JEAN MAWHIN

Présences des sommes de Riemann dans l'évolution du calcul intégral

Cahiers du séminaire d'histoire des mathématiques 1re série, tome 4 (1983), p. 117-147
<http://www.numdam.org/item?id=CSHM_1983__4__117_0>

PRESENCES DES SOMMES DE RIEMANN DANS L'EVOLUTION DU CALCUL INTEGRAL

par Jean MAWHIN*

INTRODUCTION

Le but de ce travail est de décrire le rôle joué par les sommes de Riemann dans l'évolution du calcul intégral depuis l'Antiquité grecque jusqu'à nos jours. On cherchera en particulier à mettre en évidence leurs avatars dans la lutte de préséance entre les deux aspects fondamentaux du calcul intégral, l'aspect sommatoire et l'aspect primitive. Nous ne voulons faire ni de l'histoire des sciences pour elle-même, ni de la mathématique technique. C'est pourquoi, d'une part, nous n'avons pas craint de commettre quelques anachronismes de notations ou de langage et, d'autre part, nous nous sommes limités au cas le plus simple des fonctions d'une variable réelle définies sur un intervalle borné. Nous espérons n'avoir pas trop trahi l'esprit du moment en décrivant les temps forts de cette évolution dans le langage actuel, et n'avoir pas trop banalisé les efforts intellectuels correspondants en présentant la logique des enchaînements comme presque fatale. Il faut dire que, pour ne pas trop allonger l'exposé, nous avons négligé bien des rameaux transverses à notre ligne générale. On verra d'ailleurs que les idées les plus simples se sont souvent fait attendre mais que leur beauté a très vite fait oublier ce manque de ponctualité.

1. LE COURANT GEOMETRIQUE OU LE CALCUL EXPLICITE D'INTEGRALES A L'AIDE DE SOMMES DE RIEMANN

Le calcul intégral trouve son origine et sa motivation dans des problèmes géométriques comme la détermination des longueurs de lignes courbes, des aires de figures planes ou de surfaces, des volumes de corps solides ainsi que dans

* Conférence donnée le 10 novembre 1982 au Séminaire d'Histoire des Mathématiques.

la détermination des centres de gravité de figures planes ou de corps solides. Les premières recherches de ce type remontent à l'Antiquité grecque, avec, surtout, Eudoxe, de l'école de Platon (IVe siècle av. J.C.) et Archimède, de l'école d'Alexandrie (IIIe siècle av. J.C.), qui créèrent et développèrent une méthode que le Brugeois Grégoire de Saint-Vincent devait appeler, au XVIIe siècle, la méthode d'exhaustion. Malgré la concurrence de procédés plus directs et plus efficaces, mais d'une rigueur plus discutable, cette méthode restera l'idéal des mathématiciens dans ce type de problèmes et, en 1742, soixante ans au moins après la création du calcul différentiel et intégral, Mc Laurin y fera appel dans son "Traité des Fluxions" pour donner un exposé plus rigoureux de ce calcul.

Introduisons ce qui sera le problème test de cet exposé avant de l'utiliser pour décrire les grandes lignes de la méthode d'exhaustion. Soient $a < b$ deux nombres réels, $[a,b]$ l'ensemble des réels x tels que $a \leq x \leq b$, et soit f une fonction réelle définie et positive sur $[a,b]$. Appelons $E(f)$ l'ensemble des points de coordonnées cartésiennes (x,y) du plan tels que $0 \leq y \leq f(x)$ pour chaque x appartenant à $[a,b]$. Si f est suffisamment régulière, la figure correspondante sera le quadrilatère curviligne compris entre le graphe de f, l'axe des abscisses et les parallèles à l'axe des ordonnées menées par les points de coordonnées $(a,0)$ et $(b,0)$. Jusqu'au XIXe siècle, on ne s'occupera d'ailleurs que du cas où f est donnée par une formule explicite qui est, au pire, une série de puissances. Le problème consiste à donner un sens à la notion d'aire de $E(f)$ et de la calculer et, jusqu'au XIXe siècle également, l'existence de cette aire sera acceptée comme géométriquement évidente, ainsi que ses propriétés de monotonie et d'additivité finie, et le problème se réduira au seul calcul.

La méthode d'exhaustion n'a pas pour but de fournir une méthode de calcul direct de cette aire, mais seulement un procédé de vérification pour un candidat, généralement obtenu de manière heuristique. Plus exactement on cherchera à montrer que cette aire est dans un rapport donné avec une autre aire, puisque chez les Grecs et longtemps après, il ne sera pas question de nombres réels mais de "grandeurs" et de leurs rapports. Pour ne pas alourdir trop l'exposé, nous nous en tiendrons à une description dans le langage des réels. C'est dans cette vérification qu'interviennent les premières "sommes de Riemann", traduction analytique de l'aire de figures simples susceptibles d'approcher $E(f)$, et qu'il est temps de définir de manière plus précise.

Admettons la formule donnant l'aire d'un rectangle comme produit des longueurs de sa base et de sa hauteur ainsi que le fait qu'une figure formée par un nombre fini de rectangles ne se recouvrant pas ait pour aire la somme des aires des rectangles constitutifs. Pour obtenir une valeur approchée de l'aire de $E(f)$, il est assez naturel de procéder comme suit. On divise $[a,b]$ en m intervalles (égaux ou non) $[a_{j-1},a_j]$ $(1 \leq j \leq m)$ où

$$a = a_0 < a_1 < a_2 < \dots < a_{m-1} < a_m = b$$

et l'on considère les rectangles de base $[a_{j-1},a_j]$ et de hauteur $f(x_j)$, où x_j est pris dans $[a_{j-1},a_j]$ $(1 \leq j \leq m)$. Pour un f suffisamment régulier, on peut s'attendre à ce que la somme des aires de ces rectangles, c'est-à-dire l'expression

$$\sum_{j=1}^{m} f(x_j)(a_j - a_{j-1}) \tag{1}$$

fournisse une approximation de l'aire de $S(f)$ d'autant meilleure que m sera plus grand. L'expression (1) est univoquement déterminée par la donnée de f sur $[a,b]$ et par celle de la partition pointée, ou <u>P-partition</u>

$$\Pi = \{(x_j,[a_{j-1},a_j]) \ : \ 1 \leq j \leq m\} ,$$

où l'on exige que $a = a_0 < a_1 < \dots < a_{m-1} < a_m = b$ et que $a_{j-1} \leq x_j \leq a_j$ $(1 \leq j \leq m)$. On peut donc noter (1) $S_f(\Pi)$ et l'appeler, suivant une tradition qui remonte à la fin du XIXe siècle, <u>somme de Riemann</u> associée à f et à Π, parce que Riemann en donna la formulation définitive.

Les noms de sommes d'Eudoxe, ou d'Archimède, ou d'Euler ou de Cauchy seraient également défendables.

A quelques exceptions près, où les propriétés géométriques de la figure suggèrent une approximation par des figures simples non rectangulaires, les sommes de Riemann rencontrées dans l'application de la méthode d'exhaustion correspondent à une division de $[a,b]$ en m parties égales et au choix $x_j = a_{j-1}$ $(1 \leq j \leq m)$ ou $x_j = a_j$ $(1 \leq j \leq m)$ (il y a une intéressante exception chez Fermat qui prendra les a_j en progression géométrique). Ce sont donc des expressions du type $S_f(\Pi_m)$ et $S_f(\Pi^m)$, où

$$\Pi_m = \left\{\left(a + \frac{(j-1)(b-a)}{m}, \left[a + \frac{(j-1)(b-a)}{m}, a + \frac{j(b-a)}{m}\right]\right) : 1 \leq j \leq m\right\}$$

et

$$\Pi^m = \left\{\left(a + \frac{j(b-a)}{m}, \left[a + \frac{(j-1)(b-a)}{m}, a + \frac{j(b-a)}{m}\right]\right) : 1 \leq j \leq m\right\} .$$

D'ailleurs, en décomposant si nécessaire la figure $E(f)$ en plusieurs partie, on se restreint au cas où f est monotone, disons croissante, sur $[a,b]$ ce que entraîne pour l'aire $A(f)$ de $E(f)$, quel que soit m, l'inégalité

(2) $$S_f(\Pi_m) \leq A(f) \leq S_f(\Pi^m) ,$$

et, pour la différence des sommes de Riemann, l'égalité

(3) $$S_f(\Pi^m) - S_f(\Pi_m) = m^{-1}(b-a)(f(b)-f(a)).$$

Supposons maintenant que nous ayons trouvé un réel J tel que

(4) $$S_f(\Pi_m) \leq J \leq S_f(\Pi^m)$$

quelque soit l'entier positif m ; alors, $A(f) = J$ car si, par exemple, $A(f) < J$, l'axiome d'Archimède entraîne l'existence d'un entier q tel que

$$q^{-1}(b - a)(f(b) - f(a)) < J - A(f),$$

et dès lors, par (2) et (3), on obtient la contradiction

$$J - A(f) \leq S_f(\Pi^q) - A(f) \leq S_f(\Pi^q) - S_f(\Pi_q) =$$

$$= q^{-1}(b - a)(f(b) - f(a)) < J - A(f) .$$

On exclut de même la possibilité $A(f) > J$. Telle est la méthode de démonstration par exhaustion et, pour pouvoir l'appliquer, il faut donc déterminer au préalable le réel J vérifiant (4). Dans le cas où $f(x) = (x-a)^k$ ($k \geqslant 1$, entier), on a

$$S_f(\Pi_m) = (b-a)^{k+1} m^{-(k+1)} \sum_{j=1}^{m-1} j^k ,$$

$$S_f(\Pi^m) = (b-a)^{k+1} m^{-(k+1)} \sum_{j=1}^{m} j^k ,$$

et $J = (b-a)^{k+1} L$ si L est tel que, pour tout m, on ait

$$m^{-(k+1)} \sum_{j=1}^{m-1} j^k \leqslant L \leqslant m^{-(k+1)} \sum_{j=1}^{m} j^k .$$

Lorsque $k = 1$, on calcule sans peine les expressions correspondantes, à savoir $(1/2)(1 - \frac{1}{m})$ et $(1/2)(1 + \frac{1}{m})$, et dès lors $L = 1/2$ convient, ce qui entraîne

$$A(f) = \frac{(b-a)^2}{2} .$$

Le problème se complique rapidement dès que k augmente et la formule générale

$$A(f) = \frac{(b-a)^{k+1}}{k+1}$$

sera conjecturée en 1635 par Cavalieri, qui la "démontre" pour $k \leqslant 4$ par sa méthode des indivisibles. Le cas général sera rigoureusement traité en 1636 par Fermat, dans un travail malheureusement non publié, à l'aide de la méthode d'exhaustion et de formules de récurrence sur k pour $\sum_{j=1}^{m} j^k$.

Dans la méthode des indivisibles de Cavalieri, on conçoit $E(f)$ comme la juxtaposition des droites parallèles à l'axe des ordonnées et joignant les points $(x,0)$ et $(x,f(x))$, et l'aire $A(f)$ comme la somme des ordonnées $f(x)$ entre a et b. Qu'une telle idée ait pu donner des résultats corrects provient de ce qu'on ne calcule pas directement des aires, mais des rapports d'aires. Dès lors, si g est une autre fonction définie sur $[a,b]$, on a, par exemple,

$$\frac{S_f(\Pi^m)}{S_g(\Pi^m)} = \frac{\sum_{j=1}^{m} f(a + \frac{j(b-a)}{m})}{\sum_{j=1}^{m} g(a + \frac{j(b-a)}{m})} .$$

ce qui fait, qu'en langage moderne, en passant à la limite sur m, on voit que les aires A(f) et A(g) sont entre elles comme la limite du rapport des sommes des ordonnées étendues aux points de subdivision de $[a,b]$. Citons à ce propos Pascal qui, en 1660, inspiré par la lecture des ouvrages de l'Anversois Tacquet, exprime, avec le style de polémiste qui l'a rendu célèbre par ailleurs, son sentiment sur les liens entre la méthode d'exhaustion et celle des indivisibles:

"Je ne ferai aucune difficulté d'user de cette expression la somme des ordonnées qui semble ne pas être géométrique à ceux qui n'entendent pas la doctrine des indivisibles, et qui s'imaginent que c'est pécher contre la géométrie que d'exprimer un plan par un nombre infini de lignes; ce qui ne vient que de leur manque d'intelligence puisqu'on n'entend autre chose par là sinon la somme d'un nombre indéfini de rectangles faits de chaque ordonnée avec de petites portions égales du diamètre, dont la somme est certainement un plan, qui ne diffère de l'espace du demi-cercle que d'une quantité moindre qu'aucune donnée".

La remarque de Pascal a trait au cas où E(f) est une portion de cercle, c'est-à-dire à $f(x) = (b^2-(x-a)^2)^{1/2}$. Les procédés que nous avons décrits conduisent également au calcul de A(f) pour $f(x) = x^c$ (c rationnel positif ou négatif), $f(x) = (x-a)^r(b-x)^s$ (r,s entiers positifs), sin x cos x, $\sin^r x$, $x^r \sin x$ (r entier positif), $(x-a)^2(b^2-x^2)^{1/2}$, etc. En langage moderne, tous ces résultats se ramènent au calcul explicite de la limite, pour m tendant vers l'infini, de $S_f(\Pi_m)$ ou $S_f(\Pi^m)$. Ces sommes de Riemann particulières servent donc au calcul de la valeur exacte d'une intégrale définie. On conçoit aisément qu'il s'agit là d'une méthode particulièrement ardue, que chaque nouvelle fonction f requiert de nouvelles astuces, d'autant plus que l'absence de formalisme rend bien difficile la découverte de structures communes dans des problèmes géométriques différents.

2. LE COURANT CINEMATIQUE OU L'AVENEMENT DU CALCUL D'INTEGRALES PAR LES PRIMITIVES

La cinématique a joué un rôle important dans l'évolution du calcul différentiel et intégral. Au XIVe siècle, rompant avec la tradition grecque, Suiseth (surnommé le Calculateur) à Oxford et Oresme, à Paris, réfléchissent sur le mouvement non uniforme. Ils commencent, ce qui est naturel, par considérer le cas où la vitesse varie uniformément, c'est-à-dire par le mouvement uniformément accéléré. Observant que, dans un mouvement unfforme (vitesse v constante), l'espace parcouru entre l'instant initial 0 et l'instant t, égal à vt, est représenté par l'aire du rectangle de hauteur v construit sur l'intervalle $[0,t]$, ils prolongent cette image au cas du mouvement uniformément accéléré. Si v est maintenant la vitesse atteinte à l'instant t dans un tel mouvement dont la vitesse est nulle en 0, la représentation graphique de la vitesse en fonction du temps est le segment de droite joignant l'origine des caordonnées au point (t,v). Admettant, sans dire pourquoi d'ailleurs, que l'espace parcouru à l'instant t est encore donné par l'aire de la figure située entre le graphe de la vitesse et l'axe des abscisses, ils obtiennent ainsi vt/2 dans le cas du mouvement uniformément accéléré.

En 1638, Galilée reprend de telles considérations dans ses "Discours et dissertations mathématiques relatifs à deux nouvelles sciences" lorsqu'il discute la chute des graves. Cela correspond à la loi, pour la vitesse, v(t) = gt, où g est une constante, c'est-à-dire au second cas considéré ci-dessus. Raisonnant de manière analogue, Galilée trouve que l'espace e(t) parcouru à l'instant t vaut $v(t)\frac{t}{2}$, c'est-à-dire $gt^2/2$. Torricelli, élève de Galilée, reprend cette approche en 1647 dans le cas où la vitesse est une fonction "quelconque" du temps. L'espace e(t) parcouru entre les instants 0 et t est encore donné par l'aire de la figure située entre le graphe de v(t) et le segment $[0,t]$ et, dans le cas où $v(t) = ct^2$, Torricelli montre de la sorte que $e(t) = ct^3/3$. Mais il considère alors la courbe donnée par le graphe de $e(t)$ comme la trajectoire d'un mobile soumis à des forces telles que la projection horizontale de sa vitesse soit constamment égale à 1. Si, lorsque le mobile est en $P = (t,e(t))$ on supprime l'action de la force, le mobile va décrire la tangente en P au graphe d'un mouvement uniforme avec la vitesse acquise à l'instant t, dont la projection horizontale est évidemment égale à 1. Si cette vitesse est changée de signe, le mobile décrit cette tangente dans l'autre sens et rencontre l'axe des ordonnées en un point $Q = (0,s)$ après un temps de parcours égal à t, puisque, dans son mouvement rectiligne uniforme, la composante horizontale de sa vitesse est toujours égale à -1 . Par conséquent, la vitesse en P du mobile est égale à $|QP|/t$ et sa composante verticale, qui vaut évidemment $v(t)$, à $|RQ|/t = |RQ|/|RP|$, où $R = (0,e(t))$. Dès lors $v(t)$ est égale à la tangente de l'angle $\widehat{RPQ}$, c'est-à-dire de l'angle que fait la tangente en P au graphe de $e(t)$ avec l'axe des abscisses.

Ainsi, Torricelli a entrevu le caractère inverse des opérations de quadrature (l'espace parcouru à l'instant t est donné par l'aire située entre $[0,t]$ et le graphe de la vitesse) et de tangente (la tangente en un point au graphe de l'espace parcouru fournit la vitesse à l'instant correspondant). Mais, faute d'un algorithme pour la détermination de la tangente à une courbe, c'est-à-dire faute d'un calcul différentiel, il ne voit dans ce résultat qu'un moyen de détermination de la tangente et non pas une méthode de calcul des aires. Ainsi, pour $v(t) = ct^2$, il sait que $e(t) = ct^3/3$, et dès lors la tangente à la courbe $(t,ct^3/3)$ s'obtiendra en traçant la diagonale du triangle rectangle de base 1 et de hauteur ct^2 construit sur le point $(t,ct^3/3)$.

On trouve des considérations très voisines chez Barrow mais avec une fortune bien différente : alors que Torricelli est emporté par la maladie en 1647 à 39 ans, et que ses travaux restent inédits jusqu'en 1919, Barrow, qui les a retrouvé indépendamment en 1660, a la chance d'avoir comme élève un certain Isaac Newton. Celui-ci va réaliser une synthèse magistrale de ce courant cinématique et y intégrer la composante analytique qui fait défaut chez ses prédécesseurs. Entre 1666 et 1671, quoique les publications correspondantes soient beaucoup plus tardives, Newton apporte les contributions essentielles suivantes :

a) une méthode analytique pour la détermination de la tangente à une courbe plane; celle-ci, qu'il appelle <u>fluente</u>, est supposée décrite en fonction du temps suivant la loi $(x(t),y(t))$ et avec la vitesse $(\dot{x}(t),\dot{y}(t))$ (qu'il appelle <u>fluxion</u>). Il s'agit de déterminer $\dot{y}(t)/\dot{x}(t)$ (c'est-à-dire la dérivée de y exprimée en fonction de x). Dans le cas d'une courbe algébrique, Newton retrouve le résultat donné auparavant par le Liégeois Sluse. En utilisant des développements en séries, tirés en particulier de sa série générale du binôme, Newton peut étendre son algorithme à des courbes transcendantes.

b) Ces développements en séries, combinés avec les résultats connus pour $A(f)$ lorsque $f(x) = x^k$ (k entier positif) le conduisent, par intégration terme à terme, à la détermination de $A(f)$ pour certaines fonctions f irrationnelles ou transcendantes.

c) Enfin, en considérant comme fonction $F(x)$ de x l'aire de la figure située entre le graphe de f, l'axe des abscisses et des parallèles à l'axe des ordonnées menées par $(a,0)$ et $(x,0)$, Newton montre, assez heuristiquement d'ailleurs, que le rapport des fluxions $\dot{F}(x)/\dot{x}$, c'est-à-dire notre dérivée $F'(x)$ est égale à $f(x)$. Il en conclut que la détermination de $A(f)$ peut se faire par le procédé suivant, que nous traduisons en langage actuel :

i) trouver une fonction F dont f soit la dérivée, c'est-à-dire trouver une <u>primitive</u> F de f sur $[a,b]$;

ii) calculer F(b) - F(a) .

Comme la lecture, de droite à gauche, des résultats obtenus par Newton pour la problème a) lui fournit une primitive F pour bon nombre de fonctions f (et en particulier pour toutes celles pour lesquelles A(f) avait été déterminée avec tant de peine par la méthode d'exhaustion ou des indivisibles), on voit que Newton ramène leur calcul à une soustraction. Il n'est pas surprenant qu'à partir de ce moment les sommes de Riemann vont céder la place à la notion d'intégrale considérée comme opération inverse de la dérivation.

3. LE COURANT ALGORITHMIQUE OU LE TRIOMPHE DE L'INTEGRALE COMME INVERSE DE LA DERIVEE

Indépendamment de Newton, un autre géant de la pensée, Leibniz, va proposer entre 1682 et 1693 sa propre synthèse du problème des tangentes et des aires et lui fournir des notations qui survivent encore aujourd'hui. Pour le problème des aires, son point de départ est, contrairement à Newton, proche de la notion de limite de sommes mais, dès que le lien avec l'opération inverse de la dérivation sera établi, c'est ce point de vue qui prévaudra.

Leibniz s'intéresse d'abord à des problèmes de différences finies. En utilisant un langage actuel, soit (a_k) une suite de nombres réels $(k = 0,1,\ldots)$. Si nous posons, pour $k = 1,2,\ldots$,

$$\Delta a_k = a_k - a_{k-1} ,$$

alors, trivialement,

$$\sum_{k=1}^{m} \Delta a_k = a_m - a_0$$

et

$$\Delta \left(\sum_{k=1}^{m} a_k\right) = a_m \qquad (m = 1,2,\ldots) .$$

Les opérations différence Δ et somme Σ apparaissent donc clairement comme inverses l'une de l'autre.

Dans notre problème test de détermination de A(f), divisons $[a,b]$ en m parties (égales ou non) par les points

$$a = a_0 < a_1 < \ldots < a_{m-1} < a_m = b .$$

La somme de Riemann

$$\sum_{j=1}^{m} f(a_j)(a_j - a_{j-1}) = \sum_{j=1}^{m} f(a_j)\Delta a_j$$

fournit une approximation de $A(f)$. D'autre part, au point $(a_j, f(a_j))$ du graphe, le quotient

$$\frac{f(a_j) - f(a_{j-1})}{a_j - a_{j-1}} = \frac{\Delta f(a_j)}{\Delta a_j}$$

fournit une approximation de la direction de la tangente au graphe en ce point. Si l'on pose

$$F(a_k) = \sum_{j=1}^{k} f(a_j)\Delta a_j \qquad (k = 1,2,\dots),$$

alors

$$\Delta F(a_k) = f(a_k)\Delta a_k ,$$

c'est-à-dire

$$\frac{\Delta F(a_k)}{\Delta a_k} = f(a_k) \qquad (k = 1,2,\dots) . \tag{5}$$

Leibniz "extrapole" alors ces résultats au cas où les différences Δa_j deviennent "infiniment petites". La somme de Riemann devient alors l'aire $A(f)$ et, dans un manuscrit de 1675 et un article de 1692, Leibniz la note

$$\int f(x)dx ,$$

tandis que, si l'on note $F(x)$ l'aire située entre l'axe des abscisses, le graphe de f et les parallèles à l'axe des ordonnées passant par $(a,0)$ et $(x,0)$, la relation (5) devient,

$$\frac{dF}{dx}(x) = f(x) ,$$

et redonne le résultat de Newton, que Leibniz écrit

$$d\int f(x)dx = f(x)dx .$$

Enfin, en extrapolant la relation évidente

$$\sum_{j=1}^{m} \frac{\Delta f(a_j)}{\Delta a_j}\Delta a_j = \sum_{j=1}^{m} \Delta f(a_j) = f(b) - f(a),$$

au cas où les Δa_j deviennent infiniment petits, on obtient avec nos notations,

$$A(\frac{df}{dx}) = f(b) - f(a) .$$

L'attitude de Leibniz quant à la nature de ses infiniment petits varie d'écrit en écrit, et il s'y révèle plus proche du diplomate que du mathématicien. Aussi partisans et adversaires de l'analyse non standard, c'est-à-dire d'une analyse fondée sur une notion rigoureuse d'infiniment petit, peuvent-ils de nos jours y puiser tous deux de péremptoires arguments. Notons enfin que Leibniz appelle $\int f(x)dx$ la somme de f et que ce sont les frères Bernoulli qui introduiront vers 1690 le terme <u>intégrale</u>, adopté alors également par Leibniz, pour ce que nous appelons aujourd'hui primitive. Le changement de terminologie des Bernoulli traduit d'ailleurs leur ralliement au point de vue de Newton : le calcul intégral étudie le passage de la différentielle (ou de la dérivée) à la fonction, de la partie au tout: "Les intégrales des différentielles sont ces quantités dont ces différentielles proviennent par différentiation" (Jean Bernoulli, "Lectiones mathematicae de methodo integralium", 1691), "Le calcul intégral est la méthode par laquelle, à partir d'une relation entre les différentielles, on retrouve la relation entre les quantités elles-mêmes" (Euler, "Institutiones Calculi integralis", 1768). Cette tradition se maintiendra jusqu'au début du XXe siècle. Lebesgue, dont nous reparlerons, n'écrit-il pas, dans l'introduction à sa thèse : "Le problème fondamental du calcul intégral, à savoir trouver une fonction connaissant sa dérivée,..." ("Intégrale, longueur, aire", 1902). Dans cette optique, l'<u>intégrale définie</u> de f sur [a,b] (c'est-à-dire notre A(f) ci-dessus) est, par définition, l'expression

$$\int_a^b f(x)dx = F(b) - F(a) ,$$

où F est une primitive de f. La notation de gauche fut introduite en 1816 seulement par Fourier et popularisée par Cauchy. Avant cela, Euler a utilisé la notation

$$\int f(x)dx \begin{bmatrix} ab & x=a \\ ad & x=b \end{bmatrix}$$

en parlant d'intégration étendue de la valeur a à la valeur b. Dans un mémoire de Laplace de 1779, on trouve la phrase : "Je nomme intégrale définie une intégrale prise depuis une valeur déterminée jusqu'à une autre valeur déterminée", ainsi que l'expression "limites de l'intégration".

4. LE COURANT NUMERIQUE OU DU CALCUL APPROCHE D'INTEGRALES PAR DES SOMMES DE RIEMANN AU PROBLEME DE LEUR EXISTENCE

La fin de la section précédente pourrait faire croire à la disparition complète des sommes de Riemann du calcul intégral. Il n'en est rien et Euler ("Institutiones de calculi integralis", 1768), après six premiers chapitres consacrés au calcul des primitives, leur consacre deux chapitres qui ne manquent pas d'intérêt. Mais son but est uniquement de les utiliser comme moyen de calcul approché d'intégrales qui échappent aux méthodes des chapitres précédents. Il ne s'agit plus de les utiliser, comme avant Newton et Leibniz, pour des calculs exacts d'intégrales, ni, comme plus tard chez Cauchy et Riemann, pour démontrer un théorème d'existence ou fonder une définition. Les sommes de Riemann considérées par Euler sont du type de celles apparues, dans un langage plus vague, chez Leibniz, à savoir

$$(6) \qquad \sum_{j=1}^{m} f(a_{j-1})(a_j - a_{j-1}) ,$$

où $a = a_0 < a_1 < \dots < a_{m-1} < a_m = b$. Et s'il ne calcule pas une estimation générale de l'erreur commise en remplaçant l'intégrale par (6), Euler discute différentes situations dans lesquelles l'approximation peut-être rendue aussi bonne que l'on veut en augmentant m, par exemple lorsque f est croissante. Il remarque aussi que l'expression (6) ou la quantité

$$\sum_{j=1}^{m} f(a_j)(a_j - a_{j-1})$$

donnent des valeurs d'autant plus exactes de l'intégrale que les différences $(a_j - a_{j-1})$ sont plus petites, à condition que les différences $f(a_j) - f(a_{j-1})$ deviennent elles aussi petites; dans le cas contraire, la détermination sera extrêmement incertaine. Euler fait alors l'intéressante remarque suivante, sur laquelle nous reviendrons beaucoup plus tard, et dont les notations seules sont modernisées : "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 $a_1 - a$, $a_2 - a_1$, ... 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."

L'approche d'Euler, avec des notations et une terminologie petit à petit modernisées, sera reprise dans le grand "Traité de calcul différentiel et intégral" de Lacroix (1798, 1819), tandis que Legendre et Poisson chercheront à déterminer les bornes de l'erreur commise en remplaçant les intégrales définies par les sommes de Riemann ci-dessus.

A la même époque, des situations nouvelles apparaissent où la notion d'intégrale comme inverse de la dérivée se révèle insuffisante, à savoir les théories naissantes des séries de Fourier et d'intégration des fonctions d'une variable complexe. En 1811, Fourier étudie le développement en séries trigonométriques de fonctions présentant des discontinuités de type saut (des fonctions en escalier entre autres), et il est aisé de constater qu'une fonction en escalier non constante sur un intervalle n'est pas primitivable sur cet intervalle. Le calcul des coefficients de Fourier de telles fonctions ne pourra donc se faire au moyen d'intégrales définies par la variation d'une primitive. C'est pourquoi Fourier revient à l'ancienne notion d'intégrale comme aire A(f), dont l'existence n'est pas discutée puisqu'il écrit avec optimisme, dans sa "Théorie analytique de la chaleur", présentée en 1811 et publiée en 1822 : "Quelle que soit la fonction f(x) ou la forme de la courbe qu'elle représente, l'intégrale a une valeur définie qui peut être introduite dans la formule (de calcul des coefficients de Fourier) . Les valeurs de ces intégrales définies sont analogues à celles de l'aire totale située entre la courbe et l'axe dans un intervalle donné, ou aux valeurs de quantités mécaniques, comme les ordonnées du centre de gravité de cette aire ou d'un solide quelconque. Il est évident que toutes ces quantités ont des valeurs assignables, que les figures des corps soient régulières ou non, ou que nous leur donnions une forme entièrement arbitraire.".

De la même manière, l'extrapolation pure et simple aux fonctions d'une variable complexe de la formule fondamentale du calcul intégral

$$\int_a^b f'(z)dz = f(b) - f(a)$$

où le premier membre désigne maintenant l'intégration suivant un chemin du plan complexe joignant a à b, conduit, si f n'est pas holomorphe, à des paradoxes qu'on imagine facilement et que Poisson et Cauchy chercheront à expliquer.

Toutes ces difficultés soulevées par l'approche traditionnelle du calcul intégral, ainsi que des considérations pédagogiques, conduisent Cauchy à un changement radical d'attitude face aux fondements du calcul intégral, qui se manifeste en particulier en 1823 dans un mémoire du Journal de l'Ecole polytechnique et dans son "Résumé des leçons données à l'Ecole royale polytechnique sur le calcul infinitésimal". Dans le mémoire, il écrit : "Je considère chaque intégrale comme étant juste la somme des valeurs infiniment petites de l'expression différentielle placée sous le signe intégrale, qui correspond aux différentes valeurs de la variable incluse entre les limites en question. Quand on adopte cette manière de regarder l'intégrale définie, on prouve aisément qu'une telle intégrale a une valeur unique et finie lorsque, les deux limites de la variable étant finies, les intégrands restent finis et continus entre ces limites". Il ajoute : "Il me semble que cette manière de regarder une intégrale définie devrait être adoptée de préférence, comme je l'ai fait, parce qu'elle vaut également pour tous les cas, même ceux dans lesquels nous ne pouvons pas passer généralement de la fonction sous le signe intégral à la fonction primitive". Dans l'Introduction de son "Résumé", il poursuit : "Dans le calcul intégral, il m'a paru nécessaire de démontrer généralement l'existence des intégrales ou fonctions primitives avant de faire connaître leurs diverses propriétés. Pour y parvenir, il a fallu d'abord établir la notion d'intégrales prises entre des limites données ou intégrales définies. Ces dernières pouvant être quelquefois infinies ou indéterminées, il était essentiel de rechercher dans quel cas elles conservent une valeur unique et finie".

Ce programme, Cauchy le réalise grâce à la notion de limite qu'il a mise à la base de son enseignement de l'analyse. Rompant avec la tradition, il se pose clairement le problème de l'existence des quantités que l'on veut calculer (exactement ou approximativement), avant de s'intéresser à ce calcul. Et ce sont en général des méthodes connues de calcul approché que Cauchy va transformer en méthodes de démonstration d'existence. Dans le cas de l'intégrale, Cauchy procède comme suit :

a) il démontre (21e leçon : intégrales définies) que si f est continue sur $[a,b]$, il existe un réel J tel que toutes les sommes de Riemann

$$\sum_{j=1}^{m} f(a_{j-1})(a_j - a_{j-1})$$

convergent vers J lorsque $\max\limits_{1 \leq j \leq m} (a_j - a_{j-1})$ tend vers zéro. Sa démonstration reste cependant incomplète à nos yeux puisque la notion de continuité uniforme sur $[a,b]$ n'est pas clairement dégagée et que la partie "suffisante" du critère de Cauchy est considérée comme évidente en l'absence d'une théorie des nombres réels.

b) il remarque (22e leçon : formules pour la détermination des valeurs exactes ou approchées des intégrales définies) que, pour f continue sur $[a,b]$, J est également la limite de $S_f(\Pi)$, où Π est cette fois une P-partition quelconque, lorsque $\max_{1 \leq j \leq m} (a_j - a_{j-1})$ tend vers zéro, indépendamment du choix des $x_j \in [a_{j-1}, a_j]$. Il discute également des exemples de détermination exacte de l'intégrale par des limites de sommes de Riemann lorsqu'on prend les a_j en progression arithmétique ou en progression géométrique, et renoue ainsi avec la tradition de la méthode d'exhaustion.

c) Il introduit (24e leçon : des intégrales définies dont les valeurs sont infinies ou indéterminées) la notion d'intégrale impropre ou généralisée par la formule

$$\int_a^b f(x)dx = \lim_{\substack{c \to a+ \\ d \to b-}} \int_c^d f(x)dx ,$$

lorsque $\int_c^d f(x)dx$ existe, au sens précédent, pour tout $a < c < d < b$, ainsi que la limite du membre de droite. Par additivité, il peut ainsi traiter rigoureusement l'intégrale des fonctions ayant un nombre fini de points de discontinuité.

d) il définit (26e leçon : intégrales indéfinies) pour une fonction f continue sur $[a,b]$ l'intégrale indéfinie de f comme étant la fonction définie sur $[a,b]$ par

$$x \mapsto \int_a^x f(t)dt \quad ,$$

et il prouve que cette fonction a, en chaque point x de $[a,b]$ une dérivée égale à f(x) . C'est donc une primitive de f sur $[a,b]$ et Cauchy fournit donc la première démonstration de l'existence de la primitive d'une fonction continue. Avant lui, l'existence de cette primitive était admise a priori ou résultait a posteriori du résultat du calcul lorsque celui-ci était possible. Cauchy démontre également que si F est une primitive quelconque de f sur $[a,b]$, alors

$$\int_a^b f(x)dx = F(b) - F(a) .$$

Notons pour la petite histoire que l'Ecole Polytechnique n'apprécie pas outre mesure l'enseignement de Cauchy. Ainsi, Arago, à la fin de juillet 1823, déclare que "cinq leçons sur des considérations générales sur l'intégration, cela peut être convenable à la faculté des sciences, mais non à l'école polytechnique, où les élèves sont pressés par le temps". Cauchy résistera jusqu'en 1825, mais, au procès-verbal de la séance du conseil d'instruction du 24 novembre de cette année, on peut lire : "M. Cauchy annonce que pour se conformer au voeu du conseil, il ne s'attachera plus à donner, comme il l'a fait jusqu'à présent, des démonstrations parfaitement rigoureuses".

Et ce voeu sera respecté par les grands traités d'analyse qui suivront, en France, celui de Cauchy, puisqu'il faudra attendre le "Cours d'analyse" de Jordan (et surtout sa 2e édition de 1893), et l'"Introduction à la théorie des fonctions d'une variable" de Tannery en 1886 pour renouer avec la tradition de Cauchy, avec un langage nettement affiné sous l'influence de l'école de Weierstrass.

5. L'INTEGRALE DE RIEMANN OU LES ESPOIRS ET DECEPTIONS DE LA PREMIERE DEFINITION DE L'INTEGRALE COMME LIMITE DE SOMMES

En 1854, pour sa thèse d'habilitation à l'université de Göttingen, Riemann présente un travail intitulé : "Sur la possibilité de représenter une fonction par une série trigonométrique", qui ne sera malheureusement publié qu'en 1867 par Dedekind, après la mort prématurée de Riemann. Comme l'écrit Darboux, d'ordinaire beaucoup moins tendre, "Ce mémoire de Riemann est un chef-d'oeuvre semblable à ces vieux tableaux dont quelques parties en pleine lumière nous font regretter ce que le temps a détruit ou ce que l'auteur a négligé". Après trois paragraphes consacrés à un historique des séries de Fourier, Riemann, dont la modestie n'a d'égale que le génie, débute le quatrième paragraphe par ces mots : "L'incertitude qui règne encore sur quelques points fondamentaux de la théorie des intégrales définies nous oblige à placer ici quelques remarques sur la notion de l'intégrale définie, et sur la généralité dont elle est susceptible. Et d'abord que doit-on entendre par $\int_a^b f(x)dx$?".

Ce sont ces "quelques remarques", s'étendant sur moins de sept pages, qui constituent la première théorie de l'intégration dans l'histoire des mathématiques !

Riemann part du schéma de Cauchy, celui qui utilise les sommes de Riemann générales. Avant lui, quel que soit le concept d'intégrale adopté, on avait, au mieux, cherché à voir si une fonction ou une classe de fonctions possédait une intégrale. Riemann, au contraire, comme l'écrit si bien Lebesgue, "porte son attention sur le procédé opératoire qui permet, dans le cas des fonctions continues, de calculer l'intégrale avec une telle approximation que l'on veut, et il se demande dans quels cas ce procédé appliqué à des fonctions discontinues donne un nombre déterminé. Cauchy n'appliquait son procédé de définition de l'intégrale qu'à des fonctions considérées a priori comme intéressantes : les fonctions continues; maintenant, au contraire, toute fonction sera intéressante à laquelle s'appliquera le procédé de définition". C'est, à l'époque, une attitude vraiment révolutionnaire, et c'est la rançon de son succès dans tant d'autres domaines de nous apparaître presque banale aujourd'hui.

Aux notations près, Riemann écrit : "La valeur de la somme $S_f(\Pi)$ dépendra du choix des a_j et des x_j. Si elle a la propriété , de quelque manière que les a_j et les x_j puissent être choisis, de s'approcher indéfiniment d'une limite fixe J quand les $a_j - a_{j-1}$ tendent tous vers zéro, cette limite s'appelle la valeur de l'intégrale définie $\int_a^b f(x)dx$". Avec les notations introduites plus haut, et en notant

$$p(\Pi) = \max_{1 \leq j \leq m} (a_j - a_{j-1})$$

le pas de la P-partition Π, la première définition d'intégrabilité s'énonce comme suit :

<u>Définition</u>. f, définie sur [a,b] , y est intégrable au sens de Riemann ou <u>R-intégrable</u> s'il existe un réel J tel que, pour chaque $\varepsilon > 0$, on puisse trouver un $\delta > 0$ tel que, pour toute P-partition Π de [a,b] vérifiant la condition $p(\Pi) \leqslant \delta$, on ait

$$|S_f(\Pi) - J| \leqslant \varepsilon .$$

Un tel J, s'il existe, est unique, et on le note $\int_a^b f(x)dx$.

Riemann continue alors comme suit : "Recherchons maintenant l'étendue et la limite de la définition précédente et posons-nous cette question : dans quel cas une fonction est-elle susceptible d'intégration ? Dans quel cas ne l'est-elle pas ?" . Comme la définition de Riemann est calquée sur le procédé élaboré par Cauchy pour des fonctions (uniformément) continues, on pourrait s'attendre à ce qu'elle délimite une classe de fonctions "pas trop" discontinues. Riemann donne alors un exemple de fonction R-intégrable qui possède un ensemble dense de points de discontinuité, ce qui conduit du Bois-Reymond à penser que Riemann a poussé le concept d'intégrale jusqu'à ses limites extrêmes. On va aussitôt chercher à "mesurer" l'ensemble des points de discontinuité d'une fonction R-intégrable, mais il faudra attendre Lebesgue, dont nous reparlerons, pour démontrer que du Bois-Reymond se trompe, en élaborant une théorie de la mesure des ensembles de réels permettant de caractériser les fonctions R-intégrables comme fonctions dont l'ensemble des points de discontinuité a une mesure nulle. Ce résultat ne retire rien au rôle historique important joué par l'intégrale de Riemann dans la naissance de la théorie de la mesure.

Riemann n'a pas abordé le problème du lien entre primitivation et R-intégrabilité. Il faut attendre 1875 pour que Darboux et du Bois-Reymond généralisent dans ce cadre les résultats de Cauchy sous la forme suivante :

<u>Théorème</u>. a) Si f est R-intégrable sur [a,b] , alors, en chaque point x où f est continue, son intégrale indéfinie est dérivable et a pour dérivée f(x).

b) si f est R-intégrable et primitivable sur $[a,b]$, alors, pour toute primitive F de f sur $[a,b]$, on a

$$\int_a^b f(x)dx = F(b) - F(a) .$$

Le résultat b) soulève immédiatement la question de savoir si la définition de Riemann recouvre toutes les fonctions à laquelle s'applique la définition d'intégrale définie comme variation de primitive, puisque ses hypothèses requièrent à la fois la R-intégrabilité et la primitivabilité. Volterra fournit en 1881 une réponse négative en produisant un exemple de fonction primitivable mais non R-intégrable sur un intervalle. Le prestige de l'intégrale de Riemann commence à être ébranlé, ainsi que la confiance qu'on avait mise dans la portée de sa définition. Borel en 1909 explique avec beaucoup d'à propos d'où vient la faiblesse de l'intégrale de Riemann : "Un caractère important de cette définition 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".

L'insuccès de l'intégrale de Riemann face aux fonctions primitivables provient de ce que, contrairement à la continuité, la primitivabilité de f sur $[a,b]$ n'implique pas sa primitivabilité uniforme. Notons enfin que si f est R-intégrable sur $[a,b]$, elle y est nécessairement bornée. Il faudra donc appliquer l'extension décrite en c) au paragraphe 4 pour donner un sens à l'intégrale de fonctions non bornées.

6. L'INTEGRALE DE LEBESGUE OU UN NOUVEL ART DE PARTITIONNER

Dans sa note de 1901 aux Comptes Rendus de l'Académie des Sciences de Paris qui annonce une partie des résultats de sa thèse "Intégrale, longueur, aire" de 1902, Lebesgue résume très bien la situation antérieure : "Dans le cas des fonctions continues, il y a identité entre les notions d'intégrale et de fonction primitive. Riemann a défini l'intégrale de certaines fonctions discontinues, mais toutes les fonctions dérivées ne sont pas intégrables au sens de Riemann. Le problème des fonctions primitives n'est donc pas résolu par l'intégration, et l'on peut désirer une définition de l'intégrale comprenant comme cas particulier celle de Riemann et permettant de résoudre le problème des fonctions primitives".

C'est une contribution essentielle, quoique non complète, à ce problème que Lebesgue va apporter en introduisant une nouvelle définition de l'intégrale qui, souligne ironiquement Borel en poursuivant son analogie vestimentaire : "est en quelque sorte faite sur mesure et s'adapte exactement aux propriétés de chaque fonction particulière". Dans le cas de notre problème test, l'idée de Lebesgue consiste à approcher E(f) en partitionnant l'intervalle de variation de f, c'est-à-dire $\left[\inf_{[a,b]} f \ , \ \sup_{[a,b]} f\right]$ par

$$\inf_{[a,b]} f = b_0 < b_1 < \dots < b_{m-1} < b_m = \sup_{[a,b]} f \ .$$

Soit A_j l'ensemble des x dans $[a,b]$ tels que

$$b_{j-1} \leq f(x) < b_j$$

Si A est une partie de $[a,b]$, Lebesgue définit sa mesure extérieure $\mu_e(A)$ comme l'infimum des nombres

$$\sum_{i=1}^{\infty} (d_i - c_i)$$

pour toutes les familles dénombrables $\{[c_i,d_i] \ : \ i = 1,2,\dots\}$ d'intervalles tels que $A \subset \bigcup_{i=1}^{\infty} [c_i,d_i]$, et sa mesure intérieure $\mu_i(A)$ par $\mu_i(A) = (b-a) - \mu_e([a,b] \setminus A)$. A sera dit <u>mesurable</u> si $\mu_e(A) = \mu_i(A)$, la valeur commune étant alors sa mesure $\mu(A)$. Pour une fonction f bornée sur

$[a,b]$, on dira alors que f est intégrable au sens de Lebesgue ou <u>L-intégrable</u> sur $[a,b]$ si, pour tout c et d tels que $a \leq c < d \leq b$, l'ensemble des x dans $[a,b]$ tels que

$$c \leq f(x) < d$$

est mesurable. S'il en est ainsi, alors, pour toute partition de l'intervalle de variation de f du type considéré plus haut et tout $y_j \in [b_{j-1}, b_j]$ $(1 \leq j \leq m)$, on a

$$(\inf_{[a,b]} f)(b-a) \leq \sum_{j=1}^{m} b_{j-1}\mu(A_j) \leq \sum_{j=1}^{m} y_j\,\mu(A_j) \leq$$

$$\leq \sum_{j=1}^{m} b_j\,\mu(A_j) \leq (\sup_{[a,b]} f)(b-a) ,$$

et

$$\sum_{j=1}^{m} (b_j - b_{j-1})\mu(A_j) \leq (b-a) \max_{1 \leq j \leq m} (b_j - b_{j-1}).$$

Par conséquent, il existera un réel unique J tel que, pour chaque $\varepsilon > 0$, on puisse trouver un $\delta > 0$ ayant la propriété suivante : pour chaque partition

$$\inf_{[a,b]} f = b_0 < b_1 < \dots < b_{m-1} < b_m = \sup_{[a,b]} f$$

de $[\inf_{[a,b]} f , \sup_{[a,b]} f]$ telle que $\max_{1 \leq j \leq m} (b_j - b_{j-1}) \leq \delta$, et pour chaque $y_j \in [b_{j-1}, b_j]$ $(1 \leq j \leq m)$, on a

$$\left| \sum_{j=1}^{m} y_j\,\mu(A_j) - J \right| \leq \varepsilon.$$

Ce J s'appelle l'intégrale de Lebesgue de f sur $[a,b]$. Comme on peut démontrer que toute fonction R-intégrable sur $[a,b]$ y est L-intégrable avec la même intégrale, on peut encore écrire $J = \int_a^b f(x)dx$. On voit que l'intégrale peut encore s'obtenir comme limite d'une somme mais la construction d'une telle somme est moins directe que celle d'une somme de Riemann. Lebesgue peut alors démontrer les résultats suivants :

a) si f est L-intégrable sur $[a,b]$, son intégrale indéfinie est continue en chaque point de $[a,b]$ et dérivable, avec f(x) comme dérivée, en chaque point x où f est continue.

b) si f est primitivable et bornée sur $[a,b]$, alors f est L-intégrable sur $[a,b]$ et, pour toute primitive F de f sur $[a,b]$, on a

(7) $$\int_a^b f(x)dx = F(b) - F(a) .$$

Lebesgue étend alors son intégrale aux fonctions non nécessairement bornées, mais dans ce cas, b) subsiste seulement sous la forme suivante :

b') si f est primitivable et L-intégrable sur $[a,b]$, alors, pour toute primitive F de f sur $[a,b]$, on a (7), avec l'information supplémentaire :

b") f, primitivable sur $[a,b]$, y est L-intégrable si et seulement si ses primitives sont à variation bornée sur $[a,b]$.

Il donne alors l'exemple de

$$f(x) = 2x \sin(1/x^2) - (2/x)\cos(1/x^2) \quad \text{si} \quad x \neq 0 ,$$
$$= 0 \text{ si } x = 0 ,$$

qui a comme primitive la fonction

$$F(x) = x^2 \sin (1/x^2) \quad \text{si} \quad x \neq 0 ,$$
$$= 0 \text{ si } x = 0 ;$$

comme F n'est pas à variation bornée sur $[0,1]$, f est primitivable mais non L-intégrable sur $[0,1]$. L'intégrale de Lebesgue ne contient donc pas l'intégrale définie par (7) pour les fonctions primitivables non bornées.

En 1903, Lebesgue donne la généralisation suivante de a) :

a') si f est L-intégrable sur $[a,b]$, il existe $A \subset [a,b]$ tel que $\mu(A) = 0$ et tel que l'intégrale indéfinie de f ait pour dérivée f(x) en chaque point $x \in [a,b] \setminus A$.

Il est moins connu qu'en 1909, dans un mémoire intitulé "Sur les intégrales singulières", Lebesgue s'est posé le problème de l'obtention de son intégrale comme limite de sommes de Riemann usuelles. Il prouve, pour une fonction f L-intégrable sur $[a,b]$, l'existence d'une suite de P-partitions telle que la suite des sommes de Riemann correspondantes converge vers $\int_a^b f(x)dx$. Ce résultat est amélioré par Hahn en 1914, qui démontre que, pour toute suite de partitions

$$\{ a = a_0^k < a_1^k < \dots < a_{m_k-1}^k < a_{m_k}^k = b \}$$

de $[a,b]$ telle que $\max\limits_{1 \leq j \leq m_k} (a_j^k - a_{j-1}^k) \to 0$ si $k \to \infty$ tend vers l'infini, il existe une suite de points intermédiaires

$$\{ x_1^k, x_2^k, \dots , x_{m_k}^k \}$$

avec $x_j^k \in [a_{j-1}^k, a_j^k]$ $(1 \leq j \leq m_k)$ telle que si $\Pi^{(k)}$ est la P-partition correspondante pour chaque $k = 1,2,\dots,$ on a

$$\int_a^b f(x)dx = \lim_{k \to \infty} S_f(\Pi^{(k)}) .$$

7. L'INTEGRALE DE DENJOY-PERRON OU LA SYNTHESE DU CALCUL DES PRIMITIVES ET DU CALCUL INTEGRAL

En 1912, dans une note aux Comptes Rendus de l'Académie des Sciences de Paris qui sera suivie d'une longue liste de travaux, Denjoy introduit une généralisation de l'intégrale de Lebesgue, qu'il appelle la totale de f sur $[a,b]$, et qui intègre toutes les fonctions primitivables : toute fonction f primitivable sur $[a,b]$ sera totalisable et, pour toute primitive F de f sur $[a,b]$, on aura la formule (7). Mais la définition de Denjoy est extrêmement compliquée puisqu'elle repose sur un long et délicat processus d'induction transfinie à partir d'intégrales de Lebesgue qu'il est hors de propos d'expliciter ici.

En 1914, dans une note aux Comptes Rendus de l'Académie des Sciences de Heidelberg, et qui sera son seul travail consacré à l'intégrale, Perron va obtenir le même résultat par une voie complètement différente qui, au lieu de se rattacher comme Denjoy, au courant d'idées qui va de Cauchy à Lebesgue, va prendre ses racines dans la définition de l'intégrale comme variation de primitive. L'idée de Perron consiste a remplacer l'égalité

$$F'(x) = f(x)$$

par le couple d'inégalités

$$F'(x) \geqslant f(x),\ F'(x) \leqslant f(x) .$$

et à abandonner l'exigence, qui conduit aux seules fonctions primitivables, que la même fonction F vérifie les deux inégalités. Erigeant en définition des notions auxiliaires que le Belge de la Vallée-Poussin avait introduites dans l'exposé de l'intégrale de Lebesgue contenu dans son "Cours d'Analyse infinitésimale" de 1913, Perron définit les notions de sur-fonction F_+ et de sous-fonction F_- de f sur $[a,b]$, qu'il serait peut-être plus éclairant d'appeler respectivement sur-primitive et sous-primitive de f sur $[a,b]$. Afin d'éviter l'emploi de dérivées de Dini comme chez Perron, nous nous limiterons, sans perte de généralité comme l'a montré Tolstov en 1940, au cas où F_+ et F_- sont supposées dérivables et nous dirons que F_+ (resp. F_-) , fonction dérivable sur $[a,b]$, est une <u>sur-primitive</u> (resp. <u>sous-primitive</u>) de f sur $[a,b]$ si $F_+(a) = 0$ (resp. $F_-(a) = 0$) et si, pour tout $x \in [a,b]$, on a

$$F'_+(x) \geqslant f(x) \quad (\text{resp. } F'_-(x) \leqslant f(x)) .$$

On notera que de telles fonctions existent toujours si f est bornée sur $[a,b]$.

La fonction f sera alors dite intégrable au sens de Perron ou <u>P-intégrable</u> sur $[a,b]$ si elle possède au moins une sur-primitive et une sous-primitive sur $[a,b]$ et si

(8) $$\sup_{F_-} F_-(b) = \inf_{F_+} F_+(b) ,$$

le supremum (resp. infimum) étant pris sur toutes les sous-primitives (resp. sur-primitives) de f sur $[a,b]$. S'il en est ainsi, la valeur commune de (8) sera l'intégrale de Perron de f sur $[a,b]$. Il est presque immédiat que toute fonction f primitivable sur $[a,b]$ sera P-intégrable sur $[a,b]$ avec $F(b) - F(a)$ comme intégrale pour toute primitive F de f sur $[a,b]$. Perron montre alors que toutes les fonctions bornées et L-intégrables sur $[a,b]$ y sont P-intégrables avec la même intégrale, un résultat étendu en 1915 par Bauer aux fonctions L-intégrables non bornées.

Entre 1921 et 1925, les travaux de Hake, Alexandroff et Looman établiront l'équivalence des intégrales de Denjoy et de Perron, qui s'appelleront dès lors intégrales de Denjoy-Perron.

8. L'INTEGRALE DE KURZWEIL-HENSTOCK OU LE SURPRENANT RETOUR DES SOMMES DE RIEMANN

Les théories de l'intégration esquissées aux paragraphes 6 et 7 ont été développées dans de nombreuses directions pendant la première moitié du XXe siècle, qui a vu apparaître en particulier pas mal d'autres définitions des intégrales de Lebesgue ou de Denjoy-Perron dans lesquelles les sommes de Riemann continuent à jouer un rôle plus qu'effacé.

Mais, en 1957, dans un article du Journal tchécoslovaque de mathématiques consacré à un problème lié aux équations différentielles, Kurzweil propose une définition particulièrement simple de l'intégrale de Denjoy-Perron basée sur l'emploi des sommes de Riemann. Pour introduire cette définition dans le cas de notre problème test, notons tout d'abord que la définition de Riemann est trivialement équivalente à la suivante, qui fait intervenir explicitement les x_j dans l'expression de la finesse de la P-partition : f est R-intégrable sur $[a,b]$ s'il existe un réel J ayant la propriété suivante : pour chaque $\varepsilon > 0$, il existe un $\delta > 0$ tel que pour toute P-partition Π vérifiant la condition

$$[a_{j-1}, a_j] \subset [x_j - \delta, x_j + \delta] \quad (1 \leq j \leq m) ,$$

on a $|S_f(\Pi) - J| \leq \varepsilon$. Si l'on se souvient maintenant de la remarque d'Euler relative aux points où la variation de f est violente, ou si l'on veut répondre à l'objection de Borel sur l'excès d'uniformité dans la définition de Riemann, on peut, pour chaque $\varepsilon > 0$ fixé, remplacer l'exigence de l'existence d'un "pas constant" δ par l'exigence moins forte de la seule existence d'un "pas variable" qui pourrait mieux tenir compte des caractéristiques propres de f. On est ainsi conduit à la définition suivante de Kurzweil :

<u>Définition</u>. On dit que f est intégrable au sens de Kurzweil sur $[a,b]$ s'il existe un réel J ayant la propriété sui-

vante : pour chaque $\varepsilon > 0$, il existe une fonction δ strictement positive sur $[a,b]$ telle que pour toute P-partition Π de $[a,b]$ vérifiant la condition

$$[a_{j-1},a_j] \subset [x_j - \delta(x_j),\ x_j + \delta(x_j)] \quad (1 \leq j \leq m),$$

on a $|S_f(\Pi) - J| \leq \varepsilon$.

Un tel J est nécessairement unique et comme Kurzweil établit l'équivalence de sa définition avec celle de Perron, et donc celle de Denjoy, la valeur de l'intégrale étant la même, on peut lui conserver la notation habituelle.

En 1961, dans un article des "Proceedings" de la Société Mathématique de Londres, Henstock arrive indépendamment, et à partir de motivations différentes, à une définition similaire et il démontre que f est L-intégrable sur $[a,b]$ si et seulement si f et $|f|$ sont intégrables sur $[a,b]$ au sens de Kurzweil-Henstock. Il montre aussi directement que cette intégrale contient toutes les intégrales de Riemann impropres, qu'elles soient absolument convergentes ou non. Ainsi, l'intégrabilité au sens de Riemann apparaît comme une intégrabilité uniforme au sens de Kurzweil-Henstock tandis que l'intégrabilité au sens de Lebesgue apparaît comme une intégrabilité absolue au sens de Kurzweil-Henstock. Il est intéressant également de signaler une caractérisation voisine des fonctions L-intégrables en termes de sommes de Riemann donnée par McShane dans un Mémoire de la Société Mathématique Américaine de 1969 : pour obtenir la L-intégrabilité, il suffit d'abandonner, dans la définition de Kurzweil-Henstock, la condition $x_j \in [a_{j-1},a_j]$ $(1 \leq j \leq m)$ dans la définition d'une P-partition et de considérer toutes les P-partitions telles que $x_j \in [a,b]$ $(1 \leq j \leq m)$. Notons qu'une telle modification introduite dans la définition de Riemann laisse inchangée la classe des fonctions qui la vérifient.

Les sommes de Riemann de f définissent une fonction réelle S_f sur l'ensemble $\mathcal{P}$ des P-partitions de $[a,b]$. La famille $\mathbb{F}_R$ de parties de $\mathcal{P}$ définie par $\mathbb{F}_R = \{\mathcal{P}_R(\delta) : \delta > 0\}$ où $\mathcal{P}_R(\delta)$ est l'ensemble des P-partitions $\Pi = \{(x_j, [a_{j-1},a_j]) :$

$1 \leq j \leq m \}$ telles que $[a_{j-1}, a_j] \subset [x_j - \delta, x_j + \delta]$ $(1 \leq j \leq m)$ est non vide, ne contient pas la partie vide et est telle que l'intersection de deux éléments de $\mathbb{F}_R$ contient un élément de $\mathbb{F}_R$ puisque $\mathcal{P}_R(\min(\delta_1, \delta_2)) \subset \mathcal{P}_R(\delta_1) \cap \mathcal{P}_R(\delta_2)$. C'est donc une direction ou base de filtre dans $\mathcal{P}$, et la définition de R-intégrabilité de f sur $[a,b]$ équivaut à l'existence d'une limite suivant $\mathbb{F}_R$ pour S_f. Considérons maintenant la famille $\mathbb{F}_K$ de parties de $\mathcal{P}$ définie par $\mathbb{F}_K = \{\mathcal{P}_K(\delta) : \delta$ est une application strictement positive définie sur $[a,b]\}$, où $\mathcal{P}_K(\delta)$ est l'ensemble des P-partitions $\Pi = \{(x_j, [a_{j-1}, a_j]) : 1 \leq j \leq m\}$ telles que

$$[a_{j-1}, a_j] \subset [x_j - \delta(x_j),\ x_j + \delta(x_j)] \quad (1 \leq j \leq m) .$$

Il n'est pas a priori évident que $\mathcal{P}_K(\delta)$ soit non vide pour chaque δ , mais c'est un résultat en fait équivalent au lemme de Borel-Lebesgue. Il est alors facile de montrer que $\mathbb{F}_K$ est également une base de filtre sur $\mathcal{P}$ et que S_f a une limite suivant $\mathbb{F}_K$ si et seulement si f est intégrable sur $[a,b]$ au sens de Kurzweil-Henstock. L'approche de Kurzweil-Henstock a donc conservé les sommes de Riemann traditionnelles mais a modifié le type de passage à la limite, qui n'est plus uniforme, mais peut être adapté à la fonction f.

A titre d'exemple, considérons la célèbre fonction de Dirichlet définie par $f(x) = 1$ si x est rationnel et $f(x) = 0$ si x est irrationnel. Cette fonction est partout discontinue et n'est donc R-intégrable sur aucun intervalle borné. Elle est par contre L-intégrable sur tout intervalle borné et y a une intégrale nulle. Pour prouver directement ce résultat à l'aide de la définition de Kurzweil-Henstock, notons tout d'abord que, pour une P-partition quelconque Π de $[a,b]$, $S_f(\Pi)$ ne contient que les termes où x_j est rationnel et se ramène à la somme des $a_j - a_{j-1}$ pour ces x_j. Il n'y a donc que ces intervalles qui devront être pris suffisamment petits pour que la somme de Riemann constitue une bonne approximation de 0. L'ensemble A des rationnels contenus dans $[a,b]$ est dénombrable et peut donc s'écrire $A = \{ y_1, y_2, \ldots\}$. Si $\varepsilon > 0$ est donné, définissons comme suit l'application correspondante : $\delta(x) = 1$ si x est rationnel et, si x est irrationnel, il existe un unique y_k tel que $x = y_k$, et l'on pose $\delta(x) = \varepsilon/2^{k+1}$. Il est alors facile de vérifier que si Π est une P-partition de $[a,b]$ vérifiant les conditions de la définition de Kurzweil-Henstock pour ce δ, on a $|S_f(\Pi)| = S_f(\Pi) \leq \varepsilon$, ce qui démontre le résultat.

CONCLUSION

Nous espérons avoir montré que les sommes de Riemann n'ont guère cessé de jouer un rôle important dans l'évolution des fondements du calcul intégral et que leur disgrâce s'est à chaque fois avérée passagère. D'abord moyen de calcul exact de l'aire lorsque celle-ci était considérée comme notion première, elles gardent un rôle discrèt mais indispensable pour le calcul approché d'intégrales lorsque triomphe la conception de l'intégrale comme inverse de la dérivée. Elles reviennent au premier plan chez Cauchy lorsqu'il s'agit de démontrer l'existence de l'aire ou celle de la primitive d'une fonction continue, et sont chez Riemann l'ingrédient essentiel de la première définition d'une fonction intégrable. Elles paraissent s'éclipser lors du triomphe de l'intégration selon Lebesgue et dans les généralisations de Denjoy et de Perron. Mais c'est pour fournir une définition étonnamment simple de ces intégrales qu'elles se sont récemment rappelées à notre souvenir. Auraient-elles comme devise celle de la ville de Paris ?

OUVRAGES CONSULTES

M.E. BARON, "The Origins of the Infinitesimal Calculus", Pergamon, Oxford, 1969.

N. BOURBAKI. "Eléments d'histoire des mathématiques", Hermann, Paris , 1969.

P.S. BULLEN, Non-absolute integrals : a survey, Real Analysis Exchange 5 (1980) 195-259.

G. CASTELNUOVO, "Le origini del calcolo infinitesimale nell' era moderna, Feltrinelli, Milano, 1962.

A. CAUCHY, "Equations différentielles ordinaires", cours inédit, fragment, Etudes vivantes, Johnson, New York, 1981.

J. DHOMBRES, "Nombre, mesure et continu. Epistémologie et histoire", Cedic-Fernand Nathan, Paris, 1978.

C.H. EDWARDS, Jr., "The Historical Development of the Calculus", Springer, New York, 1979.

J. GRABINER, "The Origins of Cauchy's Rigorous Calculus", MIT Press, Cambridge, Mass., 1981.

Th. HAWKINS, "Lebesgue's Theory of Integration. Its Origin and Development", Chelsea, New York, 1975.

H. LEBESGUE, "Leçons sur l'intégration et la recherche des fonctions primitives", Chelsea, New York, 1973.

J. MAWHIN, "Introduction à l'analyse", Cabay, Louvain-la-Neuve, 2e édition, 1981.

M. METIVIER, P. COSTABEL et P. DUGAC ed., "Siméon-Denis Poisson et la science de son temps", Ecole Polytechnique, Paris, 1981.

I.N. PESIN, "Classical and Modern Integration Theories", Academic Press, New York, 1970.

D. VAN DALEN and A.F. MONNA, "Sets and integration. An outline of the development", Wolters-Noordhoff, Groningen, 1972.