SUR LE DÉVELOPPEMENT DE LA NOTION D'INTÉGRALE [1]

Messieurs,

Laissant de côté tous les développements techniques, nous allons examiner ensemble les modifications successives, les enrichissements de la notion d'intégrale et comment sont apparues d'autres notions utilisées dans les recherches récentes sur les fonctions de variables réelles.

Avant Cauchy, il n'y avait pas de définition de l'intégrale, au sens actuel du mot définition. On se bornait à dire quelles étaient les aires qu'il fallait additionner ou soustraire pour obtenir l'intégrale $\int_a^b f(x)dx$.

Pour Cauchy, une définition est nécessaire ; car, avec lui, apparaît le souci de rigueur qui est la caractéristique des mathématiques modernes. Cauchy définit, à peu près comme nous le faisons maintenant, les fonctions continues et les intégrales de ces fonctions. Pour arriver à l'intégrale de $f(x)$, il lui suffit de former les sommes

$$S = \Sigma f(\xi_i)\,(x_{i+1} - x_i), \qquad (1)$$

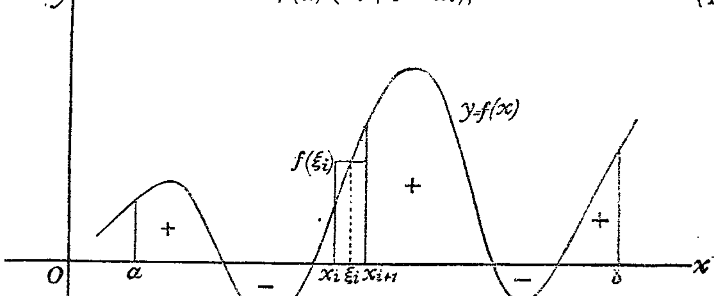

Fig. 1.

1. Conférence faite à Copenhague, le 8 mai 1926, à la Société Mathématique et communiquée par M. Bonnesen, directeur de la *Matematisk Tidsskrift*.

que les arpenteurs et les mathématiciens de tous les siècles ont utilisées pour le calcul approché des aires, et d'en déduire $\int_a^b f(x)dx$ par un passage à la limite.

Seulement, tandis que la légitimité d'un tel passage était évidente pour qui partait de la notion d'aire, Cauchy doit démontrer que les S tendent bien vers une limite dans les conditions qu'il envisage. Une nécessité analogue s'impose chaque fois qu'on remplace par une définition purement logique l'emploi d'une notion expérimentale. Il faut ajouter que l'intérêt de l'être défini n'est plus évident, il ne peut ressortir que de l'étude des propriétés de cet être.

C'est la rançon du progrès logique. Celui qu'a fait faire Cauchy est si considérable qu'il a une ampleur en quelque sorte philosophique. On dit souvent que Descartes a ramené la géométrie à l'algèbre ; je dirais plus volontiers que, par l'emploi des coordonnées, il a ramené toutes les géométries à celle de la droite et que celle-ci, en nous dotant des notions : continu et nombre irrationnel, a permis à l'algèbre d'atteindre à sa portée actuelle.

Pour que la réduction des géométries à celle de la droite fût achevée, il restait pourtant à éliminer un certain nombre de notions relatives aux géométries à plusieurs dimensions telles que longueur d'une courbe, aire d'une surface, volume d'un corps. C'est précisément là le progrès qu'a réalisé Cauchy. Après lui, il a suffi que les Arithméticiens construisent le continu linéaire à l'aide du nombre entier pour que l'arithmétisation de la science soit effectuée.

Et, maintenant, devons-nous nous borner à faire de l'Analyse ? Non. Certes, tout ce que nous ferons pourra se traduire dans le langage arithmétique, mais si l'on renonçait à avoir des vues directes, géométriques, intuitives, si l'on était réduit à la pure logique qui ne permet pas de choisir entre tout ce qui est exact, on ne penserait guère à bien des questions, et certaines notions, la plupart de celles que nous allons examiner aujourd'hui, par exemple, nous échapperaient complètement.

Depuis longtemps on avait intégré certaines fonctions discontinues ; la définition de Cauchy s'appliquait encore à ces intégrales ; aussi était-il naturel de rechercher, comme l'a fait Riemann, la portée exacte de cette définition.

Si $\underline{f_i}$ et $\overline{f_i}$ désignent les bornes inférieure et supérieure de $f(x)$ dans (x_i, x_{i+1}), S est comprise entre

$$\underline{S} = \Sigma \underline{f_i}\,(x_{i+1} - x_i) \text{ et } \overline{S} = \Sigma \overline{f_i}\,(x_{i+1} - x_i).$$

Riemann montre qu'il suffit que

$$\overline{S} - \underline{S} = \Sigma\,(\overline{f_i} - \underline{f_i})\,(x_{i+1} - x_i)$$

tende vers zéro pour une suite particulière de divisions de (a, b) en intervalles (x_i, x_{i+1}) de plus en plus petits pour que la définition de Cauchy puisse être employée. Darboux ajoute que les passages à la limite habituels effectués sur $\underline{S}$ et $\overline{S}$ donnent toujours deux nombres déterminés $\int_{\underline{a}}^{b} f(x)dx$, $\int_a^{\overline{b}} f(x)dx$; en général différents, égaux seulement lorsque l'intégrale de Cauchy-Riemann existe.

Du point de vue logique, voici des définitions très naturelles, n'est-ce pas ? Pourtant, on peut dire qu'elles n'ont pratiquement servi à rien. Celle de Riemann, en particulier, a l'inconvénient de ne s'appliquer que rarement et, en quelque sorte, par hasard.

C'est qu'en effet, s'il est bien évident que le morcellement de (a, b) en intervalles (x_i, x_{i+1}) de plus en plus petits rendra les différences $\overline{f_i} - \underline{f_i}$ de plus en plus petites si $f(x)$ est continue, en vertu de cette continuité même, s'il est clair que ce morcellement fera tendre encore $\overline{S} - \underline{S}$ vers zéro s'il n'y a que quelques points de discontinuité, on n'a aucune raison d'espérer qu'il en sera de même pour une fonction discontinue partout. Alors, en effet, prendre des intervalles (x_i, x_{i+1}) de plus en plus petits, c'est-à-dire des valeurs de $f(x)$ relatives à des valeurs de x de plus en plus voisines, ne garantit nullement que l'on prend des valeurs de $f(x)$ de moins en moins différentes.

Laissons-nous donc guider par le but à atteindre : réunir, grouper les valeurs peu différentes de $f(x)$. Il est clair alors que nous devons morceler, non pas (a, b), mais l'intervalle $(\underline{f}, \overline{f})$ limité par les bornes inférieure et supérieure de $f(x)$ dans (a, b). Faisons-le à l'aide de nombres y_i distants entre eux de moins de ε ; nous sommes conduits, par exemple, à considérer les valeurs de $f(x)$ définies par

$$y_i \leqq f(x) \leqq y_{i+1}.$$ (Voir fig. 2, p. 152.)

Les valeurs correspondantes de x forment un ensemble E_i ; dans le cas de la figure, cet ensemble E_i est constitué par quatre intervalles ; avec certaines fonctions $f(x)$ continues il pourrait être formé d'une infinité d'intervalles ; avec une fonction quelconque il pourrait être très compliqué. Mais peu importe ; c'est cet ensemble E_i qui joue le rôle analogue à celui de l'intervalle (x_i, x_{i+1}) dans

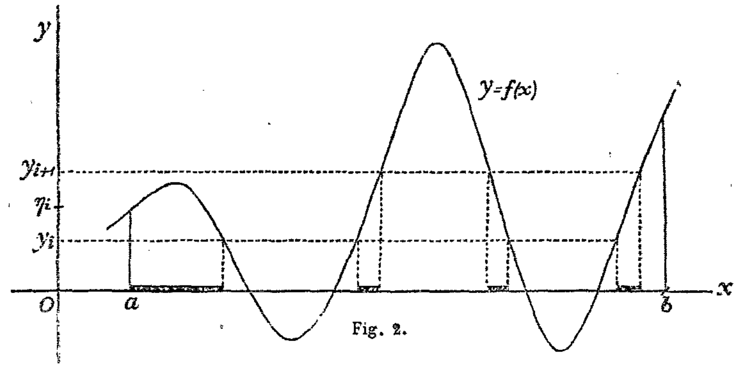

Fig. 2.

la définition de l'intégrale des fonctions continues, puisqu'il nous fait connaître des valeurs de x donnant à $f(x)$ des valeurs peu différentes.

Si η_i est un nombre quelconque pris entre y_i et y_{i+1},

$$y_i \leqq \eta_i \leqq y_{i+1},$$

les valeurs de $f(x)$ pour les points de E_i diffèrent de η_i de moins de ε. η_i va jouer le rôle que jouait $f(\xi_i)$ dans la formule (1) ; quant au rôle de la longueur ou mesure $x_{i+1} - x_i$ de l'intervalle (x_i, x_{i+1}) il sera joué par une mesure $m(E_i)$ que nous attacherons dans un instant à l'ensemble E_i. Nous formerons ainsi la somme

$$S = \Sigma \eta_i m(E_i). \qquad (2)$$

Mais, auparavant, regardons bien ce que nous venons de faire et, pour le bien comprendre, répétons-le en d'autres termes.

Les Géomètres du XVII^e siècle considéraient l'intégrale de $f(x)$, — le mot intégrale n'était pas inventé, mais peu importe, — comme la somme d'une infinité d'indivisibles dont chacun était l'ordonnée, positive ou négative, de $f(x)$. Eh bien ! nous avons tout simplement groupé les indivisibles de grandeur comparable ; nous avons, comme on dit en algèbre, fait la réunion, la réduction des termes semblables. On peut dire encore que, avec le procédé de Riemann, on essayait de sommer les indivisibles en les prenant dans l'ordre où ils étaient fournis par la variation de x, on opérait donc comme le ferait un commerçant sans méthode qui compterait pièces et billets au hasard de l'ordre où ils lui tomberaient sous la main; tandis que nous opérons comme le commerçant méthodique qui dit :

j'ai $m(E_1)$ pièces de 1 couronne valant $1 \cdot m(E_1)$,
j'ai $m(E_2)$ pièces de 2 couronnes valant $2 \cdot m(E_2)$,
j'ai $m(E_3)$ billets de 5 couronnes valant $5 \cdot m(E_2)$,

etc., j'ai donc en tout :

$$S = 1 \cdot m(E_1) + 2 \cdot m(E_2) + 5 \cdot m(E_3) + \dots\,.$$

Les deux procédés conduiront, certes, le commerçant au même résultat parce que, si riche qu'il soit, il n'a qu'un nombre fini de billets à compter; mais, pour nous, qui avons à additionner une infinité d'indivisibles, la différence entre les deux façons de faire est capitale.

Occupons-nous maintenant de la définition du nombre $m(E_i)$ attaché à E_i. L'analogie entre cette mesure et une longueur, ou même avec un nombre de billets, nous conduit naturellement à dire que, dans l'exemple de la figure (2), $m(E_i)$ sera la somme des longueurs des quatre intervalles constituant E_i; que, dans un exemple où E_i serait formé d'une infinité d'intervalles, $m(E_i)$ serait la somme des longueurs de tous ces intervalles; et, dans le cas général, à procéder comme il suit : enfermons E_i dans des intervalles, en nombre fini ou en infinité dénombrable; soient $l_1, l_2, \dots$ les longueurs de ces intervalles. Nous désirons évidemment que l'on ait :

$$m(E_i) \leq l_1 + l_2 + \dots\,.$$

Si nous cherchons la borne inférieure du second membre pour tous les systèmes possibles d'intervalles pouvant servir à enfermer E_i, cette borne sera donc une limite supérieure de $m(E_i)$. Pour cette raison nous la notons $\overline{m(E_i)}$ et nous avons

$$m(E_i) \leq \overline{m(E_i)}. \tag{3}$$

Si C est l'ensemble des points de (a, b) ne faisant pas partie de E_i, on a de même

$$m(C) \leq \overline{m(C)}.$$

Or, on désire évidemment avoir

$$m(E_i) + m(C) = m[(a, b)] = b - a;$$

donc on doit avoir

$$m(E^i) \geq b - a - \overline{m(C)}. \qquad (4)$$

Les inégalités (3) et (4) donnent donc des limites supérieure et inférieure de $m(E_i)$. On voit facilement que ces deux inégalités ne sont jamais contradictoires. Lorsque les limites inférieure et supérieure de E_i sont égales, $m(E_i)$ est définie, nous dirons alors que E_i est mesurable [1].

Une fonction $f(x)$, pour laquelle les ensembles E_i sont mesurables quels que soient les y_i, est dite mesurable. Pour une telle fonction, la formule (2) définit une somme S. On démontre facilement que, lorsqu'on fait varier le choix des y_i de façon que ε tende vers zéro, les S tendent vers une limite déterminée qui est, par définition, $\int_a^b f(x)dx$ [2].

Cette première extension de la notion d'intégrale définie en entraîne bien d'autres. Supposons qu'il s'agisse d'intégrer une fonction $f(x,y)$ de deux variables, nous procéderons exactement comme précédemment : nous lui attacherons des ensembles E_i qui seront maintenant des ensembles de points dans un plan et non plus de points sur une droite. A ces ensembles ce sera une mesure superficielle qu'il faudra maintenant attribuer; cette mesure se déduira de l'aire des rectangles

$$\alpha \leq x \leq \beta, \quad \gamma \leq y \leq \delta,$$

tout à fait de la même manière que la mesure linéaire se déduisait de la longueur des intervalles. La mesure définie, la formule (2) donnera des sommes S d'où l'intégrale se déduira par un passage à la limite.

La définition que nous avons envisagée s'étend donc immédiate-

1. Le mode de définition de la mesure des ensembles utilisé ici est celui de C. Jordan (*Cours d'Analyse de l'École Polytechnique, tome I*) ; mais avec cette modification, essentielle pour notre but, que nous enfermons l'ensemble E_i à mesurer dans des intervalles dont le nombre peut être infini, alors que C. Jordan employait toujours seulement un nombre fini d'intervalles. Cet emploi de l'infinité dénombrable à la place du nombre entier est suggéré par les travaux de M. Borel qui, d'ailleurs, avait utilisé lui-même cette idée en particulier pour une définition de la mesure. (*Leçons sur la théorie des fonctions.*)

2. *Comptes-rendus de l'Acad. des Sc.*, t. CXXIX, 1909.

Des définitions équivalentes à celle du texte ont été proposées par divers Auteurs. Les plus intéressantes sont dues à M. W.-H. Young (*Philos. Trans. of the Royal Soc. of London*, 1905 ; *Proc. of the London Math. Soc.*, 1910). Voir aussi, par exemple, des notes de M. Borel et de M. F. Riesz (*Comptes-rendus*, 1912).

ment aux fonctions de plusieurs variables; en voici une autre qui s'appliquerait également quel que soit le nombre des variables et que j'expose seulement dans le cas où il s'agit d'intégrer $f(x)$ dans (a, b).

J'ai dit qu'il s'agissait alors de faire la somme des indivisibles représentés par les diverses ordonnées des points x, $y=f(x)$; nous avons, il y a un instant, groupé ces indivisibles d'après leur grandeur. Bornons-nous maintenant à les grouper d'après leur signe; nous aurons à considérer l'ensemble superficiel des points de celles de ces ordonnées qui sont positives, E_p, et l'ensemble, E_n, des points des ordonnées négatives. Pour le cas simple où $f(x)$ est continue, on posait, avant Cauchy, ainsi que je l'ai rappelé en commençant,

$$\int_a^b f(x)\,dx = \text{aire}\,(E_p) - \text{aire}\,(E_n)\,;$$

ceci nous conduit à poser

$$\int_a^b f(x)dx = m\,(E_p) - m_s(E_n)\,;$$

m_s désignant une mesure superficielle. Cette nouvelle définition est équivalente à la précédente; elle nous ramène à la méthode intuitive d'avant Cauchy, mais la définition de la mesure lui a donné un solide fondement logique.

Nous savons donc, et de deux manières, définir l'intégrale d'une fonction d'une ou de plusieurs variables et cela sans avoir eu à considérer la forme plus ou moins compliquée du domaine d'intégration, car ce domaine D n'intervenait qu'en ceci : les ensembles E_i, dans notre première définition, les ensembles E_p et E_n, dans la deuxième, étaient formés en ne prenant les valeurs de la fonction f que pour les points de D.

Puisque le choix du domaine d'intégration D n'intervient que dans la formation des E_i, ou des E_p et E_n, il est clair que nous pourrions tout aussi bien convenir de former ces ensembles E_i, E_n, E_p en ne tenant compte que des valeurs prises par f aux points d'un ensemble E donné, et nous aurions ainsi défini l'intégrale de f étendue à l'ensemble E.

Pour préciser la portée de cette nouvelle extension de la notion d'intégrale, rappelons-nous que nos définitions exigent que f soit mesurable, c'est-à-dire que les E_i soient mesurables, pour la première définition, que E_p et E_n le soient, pour la deuxième, et cela entraîne que E soit aussi mesurable. Nous savons donc définir l'intégrale étendue à un ensemble mesurable d'une fonction mesurable dans cet ensemble, et bornée; j'ai, en effet, supposé implicitement jusqu'ici qu'il s'agissait de fonctions bornées.

Qu'y aurait-il de changé dans le premier mode de définition si la fonction f à intégrer était non bornée? L'intervalle $(\underline{f}, \overline{f})$ ne serait plus fini; il faudrait donc une infinité de nombres y_i pour le diviser en intervalles de longueur au plus égale à ϵ, il y aurait donc une infinité d'ensembles E_i et la somme S de la formule (2) serait maintenant une série. Pour ne pas être arrêté dès le début, il nous faut supposer que la série S est convergente pour le premier choix de nombres y_i que nous aurons fait; or, si S existe pour un choix des y_i, il existe pour tous les choix des y_i et la définition de l'intégrale s'applique sans modifications.

On a donné le nom de fonctions sommables à toutes les fonctions que l'on peut intégrer par les procédés indiqués, c'est-à-dire à toutes les fonctions mesurables pour lesquelles les sommes S ont un sens. Toute fonction mesurable bornée est sommable : et comme on n'a pas réussi jusqu'ici à nommer une fonction non mesurable, on peut dire que, jusqu'ici, pratiquement toute fonction bornée a une intégrale. Au contraire, il existe des fonctions non bornées très simples qui ne sont pas sommables; aussi ne doit-on pas s'étonner que notre notion d'intégrale se révèle encore insuffisante dans certaines questions.

Nous venons d'étendre la notion d'intégrale à des fonctions non bornées en partant de la première de nos définitions; la seconde conduit au même résultat. Mais il faut pour cela élargir la notion de mesure de façon à ce qu'elle s'applique non seulement aux ensembles bornés, que nous avons seuls considérés jusqu'ici, mais aussi à des ensembles de points s'étendant jusqu'à l'infini. Je ne signale cette seconde façon de procéder que par ce qu'elle est aussi en rapport avec une autre extension de l'intégrale définie dans laquelle l'intervalle, le domaine, l'ensemble auquel est étendue l'intégrale n'est plus supposé fini, comme nous l'avons fait jusqu'ici, mais peut aller jusqu'à l'infini.

Je me borne à cette indication, car je n'aurai pas à envisager, dans ce qui suit, cette extension de la notion d'intégrale. C'est pour la même raison que je me contente de signaler rapidement les recherches, pourtant si originales, entreprises par un jeune homme tué à la guerre, M. Gateaux, qui s'était proposé de définir l'opération d'intégration pour les fonctions d'une infinité de variables. Ces recherches, qui ont été prolongées par M. Paul Lévy et par M. Norbert Wiener, ne sont pas sans rapport avec les études axiomatiques entreprises par M. M. Fréchet et par M. P.-J. Daniell dans le but d'étendre la notion d'intégrale aux ensembles abstraits[1]. M. Fréchet et M. Daniell ne se sont d'ailleurs pas proposé seulement d'appliquer aux ensembles abstraits les définitions dont j'ai parlé jusqu'ici, mais aussi une autre extension de l'intégrale définie à laquelle nous conduira bientôt la notion d'intégrale indéfinie que nous allons maintenant examiner.

On appelle ordinairement intégrale indéfinie d'une fonction $f(x)$ la fonction $F(x)$ définie par

$$F(x) = C^{te} + \int_a^x f(x)\,dx\,; \tag{5}$$

ne conservons pas cette dénomination et redonnons aux mots intégrale indéfinie leur sens primitif. Primitivement, les deux dénominations « intégrale définie » et « intégrale indéfinie » s'appliquaient à la même expression $\int_a^b f(x)\,dx$; mais l'intégrale était dite définie lorsqu'il s'agissait d'un intervalle (a,b) donné, déterminé, défini, et l'intégrale était indéfinie lorsque (a,b) était variable, non déterminé, non défini ou, si l'on veut dire, indéfini.

C'est, en somme, par un véritable abus de langage qu'on appelle $F(x)$ l'intégrale indéfinie de $f(x)$; si nous remarquons, de plus, que lorsqu'on étudie $F(x)$ c'est toujours pour obtenir des propriétés de $\int_a^b f(x)dx$, que c'est au fond $\int_a^b f(x)dx$ que l'on étudie à travers $F(x)$, on sera conduit à dire : j'appelle intégrale indéfinie de $f(x)$ la fonction $\Phi(a,b)$:

1. R. Gateaux, *Bull. Soc. Math. de France*, 1919.
P. Lévy, *Leçons d'Analyse fonctionnelle*, 1922.
N. Wiener, *Proc. of the London Math. Soc.*, 1922.
M. Fréchet, *Bull. Soc. Math. de France*, 1915.
P.-J. Daniell, *Ann. of Math.*, 1918 et 1919.

$$\Phi(a, b) = \int_a^b f(x)\,dx = F(b) - F(a). \qquad (5\ bis)$$

Il y a entre une intégrale indéfinie et l'intégrale définie correspondante les mêmes relations et les mêmes différences qu'entre une fonction et une valeur particulière prise par cette fonction ; au reste, si nous représentions par D l'intervalle (a, b) d'intégration, nous pourrions dire que l'intégrale indéfinie est une fonction dont l'argument est le domaine D,

$$\Psi(D) = \Phi(a, b).$$

De ces réflexions, il résulte clairement que, relativement à une fonction de deux variables $f(x,y)$, on ne doit pas prendre pour intégrale indéfinie, comme on l'a fait quelquefois, la fonction

$$F(X, Y) = c_1(x) + c_2(y) + \int_\alpha^X \int_\beta^Y f(x, y)\,dx\,dy. \qquad (6)$$

Si l'on se bornait à la considération des domaines rectangulaires

$$a \leq x \leq b; \quad c \leq y \leq d,$$

on devrait prendre pour intégrale indéfinie la fonction de quatre variables

$$\Phi(a, b;\ c, d) = F(b, d) + F(a, c) - F(a, d) - F(b, c); \quad (7)$$

mais, si l'on veut considérer tous les domaines d'intégration, comme le domaine le plus général ne peut être déterminé par un nombre fini de paramètres, si grand que soit ce nombre, force nous est de renoncer aux fonctions ordinaires pour représenter la correspondance entre un domaine D et l'intégrale étendue à ce domaine et d'étudier directement la fonction

$$\Psi(D) = \iint_D f(x, y)\,dx\,dy,$$

dont l'argument D est un domaine. C'est cette fonction que nous appellerons l'intégrale indéfinie de $f(x,y)$; ou plutôt, puisque nous avons défini aussi l'intégrale de f étendue à un ensemble mesurable E, nous considérerons l'intégrale indéfinie comme une fonction d'ensemble qui serait définie pour tous les ensembles mesurables[1].

Dans tout ce qui précède, il n'y a certes que des questions de

1. *Annales scientifiques de l'École Normale supérieure*, 1910.

langage, de dénomination; mais ces questions de dénomination ne se seraient pas posées si nous n'avions pas acquis une notion nouvelle. C'est pourquoi on ne doit pas s'étonner que le nouveau langage ait permis de donner toute la portée possible à des faits aperçus tout d'abord à l'occasion de la fonction $F(x)$ de la formule (5). On a réussi, en particulier, à caractériser les fonctions d'ensemble qui sont des intégrales indéfinies par deux propriétés : l'additivité complète et l'absolue continuité[1].

Quand une fonction d'ensemble $\Psi(E)$ jouit de ces deux propriétés, elle est l'intégrale indéfinie d'une fonction f qui dépend de 1, 2, 3, ... variables suivant que les ensembles E sont formés à l'aide de points d'une droite, d'un plan, de l'espace ordinaire, etc. Pour avoir un langage et une notation uniformes, disons que f est une fonction de point, $f(P)$, et écrivons :

$$(\Psi E) = \int_E f(P)\, dm(E). \tag{8}$$

La fonction $f(P)$ est entièrement déterminée par $\Psi(E)$, à ceci près qu'on peut modifier arbitrairement f aux points d'un ensemble arbitraire de mesure nulle sans qu'elle cesse d'avoir $\Psi(E)$ pour intégrale indéfinie. Et l'on peut obtenir $f(P)$ à partir de $\Psi(E)$, sauf aux points d'un ensemble de mesure nulle, par le procédé suivant :

Soit P le point en lequel on veut calculer f; prenons pour domaine d'intégration Δ un intervalle de centre P, ou un cercle de centre P, ou une sphère de centre P, ... suivant qu'il s'agit du cas de la droite, du plan, de l'espace, ... et formons le rapport $\dfrac{\Psi(\Delta)}{m(\Delta)}$. Puis, faisons tendre Δ vers zéro; nous aurons :

$$\lim_{\Delta \to 0} \frac{\Psi(\Delta)}{m(\Delta)} = f(P). \tag{9}$$

Ce résultat généralise évidemment le théorème classique d'après

1. Ces dénominations sont dues respectivement à M. de la Vallée-Poussin (Intégrales de Lebesgue, Fonctions d'ensemble, Classes de Baire) et à M. G. Vitali (*R. Acc. d. Sc. di Torino*, 1908).

Une fonction d'ensemble mesurable $\Psi(E)$ est complètement additive si, de quelque manière qu'on partage E en des ensembles mesurables E_1, E_2, ..., en nombre fini ou dénombrable et sans points communs deux à deux, on a :

$$\Psi(E) = \Psi(E_1) + \Psi(E_2) + \dots$$

Une fonction d'ensemble mesurable $\Psi(E)$ est absolument continue si, quand E varie de façon que $m(E)$ tende vers zéro, $\Psi(E)$ tend aussi vers zéro.

lequel, si $f(x)$ est continue, la fonction $F(x)$ de la formule (5) admet f pour dérivée; notre procédé de calcul de $f(P)$ est bien, en effet, une sorte de dérivation de la fonction d'ensemble $\Psi(E)$.

Ce mode de dérivation avait été considéré, il y a fort longtemps. Cauchy[1] appelle grandeurs coexistantes des grandeurs déterminées en même temps, c'est-à-dire par les mêmes conditions. Si l'on a, par exemple, un corps non homogène, et comme composition et comme densité, et si l'on considère un domaine D de ce corps, le volume de D, la masse de D, la quantité de chaleur nécessaire pour élever de 1° la température de D supposé isolé, sont des grandeurs coexistantes. Ce sont des fonctions de domaine $V(D)$, $M(D)$, $Q(D)$.

Ce n'est pas par un heureux hasard que nous arrivons ici à des fonctions de domaine. Si l'on y réfléchit, on s'apercevra vite que toute grandeur de la physique est relative non à un point, mais à un corps étendu, que c'est une fonction de domaine; du moins tant qu'il s'agit de grandeurs directement mesurables. Le corps à considérer ne sera toutefois pas toujours un corps de notre espace sensible, ce pourra être un corps d'un espace de conception purement mathématique si, dans la détermination de la grandeur envisagée, interviennent des variables non spatiales comme le temps, la température, etc. Mais peu importe; les grandeurs directement mesurables, une masse, une quantité de chaleur, une quantité d'électricité, par exemple, sont des fonctions de domaine et non des fonctions de point.

La physique considère cependant aussi des grandeurs attachées à des points, comme une vitesse, une tension, une densité, une chaleur spécifique; mais ce sont des grandeurs dérivées et que l'on définit justement le plus souvent par le rapport ou la limite du rapport de deux grandeurs coexistantes :

$$\text{densité} = \frac{\text{masse}}{\text{volume}}; \quad \text{chaleur spécifique} = \frac{\text{quantité de chaleur}}{\text{masse}};$$

c'est-à-dire en prenant la dérivée d'une grandeur par rapport à une grandeur coexistante.

Ainsi la physique, et par suite la géométrie, conduit à la considération des fonctions de domaine et à leur dérivation tout aussi bien que l'analyse des fonctions de variables réelles. Même les

1. *Exercices d'Analyse et de Physique mathématique*, tome 2, p. 188-229.

fonctions de domaine ont, en physique, un rôle en quelque sorte plus primordial que les fonctions de point ; pourquoi donc les Physiciens ne parlent-ils pas de ces fonctions ? Parce que les Mathématiciens ne les ont pas encore étudiées et parce que l'algèbre n'a pas de notation ni pour les domaines ni pour les fonctions de domaine. Aussi voit-on le Physicien se borner à considérer des domaines spéciaux dépendant seulement de certains paramètres, de façon que la fonction de domaine à considérer se réduise à une fonction des paramètres ; c'est d'ailleurs exactement ce que fait le Mathématicien quand, au lieu de considérer l'intégrale définie de $f(x, y)$ dans toute sa généralité, il se borne à envisager les fonctions $F(X, Y)$, $\Phi(a, b\,;\, c, d)$ des formules (6) et (7).

Remarquons d'ailleurs que la formule (8) établit un lien entre les fonctions d'ensemble $\Psi(E)$, qui sont des intégrales indéfinies, et les fonctions de point $f(P)$, lesquelles relèvent de l'algèbre. Cette formule (8) fournit donc une sorte de notation de certaines fonctions d'ensemble. Or, quand on examine les deux conditions requises pour qu'une fonction soit une intégrale indéfinie, on ne doute pas que les grandeurs de la physique ne rentrent dans la classe des fonctions de domaine susceptibles de cette notation.

Ces réflexions sur la nature des grandeurs de la physique ont dû vous permettre de comprendre plus exactement l'intérêt et la portée des notions que nous avons rencontrées. Elles montrent, en particulier, que l'opération de dérivation qui figure dans la formule (9) n'est pas la seule à considérer ; qu'on peut toujours envisager la dérivation d'une fonction $\Psi(E)$ par rapport à une fonction coexistante $p(E)$, que celle-ci soit, ou non, la mesure $m(E)$.

Une question vient alors de suite à l'esprit : peut-on aussi remplacer la fonction $m(E)$ par une fonction donnée $p(E)$ dans la définition de l'intégrale ? Il n'y a à cela aucune difficulté ; nous remplacerons d'abord la formule (2) par

$$S = \Sigma \mu_i p\,(E_i),$$

ce qui exigera, d'abord, que les ensembles E_i appartiennent à la famille de ceux pour lesquels la fonction $p(E)$ est définie, — c'est-à-dire que la fonction à intégrer doit être mesurable par rapport à $p(E)$, — puis que la série S soit convergente, — c'est-à-dire que f doit être sommable par rapport à $p(E)$. Ceci étant supposé, la définition de l'intégrale de $f(P)$ par rapport à $p(E)$,

$$\int f(P)\, dp(E),$$

s'achèvera comme précédemment si la fonction $p(E)$ possède une certaine propriété qu'on exprime en disant que $p(E)$ doit être à variation bornée [1].

Nous venons d'arriver à une nouvelle et très considérable extension de la notion d'intégrale en nous plaçant au point de vue formel du Mathématicien ; le point de vue du Physicien y conduit bien plus naturellement encore, du moins en ce qui concerne les fonctions $f(P)$ continues. On peut même dire que les Physiciens n'ont jamais considéré que des intégrations par rapport à des fonctions de domaine.

Supposons, par exemple, que l'on veuille calculer la quantité de chaleur, $\varphi(D)$, nécessaire pour élever d'un degré la température du corps D dont nous parlions tout à l'heure ; il faudra diviser D en petits corps partiels D_1, D_2, ... de masses $M(D_1)$, $M(D_2)$, ..., prendre dans chacun d'eux un point P_1, P_2, ..., et prendre pour valeur approchée de $\varphi(D)$ la somme

$$f(P_1)\, M(D_1) + f(P_2)\, M(D_2) + \dots,$$

$f(P)$ désignant la chaleur spécifique en P. C'est-à-dire que nous calculerons $\varphi(D)$ par la formule

$$\varphi(D) = \int_D f(P)\, dM(E).$$

Sous sa forme générale la nouvelle intégrale a été définie seulement en 1913 par M. Radon ; elle était cependant connue depuis 1894

1. $p(E)$ est dite à variation bornée si, de quelque manière que l'on partage E en une infinité dénombrable d'ensembles sans points communs deux à deux, E_1, E_2, ..., la série $\Sigma \mid p(E_i) \mid$ est convergente.

La notion de fonction à variation bornée a été tout d'abord introduite par C. Jordan ; elle était alors relative aux fonctions d'une variable.

Les seules fonctions d'ensemble $p(E)$ à considérer dans ces théories sont les fonctions additives, c'est-à-dire celles pour lesquelles on a :

$$p(E_1 + E_2 + \dots) = p(E_1) + p(E_2) + \dots,$$

les E_1, E_2, ... étant sans points communs deux à deux. Si l'additivité est complète, c'est-à-dire si la suite des E_1, E_2, ... peut être prise indéfinie, $p(E)$ est nécessairement à variation bornée. En effet, l'ordre des ensembles E_1, E_2, ... étant indifférent, la série $p(E_1) + p(E_2) + \dots$ doit rester convergente quel que soit cet ordre, c'est-à-dire que la série $\Sigma \mid p(E_i) \mid$ est convergente.

On n'a pas essayé jusqu'ici de se débarrasser de la condition : $p(E)$ à variation bornée. Il faut d'ailleurs remarquer que, si $p(E)$ n'était pas à variation bornée, on pourrait trouver une fonction $f(P)$ qui soit continue et à laquelle, pourtant, notre définition de l'intégrale ne s'appliquerait pas.

pour le cas particulier d'une fonction continue d'une seule variable. Mais son premier inventeur, Stieltjès, y avait été conduit par des recherches d'analyse et d'arithmétique et il l'avait présentée sous une forme purement analytique qui masquait sa signification physique ; si bien qu'il a fallu beaucoup d'efforts pour comprendre et reconnaitre ce qui est maintenant évident. L'historique de ces efforts citerait les noms de F. Riesz, H. Lebesgue, W.-H. Young, M. Fréchet, C. de la Vallée-Poussin ; il montrerait que nous avons rivalisé en ingéniosité, en perspicacité, mais aussi en aveuglement [1].

Et pourtant, les Mathématiciens considéraient à chaque instant des intégrales de Stieltjès-Radon : l'intégrale curviligne $\int_C f(x, y)dx$ est une de ces intégrales, relative à une fonction $p(E)$ définie à partir de la longueur de la projection sur ox des arcs de C ; l'intégrale $\iint_S f(x, y, z)\,dx\,dy$ fait de même intervenir une fonction d'ensemble définie à partir des aires de S en projection sur oxy.

A la vérité ces intégrales se présentent le plus souvent groupées

$$\int_C f(x, y)\,dx + g(x, y)\,dy,$$

$$\iint_S f(x, y, z)\,dx\,dy + g(x, y, z)\,dy\,dz + h(x, y, z)\,dz\,dx.$$

Si l'on pense aussi aux intégrales considérées pour la définition des longueurs des courbes ou des aires des surfaces,

$$\int_C \sqrt{dx^2 + dy^2 + d^2z}, \quad \iint_S \sqrt{(dxdy)^2 + (dydz)^2 + (dzdx)^2},$$

on sera conduit à se dire qu'il conviendrait aussi d'étudier des modes d'intégration dans lesquels interviendraient plusieurs fonctions d'ensemble $p_1(E)$, $p_2(E)$, ... Cette étude est entièrement à faire ; MM. Hellinger et Tœplitz ont cependant utilisé certaines sommations par rapport à plusieurs fonctions d'ensemble [2].

Nous avons considéré jusqu'ici l'intégration, définie ou indéfinie, comme une opération fournissant un nombre, défini ou variable,

1. J. Radon, *Sitz. d. Kais. Ak. d. Wiss. in Wien*, Bd. CXXII, Abt. IIa, 1913. — T.-J. Stieltjès, *Ann. de la Fac. des Sc. de Toulouse*, 1894. — F. Riesz. *Comptes-Rendus de l'Ac. des Sc.*, 1909. — H. Lebesgue, *id.*, 1910. — W.-H. Young, *Proc. of the London Math. Soc.*, 1913. — M. Fréchet, *Nouv. Ann. de Math.*, 1909. — Ch. de la Vallée-Poussin. Voir, par exemple, son livre déjà cité.

2. Voir, par exemple, *Journal de Crelle*, Bd. 144.

par une sorte d'addition généralisée ; nous nous sommes placés au point de vue des quadratures. Mais on peut aussi regarder l'intégration d'une fonction continue comme fournissant une fonction, comme la plus simple des intégrations d'équations différentielles ; c'est le point de vue des fonctions primitives auquel nous allons maintenant nous placer.

Rechercher la fonction primitive $F(x)$ d'une fonction donnée $f(x)$ c'est rechercher la fonction, déterminée à une constante additive près quand elle existe, qui admet $f(x)$ pour dérivée. C'est ce problème que nous allons étudier.

Mais, auparavant, remarquons que les réflexions précédentes conduisent à formuler le problème d'une façon bien autrement générale : étant donnée une fonction $f(P)$ qui est la dérivée par rapport à une fonction connue $p(E)$ d'une fonction inconnue $\Psi(E)$, trouver la fonction primitive $\Psi(E)$ de $f(P)$.

Si, par exemple, il s'agit d'une fonction continue $f(x)$ et si $m(E)$ est la mesure, la fonction primitive ne serait plus la fonction $F(x)$ de la formule (5), mais l'intégrale indéfinie $\int_E f(x)dx$.

Je ne puis que signaler ce problème général qui n'a pas été étudié ; je me contente de faire remarquer que l'intégrale de Stieltjès sera très insuffisante pour le résoudre. Cette intégrale n'a, en effet, été définie que dans l'hypothèse où $p(E)$ est à variation bornée, et l'on peut fort bien parler de dérivée par rapport à une fonction $p(E)$ à variation non bornée.

La théorie des fonctions sommables fournit ce résultat relatif au cas où $p(E)$ est la mesure $m(E)$: lorsque la dérivée $f(P)$ est sommable, la fonction primitive de f est l'une de ses fonctions primitives. Je dis l'une de ses fonctions primitives, car on ne sait pas même encore très bien comment doit être posé ce problème général des fonctions primitives pour qu'il soit déterminé[1].

Laissons donc de côté ces questions, dont je n'ai parlé que pour montrer combien il reste à faire, et montrons combien il vient d'être fait pour la recherche de la fonction primitive $F(x)$ de $f(x)$, grâce surtout à M. Arnaud Denjoy.

Je viens de dire que, lorsque $f(x)$ est sommable, l'intégration fournit $F(x)$ par la formule (5). Supposons que, dans (a, b), $f(x)$

1. Voir à ce sujet des notes de M. Fubini et de M. Vitali parues en 1915-1916 dans les *Atti* de l'Académie de Turin et de l'Académie des Lincei.

ne cesse d'être sommable qu'au point c. Alors l'intégration nous fournit $F(x)$ dans $(a, c - \varepsilon)$, quelque petit que soit ε, donc dans tout (a, c) ; elle fournit aussi $F(x)$ dans $(c + \varepsilon, b)$, donc dans tout (c, b). Et, en tenant compte de la continuité de $F(x)$ au point c, nous aurons $F(x)$ dans tout (a, b). Par de telles considérations de continuité[1] on voit que, si l'on connaît $F(x)$ dans tout intervalle qui ne contienne aucun point d'un ensemble E ni à son intérieur ni comme extrémités, on en déduira $F(x)$ par une opération que je désignerai par A, dans tout intervalle contigu à E, c'est-à-dire dans tout intervalle ayant pour extrémités des points de E mais ne contenant pas de points de E à son intérieur.

Supposons maintenant qu'on connaisse $F(x)$ dans les intervalles (α, β) contigus à un ensemble E, que la somme $\Sigma[F(\beta) - F(\alpha)]$ soit convergente, et que $f(x)$ soit sommable sur E[2]. Alors il suffit de se dire que la fonction primitive doit résulter de la contribution de E et de celle des intervalles contigus à E, pour être conduit à la formule :

$$F(x) - F(a) = \left\{ \int_E f dx + \Sigma[F(\beta) - F(\alpha)] \right\}_a^x,$$

les accolades du second membre indiquant qu'on n'y doit utiliser que les points compris entre a et x. De cette formule résulte donc la détermination de $F(x)$, grâce à une opération que je désignerai par B.

Les deux résultats précédents marquent le point extrême auquel j'étais parvenu dans ma Thèse, et je dois dire que je ne les ai indiqués qu'en quelque sorte par hasard ; car je ne me doutais nullement de l'intérêt qu'allait leur donner M. Denjoy.

S'appuyant sur des résultats de M. Baire, M. Denjoy montre que, si $f(x)$ est une fonction dérivée dans (a, b) :

1° Les points où $f(x)$ n'est pas sommable forment un ensemble E_1 non dense dans (a, b); une opération O_1 du type A, fait connaître $F(x)$ dans les intervalles contigus à E_1 ;

1. C'est l'introduction de ces conditions de continuité qui différencie très considérablement le problème des fonctions primitives de celui des quadratures.

2. Il convient de remarquer que ces hypothèses ne sont pas contradictoires, même si E est supposé être l'ensemble des points de non sommabilité de $f(x)$ dans l'intervalle (a, b) considéré. Pour la détermination des points de non sommabilité dans (a, b) il faut, en effet, tenir compte de tous les points de (a, b), qu'ils appartiennent à E ou non, tandis que la sommabilité sur E est une condition ne faisant intervenir que les seuls points de E.

2° Ensuite, qu'il existe un ensemble E_2 formé de points de E_1 et non dense sur E_2, dans les intervalles contigus duquel on peut calculer $F(x)$ par une opération O_2 du type B ;

3° Ensuite, qu'il existe un ensemble E_3, formé de points de E_2 et non dense sur E_3, dans les intervalles contigus duquel on peut calculer $F(x)$ par une opération O_3 du type B ; ...

S'il arrive qu'après la suite infinie d'opérations O_1, O_2, ... on n'ait pas encore trouvé $F(x)$ dans tout (a, b), les points de (a, b) qui ne sont pas intérieurs à des intervalles dans lesquels on a déterminé $F(x)$ forment un ensemble E_ω et une opération de type A, l'opération O_ω fournit F dans les intervalles contigus à E_ω. On considère ensuite, s'il est nécessaire, des opérations $O_{\omega+1}$, $O_{\omega+2}$, ... du type B, puis une opération $O_{2\omega}$ du type A, puis des opérations du type B, etc.

Et M. Denjoy, utilisant des raisonnements maintenant classiques de Cantor et Bendixson, prouve que ce procédé nous fera finalement connaître $F(x)$ dans tout (a, b) après un nombre fini ou une infinité dénombrable d'opérations.

Ce procédé opératoire, certes compliqué, mais tout aussi naturel, dans son principe, que ceux antérieurement envisagés, a été appelé par M. Denjoy : la totalisation.

La totalisation résoud entièrement le problème de la recherche de la fonction primitive $F(x)$ d'une fonction donnée $f(x)$; elle permet même aussi de déterminer $F(x)$ connaissant seulement un nombre dérivé de $F(x)$ et non plus sa dérivée. Je ne m'arrête pas sur ces beaux résultats ; le fait le plus important pour nous c'est que la totalisation, par un long détour, nous fournit une extension nouvelle de la notion d'intégrale définie. Toutes les fois, en effet, que la totalisation s'appliquera à une fonction $f(x)$ et lui fera correspondre une fonction $F(x)$, nous pourrons convenir d'attacher à $f(x)$ une intégrale grâce aux formules (5) et (5 *bis*)[1].

Messieurs, je m'arrête et je vous remercie de votre courtoise attention; mais il faut un mot de conclusion. Ce sera, si vous le voulez bien, qu'une généralisation faite non pour le vain plaisir de

1. Les mémoires détaillés de M. Denjoy sont parus de 1915 à 1917 au *Journal de Mathématiques*, au *Bulletin de la Société mathématique de France* et aux *Annales scientifiques de l'École Normale supérieure*.

généraliser, mais pour résoudre des problèmes antérieurement posés, est toujours une généralisation féconde. Les divers emplois qu'ont déjà reçus les notions que nous venons d'examiner le prouveraient surabondamment.

H. Lebesgue.