *de* $\mathbb{R}$ *une application* $\delta : I \to \mathbb{R}_+^*$*.*

4. *Étant donné une jauge* δ *sur* I*, une s.p.* $\{(x_1, I_1), \ldots, (x_n, I_n)\}$ *de* I *sera dite "fine selon la jauge* δ*" ou, pour le dire plus vite,* δ*-fine si pour tout* i *dans* $\{1, \ldots, n\}$*,* $\ell(I_i) < \delta(x_i)$*.*

On voit que la jauge contrôle la taille des intervalles de la subdivision – d'où son nom. Avant d'aller plus loin, il faut vérifier une petite chose toute bête, mais sans laquelle tout cela n'aurait guère d'intérêt :

Proposition 10.1 *Pour toute jauge* δ*, il existe des s.p.* δ*-fines.*

Preuve. Borel-Lebesgue.

Maintenant, encore quelques définitions :

Définition 10.2 *Étant donné une fonction* f *sur* I*, à valeurs réelles, et une s.p.* $P = \{(x_1, I_1), \ldots, (x_n, I_n)\}$ *de* I*, on note :*

$$S(f, P) = \sum_{i=1}^{n} f(x_i)\ell(I_i)$$

("Somme de Riemann" de f *relative à* P*).*

---

*. Tout ce qui suit est assez succint. Pour plus de détails, cf. par exemple mes esquisses de cours "l'intégrale simple" et "l'intégrale de Riemann complète", ou [20], [19] ... Pour d'autres références, me contacter.

Définition 10.3 *Soit I un intervalle de $\mathbb{R}$, compact.*

1. *Une fonction $f : I \to \mathbb{R}$ est dite intégrable au sens de Riemann complété (RC-intégrable) sur I s'il existe un réel J tel que, pour tout $\varepsilon > 0$, il existe une jauge $\delta : I \to \mathbb{R}_+^*$ t.q. pour toute s.p. P δ-fine de I, on ait :*

$$|J - S(f,P)| < \varepsilon. \tag{4}$$

   J est alors unique, on le note

$$\int_I f(t)dt = J,$$

   ($\int_a^b f(t)dt$ si $I = [a,b]$) et on l'appelle l'intégrale (de Riemann complète) de f sur I. (On définit aussi alors $\int_b^a f(t)dt = -\int_a^b f(t)dt$.)

2. *Une jauge satisfaisant aux conditions ci-dessus sera dite adaptée à f sur I à la précision ε, ou, plus brièvement : ε-adaptée à f sur I.*

Cette définition est due indépendamment à Kurzweil (1957) et à Henstock (1961). On voit que la seule différence par rapport à la définition de l'intégrale de Riemann, est qu'on a remplacé le *réel* $\delta > 0$ par une *fonction* δ à valeurs > 0, et qu'on a adapté en conséquence la notion de subdivision pointée δ-fine.
Ce tout petit changement devrait répondre à nos attentes, *et il le fait.* L'intégrale RC, non seulement dépasse l'intégrale de Riemann, et permet comme prévu d'intégrer $x \to x^{-1/2}$ sur $]0,1]$ et la caractéristique des rationnels sur tout segment (cf. plus loin), mais elle contient aussi l'intégrale de Lebesgue, l'intégrale de Newton – en fait elle est *équivalente* à l'intégrale D_* de Denjoy, tout en étant beaucoup plus commode à définir : quiconque comprend la définition usuelle de l'intégrale de Riemann, comprend *ipso facto* celle de l'intégrale de Riemann complète.

10.2.1 Premières propriétés

Toutes les propriétés basiques sont vérifiées.
Pour donner juste deux exemples de démonstrations dans la théorie de Riemann complète, je vais prouver que l'intégrale contient l'antidérivation, et le théorème de convergence monotone.

Théorème 10.1 *Soit $a \in \mathbb{R}$, soit f continue et dérivable sur un intervalle ouvert $I \ni a$, borné. Alors $\forall x \in I$, f' est intégrable sur $[a,x]$ et $\int_a^x f'(t)dt = f(x) - f(a)$.*

Preuve.
Soit $\varepsilon > 0$. Pour tout $x \in I$, il existe $\delta(x) > 0$ tel que $\forall y, z \in I$, si $y \leq x \leq z$ et $z - y < \delta(x)$, alors

$$\left| f(z) - f(y) - (z-y)f'(x) \right| \leq \frac{\varepsilon(z-y)}{\ell(I)}$$

On a alors, pour toute s.p. δ-fine de $[a, x]$,

$$\begin{aligned} \left| f(x) - f(a) - S(f', P) \right| &= \left| \sum_{i=1}^{n} \big(f(a_i) - f(a_{i-1}) - (a_i - a_{i-1}) f'(x_i) \big) \right| \\ &\leq \frac{\varepsilon}{x-a} \sum_{i=1}^{n} (a_i - a_{i-1}) = \varepsilon. \end{aligned} \tag{5}$$

D'où le résultat.
Comme on le voit, la preuve n'est qu'un recopiage de la preuve classique de Cauchy. Elle est même *plus simple* que pour l'intégrale de Cauchy ou de Riemann, car on n'a *pas besoin* de minorer δ en invoquant Heine-Borel-Lebesgue (le fameux point qui avait échappé à Cauchy) : la définition de l'intégrabilité RC s'accommode très bien des δ non minorées par un nombre > 0 (elle est faite pour cela !), contrairement à la définition de l'intégrale C (ou R).

Le théorème de convergence monotone.

Théorème 10.2 *Soient I un intervalle de $\mathbb{R}$ (borné ou non), et $f_n : I \to \mathbb{R}$ $(n \geq 0)$, $f : I \to \mathbb{R}$ des applications telles que :*

1. *toutes les f_n sont RC- intégrables sur I ;*
2. *la suite (f_n) est monotone et converge (simplement) vers f sur I.*

Alors f est RC- intégrable sur I ssi *la suite des $\int_I f_n$ converge, et dans ce cas $\int_I f = \lim_{n\to\infty} \int_I f_n$.*

Preuve Quitte à changer le signe, on se ramène au cas où la suite (f_n) est croissante. L'une des implications est évidente : si f est RC-intégrable, la suite (croissante) des $\int_I f_n$ est majorée par $\int_I f$, donc converge. Pour montrer la réciproque, il suffit d'écrire, pour $P = \{(x_1, I_1), \ldots, (x_n, I_n)\}$ une s.r. P de I :

$$\begin{aligned} \left| S(f,P) - \ell \right| &\leq \sum_{i=1}^{n} \left| f(x_i) - f_{M(x_i)}(x_i) \right| \ell(I_i) \\ &+ \left| \sum_{i=1}^{n} \left(f_{M(x_i)}(x_i)\ell(I_i) - \int_{I_i} f_{M(x_i)}(t)dt \right) \right| + \left| \sum_{i=1}^{n} \int_{I_i} f_{M(x_i)}(t)dt - \ell \right|. \end{aligned} \tag{6}$$

Des trois termes du second membre, le premier est majoré par $\varepsilon|I|$, le dernier par ε et pour majorer le deuxième, on choisit pour chaque $k \in \mathbb{N}$ une jauge δ_k qui soit 2^{-k}-adaptée à f_k, et $\forall x \in I$ on prend $\delta(x) = \delta_{M(x)}(x)$. (C'est *cela* qui ne marcherait pas pour l'intégrale de Riemann : même si les jauges δ_k sont constantes, δ ne l'est pas en général, et aucune jauge constante ne pourrait la remplacer.) Alors, si P est δ- fine, pour tout k de N à μ t.q. $\sigma(k) \neq \emptyset$ la famille $(x_i, I_i)_{i\in\sigma(k)}$ est δ-fine et on voit facilement en la complétant en une subdivision que le deuxième terme est $< 2\varepsilon$, par exemple.

Quelques exemples

1. La fonction $x \to x^{-1/2}$ (complétée comme on veut en 0) est RC- intégrable sur $[0,1]$. En effet c'est la limite croissante de ses restrictions aux intervalles $[\frac{1}{n}, 1]$ (prolongées par 0 partout ailleurs), et la suite des intégrales de celles-ci converge.

2. La caractéristique des rationnels est RC- intégrable sur $[0,1]$ (par exemple), car c'est la limite croissante des $f_m(x) = \lim_{n\longrightarrow\infty} |\cos\,[2\pi m!x]|^n$, qui sont elles-mêmes RC-intégrables (et d'intégrale nulle), vu la relation de Chasles et le fait que $\forall m$, f_m est nulle sur les intervalles $]\frac{k}{2\,m!}, \frac{k+1}{2\,m!}[$.

On a le même théorème de convergence "encadrée" que pour l'intégrale de Denjoy.

Les fonctions absolument intégrables

La valeur absolue d'une fonction RC-intégrable n'est pas nécessairement RC-intégrable.

Exemple 10.1 *La fonction f définie sur $[0,1]$ par $f(0) = 0$ et*

$$\forall n \in \mathbb{N}^*, \forall x \in]\frac{1}{n+1}, \frac{1}{n}], f(x) = (-1)^{n-1}n$$

est RC-intégrable sur $[0,1]$ (et son intégrale est $\sum_{n\geq 1} \frac{(-1)^{n-1}}{n+1}$). Sa valeur absolue n'est pas RC-intégrable.

Les fonctions qui sont ainsi "non absolument RC-intégrables" sont toujours des intégrales que, dans la théorie de Lebesgue, on considérerait comme impropres et non absolument convergentes[28] – à ceci près qu'elles seraient généralement impropres en *tout point d'un ensemble nulle part dense à peu près quelconque* ! Il

28. Une fonction RC-intégrable *bornée* est absolument RC-intégrable.

faudrait donc leur appliquer le procédé de Cauchy (et, en fait, aussi celui d'Harnack) un nombre transfini de fois (souvenez-vous de la construction de Denjoy) ...
L'étude des fonctions absolument RC-intégrables a son propre intérêt, exactement comme celle des séries (ou des intégrales impropres) absolument convergentes. (En fait, *ce sont exactement les fonctions L-intégrables* : cf. [19]).

Remarque 10.1 *La définition 10.3 et les théorèmes de convergence restent valables sur un intervalle non borné (i.e. si $a = -\infty$ et/ou $b = +\infty$), à condition de ne pas tenir compte des intervalles non bornés de la subdivision dans le calcul des sommes de Riemann. Cela signifie que, dans le calcul de $S(f, P)$, on doit remplacer $\ell(I_i)$ par 0 quand I_i n'est pas borné. On appelle cela la longueur efficace de I_i.*

Mesures : comparaison avec Lebesgue. Dans la théorie RC, on dit qu'une partie E de $\mathbb{R}$ est mesurable si pour tout intervalle borné I la fonction caractéristique $\mathbf{1}|_{E\cap I}$ est RC-intégrable sur $\mathbb{R}$. L'intégrale s'appelle alors la mesure de $E \cap I$. Si on prend $I_n = [-n, +n]$ ($n \geq 1$), les fonctions caractéristiques des $E \cap I_n$ forment une suite croissante qui converge vers la caractéristique de E. Donc, par convergence monotone, la caractéristique de E est RC-intégrable (on dit : E est de mesure finie) *ssi* la suite des mesures des $E \cap I_n$ est convergente, et alors l'intégrale est la limite : on l'appelle la mesure de E. (On obtiendrait la même condition, et la même limite, avec n'importe quelle autre suite exhaustive de bornés mesurables de $\mathbb{R}$, comme le montre un simple *shuffling* des deux suites.)
Si E n'est pas de mesure finie, on dit que la mesure de E est infinie.
L'additivité complète est immédiate.
La mesure ainsi définie est exactement la mesure de Lebesgue (cf. les chapitres 5 et 8 de [19]).
Une application $f : \mathbb{R} \to \mathbb{R}$ est dite intégrable sur E *ssi* $\mathbf{1}|_E f$ est RC-intégrable sur $\mathbb{R}$, et alors on note $\int_E f = \int_{\mathbb{R}} \mathbf{1}|_E f$ Ce n'est pas parce qu'une fonction f est intégrable sur une partie E qu'elle l'est sur toute sous-partie (même mesurable) de E.
Par exemple la fonction f de l'exemple 10.1, dont on a vu qu'elle est RC-intégrable sur $[0, 1]$, ne l'est pas sur la partie mesurable $\cup_{k\geq 0}]\frac{1}{2k+2}, \frac{1}{2k+1}]$, à cause de la divergence de la série harmonique.

10.3 L'intégrale RC dans $\mathbb{R}^n$

On remplace les intervalles par des pavés. On note $\ell(I)$ la longueur efficace de I, i.e. la longueur de son côté le plus long s'il est borné, 0 sinon. Une subdivision

pointée riemannienne est $\delta-$fine si chaque I_i est inclus dans le pavé de centre x_i dont tous les côtés sont de longueur $\ell(I_i)$. La définition de l'intégrabilité RC est, à part cela, exactement la même.
Les nouveautés dans $\mathbb{R}^n$ sont les problèmes d'intégrations successives (Fubini), de changements de variables dans les intégrales multiples, et l'analogue du TF (Stokes). Pour le théorème de Fubini, on a l'énoncé le plus satisfaisant qu'on puisse raisonnablement espérer (cf. [19], [20]) : c'est le même énoncé que pour Lebesgue, mais valable même pour les fonctions non absolument intégrables. Par contre, pour le TF et la formule de changement de variables (qui étaient valable dans $\mathbb{R}$ sous des hypothèses si générales), les énoncés qu'on trouve dans la littérature ne sont pas meilleurs que pour l'intégrale de Lebesgue : la formule de changement de variables est valable pour des intégrands absolument intégrables, i.e. L-intégrables, et c'est exactement le théorème de Lebesgue ; le théorème de Stokes s'énonce pour des chaînes et des formes C^1, donc là ce n'est pas mieux qu'avec l'intégrale de Riemann.

10.4 L'intégrale RC et l'intégration abtraite

Le point de vue de Lebesgue (une intégrale définie par une théorie de la mesure) permettait des généralisations abstraites. Henstock [14] pose les bases d'une théorie générale de l'intégration dans le point de vue de Riemann.
A priori, le point de vue de Riemann est inférieur à celui de Lebesgue, car il repose de façon essentielle sur la notion particulière *d'intervalle de subdivision* ou, dans $\mathbb{R}^n$, de pavé.
Toutefois, Henstock a répertorié les propriétés des pavés qui sont utiles pour l'intégrale, et les a érigées en axiomes, obtenant la notion abstraite "d'espace de division". Cette théorie est récente (1991), mais peut-être permettra-t-elle de donner au point de vue de Riemann une force et une généralité comparables à celui de Lebesgue.

11 Pour finir : quelques variantes de l'intégrale RC

11.1 McShane

Dans [21] (cf. aussi [22]), E. J. McShane, inspiré par les idées de Kurzweil et Henstock sur l' "intégrale de Riemann complète", propose une variante : dans la définition des sommes de Riemann, on *supprime* l'hypothèse que $x_i \in [a_{i-1}, a_i]$. (Une subdivision pointée est alors dite δ-fine si $\forall i$, $d(I_i \cup \{x_i\}) < \delta(x_i)$, où $d(I_i \cup \{x_i\})$

est le diamètre de $I_i \cup \{x_i\}$.) On demande donc que l'inégalité (4) ait lieu pour une plus vaste classe de subdivisions pointées : c'est plus restrictif que pour l'intégrale RC, il y aura donc moins de fonctions intégrables. En fait, McShane prouve qu'une fonction f est intégrable en son sens *ssi* f et $|f|$ sont RC-intégrables, i.e. *ssi* f est L-intégrable : l'intégrale de McShane est équivalente à l'intégrale de Lebesgue.

Il existe d'autres variantes de l'intégrale RC, dont certaines se situent entre celle de Lebesgue et celle de Riemann complète. Je n'en citerai qu'une :

11.2 Bongiorno

Bongiorno [4] (1996) construit l'intégrale la plus faible actuellement connue parmi celles qui contiennent celle de Lebesgue et celle de Newton ; contrairement à l'intégrale RC (i.e. D_*) elle ne contient pas l'intégrale impropre de Riemann.
Sa définition est une variante de celle de McShane, mais on ne garde que les subdivisions qui possèdent une propriété supplémentaire : on est donc moins exigeant que pour l'intégrale de McShane (=Lebesgue), et il y a plus de fonctions intégrables.
Bongiorno dit qu'une subdivision pointée $P = \{(x_1, I_1), \ldots, (x_n, I_n)\}$ d'un intervalle $[a, b]$ (N.B. on ne demande pas que $x_i \in I_i$) est $1/\varepsilon$- contrôlée ($\varepsilon > 0$) si $c(P) < 1/\varepsilon$, où $c(P) = \sum_i d(I_i \cup \{x_i\})$ ($d =$ le diamètre), et :

Définition 11.1 *Une fonction f définie sur $[a, b]$ est intégrable au sens de Bongiorno (B-intégrable) s'il existe un réel J tel que, pour tout $\varepsilon \in]0, (b-a)^{-1}[$, il existe une jauge δ sur $[a, b]$ telle que pour toute subdivision pointée P de $[a, b]$ δ-fine et $1/\varepsilon$-contrôlée, l'inégalité* (4) *soit satisfaite.*

Pour bien comprendre d'où vient la condition d'$1/\varepsilon$-contrôlabilité, voyons comment elle intervient dans la preuve que toute dérivée est B-intégrable. Soit donc $F : [a, b] \to \mathbb{R}$ dérivable. $\forall x \in [a, b]$, soit $\delta(x) > 0$ t.q. $\forall y \in [a, b] \setminus \{x\}$, $|y - x| < \delta(x) \implies \left|\frac{F(y)-F(x)}{y-x} - F'(x)\right| < \varepsilon^2/2$. Alors, pour P δ-fine et $1/\varepsilon$-contrôlée,

$$\begin{aligned} \Big| \sum_i F'(x_i)\ell(I_i) - (F(b) - F(a)) \Big| &\le \sum_i \Big| F'(x_i)\ell(I_i) - (F(a_i) - F(a_{i-1})) \Big| \\ &\le \varepsilon^2 \sum_i d(I_i \cup \{x_i\}) \le \varepsilon. \end{aligned} \tag{7}$$

Tous les théorèmes usuels de l'intégrale RC sont valables pour l'intégrale B. Par exemple, pour les théorèmes de convergence monotone et encadrée, il suffit de remarquer que les fonctions $f - f_n$ (ou $f_n - g$ et $h - f_n$) et leurs limites, étant B-intégrables, sont RC-intégrables, et comme elles sont positives, elles sont mesurables et L-intégrables : on peut donc appliquer les théorèmes de la théorie de Lebesgue, et puisque les L-intégrales de ces fonctions coïncident avec leurs intégrales de McShane, et donc avec leurs intégrales de B, les deux théorèmes sont valables pour l'intégrale B !

Bongiorno donne une définition descriptive de son intégrale : disons qu'une fonction F est AC_c sur une partie E de $[a, b]$ si elle est continue et que $\forall\varepsilon$, $\exists\eta$ et une jauge δ t.q. la variation totale de F sur les partitions partielles δ-fines $1/\varepsilon$-contrôlées à points dans E et de mesure $< \eta$, est inférieure à ε. Il dit que F est ACG_c sur $[a, b]$ si $[a, b]$ est réunion d'une suite de parties mesurables sur lesquelles F est AC_c. Alors : une fonction f est B-intégrable sur $[a, b]$ *ssi* existe une fonction F ACG_c sur $[a, b]$ t.q. $\forall x \in [a, b]$, $F(x) - F(a) = \int_a^x f(t)dt$.

11.3 Une variante en dimension ≥ 2

Pfeffer (1991) [23] définit dans $\mathbb{R}^n$ une intégrale qui permet d'avoir le théorème de Gauss-Green-Ostrogradski (i.e. la formule de Stokes pour $d\omega$ de degré maximum) dans probablement tous les cas où on sait actuellement lui donner un sens. C'est une intégrale de type RC.
(Notons en passant que quand on la définit dans $\mathbb{R}$, l'intégrale de Pfeffer se place strictement entre celle de Bongiorno et la RC, et elle contient la R-intégrale impropre. En fait, la question que se posait Bongiorno était justement de descendre, plus bas que l'intégrale de Pfeffer, vers l'intégrale minimale contenant les intégrales L et N.)

Références

[1] C. ARZELÀ. Sulla integrabilità di una serie di funzioni. *rend. Accad. dei Lincei (4)*, t. I:321–326, 1885.

[2] R. BAIRE. *Leçons sur les fonctions discontinues.* Gauthier-Villars, 1905.

[3] S. K. BERBERIAN. Why there is no "Fundamental theorem of calculus" for the Riemann integral. *Expo. Math.*, 11:271–279, 1993.

[4] B. BONGIORNO. Un nuovo integrale per il problema delle primitive. *Le Matematiche*, 60:299–313, 1996.

[5] E. BOREL. La théorie des ensembles et les progrès récents de la théorie de fonctions. *Rev. gén. Sci.*, 20:315–324, 1909.

[6] E. BOREL. *Oeuvres*, volume 3. Éd. du CNRS, Paris, 1972.

[7] G. CANTOR. Fondements d'une théorie générale des ensembles. *Acta Mathematica*, 2:381–408, 1883.

[8] A. DENJOY. *L'énumération transfinie*, volume 1 à 4. Gauthier-Villars, 1946 à 1954.

[9] Arnaud DENJOY. La totalisation des nombres dérivés non sommables, II. *Ann. École Norm. Sup.*, 34:181–236, 1917.

[10] Arnaud DENJOY. *Leçons sur le calcul des coefficients d'une série trigonométrique.* Gauthier-Villard (Paris), 1941.

[11] R. DOUGHTERY and A. S. KECHRIS. The complexity of antidifferentiation. *Adv. in Math.*, 88:145–169, 1991.

[12] J. FORAN. An extension of the Denjoy integral. *Proc. of the Amer. Math. Soc.*, 49, 1975.

[13] C. FREILING. How to compute antiderivatives. *The Bulletin of Symbolic Logic*, 1:279–316, 1995.

[14] R. HENSTOCK. *The general theory of integration.* Oxford: Clarendon Press, 1991.

[15] M. KRZYŻAŃSKI. Sur les fonctions absolument continues généralisées de deux variables. *Comptes Rendus Acad. Sciences*, 198:2058–2060, 1934.

[16] Henri LEBESGUE. Sur une généralisation de l'intégrale définie. *C.R.A.S.*, 132:1025–1028, 1901.

[17] Henri LEBESGUE. *Leçons sur l'intégration et la recherche des fonctions primitives.* Gauthier-Villard (Paris), 1904 (2nde éd. augmentée 1926, nouveau tirage 1950).

[18] H. LOOMAN. Sur la totalisation des dérivées des fonctions continues de plusieurs variables indépendantes. *Fundam. Math.*, 4:246–285, 1923.

[19] R. M. MacLEOD. *The generalized Riemann integral.* Math. Assoc. Amer., The Carus Mathematical Monographs 20, 1980.

[20] Jean MAWHIN. *Analyse (Fondements, techniques, évolution)*. De Boek, 1992.

[21] E. J. McSHANE. A unified theory of integration. *Amer. Math. Monthly*, 80:349–359, 1973.

[22] E. J. McSHANE. *Unified integration*. Academic Press, New York, 1983.

[23] W. F. PFEFFER. The Gauss-Green theorem. *Adv. in Math.*, 87:93–147, 1991.

[24] S. SAKS. *Theory of the integral*. Dover (New York), 1964.