ANNALES SCIENTIFIQUES DE L'É.N.S.

HENRI LEBESGUE

Sur l'intégration des fonctions discontinues

Annales scientifiques de l'É.N.S. 3e série, tome 27 (1910), p. 361-450

<http://www.numdam.org/item?id=ASENS_1910_3_27__361_0>

SUR

L'INTÉGRATION DES FONCTIONS DISCONTINUES,

PAR M. HENRI LEBESGUE.

Introduction.

1. Je me suis proposé d'étendre aux fonctions d'un nombre quelconque de variables les résultats démontrés dans le dernier Chapitre de mes *Leçons sur l'intégration et la recherche des fonctions primitives*. C'est surtout la dérivation des intégrales multiples que j'étudie.

Je rappelle tout d'abord, en modifiant légèrement l'exposition pour mieux préparer la suite, la définition et les propriétés fondamentales des ensembles et fonctions mesurables, des fonctions sommables et de leurs intégrales définies. Je précise ensuite ce qu'il faut entendre par une intégrale indéfinie. Une fonction sommable f étant donnée, l'intégration de f dans un ensemble mesurable E permet d'attacher à E un nombre fonction de cet ensemble E. Cette fonction d'ensemble est l'intégrale indéfinie de f. On constate de suite qu'elle jouit de trois propriétés que je résume en disant que l'intégrale indéfinie est une fonction d'ensemble additive, à variation bornée et absolument continue. Dans la suite, on est conduit à voir que ces propriétés sont caractéristiques des intégrales indéfinies.

2. Pour les fonctions d'ensemble jouissant de ces propriétés on est conduit assez naturellement, en imitant les travaux de M. Volterra, à définir la dérivée de la fonction $\mathcal{F}(\mathrm{E})$ en un point P comme la limite du rapport $\dfrac{\mathcal{F}(\mathrm{E})}{m(\mathrm{E})}$, E étant un ensemble contenant P et dont on fait tendre toutes les dimensions vers zéro. Pour que la dérivation soit en général

l'opération inverse de l'intégration, comme dans le cas où la fonction intégrée est continue, il faut restreindre la catégorie des ensembles E employés, ce qui était bien à prévoir. Si l'on n'emploie que ce que j'appelle des *familles régulières d'ensembles*, la dérivation est presque en tout point l'opération inverse de l'intégration. Quant à ces familles régulières d'ensembles elles sont très étendues ; par exemple un ensemble mesurable quelconque de mesure non nulle et ses homothétiques par rapport au point P constituent une famille régulière ; dans le plan, tous les cercles contenant P, ou tous ceux vus de P sous un angle supérieur à un angle non nul donné, ou tous les secteurs de ces cercles dont l'angle au centre est supérieur à un angle donné forment une famille régulière.

3. M. Vitali (**18**) [1] s'est déjà occupé de cette question de la dérivation des intégrales. Pour lui l'intégrale indéfinie de $f(x_1, x_2, \ldots)$ est la fonction $F(x_1, x_2, \ldots)$ égale à

$$\int_0^{x_1}\int_0^{x_2}\cdots f(x_1, x_2, \ldots)\,dx_1\,dx_2\ldots.$$

C'est ce que j'appelle *l'intégrale indéfinie* exprimée à l'aide des coordonnées x_1, x_2, ... Connaissant $\mathcal{F}(E)$ on connaît en particulier $F(x_1, x_2, \ldots)$; il est facile, connaissant F, d'en déduire $\mathcal{F}$. Considérons le rapport

$$\frac{F(x_1+h, x_2+k)+F(x_1, x_2)-F(x_1+h, x_2)-F(x_1, x_2+k)}{hk},$$

en supposant qu'il n'y ait que deux variables. Si l'on fait tendre h et k vers zéro par valeurs positives ce rapport permet de définir ce que M. Vitali appelle *deux des nombres dérivés de* F. La dérivée est la valeur commune de tous les nombres dérivés lorsqu'ils sont égaux.

Avec cette définition [2] M. Vitali montre que la dérivation permet presque partout de revenir de F à f.

---

[1] Les chiffres gras indiquent des renvois à la bibliographie sommaire placée à la fin de l'Introduction.

[2] Il y a lieu cependant de réparer une inadvertance de l'auteur, laquelle s'explique d'autant mieux que M. Vitali n'a développé ses raisonnements que pour le cas d'une

La définition de M. Vitali revient à celle que j'ai donnée plus haut, à condition de prendre pour les ensembles E servant à définir la dérivée de $\mathcal{F}(E)$ en P, certains rectangles, parallélépipèdes, etc., ayant P pour sommet et des côtés parallèles aux axes. La définition que nous avons adoptée a donc cet avantage de réunir sous un même énoncé la proposition de M. Vitali et un grand nombre d'autres analogues. Il n'est que juste de dire, d'ailleurs, qu'en utilisant, comme j'ai l'occasion de le faire voir dans la suite, un raisonnement qui m'a servi dans l'étude des séries trigonométriques (**8**), on déduit facilement le théorème général du théorème de M. Vitali. Mais il y a une autre raison pour adopter le mode de langage auquel je me suis arrêté, c'est que l'intégration définie ne dépend pas des axes de coordonnées choisis; il paraît donc préférable de n'en pas faire dépendre non plus l'intégration indéfinie, d'autant que l'opération qui transforme $F(x_1, x_2, \ldots)$ en la fonction analogue relative à d'autres coordonnées est fort compliquée et en tous cas ne se réduit pas à un simple changement de variables.

4. Au reste, si l'on fixe des axes de coordonnées, on arrive presque sans raisonnement, en utilisant la dérivation des intégrales simples et la possibilité de remplacer une intégrale multiple par des intégrales simples successives, à un énoncé qui paraît de maniement beaucoup plus facile. *En laissant de côté les points d'un ensemble de mesure nulle convenablement choisi il reste un ensemble* A *sur lequel l'intégrale indéfinie* $F(x_1, x_2, \ldots)$ *de* $f(x_1, x_2, \ldots)$ *peut être dérivée une fois par rapport à chaque variable, les dérivées partielles ainsi obtenues étant indépendantes de l'ordre des dérivations et pouvant s'exprimer par certaines*

---

variable; pour le cas général il s'est borné à indiquer les changements de mots qu'il faudrait faire dans la démonstration. Pour que les changements qu'il indique suffisent il faut ajouter à la définition des nombres dérivés cette restriction que $h = k$ ou tout au moins que $\frac{k}{h}$ reste compris entre deux limites finies non nulles.

Cette restriction est exactement celle à laquelle conduit la définition de la dérivée de $\mathcal{F}(E)$ à l'aide de familles régulières d'ensembles; mais, tandis que nous savons qu'une restriction apportée aux familles d'ensembles mesurables employées pour la définition de la dérivée de $\mathcal{F}(E)$ est indispensable, il n'est nullement prouvé qu'il soit indispensable d'apporter une restriction à l'ordre de grandeur comparée de h et k pour l'exactitude du théorème de M. Vitali. Ce théorème lui-même peut fort bien être exact; je n'ai réussi ni à le démontrer, ni à montrer son inexactitude.

intégrales portant sur f comme dans le cas où f est continue. On peut même choisir A tel que, *sur* A, F *admette une différentielle totale première permettant le calcul de la dérivée de* F *dans une direction quelconque ou le calcul des dérivées partielles premières de* F *après un changement de variables quelconques*.

5. Je veux dire maintenant quelques mots du mode de démonstration que j'ai adopté. Dans mes *Leçons sur l'intégration*, où je traitais le cas d'une variable, pour comparer l'intégrale indéfinie de $f(x)$ à une fonction $F(x)$ supposée existante et admettant f pour dérivée, je cherchais à évaluer $F(b) - F(a) - \int_a^b f(x)\,dx$, a et b fixes quelconques, en remplaçant la courbe $F(x) = y$ par une ligne polygonale circonscrite et dont, par conséquent, on peut définir les côtés à l'aide de $f(x)$. C'est à cette même méthode que se rattachent les démonstrations de M. Beppo Levi (**10, 11**) et de M. Giuseppe Vitali (**16, 18**). M. de la Vallée-Poussin (**15**) opère autrement; il compare la fonction $F(x)$, x variable, à l'intégrale indéfinie $\int f(x)\,dx$ à l'aide de théorèmes, généralisations du théorème fondamental : *deux fonctions ayant partout la même dérivée ne diffèrent que par une constante*, ou du théorème de Ludwig Scheeffer. Pour cela M. de la Vallée-Poussin utilise des fonctions approchées de $\int f(x)\,dx$ choisies de manière que l'on connaisse leur dérivée presque partout. C'est une méthode imitée de celle de M. de la Vallée-Poussin que j'emploie ici. Je considère comme un avantage de cette méthode le fait d'exiger une petite étude des fonctions de plusieurs variables.

Il semble d'ailleurs qu'on va être arrêté dès le début de cette étude dans les généralisations des propriétés les plus simples des fonctions d'une variable. Définissons, par exemple, les nombre dérivés à droite de $F(x_1, x_2)$, par la méthode de M. Vitali en prenant h et k positifs. Et considérons la fonction continue

$$F = (x + y) + |(x + y)|;$$

cette fonction a sa dérivée à droite constamment nulle et cependant elle n'est pas la somme d'une fonction de x et d'une fonction de y. Ainsi, on ne peut pas généraliser cette proposition immédiate et fondamentale ; une fonction continue d'une variable dont un nombre dérivé est constamment nul est constante. On se tire d'affaire en restreignant la classe des fonctions F considérées.

J'avais, dans mes *Leçons*, tout à fait incidemment et sans démonstration, fait connaître que la condition nécessaire et suffisante pour qu'une fonction $F(x)$ soit une intégrale indéfinie est que la somme $\Sigma|F(x_i+\delta_i)-F(x_i)|$, relative à des intervalles $(x_i, x_i+\delta_i)$ non empiétant, tende vers zéro avec $\Sigma|\delta_i|$ (**4**, **7**). M. Vitali (**16**) qui appelle *fonctions absolument continues* les fonctions jouissant de ces propriétés, a retrouvé cette proposition. Il en a le premier publié une démonstration et il l'a étendue aux fonctions de plusieurs variables (**18**). De plus, il en a montré l'intérêt en s'en servant d'une part (**17**), pour obtenir la condition nécessaire et suffisante pour qu'une série soit intégrable terme à terme, et qu'elle le reste si on la multiplie par une fonction sommable bornée quelconque et, d'autre part (**18**), dans l'étude de la dérivation des intégrales. J'ai imité son exemple en me bornant tout d'abord aux fonctions F absolument continues, ce qui supprime les difficultés du genre de celle que je signalais plus haut ; mais, bien entendu, les démonstrations ne peuvent pas garder la simplicité qu'elles avaient dans le cas d'une variable.

6. Je base ici ces démonstrations sur une proposition géométrique due à M. Vitali : *Soit un ensemble mesurable* E ; *supposons que chacun de ses points appartiennent à une infinité d'intervalles arbitrairement petits tirés d'une famille* F *d'intervalles*. Le théorème de M. Vitali affirme *la possibilité d'extraire de* F *une suite dénombrable ou finie d'intervalles non empiétant dont la somme des mesures est au moins la mesure de* E. Une proposition analogue a été démontrée pour le plan, les intervalles étant remplacés par des carrés ; pour l'espace, les intervalles étant remplacés par des cubes, etc. Ici je la démontre pour toutes les familles F régulières de domaines ; mais, et c'est là une différence essentielle avec un théorème analogue de M. Borel, le théorème n'est pas vrai pour une famille quelconque de domaines, comme on le verra dans ce Mémoire.

Dans mes *Leçons* je m'étais servi d'un procédé de démonstration que j'appelais *l'emploi de chaînes d'intervalles*. Ce procédé réussissait parce qu'il constituait, dans les cas où je l'employais, une façon d'effectuer le choix que le théorème de M. Vitali affirme être possible. Je ne me sert pas ici de ce procédé qui ne semble pas pouvoir être étendu au cas de plusieurs variables. Il est vrai qu'au moment où j'écrivais le dernier Chapitre de mes *Leçons* j'avais aperçu sa généralisation au cas de plusieurs variables en utilisant le procédé des chaînes, mais c'était grâce à un raisonnement qui m'a paru très artificiel quand j'ai connu le théorème de M. Vitali permettant d'arriver plus directement au but. On trouvera plus loin le résumé de cette démonstration qui conduisait à utiliser provisoirement une définition de la dérivée de $\mathscr{F}(\mathrm{E})$ à l'aide des ensembles E formés des points d'un arc d'une courbe de Peano ou d'Hilbert remplissant l'espace que l'on considérait.

7. Les fonctions qui, sans être absolument continues, sont à variation bornée semblent devoir se présenter dans beaucoup de questions; aussi ai-je essayé d'étendre à ces fonctions quelques-uns des énoncés obtenus pour les fonctions absolument continues. Je leur ai consacré un long Chapitre qui n'épuise pas la question.

Ce Mémoire se termine par la généralisation de quelques théorèmes classiques, relatifs à l'intégration par substitution ou par partie et aux théorèmes de la moyenne. On verra en particulier qu'on peut donner au second théorème de la moyenne une portée bien plus grande qu'on ne le fait ordinairement et cela sans compliquer notablement son énoncé.

J'énumère ici quelques travaux sur l'intégration que j'ai l'occasion de citer :

G. Fubini. **1**. Sugli integrali multipli (*Rend. del. R. Acc. dei Lincei*, 1er semestre 1907).

E.-W. Hobson. **2**. On the second mean-value theorem of the integral calculus (*Proc. of the London Math. Soc.*, 1909).

H. Lebesgue. **3**. Integrale, Longueur, Aire (Thèse et *Annali di Mat.*, 1902). — **4**. Leçons sur l'intégration et la recherche des fonctions primitives, Paris, 1904. — **5**. Sur les fonctions dérivées (*Rend. del R. Acc. dei Lincei*, 2e semestre 1906). — **6**. Encore une observation sur les fonctions dérivées (*Ib.*, 1er sem. 1907). — **7**. Sur la recherche des fonctions primitives par l'intégration (*Ib.*). — **8**. Recherches sur la convergence des séries de Fourier (*Math. Ann.*, Bd LXI, 1905. — **9**. Sur les intégrales singulières (*Ann. de la Fac. des Sc. de Toulouse*, 1909).

B. Levi. **10**. Richerche sulle funzioni derivate (*Rend. del. R. Acc. dei Lincei*, 1^er^ semestre 1906 ; 3 notes). — **11**. Ancora alcune osservazioni sulle funzioni derivate (*Ib.*, 2^e^ sem. 1906). — **12**. Sopra l'integrazione delle serie (*Rend. del R. Ist. lomb. di Sc. e Lett.*, 1906).

L. Tonelli. **13**. Sull'integrazione per parti (*Rend. del R. Acc. dei Lincei*, 2^e^ sem. 1909). — **14**. Sulla rappresentazione analitica delle funzioni di più variabili reali (*Rend. del Circ. mat. di Palermo*, 1^er^ sem. 1910).

Ch. de la Vallée-Poussin. **15**. Cours d'Analyse infinitésimale (2^e^ édit., t. I, Paris-Louvain, 1909).

G. Vitali. **16**. Sulle funzioni integrali (*Rend. del R. Acc. del. Sc. di Torino*, 1900). — **17**. Sull'integrazione per serie (*Rend. del Circ. math. di Palermo*, 1^er^ sem. 1907). — **18**. Sui gruppi di punti e sulle funzioni di variabili reali (*Rend. del. R. Acc. del. Sc. di Torino*, 1908).

I. — L'intégrale définie.

8. Pour la simplicité du langage je dirai, suivant l'usage, qu'un ensemble de k valeurs données aux nombres $x_1, x_2, \ldots, x_k$ est un point P de l'espace à k dimensions et je désignerai par $f(\mathrm{P})$ une fonction des k coordonnées de P. Les fonctions que je vais avoir à considérer sont définies pour tous les points d'un ensemble ou d'un *domaine*. J'appelle ainsi un ensemble borné, formé de points de l'espace à k dimensions considéré, qui est parfait, d'un seul tenant, et dont tous les points sont, ou bien points intérieurs de l'ensemble, ou bien limites de tels points. Parmi les domaines se trouvent la variété limitée par l'hypersphère et les variétés homéomorphes, c'est-à-dire les ensembles de points définis comme il suit : considérons une transformation univoque et continue de l'espace à k dimensions en lui-même ; cette transformation est définie par un ensemble de k fonctions $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_k$ des k variables $x_1, x_2, \ldots, x_k$, continues par rapport à l'ensemble de ces variables et dont l'inversion est possible d'une manière unique.

Un domaine homéomorphe à l'hypersphère est celui formé des points $x_1, x_2, \ldots, x_k$ qui vérifient l'inégalité

$$X_1^2 + X_2^2 + \ldots + X_k^2 \leqq 1.$$

On peut, si l'on veut, se borner dans la suite à la considération de ces domaines-là ([1]) et de ceux qu'on obtient en leur retranchant les points intérieurs d'un ou plusieurs domaines analogues, c'est-à-dire homéomorphes à l'hypersphère.

Un domaine D est dit la somme des domaines $D_1, D_2, \ldots,$ si deux quelconques des domaines $D_1, D_2, \ldots$ n'ont pas de points intérieurs communs et si tout point intérieur à D appartient au moins à l'un des D_i. Cette définition est universellement adoptée lorsque les D_i sont en nombre fini; s'ils étaient en nombre infini d'autres conventions pourraient être adoptées. Il importe de ne pas confondre avec d'autres celle qui est faite ici.

J'appellerai *domaine unité* celui qui est défini par les inégalités

$$0 \leqq x_i \leqq 1 \qquad (i = 1, 2, \ldots, k).$$

On suppose dans tout ce qui suit les axes rectangulaires, c'est-à-dire qu'on définit les déplacements par l'invariance des fonctions $\Sigma(x_i - x'_i)^2$ relatives à tous les couples de points $(x_1, x_2, \ldots)$, $(x'_1, x'_2, \ldots)$ de la figure.

9. *Le problème d'intégration.* — On se propose d'attacher à toute fonction bornée $f(P)$, définie pour les points P d'un ensemble borné E, un nombre fini, qu'on appellera *l'intégrale définie de* $f(P)$ dans E et qu'on notera $\int_E f(P)\,dP$, satisfaisant aux conditions suivantes :

1° (premier théorème d'addition)

$$\int_E [f(P) + \varphi(P)]\,dP = \int_E f(P)\,dP + \int_E \varphi(P)\,dP;$$

([1]) Nous les appellerons *domaines simplement connexes* dans ce qui suit.

2° (deuxième théorème d'addition)

$$\int_{E_1+E_2+\ldots} f(P)\,dP = \int_{E_1} f(P)\,dP + \int_{E_2} f(P)\,dP + \ldots,$$

les $E_1, E_2, \ldots$ étant supposés former au plus une suite dénombrable et être sans point commun deux à deux;

3° Si l'on fait subir à E un déplacement qui l'amène en E′ et si l'on désigne par $f'(P')$ la fonction définie sur E′ par la condition qu'en deux points homologues P et P′ de E et E′ les valeurs de f et f' soient égales, on doit avoir

$$\int_{E} f(P)\,dP = \int_{E'} f'(P')\,dP';$$

en d'autres termes l'intégrale ne doit pas dépendre de la position de E dans l'espace ;

4° Si f n'est jamais négative dans E, son intégrale dans E est positive ou nulle ;

5° Si D_0 est le domaine unité, on a

$$\int_{D_0} 1\,dP = 1.$$

10. *Le problème de la mesure.* — Supposons d'abord $f(P) = 1$ en tout point P de E; alors il n'y a plus à tenir compte de la condition 1° et l'on peut énoncer le problème à résoudre ainsi :

On se propose d'attacher à chaque ensemble borné E un nombre positif ou nul, qu'on appellera la mesure de E et qu'on notera $m(E)$, satisfaisant au conditions suivantes :

1° Deux ensembles égaux ont des mesures égales ;

2° L'ensemble somme d'un nombre fini ou dénombrable d'ensembles, sans point commun deux à deux, a pour mesure la somme des mesures ;

3° D_0 étant le domaine unité, on a $m(D_0) = 1$.

Je rappelle la marche de la solution, en me plaçant pour la commodité du langage dans le cas de deux variables. On reconnaît de suite qu'un point, qu'un segment de droite doivent avoir des mesures nulles; il en résulte que si l'on divise un polygone en deux autres la mesure du polygone total est la somme des mesures des deux polygones composants et cela permet de montrer que la mesure d'un polygone est son aire, au sens élémentaire du mot.

Soit maintenant un ensemble borné E; on peut, d'une infinité de manières, enfermer les points de E dans un nombre fini ou une infinité dénombrable de polygones; soit $m_e(\mathrm{E})$ la limite inférieure de la somme des aires de ces polygones. Ce nombre $m_e(\mathrm{E})$, qu'on appelle la *mesure extérieure de* E, est évidemment au moins égal à la mesure de E. D'autre part, si F est l'ensemble des points qui n'appartiennent pas à E et intérieurs à un certain polygone G contenant tout E à son intérieur, la quantité $m(\mathrm{G}) - m_e(\mathrm{F})$ qu'on appelle la *mesure intérieure de* E et qu'on note $m_i(\mathrm{E})$ doit être au plus égale à la mesure de E.

On démontre qu'on a toujours

$$m_i(\mathrm{E}) \leqq m_e(\mathrm{E});$$

il n'y a donc jamais contradiction entre les deux conclusions précédentes; c'est-à-dire que cela ne nous permet pas de définir une classe d'ensembles pour laquelle le problème de la mesure est impossible, mais nous voyons que pour les ensembles tels que

$$m_i(\mathrm{E}) = m_e(\mathrm{E})$$

le problème de la mesure est déterminé, s'il est possible, et qu'on doit avoir

$$m(\mathrm{E}) = m_i(\mathrm{E}).$$

Les ensembles ainsi caractérisés sont dits *mesurables*. On vérifie que le problème de la mesure est bien possible pour eux en vérifiant que la condition 2° de ce problème est bien remplie.

On peut démontrer aussi que si l'on effectue, à partir d'ensembles mesurables, les deux opérations suivantes :

I. Faire la somme d'une infinité dénombrables d'ensembles;

II. Prendre la partie commune à tous les ensembles d'une famille contenant une infinité dénombrable d'ensembles,

on obtient des ensembles mesurables.

Or il est facile de se rendre compte que tous les ensembles plans qu'on a pu définir effectivement jusqu'ici s'obtiennent par la répétition des opérations 1° et 2° à partir de polygones; par suite tous les ensembles effectivement définis jusqu'ici sont mesurables [1].

11. *La définition de l'intégrale.* — Il est donc pratiquement suffisant de se poser le problème de la mesure pour les seuls ensembles mesurables et le problème d'intégration pour les seules *fonctions mesurables*, c'est-à-dire celles pour lesquelles, quels que soient a et b, l'ensemble des points en lesquels on a $a < f(\mathrm{P}) < b$ est mesurable. Par l'emploi des opérations I et II on reconnaît d'ailleurs de suite que, pour ces fonctions, on peut remplacer si l'on veut un signe $<$ par le signe $=$ ou $\leqq$ sans que l'ensemble défini cesse d'être mesurable; on voit de même que l'ensemble formé par les points en lesquels une fonction mesurable est définie est nécessairement mesurable.

Ceci posé, soit f une fonction mesurable bornée définie dans un ensemble mesurable borné E. Soient m et M les bornes inférieure et supérieure de f; soient $a_1, a_2, \ldots, a_{n-1}$ des nombres tels qu'on ait

$$m = a_0 < a_1 < a_2 < \ldots < a_{n-1} < a_n = \mathrm{M} = a_{n+1}$$

et que $(a_i - a_{i-1})$ ne surpasse jamais ε arbitrairement choisi ou positif. Soit $\mathcal{E}_i$ l'ensemble des points P en lesquels on a

$$a_i \leqq f(\mathrm{P}) < a_{i+1} \qquad (i = 0, 1, \ldots, n-1);$$

$\mathcal{E}_i$ est mesurable; $\mathcal{E}_n$ sera l'ensemble mesurable des points où $f(\mathrm{P}) = a_n$.

Des conditions du problème d'intégration il résulte qu'on a

$$\sum_{i=0}^{i=n} a_i\, m(\mathcal{E}_i) \leqq \int_{\mathrm{E}} f(\mathrm{P})\, d\mathrm{P} \leqq \sum_{i=0}^{i=n} a_{i+1}\, m(\mathcal{E}_i),$$

---

[1] Pour plus de détails sur ce point, *voir* mon Mémoire *Sur les fonctions représentables analytiquement* (*Journal de Math.*, 1905). Il est nécessaire d'ajouter que divers auteurs [VITALI, *Sul problema della misura dei gruppi di punti di una retta*, Bologne, 1905; LEBESGUE, *Sur les correspondances ponctuelles*, etc. (*Atti d. R. Acc. d. Sc. di Torino*, 1907); *Contribution à l'étude des correspondances de M. Zermelo* (*Bul. de la Soc. math. de France*, 1907); Ed. VAN VLECK, *On non mesurable sets of points* (*Trans. of the Am. mat. Soc.*, 1908)] ont indiqué des procédés de formation d'ensembles non mesurables, mais ces procédés supposent qu'on emploie des opérations qu'on ne sait ni effectuer, ni même caractériser logiquement.

si le problème d'intégration est possible. Or un raisonnement exactement semblable à celui qu'on fait dans la théorie classique de l'intégration montre que, lorsqu'on fait tendre ε vers zéro, les deux membres extrêmes de cette double inégalité tendent vers une même limite indépendante du choix des a_i remplissant les conditions indiquées. C'est cette limite commune qu'on doit prendre pour valeur de $\int_E f(P)\,dP$. Reste à vérifier qu'elle satisfait aux conditions du problème d'intégration.

Cela ne peut faire quelque doute que pour la condition 2°. Or, supposons E formé par la réunion des ensembles E_1, E_2, ..., sans point commun deux à deux, et plaçons-nous dans le cas où la suite des E_i est infinie. $\mathcal{E}_{i,p}$ désigne la partie commune à $\mathcal{E}_i$ et à E_p. Alors le premier membre de l'inégalité précédente peut s'écrire

$$\sum_{i=0}^{i=n} a_i \left[\sum_{j=0}^{j=\infty} m(\mathcal{E}_{i,j})\right].$$

Or les séries qui figurent dans cette expression sont absolument convergentes, donc elle peut être mise sous la forme

$$\sum_{=0}^{j=\infty} \left[\sum_{i=0}^{i=n} a_i\, m(\mathcal{E}_{i,j})\right].$$

Sous cette forme on voit que l'expression considérée tend vers

$$\sum_{j=0}^{j=\infty} \int_{E_j} f(P)\,dP$$

et la proposition est démontrée.

12. Supposons maintenant que $f(P)$ mesurable ne soit pas bornée et considérons des nombres a_i allant en croissant de $-\infty$ à $+\infty$, quand i parcourt la suite des entiers de $-\infty$ à $+\infty$, et tels que la différence $a_{i+1} - a_i$ soit bornée uniformément quel que soit i. Alors

les quantités analogues aux deux membres extrêmes de l'inégalité qui nous a servi sont les séries

$$\sum_{i=-\infty}^{i=+\infty} a_i m(\mathcal{E}_i), \qquad \sum_{i=-\infty}^{i=+\infty} a_{i+1} m(\mathcal{E}_i).$$

Il est bien évident que si l'une des séries est absolument convergente il en est de même de l'autre et l'on voit facilement qu'il en est encore de même de toute autre série analogue formée à l'aide d'un autre choix des a_i vérifiant les mêmes conditions. Si donc on se posait le problème d'intégration pour les fonctions mesurables, qu'elles soient bornées ou non, auxquelles correspondent une série $\sum_{i=-\infty}^{i=+\infty} a_i m(\mathcal{E}_i)$, qui soit absolument convergente, fonctions que nous appellerons *fonctions sommables*, nous pourrions reprendre sans changements les raisonnements faits sur les fonctions bornées et nous serions conduits aux mêmes conclusions.

Notons, en passant, cette conséquence essentielle de notre définition : si une fonction est sommable il en est de même de sa valeur absolue. Il résulte de là que certaines fonctions non bornées, intégrables par la méthode classique, ne sont pas sommables; mais on verra que les fonctions sommables forment une classe bien homogène à propriétés simples et intéressantes, ce qui ne serait plus vrai si l'on étendait la définition de l'intégrale. Il y a donc, en un certain sens, grand avantage à ne pas élargir davantage la définition de l'intégrale.

13. Il reste à donner une idée de l'étendue de la classe des fonctions sommables. Il a déjà été dit que toutes les fonctions qu'on a jusqu'ici nommées sont mesurables (de sorte que celles qui sont bornées sont sommables); voici comment on s'en rend compte.

En utilisant les opérations I et II dont il a été parlé précédemment on voit facilement que la somme et le produit de deux fonctions mesurables sont des fonctions mesurables, que la limite d'une suite convergente de fonctions mesurables est une fonction mesurable. Il en résulte, comme 1 et x_i sont des fonctions mesurables, que tout polynome est une fonction mesurable; que toute limite de polynomes et en particulier toute fonction continue est mesurable, que toute limite de limites de polynomes est mesurable, etc.

En ce qui concerne les fonctions non bornées, il est facile de voir que la somme de deux fonctions sommables est toujours une fonction sommable tandis qu'il n'est pas toujours vrai que le produit de deux fonctions sommables soit sommable. Il en est ainsi cependant quand l'un des deux facteurs est borné. On peut rattacher ce dernier fait à une remarque immédiate : *une fonction mesurable, inférieure en valeur absolue à une fonction sommable positive, est nécessairement sommable.*

14. Il est maintenant nécessaire d'examiner les différences qu'il y a entre l'énoncé du problème d'intégration tel que je l'ai formulé dans mes *Leçons sur l'intégration* et tel que je le formule ici. Dans mes *Leçons*, où je ne m'occupais que de fonctions d'une seule variable, le mot *ensemble* était partout remplacé dans l'énoncé par le mot *intervalle;* de plus la condition 2° n'était énoncée que pour l'intervalle somme de deux autres (et non d'un nombre fini ou d'une infinité dénombrable d'autres), mais par contre je formulais une sixième condition :

6° *Si f_n tend en croissant vers f, l'intégrale de f_n tend vers celle de f.*

Il est tout d'abord évident que si, dans la condition 2°, nous nous étions restreints à la considération de l'ensemble somme de deux autres, c'est-à-dire si nous avions restreint l'énoncé du deuxième théorème d'addition, mais que par contre nous ayons donné au premier théorème d'addition une portée plus étendue en formulant la condition 6°, nous aurions été conduits à la même définition de l'intégrale. Seulement il aurait fallu vérifier que cette sixième condition est remplie.

Le théorème exprimé par cette sixième condition étant important, je vais le démontrer. Pour cela, et parce que cette propriété nous sera utile dans la suite, je montre d'abord que : *si une fonction f, sommable, est définie dans un ensemble* E *et si e est un ensemble mesurable* *quelconque contenu dans* E, *l'intégrale* $\int_e f(\mathrm{P})\,d\mathrm{P}$ *tend vers zéro avec la mesure de e.*

Conservons les notations précédentes; la série $\sum_{i=-\infty}^{i=+\infty} a_i m(\mathcal{E}_i)$ représente $\int_{\mathrm{E}} f(\mathrm{P})\,d\mathrm{P}$ à $\varepsilon m(\mathrm{E})$ près au plus. Si q est choisi assez grand, en ne conservant dans la série que les termes $\sum_{i=-q}^{i=+q} a_i m(\mathcal{E}_i)$, on commettra une nouvelle erreur qui sera inférieure à ε. Cela revient à ne tenir compte que de certains $\mathcal{E}_i$; soit U l'ensemble de points formé par les $\mathcal{E}_i$ utilisés. Dans U, f est bornée; soit A la borne supérieure de $|f|$ dans V.

Ceci posé, un ensemble e de mesure l contenu dans E fournit dans $\sum_{i=-\infty}^{i=+\infty} a_i m(\mathcal{E}_i)$ une contribution au plus égale à $\varepsilon + l\mathrm{A}$, donc $\int_e f(\mathrm{P})\,d\mathrm{P}$ est au plus égae à

$$\varepsilon + l\mathrm{A} + \varepsilon m(\mathrm{E})$$

et comme ε est quelconque la proposition est démontrée.

15. Si des *fonctions sommables* f_n *forment une suite convergente et sont toutes, en valeur absolue, inférieures à une fonction sommable positive* F, *la limite* f *des* f_n *est sommable et son intégrale est la limite de l'intégrale de* f_n.

Il est tout d'abord évident que f, mesurable, est sommable puisque l'on a $|f| \leqq \mathrm{F}$. Désignons maintenant par E_n l'ensemble des points en lesquels on a, quel que soit $p > 0$,

$$|f_{n+p} - f| < \varepsilon.$$

E_n est mesurable, il est contenu dans E_{n+1}, et tout point de l'ensemble E, où les f et les f_n sont définis, appartient à E_n pour n assez grand. En d'autres termes on a

$$\mathrm{E}_1 + (\mathrm{E}_2 - \mathrm{E}_1) + (\mathrm{E}_3 - \mathrm{E}_2) + \ldots = \mathrm{E};$$

donc

$$\mathrm{E} - \mathrm{E}_n = (\mathrm{E}_{n+1} - \mathrm{E}_n) + (\mathrm{E}_{n+2} - \mathrm{E}_{n+1}) + \ldots,$$

Or il résulte de là qu'on a

$$m(E_1)+m(E_2-E_1)+\ldots=m(E),$$

donc que

$$m(E-E_n)=m(E_{n+1}-E_n)+m(E_{n+2}-E_{n+1})+\ldots$$

tend vers zéro quand n croît indéfiniment [1].

Or, on a

$$\left|\int_E f\,dP-\int_E f_n\,dP\right|\leqq\int_{E_n}|f-f_n|\,dP+\int_{E-E_n}(|f|+|f_n|)\,dP$$
$$\leqq\varepsilon m(E_n)+2\int_{E-E_n}F\,dP,$$

et cela démontre le théorème.

Remarquons que dans l'énoncé précédent la condition que f_n est inférieure à une fonction sommable positive est équivalente à celle-ci : les différences $f-f_n$ sont toutes inférieures à une même fonction sommable positive.

Du théorème général il en résulte en particulier que, *si f_n sommable tend en croissant vers f sommable, $\int f\,dP$ est la limite de $\int f_n\,dP$* et que *si les restes d'une série convergente sont uniformément bornés la série est intégrable terme à terme* [2].

Il résulte immédiatement de la définition de l'intégrale que les énoncés précédents ne cesseraient pas d'être vrais si les conditions qu'ils contiennent n'étaient pas vérifiées pour des points exceptionnels formant des ensembles de mesure nulle.

Dans la suite je dirai qu'une condition est remplie *presque partout* lorsqu'elle est vérifiée en tout point, sauf aux points d'un ensemble de mesure nulle.

Avant de passer à l'autre point, remarquons encore que si l'on avait restreint comme il a été dit la condition 2° du problème d'intégration

---

(1) Ces développements ont en somme comme résultat de légitimer une conséquence de la définition de la mesure que je n'ai pas eu l'occasion de citer plus haut mais qui est très importante. Si l'on a une suite d'ensembles mesurables E_n et si chacun d'eux est contenu dans ceux d'indice supérieur (ou contient tous ceux d'indice supérieur) l'ensemble somme (ou l'ensemble des points communs) a pour mesure la limite de $m(E_n)$.

(2) Pour plus de détails sur ces questions, *voir* mon Mémoire *Sur les intégrales singulières* (**9**).

sans énoncer la condition 6°, ce qui eût entraîné une modification corrélative dans la condition 2° du problème de la mesure, les seuls ensembles auxquels nos raisonnements nous auraient conduits à attacher une mesure déterminée, auraient été les ensembles mesurables de M. Jordan et les fonctions mesurables correspondantes auraient été les fonctions intégrables au sens de Riemann. Retenons de cela que ces fonctions sont sommables et qu'il n'y a jamais désaccord entre la définition de Riemann et la mienne ([1]).

16. Supposons maintenant que, dans l'énoncé du problème d'intégration, on ait partout remplacé *ensemble* par *domaine*, mais qu'on n'ait pas fait d'autres changements. Des raisonnements fort analogues à ceux qui nous ont servi montrent que la définition de l'intégrale reste la même, seulement on se trouve conduit à résoudre un problème de la mesure des domaines dont l'énoncé ne diffère de celui de la mesure des ensembles que par le changement d'*ensemble* en *domaine;* et c'est seulement aux domaines qu'on sait mesurer qu'on peut attacher une intégrale.

La différence entre la mesure des domaines et la mesure des ensembles provient de ce que, tandis que dans le second cas on ne s'occupe que de la somme de deux ensembles sans point commun, dans le premier, où l'on s'occupe du domaine somme de deux autres, ces deux autres ont nécessairement des points frontières communs ([2]).

La mesure d'un domaine, en tant que domaine, ne pouvant évidemment différer de sa mesure, en tant qu'ensemble, les seuls domaines pour lesquels le problème de la mesure des domaines est possible sont ceux dont la frontière est de mesure nulle. On les appelle *domaines quarrables*.

([1]) Bien entendu il s'agit ici de fonctions bornées, la définition classique de l'intégrale des fonctions non bornées étant de tout autre nature que celle étudiée ici.

([2]) Avec la définition adoptée ici, il se pourrait d'ailleurs que les domaines composants n'aient qu'un point commun. Si l'on veut que cette singularité ne se présente pas, il faut, dans la définition des domaines, remplacer la condition d'être d'un seul tenant par la condition que deux points intérieurs du domaine puissent toujours être joints par un chemin continu formé tout entier de points intérieurs.

Ces distinctions sont sans importance pour la suite.

Or, tandis qu'on ne sait pas nommer un ensemble non mesurable, il est facile de nommer un domaine non quarrable (¹), c'est pourquoi il me paraît préférable d'énoncer le problème d'intégration en employant le mot *ensemble*.

La différence entre les deux énoncés ne se manifeste naturellement pas dans le cas d'une seule variable car alors tous les domaines sont quarrables : ce sont des intervalles.

Dans ma Thèse, où j'examinais le cas de deux variables, mais le cas de k variables ($k > 2$) est identique à celui de $k = 2$, j'avais aussi examiné ce que devenait le problème de la mesure des domaines, problème que j'appelais *problème des aires*, quand dans la condition 2° de ce problème on ne s'occupait que d'un domaine somme de deux autres. Cette formulation correspond à celle du problème de l'intégration utilisant le mot *domaine* et n'utilisant que la condition 2° restreinte sans utiliser la condition 6°.

On voit sans peine que, du moins si l'on se restreint à la considération des domaines simplement connexes, le problème est possible et déterminé pour les domaines quarrables mais qu'il est indéterminé pour les autres (²); il en résulte qu'avec la dernière formulation le problème d'intégration serait indéterminé pour certaines fonctions non intégrables au sens de Riemann. Il importe de se rappeler que ces conclusions supposent expressément qu'on a défini la somme des domaines comme je l'ai fait (³).

17. Jusqu'ici tous les ensembles que nous avons considérés étaient bornés. Un ensemble non borné E sera dit *mesurable* si la partie

---

(¹) J'ai indiqué le principe de cette construction dans ma Thèse (p. 17, en note). Plus tard, M. Osgood a été conduit à développer une construction basée sur le même principe (*Trans. Am. math. Soc.*, 1903, p. 107).

(²) La preuve que j'en avais donnée dans ma Thèse était inexacte, j'ai réparé mon erreur dans deux Notes du *Bulletin de la Société mathématique de France*, t. XXXI, p. 97; t. XXXIII, p. 273.

(³) La lecture de la note 4 de la page 253 du Tome II de l'Ouvrage de M. Schœnflies (*Die Entwickelung der Lehre von den Punktmannigfaltigkeiten*) pourrait faire croire que les conclusions rapportées ci-dessus sont erronées. M. Schœnflies a bien voulu m'expliquer qu'on devait interpréter sa note ainsi : avec les définitions adoptées par M. Schœnflies mes énoncés ne sont plus exacts.

commune à E et à une hypersphère quelconque est mesurable; si, quand l'hypersphère S varie de façon que tout point de l'espace soit à partir d'un certain moment intérieur à S, la mesure de cette partie commune n'augmente pas indéfiniment, E est dit *mesurable* et *de mesure finie* et sa mesure $m(E)$ est la limite évidemment unique et déterminée de la mesure de la partie commune à E et à S. Si E est de mesure α on peut enfermer E dans des domaines bornés dont la somme des mesures surpasse α d'aussi peu qu'on le veut : E sera dit un *domaine* s'il est parfait, d'un seul tenant et si tout point de E fait partie d'un domaine borné dont tous les points sont points de E.

Une fonction f définie pour les points d'un ensemble E est dite *mesurable* dans les mêmes conditions que si E était borné; f sera dite *sommable* si, quand S varie comme il a été dit, $\int |f(P)|\,dP$ étendue à la partie commune e à S à E n'augmente pas indéfiniment. Alors $\int_e f(P)\,dP$ tend vers une limite déterminée qui est l'intégrale

$$\int_E f(P)\,dP.$$

On peut d'ailleurs encore dire : soient des nombres a_i croissant avec i entier de $-\infty$ à $+\infty$, supposons a_0 négatif, a_{+1} positif. Soit $\mathcal{E}_i$ l'ensemble des points où l'on a

$$a_i \leqq f(P) < a_{i+1} \qquad (i = 1, 2, \ldots; -1, -2, \ldots).$$

Supposons que ces $\mathcal{E}_i$ soient mesurables et de mesure finie. Nous formons la somme $\Sigma a_i m(\mathcal{E}_i)$ étendue à toutes les valeurs non nulles de i; nous la supposons absolument convergente; alors, si elle tend vers une limite déterminée finie, ainsi que la somme de ses valeurs absolues, quand on fait varier les a_i de façon que le maximum de $a_{i+1} - a_i$ tende vers zéro, cette limite est l'intégrale de f qui est dite *sommable*. On étendra sans peine aux ensembles non bornés les propositions indiquées précédemment.

18. Considérons maintenant une fonction f non partout finie sur E, si l'*ensemble des points où elle est infinie est de même nulle;* quelques auteurs disent qu'elle est *sommable* si elle est sommable sur l'ensemble

E_1 des points où elle est finie et lui attribue pour intégrale sur E son intégrale sur E_1. Ce mode de langage est souvent avantageux; mais comme il présente parfois des inconvénients et qu'il peut prêter à des malentendus, j'ai évité en général de l'employer.

Complétons maintenant avec M. Beppo Levi (**12**) un théorème précédent. *Soit une suite croissante f_1, f_2 de fonctions ne devenant infinies sur* E *qu'en des ensembles de mesure nulle de points et sommable quand on excepte ces points. Alors la suite f_1, f_2, ... est toujours intégrable terme à terme.* Ce qui signifie que la limite f ne devient infinie qu'en un ensemble de mesure nulle de points, qu'elle est sommable sur l'ensemble restant et que son intégrale est la limite de celles des f_i si cette limite existe et qu'inversement cette limite existe toutes les fois que f ne devient infinie que pour points d'un ensemble de mesure nulle et est sommable dans l'ensemble restant.

Cela résulte de suite de ce que, si l'on désigne par f_i^n ou f^n des fonctions égales à f_i ou f lorsque f_i ou f ne dépasse pas n et à n dans le cas contraire, on a

$$\lim_{i=\infty} \int f_i^n \, dP = \int f^n \, dP,$$

et de ce que, quand n croît, $\int f_i^n \, dP$ tend vers $\int f_i \, dP$, tandis que $\int f^n dP$ tend vers $\int f \, dP$ si f a les propriétés indiquées et au contraire augmente indéfiniment si f ne les a pas.

II. — L'intégrale indéfinie.

19. Dans ce qui précède on a été amené, une fonction sommable f étant donnée dans un ensemble mesurable E, à prendre l'intégrale de f dans des ensembles mesurables e. Cette opération conduit à attacher un nombre à chacun de ces ensembles et nous allons étudier un peu cette correspondance définissant ce qu'on peut appeler une *fonction d'ensemble mesurable*.

Pour la brièveté du langage je supprimerai toujours le qualificatif *mesurable* et je supposerai toujours que f est partout définie, ce qu'on réalise d'ailleurs sans changer l'intégrale de f dans les ensembles faisant partie de E, en attribuant à f la valeur zéro aux points ne faisant pas partie de E. La fonction d'ensemble que l'intégration permet d'attacher à une fonction f sera dite l'*intégrale indéfinie* de f.

D'après ce qui précède, cette intégrale indéfinie jouit des deux propriétés que je rappelle :

1° *Elle est absolument continue*, c'est-à-dire que si l'on considère une suite d'ensembles $e_1, e_2, \ldots$ dont les mesures tendent vers zéro, les valeurs correspondantes $\int_{e_1}, \int_{e_2}, \ldots$ pour l'intégrale indéfinie tendent vers zéro.

2° *Elle admet un théorème d'addition*, celui qu'exprime l'égalité

$$\int_{e_1+e_2+\ldots} = \int_{e_1} + \int_{e_2} + \ldots,$$

les e_i étant en nombre fini ou dénombrable et sans point commun deux à deux.

A cause de la condition de continuité, cette seconde propriété peut d'ailleurs, si l'on veut, n'être énoncée que pour le cas de deux ensembles e_i. Si l'on conserve sans changement son énoncé, il est évident que cette propriété contient la suivante : la série $\int_{e_1} + \int_{e_2} + \ldots$ est absolument convergente quels que soient les ensembles $e_1, e_2, \ldots$ sans point commun deux à deux, sans quoi cette série, dans lequel l'ordre des termes est indifférent, n'aurait pas de sens. C'est ce que nous exprimerons en disant que l'intégrale indéfinie est *à variation totale bornée*.

Considérons un ensemble donné e comme la somme d'une infinité dénombrable d'ensembles composants sans point commun de toutes les manières possibles. A chacune de ces manières correspond une série telle que

$$\int_{e_1} + \int_{e_2} + \int_{e_3} + \ldots.$$

Soient π la somme de ceux des termes de cette série qui sont positifs, $-\nu$ la somme de ceux qui sont négatifs ; comme $\pi - \nu = \int_e$, les bornes supérieures P et N des nombres π et ν vérifient la relation $P - N = \int_e$. P s'appelle la relation *positive* de la fonction dans e, $-N$ est sa variation négative, $V = P + N$ sa variation totale. Il est évident que P, N, V sont des fonctions d'ensemble absolument continues, possédant un théorème d'addition et positives. De sorte que *toute fonction absolument continue et possédant un théorème d'addition est la différence de deux fonctions de même nature qui, de plus, ne sont jamais négatives.*

Les définitions des fonctions à variation bornée et des variations auraient bien entendu pu être énoncées en utilisant seulement la décomposition d'un ensemble en un nombre *fini* d'ensembles partiels; car si, par exemple, la série $\int_{e_1} + \int_{e_2} + \ldots$ conduisait à des nombres π et $-\nu$ pour n assez grand, la suite finie

$$\int_{e_1} + \int_{e_2} + \ldots + \int_{e_n} + \int_{e-e_n}$$

aurait donné pour quantités analogues des nombres aussi voisins qu'on le voudrait de π et ν.

Dans le cas de l'intégrale indéfinie, les quantités P, N, V se calculent facilement. Divisons e en l'ensemble e_+ des points où f est positive ou nulle et l'ensemble e_- des autres points. A cette division correspondent des nombres π_1 et ν_1 donnés par

$$\pi_1 - \nu_1 = \int_e f(P)\,dP,$$

$$\pi_1 + \nu_1 = \int_e |f(P)|\,dP.$$

Soit maintenant une division de e en ensembles partiels $e_1, e_2, \ldots$ sans point commun; soient e_{+n}, e_{-n} les ensembles communs à e_+ et e_n, à e_- et e_n. La décomposition de e en les e_i donne évidemment des nombres π et ν au plus égaux à ceux que donne la décomposition de e en les e_{+i}, e_{-i}. Or ces derniers nombres ne sont autres que π_1 et ν_1, donc P et N sont égaux à π_1 et ν_1.

Par suite

$$P = \frac{1}{2}\int_e [\,f(P) + |f(P)|\,]\,dP,$$

$$N = \frac{1}{2}\int_e [\,|f(P)| - f(P)\,]\,dP,$$

$$V = \int_e |f(P)|\,dP.$$

Cette dernière égalité, qui entraîne les deux premières, exprime que *la variation totale d'une intégrale indéfinie dans un ensemble est l'intégrale définie, étendue à cet ensemble, de la valeur absolue de la fonction proposée.*

En considérant V comme une fonction d'ensemble, on peut dire aussi que *la variation totale de l'intégrale indéfinie de f est l'intégrale indéfinie de $|f|$.*

20. Si nous nous étions astreints à ne considérer que des intégrales étendues à des domaines quarrables, l'intégration, appliquée à une fonction fixe f, aurait conduit à une *fonction de domaine*, je sous-entends *quarrable*, qu'on appelle encore *l'intégrale indéfinie de f.*

Cette fonction aurait possédé les propriétés suivantes :

1° Elle est absolument continue, en ce sens que $\int_\delta$ tend vers zéro avec $m(\delta)$, δ étant la somme d'une infinité dénombrable de domaines;

2° Elle possède le théorème d'addition exprimé par l'égalité

$$\int_D = \int_{D_1} + \int_{D_2} + \ldots,$$

quand

$$D = D_1 + D_2 + \ldots,$$

D étant un domaine et le second membre étant indéfini et formé d'ensembles sans point commun deux à deux.

Il résulte de là que la fonction de domaine en question est à variation bornée. Cette propriété était déjà supposée d'ailleurs dans la définition de l'absolue continuité, puisque, l'intégrale indéfinie étant considérée comme une fonction de domaine, $\int_\delta$ ne peut représenter que la somme des intégrales relatives aux domaines composants δ. Les variations d'une fonction de domaine se définissent comme celles d'une fonction d'ensemble et en utilisant une décomposition du domaine en un nombre fini ou dénombrable de domaines composants.

Les variations relatives à l'intégrale d'une fonction donnée, et à un certain domaine peuvent être prises, soit en considérant l'intégrale comme une fonction de domaine, et alors on obtient V', P', N', soit en la considérant comme une fonction d'ensemble, ce qui donne V, P, N.

Il est d'abord évident que V′, P′, N′ sont au plus égaux respectivement à V, P, N. Or, divisons le domaine considéré D en D_+ et D_- comme il a été dit. Enfermons D_+ dans une infinité dénombrable de domaines simples, des polygones s'il n'y a que deux dimensions, contenus dans D et dont la somme des mesures surpasse très peu celle de D_+. Si A est cet ensemble de domaines, on pourra toujours supposer que $\int_{A-D_+} |f(P)|\,dP$ soit inférieure à ε; donc la contribution de $A - D_+$ dans V ne surpasse pas ε. Ne conservons qu'un nombre fini de ces domaines, mais assez pour que a étant l'ensemble de ceux conservés, $\int_{A-a} |f(P)|\,dP$ soit inférieur à ε. Associons de même à D_-, B puis b. Les parties communes à a et b fournissent dans $\int |f(P)|\,dP$ une contribution au plus égale à 2ε; supprimons-les. L'ensemble des a et b restants fournit dans $\int |f(P)|\,dP$ une contribution supérieure à $V - 4\varepsilon$; or cet ensemble est formé d'un nombre fini de domaines sans point commun deux à deux, qu'on peut faire rentrer dans une décomposition de D en domaines partiels et par suite V′ surpasse $V - 4\varepsilon$.

On a donc

$$V = V', \qquad P = P', \qquad N = N'$$

et, *pour la définition des variations d'une intégrale indéfinie dans un domaine, il n'y a pas lieu de distinguer si l'on considère cette intégrale indéfinie comme fonction de domaine ou d'ensemble.*

L'intégrale indéfinie, en tant que fonction d'ensemble ou de domaine, jouit des deux propriétés mises en évidence qu'on peut rappeler en disant que c'est *une fonction absolument continue et additive.* Se donner une fonction d'ensemble absolument continue et additive, c'est, en particulier, se donner une fonction de domaine absolument continue et additive. Réciproquement, si l'on se donne une fonction $\mathcal{F}$ de domaine absolument continue et additive, on définira une fonction d'ensemble absolument continue et additive par le procédé suivant : soit E un ensemble mesurable, enfermons-le dans une infinité dénombrable de domaines δ sans point commun deux à deux et dont la somme des mesures ne surpasse $m(E)$ que de ε au plus; alors on prendra pour $\mathcal{F}(E)$ la limite de $\Sigma\mathcal{F}(\delta)$, quand ε tend vers zéro.

21. Particularisons encore l'ensemble auquel on étend l'intégration; supposons que l'intégrale soit toujours étendue à un *intervalle*, c'est-à-dire à un domaine défini, dans l'espace considéré à k dimensions, par k inégalités de la forme

$$a_\alpha \leqq x_\alpha \leqq b_\alpha \qquad (\alpha = 1, 2, \ldots, k),$$

les a_k et les b_k étant des constantes données.

L'intégration conduit alors à une *fonction d'intervalle* que nous appellerons encore une *intégrale indéfinie* et qui sera absolument continue et additive, donc à variation bornée. Tous ces mots se définissent comme précédemment, le mot *intervalle* remplaçant partout le mot *domaine*. Il suffit de connaître une fonction d'intervalle absolument continue et additive pour pouvoir en déduire une fonction d'ensemble continue et additive.

Soit maintenant $\alpha_1, \alpha_2, \ldots, \alpha_k$ un point quelconque α de l'espace considéré; les inégalités $0 \leqq x_i \leqq \alpha_i$ $(i = 1, 2, \ldots, k)$ définissent un intervalle que j'appelle $\mathrm{I}(\alpha_1, \alpha_2, \ldots, \alpha_k)$ ou plus brièvement $\mathrm{I}(\alpha)$ et je pose

$$\mathcal{F}(\alpha) = \mathcal{F}(\alpha_1, \alpha_2, \ldots, \alpha_k) = \int_{\mathrm{I}(\alpha)} f(\mathrm{P})\, d\mathrm{P}.$$

La fonction

$$\varphi(\alpha) = \mathcal{F}(\alpha) + f_1(\alpha_2, \alpha_3, \ldots, \alpha_k) + f_2(\alpha_1, \alpha_3, \ldots, \alpha_k) + \ldots + f_k(\alpha_1, \alpha_2, \ldots, \alpha_{k-1}),$$

dans laquelle $f_1, f_2, \ldots, f_k$ sont des fonctions arbitraires de $k-1$ variables [1], est appelée l'*intégrale indéfinie de f exprimée à l'aide des k variables $x_1, x_2, \ldots x_k$*.

Voici comment on peut en déduire l'intégrale définie comme fonction d'intervalle. Soit un intervalle $\alpha_i \geqq x_i \geqq \beta_i$; j'appelle A le point $\alpha_1, \alpha_2, \ldots, \alpha_k$; B, le point $\beta_1, \ldots, \beta_k$, et $\mathrm{I}(\mathrm{A}, \mathrm{B})$ l'intervalle considéré, de sorte que $\mathrm{I}(\mathrm{A})$ est identique à $\mathrm{I}(0, \mathrm{A})$. On a évidemment

$$\begin{aligned}\mathrm{I}(\mathrm{A},\mathrm{B}) = {} & \mathrm{I}(\alpha_1, \alpha_2, \ldots, \alpha_k) \\ & -[\mathrm{I}(\beta_1, \alpha_2, \alpha_3, \ldots, \alpha_k) + \mathrm{I}(\alpha_1, \beta_2, \alpha_3 \ldots, \alpha_k) + \ldots + \mathrm{I}(\alpha_1, \ldots, \alpha_{k-1}, \beta_k)] \\ & +[\mathrm{I}(\beta_1, \beta_2, \alpha_3, \ldots, \alpha_k) + \mathrm{I}(\beta_1, \alpha_2, \beta_3 \ldots, \alpha_k) + \ldots + \mathrm{I}(\alpha_1, \ldots, \alpha_{k-2}, \beta_{k-1}, \beta_k)] \\ & - \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & + (-1)^k \mathrm{I}(\beta_1, \beta_2, \ldots, \beta_k).\end{aligned}$$

---

[1] Si $k = 1$, il n'y a qu'une quantité f_1, qui est alors constante.

De là et de la définition de $\varphi(\alpha)$ il résulte qu'on a

$$\begin{aligned}\int_{\mathrm{I(A,B)}} f(\mathrm{P})\,d\mathrm{P} = {} & \varphi(\alpha_1, \alpha_2, \ldots, \alpha_k) \\ & -[\varphi(\beta_1, \alpha_2, \ldots, \alpha_k) + \varphi(\alpha_2, \beta_1, \alpha_3, \ldots, \alpha_k) + \ldots + \varphi(\alpha_1, \alpha_2, \ldots, \alpha_{k-1}, \beta_k)] \\ & + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & + (-1)^k \varphi(\beta_1, \beta_2, \ldots, \beta_k) \\ = {} & \psi(\mathrm{A}, \mathrm{B}).\end{aligned}$$

De sorte que la connaissance de l'intégrale indéfinie de f exprimée à l'aide des variables, $x_1, \ldots, x_k$ suffit pour qu'on déduise l'intégrale indéfinie de f comme fonction d'ensemble.

Cette intégrale indéfinie, exprimée à l'aide de $x_1, \ldots, x_k$, jouit des deux propriétés suivantes :

1° Si l'on considère un nombre fini ou une infinité dénombrable d'intervalles non empiétant $\mathrm{I}(\mathrm{A}_1, \mathrm{B}_1)$, $\mathrm{I}(\mathrm{A}_2, \mathrm{B}_2)$, ..., la somme

$$s = \psi(\mathrm{A}_1, \mathrm{B}_1) + \psi(\mathrm{A}_2, \mathrm{B}_2) + \ldots$$

a un sens, ce qu'on exprime en disant que l'intégrale indéfinie $\varphi(\alpha)$ est à variation bornée ;

2° Si la mesure de $\mathrm{I}(\mathrm{A}_1, \mathrm{B}_1) + \mathrm{I}(\mathrm{A}_2, \mathrm{B}_2) + \ldots$ tend vers zéro, cette somme tend vers zéro, ce qu'on exprime en disant que $\varphi(\alpha)$ est absolument continue.

Il faut d'ailleurs remarquer que cette absolue continuité n'entraîne pas la continuité ordinaire, sauf dans l'hypothèse $k = 1$, mais les seules singularités que peut présenter φ sont en quelque sorte artificielles, car elles ne peuvent provenir que des fonctions $f_1, f_2, \ldots, f_k$ ajoutées à $\mathcal{F}(\alpha)$. $\mathcal{F}(\alpha)$ est continue.

On pourra donc convenir, si on le veut, de ne prendre pour les $f_1, \ldots, f_k$ que des fonctions continues et l'intégrale indéfinie exprimée à l'aide des coordonnées sera alors continue.

Au lieu de l'intégrale indéfinie, exprimée à l'aide des coordonnées données, on pourrait de même, du moins sous de très larges hypothèses, parler de l'intégrale indéfinie exprimée à l'aide de coordonnées curvilignes quelconques. Je laisse cela de côté pour le moment.

En résumé se donner une fonction d'ensemble additive et absolument continue, ou se donner une fonction des k coordonnées à variation bornée et absolument continue sont deux opérations équivalentes. Il nous suffira de connaître par suite l'intégrale indéfinie sous l'une de ses formes pour pouvoir la construire sous un autre quelconque de ses aspects ; aussi le plus souvent parlera-t-on de l'intégrale indéfinie sans préciser comment on la considère.

III. — Les familles régulières de domaines.

22. Nous venons de voir comment, une fonction sommable étant donnée, on lui attache une intégrale indéfinie ; nous allons maintenant nous proposer de trouver la fonction sommable connaissant son intégrale indéfinie. Comme dans cette recherche il n'interviendra pas d'autres propriétés des intégrales indéfinies que celles énoncées ci-dessus, il en résultera en particulier que ces propriétés sont caractéristiques des fonctions qu'on peut regarder comme des intégrales indéfinies.

Tout d'abord, le problème que nous nous proposons a-t-il un sens ? Deux fonctions sommables f_1 et f_2 égales presque partout ont la même intégrale indéfinie; donc il est indispensable de ne pas considérer ces deux fonctions comme différentes. Mais, si l'on fait cette convention, alors *une fonction est déterminée par son intégrale indéfinie ;* ce qui veut dire que deux fonctions sommables f_1, f_2 qui diffèrent aux points d'un ensemble de mesure non nulle ont des intégrales indéfinies différentes. En effet, on peut alors trouver un nombre $\varepsilon > 0$ assez petit pour que l'un des ensembles de points en lesquels on a soit $f_1 > f_2 + \varepsilon$, soit $f_1 < f_2 + \varepsilon$, soit de mesure différente de zéro, et dans un tel ensemble nos deux fonctions n'ont pas la même intégrale définie.

Il résulte de là que, quel que soit le procédé qu'on emploiera pour repasser d'une intégrale indéfinie à la fonction qui lui a donné naissance, ce procédé sera en général en défaut pour un ensemble de points de mesure nulle. Quant à ce procédé on y est conduit en cherchant à imiter le procédé de dérivation qui réussit pour les intégrales indéfinies de fonctions continues.

23. Une fonction d'ensemble $\mathcal{F}(E)$ étant donnée, nous appellerons

dérivée de cette fonction en un point P la limite, si elle existe, du rapport $\frac{\mathfrak{F}(D)}{m(D)}$, D étant un domaine contenant P et dont on fait tendre toutes les dimensions vers zéro. Lorsque cette limite n'existe pas, les plus petite et plus grande limites d'indétermination de ce rapport sont les nombres dérivés inférieur et supérieur en P. Ainsi la dérivée et les nombres dérivés d'une fonction d'ensemble sont des fonctions de points qui résultent de la comparaison de la fonction proposée à une fonction type, la fonction $m(E)$.

Mais ces définitions sont, on va le voir de suite, trop larges ; recherchons, en effet, dans quel cas, avec elles, la dérivée d'une intégrale indéfinie est la fonction intégrée ? Seulement dans le cas où, en négligeant les points d'un certain ensemble de mesure nulle, la fonction f est continue au point P étudié.

Il est d'abord évident que, si cette condition est remplie, f est en P la dérivée de son intégrale indéfinie; mais supposons que la condition énoncée ne soit pas vérifiée. Alors, dans tout domaine Δ entourant P, si petit soit-il, on pourra trouver un ensemble e de points de mesure non nulle en lesquels on a soit $f > f(P) + \varepsilon$, soit $f > f(P) - \varepsilon$, pourvu que $\varepsilon > 0$ ait été choisi assez petit. Supposons que ce soit la première inégalité qui est vérifiée aux points de e et soit σ un nombre tel que $\int |f(\varphi)| d\varphi$ étendue à un ensemble quelconque de mesure σ soit toujours inférieure à $\frac{1}{2}\varepsilon m(e)$. Enfermons e dans des domaines δ intérieurs à Δ et dont la mesure surpasse celle de e de moins de σ. Enfin soit D un domaine intérieur à Δ contenant tous les δ et le point P et dont la mesure surpasse celle de e de σ au plus. Alors on a

$$\mathfrak{F}(D) = \int_D f(Q)\, dQ > [f(P) + \varepsilon]\, m(e) - \frac{1}{2}\varepsilon m(e) = \left[f(P) + \frac{\varepsilon}{2}\right] m(e).$$

Donc l'emploi de ces domaines D conduirait pour valeur limite de $\frac{\mathfrak{F}(D)}{m(D)}$ à une quantité supérieure à $f(P) + \frac{\varepsilon}{2}$.

24. Pour notre but, il est donc nécessaire de restreindre la classe des domaines D employés. L'étude des fonctions d'une seule variable montre qu'une telle restriction n'est pas nécessaire dans ce cas, et en

effet notre objection ne porte que s'il y a plusieurs dimensions, puisque c'est alors seulement qu'on peut affirmer l'existence d'un domaine D contenant les δ et de mesure aussi voisine qu'on le veut de la somme $\Sigma m(\delta)$. Il y a là une différence essentielle entre le cas de plusieurs dimensions et celui d'une seule.

Notre raisonnement montre par exemple que, dans le cas de plusieurs variables, on ne modifie en rien la portée de la définition provisoire de la dérivée, qui a été donnée plus haut, en y remplaçant le domaine D par un domaine ne contenant pas P ou par un ensemble mesurable quelconque ; il est bien évident que cela n'est plus vrai dans le cas d'une seule dimension. Cette différence subsistera après que nous aurons restreint la classe des domaines D, de sorte que notre définition sera, dans le cas d'une seule dimension, plus particulière que la définition classique.

La différence qui apparaît là entre le cas d'une seule ou de plusieurs dimensions aurait pu être mise en évidence dans le Chapitre précédent. Dans le cas d'une seule variable, une fonction additive de domaine $\mathfrak{F}(D)$ continue en ce sens que $\mathfrak{F}(D)$ tend vers zéro quand $m(D)$ tend vers zéro n'est pas nécessairement absolument continue ; on en trouvera un exemple dans les dernières lignes de la page 129 de mes *Leçons sur l'Intégration*. Au contraire, ce genre de continuité ne diffère pas de l'absolue continuité lorsqu'il y a plusieurs dimensions. On peut encore énoncer la différence entre les deux cas en remarquant que, dans le cas de plusieurs dimensions, un domaine D peut avoir une mesure qui tend vers zéro sans qu'il en soit de même de toutes ses dimensions, ce qui n'est évidemment pas possible s'il n'y a qu'une seule dimension. Si l'on disait d'une fonction additive de domaine $\mathfrak{F}(D)$ qu'elle est continue si $\mathfrak{F}(D)$ tend vers zéro, quand toutes les dimensions de D tendent vers zéro, ce genre de continuité n'entraînerait pas l'absolue continuité.

25. Revenons à la définition de la dérivée. Si l'on se reporte au raisonnement que nous avons fait pour montrer que notre définition provisoire est trop large, on voit que, avec notre définition provisoire, pour la fonction égale à m partout, sauf aux points d'un ensemble mesurable E, en lesquels elle égale 1, la dérivée de l'intégrale indé-

finie n'existera pas en tout point P tel que, dans toute hypersphère de centre P, E et son complémentaire aient des parties de mesures nulles α et β. Et notre définition ne fait aucune différence entre le cas où $\frac{\alpha}{\beta}$ tend vers zéro avec le rayon de l'hypersphère et les autres. Cependant, si $\frac{\alpha}{\beta}$ tend vers zéro, l'influence de E dans l'intégrale indéfinie doit, il semble, être considérée comme négligeable autour de P vis-à-vis de celle de son complémentaire. L'influence de E dans la dérivée en P sera négligeable si l'on convient de n'utiliser pour la définition de cette dérivée que des domaines ou ensembles D tels que le rapport de la mesure D à celle de l'hypersphère de centre P contenant D ne tende pas vers zéro avec le rayon de l'hypersphère.

Une classe de tels domaines sera formée de ceux contenant P et tels que le rapport entre le plus petit rayon vecteur issu de P à la frontière et le plus grand surpasse une quantité positive fixe.

Comme classe plus particulière on peut citer, dans le cas de deux dimensions, celle formée à l'aide de carrés quelconques contenant P à son intérieur ou sur sa frontière, de cercles quelconques placés de même ou encore vus de P sous un angle supérieur à un angle donné, de secteurs circulaires de centre P et d'ouverture supérieure à un angle donné, etc.

Relativement à ces domaines nous démontrerons d'abord un théorème capital que M. Vitali a établi pour le cas particulier des intervalles dont toutes les dimensions sont égales (carrés, cubes, etc.); mais sa démonstration s'applique sans changement au cas général ([1]). Pour énoncer simplement le théorème, *nous conviendrons d'appeler* ensemble *ou* domaine régulier *pour une valeur* μ *tout ensemble mesurable ou domaine* D *tel que, si la plus petite hypersphère qui le contient est* S, *on ait la relation*

$$m(D) > \mu m(S),$$

μ *étant une constante positive fixe.* Le diamètre de S sera ce que nous appellerons le *diamètre* de D; quand $m(D)$ tend vers zéro, ce diamètre tend nécessairement vers zéro. *Dire qu'une famille de domaines ou*

([1]) La démonstration qu'on lira plus loin est presque copiée sur celle de M. Vitali.

d'ensembles est régulière, c'est dire qu'on peut trouver une valeur $\mu > 0$ telle que les domaines ou ensembles soient réguliers pour cette valeur μ du paramètre.

26. Nous pouvons maintenant démontrer le théorème de M. Vitali :

Supposons donnée une famille régulière $\mathcal{F}$ de domaines quarrables et un ensemble mesurable E *de mesure finie tel que chacun de ses points est intérieur à une infinité de domaines de $\mathcal{F}$ de diamètres aussi petits que l'on veut. Alors on peut trouver dans la famille $\mathcal{F}$ un nombre fini ou une infinité dénombrable de domaines sans point commun deux à deux et dont la somme des mesures n'est pas inférieure à* m(E).

Il suffira évidemment de supposer que E est borné, sinon l'on décomposerait E en une somme de parties bornées et l'on opérerait sur chacune de ces parties comme il va être indiqué.

Soit μ la valeur par rapport à laquelle les domaines sont réguliers. Considérons les domaines de $\mathcal{F}$ dont le diamètre est compris entre $\frac{\varepsilon}{2^p}$ et $\frac{\varepsilon}{2^{p-1}}$, ε ayant été arbitrairement choisi. Soit E_p la partie commune à E et à ces domaines.

Il est évident que la mesure de $E_0 + E_1 + \ldots + E_p$ tend en croissant vers celle de E, quand p croît indéfiniment.

Soit P_1 un point quelconque de E_0 ; soit D_1^0 un domaine faisant partie de $\mathcal{F}$, contenant P_1 et de diamètre compris entre ε et 2ε; soient S_1^0 la plus petite hypersphère contenant D_1^0 et Σ_1^0 l'hypersphère concentrique de rayon 3ε. On a

$$m(D_1^0) > \mu m(S_1^0) \geqq \frac{\mu}{6^k} m(\Sigma_1^0).$$

Soit P_2 un point de E_0 extérieur à Σ_1^0, s'il en existe. A P_2 on attache de même D_2^0 et Σ_2^0. Puis, si l'on peut trouver un point P_3 de E_0 extérieur à $\Sigma_1^0 + \Sigma_2^0$, à P_3 on attachera D_3^0 et Σ_3^0. Et ainsi de suite.

On ne sera arrêté que lorsque $\Sigma_0 = \Sigma_1^0 + \Sigma_2^0 + \ldots$ contiendra tout E_0. Or cela arrivera au bout d'un nombre fini d'opérations, car les domaines D_1^0, D_2^0, ... sont évidemment sans point commun deux à deux d'après leur construction, de sorte que s'ils sont en nombre p leur ensemble d_0 a une mesure au moins égale à $\frac{p\mu}{6^k}\, m(\Sigma_1^0)$; or ils sont tous dans la même région bornée de l'espace.

Σ_0 couvrant ainsi tout E_0, désignons par $\mathcal{E}_0$ la partie de E contenu dans Σ_0 ; $\mathcal{E}_0$ contient E_0 et l'on a

$$m(d_0) \geqq \frac{\mu}{6^k}\, m(\Sigma_0) \geqq \frac{\mu}{6^k}\, m(\mathcal{E}_0).$$

Soit E'_1 la partie de $E_0 + E_1$ extérieure à Σ_0 ; on opère sur elle comme sur E_0, ε étant remplacée par $\frac{\varepsilon}{2}\cdot$ On trouve ainsi des domaines D_1^1, $D_2^1, \ldots$, en nombre fini sans point commun entre eux ni avec les précédents; soit d_1 l'ensemble total des D_i^0 et D_j^1. On a aussi des sphères Σ_j^1 qui avec les précédentes couvrent Σ_1. Soit $\mathcal{E}_1$ la partie de E contenue dans Σ_1 ; $\mathcal{E}_1$ contient $E_0 + E_1$ et l'on a

$$m(d_1) \geqq \frac{\mu}{6^k}\, m(\Sigma_1) \geqq \frac{\mu}{6^k}\, m(\mathcal{E}_1).$$

Puis on opérera de même sur la partie E''_2 de $E_0 + E_1 + E_2$ extérieure Σ_1, ε étant cette fois remplacée par $\frac{\varepsilon}{2^2}$, et ainsi de suite.

Nous choisissons ainsi dans la famille $\mathcal{F}$ une suite finie ou dénombrable de domaines sans point commun deux à deux et qui a une mesure totale au moins égale à $\frac{\mu}{6^k}\, m(E)$. Si donc λ est un nombre quelconque inférieur à $\frac{\mu}{6^k}$, on sait trouver un nombre fini de domaines sans point commun deux à deux et dont la mesure totale surpasse $\lambda\, m(E)$.

Complétons cette première partie de la démonstration par une remarque (¹). Supposons que des domaines quarrables $\Delta_0, \Delta_1, \ldots, \Delta_m$ en nombres finis étant donnés et E étant extérieur à ces domaines, on désire ne se servir que de domaines sans points communs avec les Δ_i;

(¹) Relativement à cette première partie, on peut aussi remarquer avec M. Vitali qu'elle suppose seulement chaque point de E intérieur à l'un des domaines de $\mathcal{F}$.

La limite trouvée $\frac{\mu}{6^k}$ est trop grande en remplaçant la progression des $\frac{\varepsilon}{2^p}$ par une progression moins rapidement décroissante; on trouvera $\frac{\mu}{3^k}$ pour limite supérieure de λ.

rien ne sera changé à notre conclusion. Il n'y a pas de difficulté, en effet, si l'on opère, non sur E, mais sur les points de E distants des Δ_i de plus de 2ε, et la mesure des points de E ainsi laissés de côté tend évidemment vers zéro avec ε. Il suffit donc de prendre ε suffisamment petit et notre conclusion subsiste.

Ceci posé, soit δ_0 l'ensemble fini des domaines obtenus comme il a été dit plus haut; on a $m(\delta_0) > \lambda m(\mathrm{E})$.

D'autre part, si E' est la partie de E extérieure à δ_0, on a

$$m(\mathrm{E}') \geqq m(\mathrm{E}) - m(\delta_0);$$

nous supposons le second membre positif, sans quoi le théorème serait démontré.

Or, d'après ce qui précède, on peut dans $\mathcal{F}$ choisir des domaines δ_1 en nombre fini sans point commun deux à deux ni avec les δ_0 et tels que l'on ait $m(\delta_1) > \lambda m(\mathrm{E}')$. Donc

$$m(\delta_0) + m(\delta_1) > \lambda m(\mathrm{E}) + \lambda[m(\mathrm{E}) - m(\delta_0)];$$

et la partie E'' de E extérieure à $\delta_1 + \delta_0$ a une mesure au moins égale à $m(\mathrm{E}) - m(\delta_0) - m(\delta_1)$, quantité qu'on supposera positive; de sorte que l'opération analogue sur E'' conduira à δ_2 tel que

$$\begin{aligned} &m(\delta_0) + m(\delta_1) + m(\delta_2) \\ &\quad > \lambda m(\mathrm{E}) + \lambda[m(\mathrm{E}) - m(\delta_0)] + \lambda[m(\mathrm{E}) - m(\delta_0) - m(\delta_1)] \end{aligned}$$

et par suite, en continuant ainsi, on fera bien le choix demandé.

27. On peut étendre un peu la portée du théorème précédent en l'énonçant sous la forme :

Soit un ensemble mesurable E *dont chaque point appartient à une infinité de domaines de diamètres aussi petits qu'on veut, formant une famille régulière et extraits d'un ensemble donné $\mathcal{F}$ de domaines. Alors on peut trouver dans $\mathcal{F}$ un nombre fini ou une infinité dénombrable de domaines, sans point commun deux à deux et dont la somme des mesures n'est pas inférieure à m* (E).

En effet, soit $\mathcal{F}_p$ la famille formée des domaines de $\mathcal{F}$ qui sont réguliers pour la valeur $\frac{1}{p}$ du paramètre. Soit d'ailleurs E_p l'ensemble des points de E contenus dans une infinité de domaines de diamètres aussi petits qu'on veut et appartenant à $\mathfrak{F}_p$. E_p est mesurable et $m(E - E_p)$ tend vers zéro avec $\frac{1}{p}$. Or, d'après ce qui précède, on peut de $\mathfrak{F}_1$ tirer un nombre fini de domaines de mesure supérieure à $m(E_1) - 1$; puis de $\mathfrak{F}_2$ un nombre fini de domaines, sans point commun avec les précédents ni entre eux, de mesure supérieure à $m(\mathcal{E}_2) - \frac{1}{2}$, $\mathcal{E}_2$ étant la partie de E_2 extérieure aux domaines déjà employés. Et ainsi de suite.

Comme il est évident que E_p contient E_{p-1} il s'ensuit qu'on arrive bien à effectuer le choix demandé.

Supposons, comme précédemment, que nous ne sortions pas d'une certaine région bornée R; alors il est évident que $m(\mathcal{E}_p)$ doit tendre vers zéro, sans quoi les domaines successifs choisis, qui ont une mesure supérieure à $\Sigma\left[m(\mathcal{E}_p) - \frac{1}{p}\right]$, aurait une mesure totale indéfinie. Donc *les domaines choisis, comme il vient d'être dit, couvrent tout* E *aux points d'un ensemble de mesure nulle près.*

Cette région R, qui n'est pas nécessairement connexe, peut être prise de mesure aussi peu supérieure à $m(E)$ qu'on le veut; donc on peut faire en sorte que *la mesure totale des domaines choisis surpasse* $m(E)$ *d'aussi peu qu'on le veut.*

Cela suggère un énoncé qu'on peut étendre au cas des ensembles réguliers et que je me borne à formuler :

Un ensemble mesurable E *et une famille* $\mathfrak{F}$ *d'ensembles mesurables étant donnés, supposons qu'à tout point* P *de* E *corresponde une famille régulière d'ensembles, extraite de* $\mathfrak{F}$ *et contenant des ensembles de diamètre aussi petit qu'on le voudra, qui tous contiennent* P. *Il est alors possible d'extraire de* $\mathfrak{F}$ *un nombre fini ou dénombrable d'ensembles qui couvrent tout* E, *à un ensemble de mesure nulle près, et dont la somme des mesures surpasse* $m(E)$ *d'aussi peu qu'on le désire.*

Dans tous ces énoncés, la condition, qu'il s'agisse de domaines ou d'ensembles réguliers, est essentielle. Je ne veux pas dire qu'il est impossible de la remplacer par une condition plus large, mais on ne pourrait pas la supprimer simplement sans que les énoncés ne soient faux. On verra en effet que la condition de régularité n'interviendra plus dans la suite que par ces énoncés et nous serons conduits à des théorèmes sur la dérivation des intégrales qui seraient faux si l'on considérait des domaines quelconques au lieu de domaines réguliers ([1]).

IV. — La dérivation des fonctions absolument continues.

28. Précisons les définitions des dérivées et nombres dérivés. Une fonction d'ensemble $\mathcal{F}(E)$ étant donnée, pour définir ses nombres dérivés en un point P, on considère un ensemble mesurable E contenant P ([2]) et l'on forme le rapport de $\frac{\mathcal{F}(E)}{m(E)}$. Puis on fait varier E de façon qu'il ne cesse pas d'appartenir à une famille régulière, d'ailleurs quelconque, et que son diamètre tende vers zéro. Les plus petites et plus grandes limites correspondantes de $\frac{\mathcal{F}(E)}{m(E)}$ sont les deux nombres dérivés. La dérivée est la valeur commune de ces deux nombres lorsqu'ils sont égaux.

Cette définition, appliquée à l'intégrale indéfinie $\mathcal{F}(x)$ d'une fonction d'une variable exprimée à l'aide de cette variable x, apparaît alors comme assez différente de la définition classique. Toutes les fois que $\mathcal{F}(x)$ a une dérivée au sens du mot qui vient d'être défini, elle en a une également au sens habituel, et ces deux nombres sont égaux; mais la réciproque n'est pas vraie. D'ailleurs la définition proposée ici conduit à effectuer sur $\mathcal{F}(x)$ un calcul qui s'appuie essentiellement sur le fait que $\mathcal{F}(x)$ est à variation bornée.

On reconnaîtra facilement que si $\mathcal{F}(x)$ intégrale indéfinie de $f(x)$ a, au sens actuel, une dérivée égale à $f(x_0)$ au point x_0; de même, l'in-

---

([1]) On ne sait pas si les énoncés précédents subsistent pour le cas où la famille $\mathcal{F}$ est formée d'intervalles quelconques. C'est pour ce cas qu'il faudrait les justifier pour qu'on puisse utiliser la définition de la dérivée qui, pour le cas de deux variables, consiste à prendre la limite de

$$\frac{f(x+h,y+k)+f(x,y)-f(x+h,y)-f(x,y+k)}{hk}$$

quand h et k tendent vers zéro, *d'une façon quelconque*.

([2]) La condition que E contient P ne modifie ni $\mathcal{F}(E)$, ni $m(E)$, mais quand on en tient compte la plus petite hypersphère contenant E contient P, et par suite c'est à une hypersphère contenant P que l'on compare E pour savoir si la famille des E est régulière.

tégrale indéfinie de $|f(x) - \alpha|$ a, en x_0, une dérivée égale à $|f(x_0) - \alpha|$, quelle que soit la constante α. D'ailleurs si, quelle que soit la constante α, l'intégrale indéfinie de $|f(x) - \alpha|$ a une dérivée égale à $|f(x_0) - \alpha|$ au point x_0 lorsqu'on définit la dérivée à la façon habituelle, il en est encore de même quand on adopte la définition proposée ici.

Cette remarque nous servira dans ce qui suit. Certaines de nos démonstrations ne sont en effet pas exactes lorsqu'il s'agit de domaines à une seule dimension parce que, dans ce cas, comme je l'ai déjà dit, suivant qu'on emploie des ensembles quelconques ou seulement des domaines, la définition de la dérivée prend des sens différents et les démonstrations supposent qu'il est indifférent de raisonner sur des ensembles ou des domaines. Pour nos raisonnements, c'est donc la définition classique qui devrait être adoptée dans le cas d'une seule dimension; mais la remarque faite précédemment permet d'utiliser un raisonnement, qui m'a servi dans des recherches sur les séries trigonométriques et que j'aurai plus loin l'occasion de rappeler, et l'on étend ainsi les résultats au cas où l'on définit la dérivée à l'aide d'ensembles.

Au reste, en utilisant la dernière forme de l'énoncé du théorème de M. Vitali, qui s'applique aux ensembles aussi bien qu'aux domaines, on pourrait exposer ces démonstrations de façon qu'elles conviennent de suite à tous les cas.

29. *Les nombres dérivés d'une fonction d'ensemble absolument continue, additive sont des fonctions mesurables.* — Pour démontrer, par exemple, que le nombre dérivé supérieur est mesurable, je m'appuierai sur ce fait évident qu'une hypersphère S étant donnée, si l'on choisit dans S un ensemble E de mesure supérieure à $\mu\, m(\mathrm{S})$, μ étant positif fixe, la limite supérieure de $\frac{\mathcal{F}(\mathrm{E})}{m(\mathrm{E})}$, pour E variant dans S fixe, varie d'une manière continue avec S, μ étant fixe, quand $\mathcal{F}$ est absolument continue. De l'absolue continuité de $\mathcal{F}$ il résulte en effet que si E désigne la partie commune à un ensemble quelconque $\mathcal{E}$ et à S, et si S varie assez peu pour que $m(\mathrm{E})$ varie de moins de ε, $\mathcal{F}(\mathrm{E})$ variera de moins d'une quantité η qui tend vers zéro avec ε, et la proposition énoncée s'en déduit. Je désigne par $f(\mathrm{P}, \mu, r)$ cette limite supérieure

pour l'hypersphère de centre P et de rayon r; c'est une fonction continue de $k+1$ variables qui sont les k coordonnées du centre et le rayon de l'hypersphère S. En tant que fonction de μ, c'est une fonction non croissante.

Si donc on donne à r une infinité dénombrable de valeurs parmi lesquelles se trouvent les valeurs $r=\frac{1}{p}$ et entre $\frac{1}{p}$ et $\frac{1}{p+1}$ assez de valeurs de r pour que, entre deux d'entre elles, l'oscillation de $f(\mathrm{P}, \mu, r)$ en tant que fonction de r soit inférieure à $\frac{1}{p}$, P étant variable quelconque et μ étant fixe, les fonctions de P ainsi déterminées ont même limite supérieure pour $r=0$ que $f(\mathrm{P}, \mu, r)$. Soient $f_1, f_2, \ldots$ ces fonctions ; elles sont continues en tant que fonction de P.

Formons maintenant la suite non décroissante $s_1, s_2, s_3, \ldots$ dans laquelle on a

$$s_1 = f_1, \qquad \ldots, \qquad s_p = s_{p-1} + \frac{1}{2}(f_p - s_{p-1}) + \frac{1}{2}|f_p - s_{p-1}|.$$

Elle a une limite F_1 qui est une fonction mesurable de P. Soit F_i la fonction analogue relative à la suite $f_i, f_{i+1}, \ldots$. La suite des F_i est non croissante, donc convergente; sa limite, évidemment mesurable, sera désignée par f^μ.

f^μ est une fonction non croissante de μ. De sorte que la suite f^1, $f^{\frac{1}{2}}, f^{\frac{1}{3}}, \ldots$ est croissante; elle a donc une limite finie ou non, laquelle est mesurable. Or cette limite est le nombre dérivé supérieur cherché.

La proposition est donc démontrée. Mais la démonstration suppose expressément que les nombres dérivés aient été définis comme il a été dit; si pour leur définition on s'était astreint à n'utiliser qu'une famille particulière d'ensembles ou domaines réguliers, si par exemple on définissait les nombres dérivés par la considération de familles régulières d'intervalles, il faudrait reprendre la démonstration. Des modifications presque insignifiantes suffisent le plus souvent; on verra d'ailleurs que le choix d'une famille particulière d'ensembles réguliers ne peut modifier la valeur qu'il conduit à attribuer aux nombres dérivés qu'en un ensemble de points de mesure nulle.

30. *Lorsqu'une fonction d'ensemble additive et absolument continue*

est positive pour un domaine quarrable D, *il y a dans* D *un ensemble de mesure non nulle de points où ses deux nombres dérivés sont positifs.*

Si cela n'était pas, à chaque point P de D, sauf aux points d'un ensemble de mesure nulle, on pourrait attacher une suite régulière de domaines δ, intérieurs à D et contenant P, dont les diamètres tendent vers zéro et tels que l'on ait toujours, $\mathcal{F}$ étant la fonction donnée,

$$\frac{\mathcal{F}(\delta)}{m(\delta)} < \frac{\mathcal{F}(D)}{m(D)}.$$

Le théorème de M. Vitali nous permet d'affirmer l'existence d'une suite de δ, non empiétants les uns sur les autres, et couvrant tout D aux points d'un ensemble E de mesure nulle près. En considérant D comme la somme de ces δ et de E et en appliquant les propriétés d'additivité et d'absolue continuité, on voit que $\mathcal{F}(D)$ est égale à la forme $\Sigma\mathcal{F}(\delta)$ étendue aux δ employés. Mais d'autre part on a

$$\Sigma\mathcal{F}(\delta) < \frac{\mathcal{F}(D)}{m(D)}\Sigma m(\delta) = \mathcal{F}(D),$$

d'où une contradiction qui justifie notre énoncé.

La démonstration prouve même plus : *il y a dans* D *un ensemble de mesure non nulle de points en lesquels aucun des deux nombres dérivés n'est inférieur à* $\frac{\mathcal{F}(D)}{m(D)}$ *et, de même, un ensemble de mesure non nulle de points en lesquels aucun des deux nombres dérivés n'est supérieur à* $\frac{\mathcal{F}(D)}{m(D)}$. A partir de là on généraliserait facilement la théorie classique des nombres dérivés des fonctions d'une variable.

Nous donnerons encore à la proposition qui nous occupe la forme suivante :

Une fonction d'ensemble additive et absolument continue, dont un nombre dérivé est presque partout positif ou nul (négatif ou nul), est constamment positive ou nulle (négative ou nulle). Si d'ailleurs le nombre dérivé était presque partout nul, la fonction serait constamment nulle.

Si en effet la fonction considérée $\mathcal{F}$ est positive pour un ensemble E, elle l'est aussi pour un domaine Δ contenant E et assez petit et, par suite, il y a dans Δ un ensemble de points dont la mesure est positive et en lequel les deux nombres dérivés de $\mathcal{F}$ sont positifs.

31. *Une fonction d'ensemble absolument continue et additive a une dérivée finie et déterminée presque partout et elle est l'intégrale indéfinie d'une fonction égale à cette dérivée, là où elle est finie et déterminée, ayant une valeur quelconque ailleurs.*

Une fonction de la nature considérée étant la différence de deux fonctions positives, il nous suffira de raisonner sur une fonction $\mathcal{F}(E)$ non négative. Soit $D(\mathcal{F})$ le nombre dérivé que l'on étudie.

ε étant pris arbitrairement positif, soit E_i $(i = 0, 1, 2, \ldots)$ l'ensemble des points en lesquels on a

$$i\varepsilon \leqq D(\mathcal{F}) < (i+1)\varepsilon.$$

Soit E_∞ l'ensemble des points où $D(\mathcal{F}) = \infty$.

Nous nous occupons d'un certain domaine borné Δ par rapport auquel nous prendrons les complémentaires. Enfermons le complémentaire $C(E_i)$ de E_i à l'intérieur de domaines dont la mesure surpasse celle de $C(E_i)$ de $\frac{\varepsilon}{2^i}$ au plus. Soit $\mathcal{E}_i$ la partie de E_i extérieure à ces domaines; on a

$$m(\mathcal{E}_i) \geqq m(E_i) - \frac{\varepsilon}{2^i}.$$

Enfermons d'autre part $C(E_\infty)$ à l'intérieur de domaines dont la mesure ne surpasse celle de $C(E_\infty)$ que de ε au plus et soit $\mathcal{E}_\infty$ la partie de E_∞ extérieure à ces domaines; on a

$$m(\mathcal{E}_\infty) \geqq m(E_\infty) - \varepsilon.$$

Posons $f_i(P) = 0$ quand P n'appartient pas à $\mathcal{E}_i$, et $f_i(P) = i\varepsilon$ quand P est point de $\mathcal{E}_i$ $(i = 0, 1, 2, \ldots)$. Posons $f_\infty(P) = 0$ si P n'est pas point de $\mathcal{E}_\infty$, et $f_\infty(P) = n\varepsilon$ si P est point de $\mathcal{E}_\infty$. Enfin soit

$$\varphi(E) = \sum_{i=0}^{i=n-1} \int_E f_i(P)\, dP + \int_E f_\infty(P)\, dP.$$

Cette fonction dépend de deux paramètres n et ε.

Par construction, que i soit fini ou non, $\int_E f_i(P)\,dP$ a une dérivée nulle pour les points ne faisant pas partie de E_i et a, en tout point, des nombres dérivés au plus égaux à $i\varepsilon$ si i est fini, au plus égaux à $n\varepsilon$ si $i=\infty$. $\varphi(E)$ a donc en tous points ses nombres dérivés inférieurs $D(\mathcal{F})$. Or si $\Lambda(\varphi)$ désigne le nombre dérivé supérieur de φ, on a évidemment

$$D(\mathcal{F}-\varphi) \geqq D(\mathcal{F}) - \Lambda(\varphi),$$

donc $D(\mathcal{F}-\varphi)$ n'est jamais négatif ; $\mathcal{F}-\varphi$ est donc positive ou nulle.

Examinons en particulier la valeur $\varphi(\Delta)$; on a

$$\varphi(\Delta) \geqq \int_{\mathcal{E}_1+\mathcal{E}_2+\ldots+\mathcal{E}_{n-1}} [D(\mathcal{F})-\varepsilon]\,dP + n\varepsilon\, m(\mathcal{E}_\infty).$$

Cela est vrai, si petit que soit ε, si grand que soit $n\varepsilon$; donc $m(\mathcal{E}_\infty)$ est nulle et par suite il en est de même de $m(E_\infty)$; par suite, quand ε et $\frac{1}{n\varepsilon}$ tendent vers zéro, $m(\mathcal{E}_1+\mathcal{E}_2+\ldots+\mathcal{E}_{n-1})$, quantité supérieure à $m(\mathcal{E}_1+\ldots+\mathcal{E}_{n-1})-\varepsilon$, tend vers $m(\Delta)$ et par suite l'intégrale du second membre de l'inégalité précédente tend vers $\int_{\Delta-m(E_\infty)} D(\mathcal{F})\,dP$ ou, si l'on désigne par $D'(P)$ la fonction égale à $D(\mathcal{F})$ sauf aux points de E_∞ où $D'(P)=0$, vers

$$\int_\Delta D'(P)\,dP,$$

quantité qui par suite est finie.

On a donc

$$\mathcal{F}(\Delta) \geqq \varphi(\Delta) \geqq \int_\Delta D'(P)\,dP.$$

Une conclusion analogue étant valable aussi quel que soit le domaine qu'on met à la place de Δ, on voit qu'on a, pour tout ensemble E,

$$\mathcal{F}(E) \geqq \int_E D'(P)\,dP.$$

Enfermons maintenant les points de E_i à l'intérieur de domaines D_i dont la mesure ne surpasse celle de E_i que de $\frac{1}{(i+1)2^i}$ au plus, et soit F_i

la fonction égale à $(i+1)\varepsilon$ aux points de D_i et nulle ailleurs; i est entier et fini quelconque. Posons

$$\Phi(E)=\sum\int_E F_i(P)\,dP,$$

la somme étant étendue à toutes les valeurs finies entières de i. Il est évident qu'aux points de E_i, $\Phi(E)$ a son plus petit nombre dérivé $\lambda(\Phi)$ au moins égal à $(i+1)\varepsilon$, donc supérieur à $D(\mathcal{F})$ et, comme on a

$$D(\mathcal{F}-\Phi)\leqq D(\mathcal{F})-\lambda(\Phi),$$

on voit que $D(\mathcal{F}-\Phi)$ est négatif partout, sauf peut-être aux points de l'ensemble de mesure nulle E_∞; donc $\mathcal{F}-\Phi$ est partout négative, car Φ étant évidemment absolument continue $\mathcal{F}-\Phi$ l'est aussi. Or on a évidemment

$$\int_E D'(P)\,dP<\Phi(E)\leqq\int_E[D'(P)+\varepsilon]\,dP+\sum(i+1)\varepsilon\frac{1}{(i+1)2^i},$$

d'où il résulte

$$\mathcal{F}(E)\leqq\int_E D'(P)\,dP.$$

En rapprochant les deux parties du raisonnement, on conclut enfin

$$\mathcal{F}(E)=\int_E D'(P)\,dP.$$

Mais comme la conclusion analogue s'applique à l'autre nombre dérivé de $\mathcal{F}$, on voit que ce second nombre est presque partout égal à $D(\mathcal{F})$, ou à D', et le théorème est démontré.

D'après ce qui a été dit, la définition de la dérivée est la définition ordinaire lorsqu'il n'y a qu'une dimension. D'ailleurs, dans ce cas, les théorèmes sur les nombres dérivés qui ont été utilisés sont bien connus et par suite le théorème précédent est, pour ce cas, démontré sans l'emploi du théorème de M. Vitali.

32. Dans mes *Leçons sur l'Intégration*, j'ai donné, pour le cas d'une seule variable, un raisonnement utilisant ce que j'ai appelé des *chaînes d'intervalles*. Le raisonnement que je vais indiquer pour le cas général

est formé d'une partie exactement calquée sur les démonstrations que je viens de rappeler, utilisant par suite le procédé des chaînes, mais un raisonnement complémentaire est indispensable et, au cours de ce raisonnement, intervient sous une forme plus ou moins détournée le théorème de M. Vitali.

Je suppose, pour simplifier, qu'il n'y ait que deux variables x et y et qu'on ne s'occupe pas de l'extérieur du carré C, $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$. Soit Γ la courbe de M. Hilbert (¹) couvrant tout C et définie à l'aide d'un paramètre t variant de 0 à 1. Cette courbe a des points doubles et quadruples, qui sont parmi ceux dont une ou deux des coordonnées sont de la forme $\frac{a}{3^\alpha}$, a et α étant entiers. Tout carré de la forme $\frac{a}{3^p} \leqq x \leqq \frac{a+1}{3^p}$, $\frac{b}{3^p} \leqq y \leqq \frac{b+1}{3^p}$, a, b, p étant entiers, est un arc de Γ qui correspond à un segment de longueur $\frac{1}{9^p}$ de l'axe des t.

A un segment (a, b) de l'axe des t correspond un arc de Γ couvrant un domaine $\delta(a, b)$ qui est quarrable, car ses points frontières sont des points multiples de Γ et les points correspondant à $t = a$ et $t = b$.

Si l'on exprime x, y en fonction de t, à une fonction $f(x, y)$ correspond une fonction de t qu'on peut noter $f(t)$ et à l'intégrale indéfinie $\mathcal{F}(\mathrm{E})$ de $f(x, y)$ correspond la fonction $\mathcal{G}(t) = \mathcal{F}[\delta(0, t)]$, et ce sont les relations entre $f(t)$ et $\mathcal{G}(t)$ qu'il s'agit d'obtenir. Pour cela, on peut calquer les raisonnements de mes *Leçons sur l'Intégration* (**4**), et des Notes des Lincei (**5**, **6**, **7**), en effectuant quelques changements de mots.

A un intervalle (en x) on fera correspondre un intervalle (a, b) en t, donc un arc de Γ; à la mesure $b - a$ de cet intervalle, on fait correspondre le nombre $m[\delta(a, b)]$, de sorte qu'à un rapport incrémentiel $\frac{\mathcal{F}(x+h) - \mathcal{F}(x)}{h}$ on fera correspondre une quantité telle que

$$\frac{\mathcal{G}(t+h) - \mathcal{G}(t)}{m[\delta(t, t+h)]}.$$

A l'aide de ce rapport on définira 4, 8 ou 16 nombres dérivés de $\mathcal{G}(t)$ aux points simples, doubles ou quadruples de Γ.

(¹) *Math. Annalen*, Bd. XXXVIII.

Lorsque tous ces nombres dérivés sont égaux, leur valeur est la dérivée au point considéré.

Moyennant ces changements (1) mes raisonnements s'appliquent sans difficultés et l'on généralise toutes les propositions de mes *Leçons sur l'Intégration*, en particulier le théorème énoncé précédemment; mais la définition de la dérivée n'est pas celle que nous avons adoptée dans ce Mémoire. Pour revenir à cette définition il faut un raisonnement complémentaire au cours duquel (je l'ai reconnu depuis en étudiant la Note de M. Vitali) j'utilisais en somme une démonstration d'un cas particulier du théorème de M. Vitali. Ici j'exposerai le raisonnement en m'appuyant explicitement sur ce théorème lui-même.

Il est démontré qu'une fonction d'ensemble $\mathcal{F}(\mathcal{E})$ additive et continue a une dérivée presque partout et que l'intégration permet de remonter de cette dérivée à $\mathcal{F}$ lorsque la dérivée est définie à l'aide de la courbe Γ de M. Hilbert. Or tout carré

$$\frac{a}{3^c} \leqq x \leqq \frac{a+1}{3^c}, \qquad \frac{b}{3^c} \leqq y \leqq \frac{b+1}{3^c},$$

a, b, c étant des entiers, est un arc de cette courbe; donc nous pouvons dire que la proposition est démontrée quand on définit la dérivée en un point P par le procédé indiqué précédemment dans ce Mémoire en utilisant les carrés arcs de Γ contenant P au lieu des domaines réguliers quelconques.

Démontrons maintenant que l'emploi de deux systèmes de carrés différents contenant P, pour définir des nombres dérivés, donnera presque partout le même résultat. Si cela n'était pas on pourrait, à chaque point P d'un certain ensemble E de mesure non nulle, attacher deux systèmes différents de carrés contenant P, dont les diamètres

---

(1) Il faut bien avoir soin de faire ces changements partout; par exemple, si l'on ne s'écarte pas de mon raisonnement primitif, on est conduit à parler de la longueur de l'arc de la courbe $y = f(x)$ qu'on associe à une fonction à variation bornée $f(x)$. Il ne faut pas oublier alors que cette longueur d'arc étant la limite d'une somme de quantités de la forme $\sqrt{[f(x_1)-f(x_2)]^2+(x_1-x_2)^2}$ devra, maintenant qu'il s'agira d'une fonction $f(t)$, être remplacée par une limite de sommes de quantités de la forme

$$\sqrt{[f(t_1)-f(t_2)]^2+\overline{m[\delta(t_1,t_2)]^2}}$$

tendent vers zéro et tels que l'emploi du premier système donne pour la limite du rapport incrémentiel un résultat supérieur de ε au moins à celui que fournit l'emploi du second système; ε est un nombre choisi positif assez petit pour que E existe tel qu'il a été dit. Alors enfermons E dans un domaine Δ dont la mesure ne surpasse $m(E)$ que de ε au plus et faisons, d'une part dans la famille des carrés des premiers systèmes, d'autre part dans la famille des carrés des seconds systèmes, les choix que légitime le théorème de M. Vitali, en ayant soin de n'utiliser que des carrés intérieurs à Δ.

Soient S_1, S_2 les deux suites de carrés choisies. Elles couvrent tout E à un ensemble de mesure au plus égale à ε près et n'ont en dehors de E que des points formant deux ensembles de mesure ε au plus. Donc, quand ε tend vers zéro, $\mathcal{I}(S_1)$ et $\mathcal{I}(S_2)$ tendent vers $\mathcal{I}(E)$; or cela est impossible puisque $\mathcal{I}(S_1)$ surpasse $\mathcal{I}(S_2)$ d'au moins $\varepsilon\, m(E)$.

Ainsi notre proposition est justifiée quand la dérivée en P est définie à l'aide de carrés quelconques contenant P. Ensuite, par un raisonnement exactement identique à celui que j'ai donné pour les fonctions d'une variable (**8**) et que M. Tonelli (**14**) vient d'ailleurs de reprendre pour les fonctions de deux variables, on montrera qu'avec cette définition de la dérivée les points P, tels que $\int |f(Q) - f(P)|\, dQ$ n'a pas une dérivée nulle au point P, forment un ensemble de mesure nulle.

Ceci étant, supposons P pris hors de cet ensemble de mesure nulle. Soit $C(a)$ le carré de côté a et de centre P :

$$\varphi(a) = \frac{1}{a^2} \int_{C(a)} |f(Q) - f(P)|\, dQ$$

tend vers zéro avec a. Soit E un ensemble contenant P et régulier par rapport au module μ, $\frac{a}{4}$ le rayon du cercle contenant E; on a par définition

$$m(E) > \mu\pi \frac{a^2}{4^2} = \frac{\mu\pi}{4^2} m[C(a)].$$

Or on a

$$\int_E |f(Q) - f(P)|\, dQ \leqq \int_{C(a)} |f(Q) - f(P)|\, dQ,$$

donc

$$\frac{1}{m(E)}\int_E |f(Q)-f(P)|\,dQ < \frac{4^2}{\mu\pi}\varphi(a),$$

et par suite, au point P, $|f(Q)-f(P)|$ est la dérivée de son intégrale indéfinie prise par rapport à Q, même quand on utilise la définition générale de la dérivée à l'aide d'ensembles réguliers. Cette dérivée en P étant nulle, il en est *a fortiori* de même de la dérivée en P de l'intégrale indéfinie de $f(Q)-f(P)$ prise par rapport à Q. Donc enfin en P la dérivée de $\int f(Q)\,dQ$ est égale à $f(P)$.

La légitimation du théorème est ainsi achevée. La démonstration qu'on vient de lire est certainement très artificielle et très longue à exposer complètement; on conçoit cependant que, parce qu'elle n'utilise presque aucun raisonnement nouveau, elle m'ait permis de m'assurer très rapidement de l'exactitude des généralisations des théorèmes sur les fonctions d'une variable.

33. J'indique maintenant un raisonnement (¹) qui, dans le cas des fonctions d'une seule variable, permet aussi l'étude de la dérivation des intégrales.

Soit E un ensemble dénombrable d'intervalles non empiétants; soient (a_1, b_1), (a_2, b_2), ... les intervalles de E.

Prolongeons (a_1, b_1) vers la gauche, ce qui donne (α_1, b_1) de façon qu'on ait

$$(b_1-\alpha_1) = K(b_1-a_1),$$

K étant un nombre positif fixe supérieur à 1. Faisons de même sur (a_2, b_2), ce qui donnera (α_2, b_2); si (α_1, b_1) et (α_2, b_2) ont une partie commune, on reporte cette partie commune vers la gauche de façon à obtenir un intervalle (λ, μ) contenant (α_1, b_1) et (α_2, b_2) et de longueur égale à $(b_1-\alpha_1)+(b_2-\alpha_2)$. On obtient donc un ou deux intervalles que je désigne par (A_1, B_1), (A_2, B_2).

(¹) J'ai fait allusion à ce raisonnement dans une Note des *Lincei* (5) en le comparant à un raisonnement de M. Vitali concernant les fonctions à intégrales indéfinies constamment nulles, parce qu'il repose comme celui-ci sur la constatation d'une propriété vraie pour les ensembles d'intervalles, propriété d'où l'on déduit un théorème sur les ensembles mesurables, cas particulier du théorème à démontrer.

Opérons ensuite sur (a_3, b_3) comme sur (a_1, b_1), ce qui donne (α_3, b_3); si (α_3, b_3) a une partie commune avec (A_1, B_1), on reporte cette partie commune vers la gauche, ce qui donne (A', B'); si (A', B') a une partie commune avec (A_2, B_2), on reporte aussi cette partie vers la gauche; si (α_3, b_3) n'a pas de partie commune avec (A_1, B_1), mais en a une commune avec (A_2, B_2), on opérera de même en commençant par (A_2, B_2) au lieu de (A_1, B_1). Cela donnera donc 1, 2 ou 3 intervalles. Et l'on continuera ainsi. On arrivera finalement à un nombre fini ou dénombrable d'intervalles (λ_1, μ_1), (λ_2, μ_2), ... formant l'ensemble $\mathcal{E}$ de mesure totale $K m(E)$.

Ces intervalles jouissent de la propriété suivante : quel que soit (λ_i, μ) intérieur à (λ_i, μ_i), la mesure des points de (λ_i, μ) ne faisant pas partie de E est au moins $(K-1)$ fois la mesure des points communs à E et à (λ_i, μ). Il en résulte que dans un intervalle (λ, μ), dont l'origine n'appartient pas à $\mathcal{E}$, la mesure des points de E est au plus $\frac{\mu-\lambda}{K}$.

Ceci étant, posons quelques définitions. Soit un ensemble quelconque E, on appelle *densité moyenne* de E dans (α, β) le quotient par $(\beta-\alpha)$ de la mesure des points de E situés dans (α, β). On appelle *densité à droite* de E en un point α, la limite, si elle existe, de la densité moyenne dans (α, β) quand β tend vers α en décroissant. La densité à gauche se définit de même; la densité en un point est la valeur, supposée commune, des densités à droite et à gauche.

Revenant au cas de l'ensemble d'intervalles E, on voit que, s'il existe des points en lesquels E a une densité à droite supérieure à $\frac{1}{K}$, ils font partie de $\mathcal{E}$. Or remarquons que si dans la construction de $\mathcal{E}$ nous n'avions tenu compte que des intervalles de E à partir du $p^{\text{ième}}$, intervalle formant ce que nous désignerons par E_p, cela nous aurait conduit à un $\mathcal{E}_p$ tel que dans l'ensemble $E+\mathcal{E}_p$ se trouvent encore tous les points en lesquels la densité à droite de E est supérieure à $\frac{1}{K}$.

Or soit s_p la longueur totale des $p-1$ premiers intervalles de E, on a

$$m(E+\mathcal{E}_p) = s_p + K[m(E) - s_p],$$

quantité qui surpasse $m(E)$ d'aussi peu qu'on le veut. Concluons de là qu'en dehors des points de E les points en lesquels E a une densité à droite supérieure à $\frac{1}{K}$ (donc différente de zéro puisque K est quelconque) forment un ensemble de mesure nulle.

Soient maintenant un ensemble quelconque A, E un ensemble d'intervalles contenant A; les points extérieurs à E en lesquels A n'a pas une densité à droite nulle forment un ensemble de mesure nulle; et cela étant vrai quel que soit E contenant A est vrai aussi des points extérieurs à A. Soit B le complémentaire de A par rapport à un certain intervalle; les points de B en lesquels A n'a pas une densité à droite nulle forment, on vient de le voir, un ensemble de mesure nulle. Permutons A et B dans l'énoncé de ce résultat, il reste vrai; or $A + B$ a une densité égale à 1 en tout point; donc, les points d'un ensemble de mesure nulle étant exceptés, la densité à droite de A est égale à *un* en tout point de A, égale à *zéro* en tout point de B.

Raisonnant de même sur la densité à gauche, on voit finalement que *la densité d'un ensemble mesurable est égale à* un *en presque tous les points de cet ensemble.*

34. C'est-à-dire qu'il est démontré qu'une fonction est presque partout la dérivée de son intégrale indéfinie, lorsqu'il s'agit d'une fonction ne prenant que les valeurs 0 ou 1.

Par suite, ce théorème s'en déduit quand il s'agit de fonctions ne prenant qu'un nombre fini de valeurs différentes. Passons au cas général et supposons qu'il s'agisse d'une fonction qui n'est jamais négative f. Soit E_p l'ensemble des points en lesquels on a

$$p\varepsilon \leqq f < (p+1)\varepsilon \qquad (p = 0, 1, \ldots);$$

ε est une quantité positive arbitraire. Soit φ_n la fonction égale à $p\varepsilon$ dans E_p, pour $p = 0, 1, \ldots, n$ et égale à zéro ailleurs. Les nombres dérivés de l'intégrale de f sont au moins égaux à ceux de l'intégrale de φ_n; donc, presque partout, les nombres dérivés de l'intégrale de f sont égaux ou supérieurs à f.

Soit Φ la fonction égale à $(p+1)\,\varepsilon$ dans chaque E_p. Les nombres dérivés de $\int f\,dx$ ne surpassant pas ceux de $\int \Phi\,dx$, il suffirait de démontrer le théorème pour la fonction Φ. Or, si Φ_n est la fonction définie comme étant égale à zéro dans $E_1 + E_2 \ldots + E_n$ et égale à Φ

ailleurs, il suffit de démontrer la proposition pour Φ_n, c'est-à-dire qu'il suffit de démontrer que Φ_n a une dérivée nulle presque en tous les points de $E_1 + E_2 + \ldots + E_n$. Enfermons E_{n+p} dans des intervalles A_p dont la mesure ne surpasse celle de E_{n+p} qu'au plus de $\frac{\theta}{2^p(n+p+1)\varepsilon}$ et attachons à ces intervalles un poids $\pi_p = (n+p+1)\varepsilon$.

La somme $\Sigma\pi_p m(A_p)$ est au plus égale à θ, plus l'intégrale de f dans $E_{n+1} + E_{n+2} + \ldots$; elle est donc finie.

Soit E l'ensemble de tous les A_p; convenons de définir la densité moyenne de (α, β) comme le quotient par $(\beta - \alpha)$ de la somme des parties de E situées dans (α, β) multipliées respectivement par leur poids; et définissons les densités à partir de là. Notre raisonnement du début s'appliquera; seulement, pour former l'ensemble $\mathcal{E}$ contenant les points en lesquels la dérivée à droite de E surpasse $\frac{1}{K}$, il faudra remplacer un intervalle (a_i, b_i) de poids p_i par (α_i, β_i) tel que

$$b_i - \alpha_i = K p_i (b_i - a_i).$$

Mais rien d'essentiel ne sera changé au raisonnement, parce que la série $\Sigma p_i(b_i - a_i) = \Sigma\pi_p m(A_p)$ est convergente.

Donc il est prouvé d'une façon nouvelle qu'une fonction sommable d'une variable est presque partout la dérivée de son intégrale indéfinie; ce théorème étant établi, les autres propriétés s'en déduisent facilement.

V. — La dérivation des fonctions à variation bornée.

35. Revenons à notre première démonstration. Elle nous a prouvé l'identité des fonctions d'ensemble additives et absolument continues et des intégrales indéfinies; mais en un sens elle prouve moins que le second mode de démonstration qui, généralisant tous les résultats relatifs aux fonctions d'une variable, permet par suite d'obtenir des résultats sur des fonctions plus générales que les fonctions d'ensemble additives et absolument continues. Pour obtenir ces résultats par notre première méthode il faudrait, se rappelant que c'est la même chose de se donner une fonction d'ensemble additive et absolument

continue, ou une fonction de domaine ou d'intervalle additive et absolument continue, ou une fonction des k coordonnées à variation bornée et absolument continue, examiner en quoi sont changés nos résultats quand on abandonne quelques-unes des hypothèses faites sur les fonctions considérées. Nous ferons ici la partie la plus élémentaire de cette étude en nous plaçant dans le cas de deux variables.

Soit $f(x, y)$ une fonction de deux variables; à l'intervalle $I(A, B)$ dont deux sommets opposés sont les points A et B de coordonnées (a, b), $(a+h, b+k)$, $(h > 0, k > 0)$, nous attachons la fonction

$$\psi(A, B) = f(a, b) + f(a+h, b+k) - f(a+h, b) - f(a, b+k).$$

Dire que f ou ψ est *à variation bornée*, c'est dire que si l'on considère des intervalles $I(A, B)$ non empiétants la somme des valeurs absolues des $\psi(A, B)$ correspondantes est inférieure à un nombre fixe. Voyons les conséquences de cette seule hypothèse. Tout d'abord, si l'on divise un intervalle I en un nombre fini d'intervalles partiels I_p, on a $\psi(I) = \Sigma\,\psi(I_p)$, ce qu'on exprime en disant que ψ est additive. Puis si, divisant un intervalle I en un *nombre fini* d'intervalles partiels, on fait la somme des ψ positifs ou nuls d'une part, des ψ négatifs d'autre part, et si l'on cherche les bornes supérieures de ces sommes pour toutes les divisions possibles on arrive à former deux fonctions additives ψ_1, ψ_2 d'intervalles qui ne sont jamais négatives et dont ψ est la différence. Si ψ est tel qu'on puisse trouver des intervalles dont les mesures tendent vers zéro sans que les valeurs correspondantes de ψ tendent vers zéro, il en est de même de l'une au moins des deux fonctions ψ_1 et ψ_2. Soient x_0 une valeur quelconque, ab un segment de $x = x_0$, I un intervalle ayant ab pour côté, et dont tous les points intérieurs sont d'abscisse inférieure à x_0; soit l la limite, nécessairement bien déterminée, de $\psi_1(I)$ quand $m(I)$ tend vers zéro; l est une fonction de x_0 et des ordonnées y_1, y_2 de a et b; posons $l = \sigma_{x_0-0}(y_1, y_2)$. On définira de même $\sigma_{x_0+0}(y_1, y_2)$ et aussi deux fonctions analogues $\sigma_{y_0-0}(x_1, x_2)$, $\sigma_{y_0+0}(x_1, x_2)$. Ces fonctions sont les analogues des sauts des fonctions d'une variable à points de discontinuité de première espèce. Prenons l'une d'elles $\sigma_{x_0-0}(y_1, y_2)$. C'est évidemment, quand x_0 est fixe, une fonction positive et additive de l'intervalle (y_1, y_2); on pourra donc recommencer de même sur cette fonction et définir des nombres σ_{x_0-0, y_0-0}; σ_{x_0-0, y_0+0}, qui sont les sauts de la fonction croissante en y qui s'écrit $\sigma_{x_0-0}(o, y)$. On arrive ainsi à quatre sauts relatifs au point (x_0, y_0), les nombres σ_{x_0-0, y_0+0}, σ_{x_0-0, y_0-0}, σ_{x_0+0, y_0-0}, σ_{x_0+0, y_0+0}; dans ces symboles l'ordre des indices n'intervient pas comme on s'en aperçoit facilement ([1]). Un point (x_0, y_0) est singulier si l'un de ses quatre sauts est différent de zéro.

Enlevons maintenant de $\sigma_{x_0-0}(y_1, y_2)$, $y_2 > y_1$, les sauts $\sigma_{x_0-0, y-0}$ relatifs aux points singuliers (x_0, y) intérieurs au segment (y_1, y_2) de $x = x_0$ et le saut σ_{x_0-0, y_0+0}. Cette opération a un sens, car il est évident qu'il n'y a qu'une infinité dénombrable de points singuliers et que la série totale des sauts est convergente, nous obtenons ainsi une quantité qu'on appellera $s_{x_0-0}(y_1, y_2)$. On aura de même $s_{x_0+0}(y_1, y_2)$, etc. Un segment (y_1, y_2) de $x = x_0$ sera dit *singulier* si, quelle que soit la partie (a, b) de (y_1, y_2), l'un des deux sauts $s_{x_0-0}(a, b)$, $s_{x_0+0}(a, b)$ est différent de zéro. On voit facilement là encore que les segments singuliers différents sont au plus en infinité dénombrable et que la série des sauts correspondants est convergente.

On attachera de même à ψ_2 des nombres σ'_{x_0-0, y_0-0}, ..., puis à ψ on attachera les différences $\sigma_{x_0-0, y_0-0} - \sigma'_{x_0-0, y_0-0}$, ... Or, il est évident que les deux nombres σ_{x_0-0, y_0-0}, σ'_{x_0-0, y_0-0} ne peuvent être différents de zéro à la fois, car s'ils étaient supérieurs à $K > o$ tous deux la fonction ψ, égale à K dans tout intervalle $a \leqq x \leqq A$, $b \leqq x \leqq B$ tel qu'on ait $a < x_0 \leqq A$, $b < x_0 \leqq A$, et égale à zéro ailleurs, pourrait être soustraite et de ψ_1 et de ψ_2 sans que ces fonctions deviennent négatives, ce qui est en contradiction avec leur définition même. De même on verrait que deux quelconques des sauts correspondants de ψ_1 et ψ_2 ne peuvent être différents de zéro à la fois, il en résulte que la définition des points et segments singuliers de ψ et des sauts correspondants aurait pu être donnée directement pour ψ comme pour ψ_1 sans passer par l'intermédiaire de ψ_1 et ψ_2. L'avantage que nous avons eu à utiliser des fonctions non négatives, c'est que nous étions assurés de l'existence des limites que nous avons appelées *sauts*.

Dans le cas général de k coordonnées on définirait de même des

---

([1]) Je veux dire qu'il importe peu de déduire σ_{x_0-0, y_0+0} de $\sigma_{x_0-0}(y_1, y_2)$ ou de $\sigma_{y_0+0}(x_1, x_2)$.

points singuliers auxquels seraient attachés 2^k sauts et, pour $0 < p < k$, des variétés singulières obtenues en faisant varier p coordonnées seulement et auxquels sont attachés 2^{k-p} sauts qui, en tant que fonctions d'intervalle dans l'espace à $k-p$ coordonnées variables, sont à variation bornée et même sont sans singularités, au sens qui vient d'être dit, d'après ce qu'on va voir.

36. Revenons à l'hypothèse $k=2$; les sauts en chaque point singulier peuvent être alors considérés comme attachés à chacun des quadrants ayant pour sommet le point considéré, les sauts suivant un segment singulier sont de même attachés aux deux demi-plans déterminés par la droite portant le segment. Soit I un intervalle. Désignons par S(I) la somme des sauts relatifs à tout point ou segment singulier intérieur à I; cependant, s'il s'agit d'un point singulier ou d'un segment singulier situé sur la frontière de I, on ne s'occupera que des sauts relatifs aux quadrants ou demi-plans contenant des points de I. S(I) est la fonction des sauts; elle est additive et à variation bornée.

Or on voit de suite que $\psi(\mathrm{I}) - \mathrm{S}(\mathrm{I})$ est une fonction pour laquelle on ne peut pas trouver une suite d'intervalles δ dont les mesures tendent vers zéro et tels que $\psi(\delta) - \mathrm{S}(\delta)$ ne tende pas vers zéro, sans quoi : ou bien on pourrait trouver un point A limite d'une infinité de δ, ou bien on pourrait trouver un segment $\alpha\beta$ parallèle à un des axes limite d'une telle infinité de δ. Or, on voit facilement, dans le premier cas, que A est singulier pour $\psi - \mathrm{S}$ et, dans le second, qu'une partie ou un point de $\alpha\beta$ est singulier pour $\psi - \mathrm{S}$, et cela est évidemment impossible d'après la construction de S.

Remarquons que s_{x_0-0} par exemple a été déduite de σ_{x_0-0} par l'opération analogue à celle qui vient d'être indiquée; donc s_{x_0-0} et les fonctions analogues sont sans point singulier.

A $\psi - \mathrm{S}$ on peut faire correspondre une fonction $f_1(x, y)$, laquelle n'est déterminée qu'à des fonctions, de x seul et de y seul, additives près, qui correspond à $\psi - \mathrm{S}$ comme f correspondait à ψ. En choisissant convenablement ces fonctions additives $f_1(x, y)$ est continue. Ainsi *une fonction des k coordonnées à variation bornée est la somme d'une fonction à variation bornée et continue et d'une fonction corres-*

pondante à une fonction des sauts définie par la considération d'une infinité dénombrable de singularités.

Bien que la décomposition fournie par l'énoncé précédent ne soit possible que d'une seule manière on peut la considérer comme étant en partie arbitraire parce que si, faisant un changement de coordonnées, on remplaçait $f(x,y)$ par $\varphi(\xi,\eta)$ la décomposition faite sur φ ne correspondrait pas à celle faite sur f. Précisons cela par un exemple.

Prenons pour $f(x,y)$ la fonction à laquelle correspond une fonction d'intervalle ψ égale à la longueur de la partie de la droite $x-y=0$ contenue dans l'intervalle considéré. Prenons maintenant la droite $x=y$ pour axe des η; nous ne savons même plus ce qu'on doit appeler $\varphi(\xi,\eta)$, ni ce qu'on doit entendre par la fonction d'intervalle correspondante. Prenons une fonction d'intervalle en ξ, η se réduisant à sa fonction des sauts, laquelle sera caractérisée par le fait que seules existent les deux fonctions σ_{ξ_0-0}, σ_{ξ_0+0} et pour la seule valeur $\xi_0=0$. Ces deux fonctions seront d'ailleurs des fonctions à variation bornée arbitraires de l'intervalle (η_1,η_2) et dont la somme est égale à $\eta_2-\eta_1$. Soit X cette fonction d'intervalle en (ξ,η); on pourra, si l'on veut, la faire correspondre à ψ, mais à X correspondra toujours une fonction discontinue de (ξ,η).

Si au contraire, dans la formation de la fonction S des sauts, on n'avait tenu compte que des points singuliers et non des segments singuliers, la décomposition de ψ en S et $\psi-S$ aurait eu un certain caractère d'invariance par rapport aux changements de coordonnées. Dans ce cas $\psi-S$ aurait été une fonction d'intervalle *continue en ce sens étroit* qu'à un intervalle dont toutes les dimensions tendent vers zéro correspond une valeur de la fonction qui tend elle aussi vers zéro. Or, si par rapport à d'autres axes (ξ,η) on peut trouver une fonction d'intervalles qui ne soit pas en désaccord avec $\psi-S$, cette nouvelle fonction est elle-même continue au même sens étroit du mot.

37. Quoi qu'il en soit, partant d'une fonction des coordonnées, additive et à variation bornée, nous l'avons mise sous la forme $S+S_1+F$, S étant la partie de la fonction des sauts relative aux points singuliers, S_1 la partie relative aux variétés singulières, et F une fonction continue. $S+S_1$ est la fonction des sauts. A cette décompo-

sition correspond une décomposition analogue en $s + s_1 + \mathcal{f}$ d'une fonction d'intervalle additive et à variation bornée. $s_1 + \mathcal{f}$ est alors continue au sens étroit et $\mathcal{f}$ est *continue au sens large*, ce qui signifie que $\mathcal{f}(I)$ tend vers zéro avec $m(I)$.

Si $\mathcal{f}(I)$ n'est pas absolument continue, c'est qu'on peut trouver des suites Σ_p formées chacune d'un nombre fini d'intervalles non empiétant les uns sur les autres tels que, lorsque p augmente indéfiniment, $m(\Sigma_p)$ tende vers zéro, et que la somme des valeurs correspondantes de $\mathcal{f}$, que je désignerai par $\mathcal{f}(\Sigma_p)$, ne tende pas vers zéro.

Soit A une des limites différentes de zéro de $\mathcal{f}(\Sigma_p)$.

Supposons d'abord $\mathcal{f}(I)$ positive ou nulle; les nombres A qui sont tous positifs remplissent un intervalle $(0, \mathcal{A})$, et il existe des suites Σ_p pour lesquelles $A = \mathcal{A}$. Soit $s_2(I)$ la fonction égale au nombre analogue à $\mathcal{A}$ dans chaque intervalle ; cette fonction évidemment additive, à variation bornée et continue au sens large, sera dite *la fonction des singularités de* $\mathcal{f}(I)$, supposée positive ou nulle.

Si $\mathcal{f}(I)$ n'est pas toujours positive ou nulle, elle est la différence de deux fonctions $\mathcal{f}'$, $\mathcal{f}''$ jouissant de cette propriété. A chacune d'elles nous attachons des fonctions $s'_2(I)$, $s''_2(I)$, et c'est $s'_2(I) - s''_2(I)$ qui est *la fonction des singularités de* $\mathcal{f}(I)$; elle sera additive, à variation bornée et continue au sens large.

Pour la fonction $\mathcal{f}(I) + s(I) + s_1(I)$, la fonction des singularités sera $s(I) + s_1(I) + s_2(I)$; elle sera additive et à variation bornée.

38. *Si* $s(I)$ *est la fonction des singularités d'une fonction* $f(I)$ *additive et à variation bornée, la différence* $f(I) - s(I)$ *est absolument continue et* $s(I)$ *a une variation totale plus petite que celle de toute autre fonction* $\sigma(I)$ *telle que* $f(I) - \sigma(I)$ *soit absolument continue.*

Si l'on pose, comme précédemment,

$$f(I) = \mathcal{f}(I) + s(I) + s_1(I),$$

il suffit évidemment de démontrer le théorème pour $\mathcal{f}(I)$.

Supposons d'abord $\mathcal{f}(I)$ positive ou nulle partout. Si $\mathcal{f}(I) - s_2(I)$ n'était pas absolument continue, on pourrait trouver des suites Σ'_p formées chacune d'un nombre fini d'intervalles non empiétants tels que

$m(\Sigma'_p)$ tende vers zéro et que $\mathcal{F}(\Sigma'_p) - s_2(\Sigma'_p)$ tende vers une limite non nulle α; α sera positif, car $\mathcal{F}(\mathrm{I}) - s_2(\mathrm{I})$ n'est jamais négative. Les Σ_p désignant la suite précédemment indiquée, la suite $\Sigma_p + \Sigma'_p$ qu'on pourra encore regarder comme formée d'intervalles non empiétants et en nombre fini sera telle que $m(\Sigma_p + \Sigma'_p)$ tende vers zéro et que $\mathcal{F}(\Sigma_p + \Sigma'_p)$ tende au moins vers $\mathcal{A} + \alpha$, ce qui est contraire à la définition de $\mathcal{A}$.

Si d'ailleurs $\mathrm{T}(\mathrm{I})$ est tel que $\mathcal{F}(\mathrm{I}) - \mathrm{T}(\mathrm{I})$ soit absolument continue et si $\mathrm{T}(\mathrm{I}) = \mathrm{T}_1(\mathrm{I}) - \mathrm{T}_2(\mathrm{I})$, T_1 et T_2 étant les variations positive et négative de $\mathrm{T}(\mathrm{I})$, il faut que $\mathcal{F}(\mathrm{I}) - \mathrm{T}_1(\mathrm{I})$ soit absolument continue. Or il est évident que $\mathrm{T}_1(\mathrm{I})$ ne peut être inférieur à $s_2(\mathrm{I})$; sans quoi, si pour I_0 on avait $\mathrm{T}_1(\mathrm{I}_0) < s_2(\mathrm{I}_0)$, quelle que soit la suite de familles Σ_p formées chacune d'intervalles non empiétants compris dans I_0 et telles que $m(\Sigma_p)$ tende vers zéro, comme on a

$$|\mathcal{F}(\Sigma_p)| \leqq |\mathcal{F}(\Sigma_p) - \mathrm{T}_1(\Sigma_p)| + \mathrm{T}_1(\mathrm{I}_0),$$

on aurait

$$\limsup_{p=\infty} \mathcal{F}(\Sigma_p) \leqq \mathrm{T}_1(\mathrm{I}_0) < s_2(\mathrm{I}_0),$$

ce qui est contraire à la définition même de s_2.

Notre proposition est donc démontrée pour $\mathcal{F}(\mathrm{I})$ non négative. Remarquons que, dans ce cas, quelle que soit la fonction $\mathrm{A}(\mathrm{I})$ absolument continue et non négative sans être identiquement nulle, il existe quelque intervalle dans lequel $\mathrm{A}(\mathrm{I})$ surpasse $s_2(\mathrm{I})$. Sans quoi, en effet, $s_2(\mathrm{I}) - \mathrm{A}(\mathrm{I})$ serait non négative, au plus égale à $s_2(\mathrm{I})$ sans lui être identiquement égale, et cependant $\mathcal{F}(\mathrm{I}) - [s_2(\mathrm{I}) - \mathrm{A}(\mathrm{I})]$ serait absolument continue, ce qui est impossible.

Pour le cas général, nous avons posé

$$\mathcal{F} = \mathcal{F}' - \mathcal{F}'', \qquad s_2 = s'_2 - s''_2.$$

De là il résulte que $\mathcal{F} - s_2$ est absolument continue et que $s'_2 + s''_2$, qui est la fonction des singularités de $\mathcal{F}' + \mathcal{F}''$, est telle que toute fonction absolument continue, non négative sans être identiquement nulle, la surpasse dans quelque intervalle. Cela est vrai *a fortiori* de la variation totale de s_2. Nous appellerons s'_2, $-s''_2$ et $s'_2 + s''_2$ les variations de s_2.

Soit $\mathrm{B}(\mathrm{I})$ une fonction absolument continue, dont les variations sont b', $-b''$ et $b'+b''$. Les variations de $\mathcal{S}_2(\mathrm{I})+\mathrm{B}(\mathrm{I})$ seront au plus s'_2+b', $-(s''_2+b'')$, $(s'_2+b')+(s''_2+b'')$; on va voir que ces valeurs sont exactement celles des variations. S'il n'en était pas ainsi, en effet, on pourrait soustraire de s'_2+b' et de s''_2+b'' une même fonction m, continue au sens large, non négative et non identiquement nulle sans obtenir des résultats négatifs. Or posons $m=\sigma_2+\mu$, σ_2 étant la fonction des singularités de m. Je dis que $s'_2-\sigma_2$, $s''_2-\sigma_2$, $b'-\mu$, $b''-\mu$ doivent être non négatifs. Soit I un intervalle quelconque; on peut dans I trouver un ensemble Σ formé d'un nombre fini d'intervalles non empiétants dont la mesure totale ne surpasse pas ε et tel que $\sigma_2(\Sigma)$ soit aussi voisine qu'on veut de $\sigma_2(\mathrm{I})$; $b'(\Sigma)$, $b''(\Sigma)$, $\mu(\Sigma)$ tendent vers zéro avec ε; or $s'_2(\Sigma)+b'(\Sigma)-m(\Sigma)$ et $s''_2(\Sigma)+b''(\Sigma)-m(\Sigma)$ ne sont pas négatifs, il en résulte que $s'_2(\mathrm{I})-\sigma_2(\mathrm{I})$ et $s''_2(\mathrm{I})-\sigma_2(\mathrm{I})$ ne sont pas négatifs. Mais on peut aussi trouver dans I un ensemble Σ' formé d'un nombre fini d'intervalles non empiétants dont la mesure totale ne surpasse pas ε et tel que $s'_2(\Sigma')$, $s''_2(\Sigma'')$, $\sigma'_2(\Sigma')$ soient aussi voisins qu'on veut de $s'_2(\mathrm{I})$, $s''_2(\mathrm{I})$, $\sigma''_2(\mathrm{I})$; or $b'(\Sigma')$, $b''(\Sigma')$, $\mu(\Sigma')$ tendent vers zéro avec ε et

$$s'_2(\mathrm{I}-\Sigma')+b'(\mathrm{I}-\Sigma')-m(\mathrm{I}-\Sigma'),\qquad s''_2(\mathrm{I}-\Sigma')+b''(\mathrm{I}-\Sigma')-m(\mathrm{I}-\Sigma')$$

ne sont pas négatifs, il en résulte que $b'(\mathrm{I})-\mu(\mathrm{I})$ et $b''(\mathrm{I})-\mu(\mathrm{I})$ ne sont pas négatifs.

Mais ceci entraîne $\sigma_2\equiv 0$, $\mu\equiv 0$ et par suite la variation totale de $\mathcal{S}_2+\mathrm{B}$ est bien la somme des variations totales de $\mathcal{S}_2$ et de B.

Or soit σ une fonction telle que $\mathcal{F}-\sigma$ soit absolument continue, puisqu'il en est de même de $\mathcal{F}-\mathcal{S}_2$, $\mathcal{S}_2-\sigma$ sera absolument continue; donc la variation totale de

$$\sigma=\mathcal{S}_2-(\mathcal{S}_2-\sigma)$$

sera supérieure à celle de $\mathcal{S}_2$ sauf si σ est identique à $\mathcal{S}_2$.

39. Notre théorème est donc entièrement démontré (1); il est bien

(1) Dans le cas d'une fonction $\mathcal{F}$ de signe variable, on aurait pu craindre que la définition que nous avons donnée pour la fonction $\mathcal{S}_2$ des singularités dépende de la décomposition $\mathcal{F}=\mathcal{F}'-\mathcal{F}''$ utilisée. Le théorème précédent prouve qu'il n'en est rien.

certain qu'il ne met pas en évidence la propriété caractéristique la plus nette de la fonction des singularités. Cette propriété, qui transparaissait à chaque instant dans nos raisonnements, peut être énoncée ainsi : La fonction s des singularités peut, comme la fonction des sauts, être considérée comme attachée à un certain ensemble de points de mesure nulle.

Pour préciser cela j'aurai besoin des nombres dérivés d'une fonction d'intervalle $f(\mathrm{I})$ additive et à variation bornée quelconque. Tout d'abord ces nombres dérivés ne peuvent être définis que par la considération d'intervalles réguliers et non d'ensembles. Il me suffira de définir les nombres dérivés en P par la considération du rapport $\frac{f(\mathrm{I})}{m(\mathrm{I})}$, I étant un intervalle de centre P dont toutes les dimensions sont égales et tendent vers zéro.

Ces nombres dérivés sont mesurables. — Plaçons-nous dans le cas de deux coordonnées, et considérons le rapport $\frac{f(\mathrm{I})}{m(\mathrm{I})}$ pour le carré I de centre $\mathrm{P}(\alpha, \beta)$ et de demi-côté égal à r. Si f était continue au sens large, ce rapport serait une fonction continue $\rho(\alpha, \beta, r)$ et il n'y aurait rien à changer au raisonnement du Chapitre précédent. Dans le cas général, le rapport ρ n'a comme points de discontinuité que ceux pour lesquels on a $\alpha \pm r = x_0$, ou $\beta \pm r = y_0$, x_0 ou y_0 étant l'une des coordonnées d'un point singulier ou la coordonnée fixe d'un segment singulier. Encore, tous ces points-là ne sont-ils pas singuliers. Quoi qu'il en soit, supposons tracée cette infinité de plans de discontinuité dans l'espace α, β, r; si P et r varient, cela revient à faire décrire au point α, β, r une certaine courbe ; à chaque traversée d'une de ces surfaces de discontinuité il pourra y avoir deux sauts brusques pour le rapport, mais il n'y a qu'un nombre fini de surfaces pour lesquelles l'un de ces sauts brusques est, en certains points, supérieur à un nombre positif choisi arbitrairement.

Nous supposons P dans un certain rectangle S et nous considérons le parallélépipède Π de section droite S et compris entre les plans $r = \frac{1}{p}$, $r = \frac{1}{p+1}$. Dans ce parallélépipède marquons les plans de discontinuité qui, en certains points, correspondent à des sauts brusques supérieurs à $\frac{1}{p}$. Π est ainsi divisé en un nombre fini de parties dans chacune desquelles il n'y a plus de sauts supérieurs à $\frac{1}{p}$. Donc, d'après une généralisation d'un théorème de M. Baire qu'on justifiera facilement, on pourra diviser chacune de ces parties par un nombre fini de plans $r =$ const. en morceaux dans lesquels l'oscillation pour α, β constants est inférieure à $\frac{3}{p}$. Π est donc divisé en polyèdres avec les frontières desquels on peut constituer un nombre fini de surfaces $r = \varphi(\alpha, \beta)$, φ étant à une seule détermination. Ces surfaces se coupent et ont des parties communes, peu importe. A chacune de ces surfaces nous attachons les trois fonctions

$$\rho[\alpha, \beta, \varphi(\alpha, \beta)], \qquad \rho[\alpha, \beta, \varphi(\alpha, \beta) - 0], \qquad \rho[\alpha, \beta, \varphi(\alpha, \beta) + 0].$$

Nous avons ainsi un nombre fini de fonctions de (α, β) évidemment mesurables et telles que, pour α, β, r dans Π, $\rho(\alpha, \beta, r)$ diffère de l'une d'elles de moins de $\frac{3}{p}$.

Rangeons dans un ordre quelconque celles de ces surfaces correspondant à $p = 1$, puis celles pour lesquelles $p = 2$, etc., on a une suite simplement infinie de fonctions mesurables dont les limites d'indétermination sont les nombres dérivés cherchés, lesquels sont donc bien mesurables (¹).

40. *La fonction des singularités a une dérivée nulle presque partout.* — Soit $\sigma(I)$ cette fonction ; si l'énoncé n'était pas vrai, on pourrait, au besoin, en changeant le signe de $\sigma(I)$, trouver un ensemble E de mesure non nulle aux points duquel un des nombres dérivés de $\sigma(I)$ surpasse un nombre positif fixe ε choisi assez petit. A chacun des points P de E on peut donc attacher une famille d'intervalles dont toutes les dimensions sont égales et tendent vers zéro, qui ont P pour centre et pour chacun desquels le rapport $\frac{s(I)}{m(I)}$ surpasse ε.

(¹) On aurait pu se passer ici, comme d'ailleurs au Chapitre précédent, de la démonstration de cette propriété moyennant l'emploi d'une forme d'énoncé plus générale pour le théorème de M. Vitali. Dans cette forme d'énoncé, on n'aurait plus supposé l'ensemble E mesurable et les domaines à choisir devraient couvrir une partie de E de même mesure extérieure que E.

D'après le théorème de M. Vitali, on peut couvrir tout E, à un ensemble de mesure nulle près, à l'aide d'une suite formée de certains de ces intervalles; il en résulte que la variation totale de $s(\mathrm{I})$, dans l'intervalle I dont on s'occupe, est supérieure à $\varepsilon\Sigma m(\mathrm{I}) \leqq \varepsilon m(\mathrm{E})$; donc supérieure à $\varepsilon \int_{\mathrm{I}} e(\mathrm{P})\,d\mathrm{P}$, $e(\mathrm{P})$ étant une fonction égale à 1 aux points de E et nulle ailleurs.

Mais ce qui vient d'être démontré pour I peut être prouvé pour un intervalle quelconque; donc la variation totale de $s(\mathrm{I})$ surpasserait partout celle de $\varepsilon \int e(\mathrm{P})\,d\mathrm{P}$, fonction qui est absolument continue, ce qui est impossible, d'après ce qui précède.

La propriété démontrée est vraie encore si l'on adopte la définition générale de la dérivée; il est nécessaire de préciser ce que cela veut dire.

Soit $f(\mathrm{I})$ une fonction d'intervalle additive et à variation bornée; pour définir ses nombres dérivés (au sens large) en P, je considère les valeurs du rapport $\frac{f(\mathrm{E})}{m(\mathrm{E})}$, E étant un ensemble contenant P et faisant partie d'une famille régulière d'ensembles dont la dimension tend vers zéro. Par $f(\mathrm{E})$ je désigne l'une quelconque des limites de $f(\Sigma_p)$, Σ_p étant une suite d'intervalles non empiétants, contenant E, dont la mesure tend vers celle de E quand p augmente, et qui sont intérieurs à la plus petite hypersphère qui contient E.

Ceci étant, si en P la dérivée au sens restreint de $f(\mathrm{I})$ non négative est nulle, il en est aussi de même de la dérivée au sens large. Car si I est le plus petit intervalle de centre P dont toutes les dimensions sont égales et qui contient E, on a évidemment $m(\mathrm{E}) > k m(\mathrm{I})$, k étant une certaine constante positive. Donc

$$\frac{f(\mathrm{E})}{m(\mathrm{E})} \leqq \frac{f(\mathrm{I})}{m(\mathrm{E})} < \frac{1}{k}\,\frac{f(\mathrm{I})}{m(\mathrm{I})};$$

si ce dernier rapport tend vers zéro, il en est de même du premier.

Dans le cas de $s(\mathrm{I})$, fonction des singularités de $f(\mathrm{I})$, la dérivée est en valeur absolue au plus égale à celle de la variation totale de $s(\mathrm{I})$, laquelle, étant la fonction des singularités de la variation totale de $f(\mathrm{I})$, a une dérivée, au sens restreint, nulle presque partout; mais, comme la variation totale est une fonction non négative, d'après ce qui précède, sa dérivée au sens large sera aussi nulle presque partout, donc la fonction des singularités a, avec la définition générale de la dérivée, une dérivée nulle presque partout.

41. Effectuons sur la fonction $f(I)$ la décomposition $f(I) = s(I) + \varphi(I)$, $s(I)$ étant la fonction des singularités et $\varphi(I)$ une fonction absolument continue ; on voit que la dérivée de $f(I)$ est presque partout égale à celle de $\varphi(I)$, donc :

Une fonction d'intervalle additive et à variation bornée a une dérivée finie et déterminée presque partout ; la différence entre cette fonction et l'intégrale indéfinie d'une fonction égale à la dérivée là où elle existe et est finie et nulle ailleurs est la fonction des singularités.

Il reste à préciser le moyen d'obtenir cette fonction des singularités ; cela nous donnera un énoncé analogue à l'un de ceux qu'a donnés M. de la Vallée-Poussin (**15**), pour le cas d'une seule variable.

La fonction des singularités est définie grâce à certaines suites d'intervalles Σ_p dont les mesures tendent vers zéro, et telles que la somme des valeurs absolues de f pour les intervalles correspondants aient la plus grande limite possible ; laissant au besoin de côté certaines des Σ_p, on peut donc supposer que la série des $m(\Sigma_p)$ est convergente, auquel cas, en posant $\sigma_p = \Sigma_p + \Sigma_{p+1} + \ldots$, on pourra remplacer les Σ_p par les σ_p. Chaque σ_p est maintenant contenu dans le précédent, de sorte que l'ensemble des points communs à une infinité de σ_p est l'ensemble de mesure nulle commun à tous les σ_p. Cet ensemble est ce que j'appellerai *l'ensemble des singularités de* $f(I)$. On pourrait démontrer que cet ensemble existe bien ; mais cela est inutile, car les Σ_p ne sont pas entièrement déterminées ; on peut étendre chaque Σ_p, pourvu que la série des $m(\Sigma_p)$ reste convergente, et, par suite, on pourra toujours faire rentrer un ensemble de mesure nulle quelconque dans l'ensemble des singularités. Cet ensemble de singularités n'est donc défini qu'à un ensemble additif près, ce qui ne veut nullement dire d'ailleurs qu'on puisse retrancher un ensemble de mesure nulle quelconque de l'ensemble des singularités sans en changer les propriétés.

La fonction des singularités peut être considérée comme attachée à cet ensemble de singularités en ce sens que, si un intervalle ne contient aucun point de cet ensemble, la fonction des singularités y est identiquement nulle, et que, pour la définition de la fonction des singularités, on peut remplacer les Σ_p par n'importe quelles suites formées d'intervalles non empiétants contenant l'ensemble des singularités et dont la somme des mesures tend vers zéro. Pour la légitimation de cette dernière propriété, il suffit évidemment de considérer le cas d'une fonction $\mathcal{F}(\mathrm{I})$, non négative, continue au sens large et additive en ce sens étroit que, I étant la somme de $\mathrm{I}_1 + \mathrm{I}_2 + \ldots + \mathrm{I}_p$ en nombre fini, on a

$$\mathcal{F}(\mathrm{I}) = \mathcal{F}(\mathrm{I}_1) + \ldots + \mathcal{F}(\mathrm{I}_p).$$

Mais il est évident que ces propriétés entraînent l'additivité au sens large, où p est infini [1]. Enfermons en effet les frontières de I et des I_k dans des intervalles assez petits pour que la somme des valeurs de $\mathcal{F}$ correspondantes ne surpasse pas ε. Avec un nombre fini des I_k et de ces nouveaux intervalles on peut couvrir tout I; donc on a

$$\mathcal{F}(\mathrm{I}) \leqq \Sigma\, \mathcal{F}(\mathrm{I}_k) + \varepsilon.$$

Mais on a évidemment aussi

$$\mathcal{F}(\mathrm{I}) \geqq \Sigma\, \mathcal{F}(\mathrm{I}_k),$$

et la proposition d'additivité au sens large est démontrée.

De là il résulte qu'un ensemble quelconque E, formé des points communs à une suite d'ensembles σ_p d'intervalles non empiétants et contenus chacun dans le précédent, étant donné, la suite des $\mathcal{F}(\sigma_p)$ tend vers une valeur bien déterminée qui est au plus égale à la valeur $\mathcal{F}(\Sigma)$, Σ étant une famille d'intervalles non empiétants contenant E. Or, dans le cas où E est l'ensemble des singularités de $\mathcal{F}(\mathrm{I})$, on sait que, si la mesure de Σ tend vers zéro, $\mathcal{F}(\Sigma)$ ne peut avoir une limite supérieure à celle des $\mathcal{F}(\sigma_p)$, donc ces limites sont égales, et l'on conçoit qu'on puisse dire que la fonction des singularités est égale à la valeur de $\mathcal{F}(\mathrm{I})$ sur l'ensemble des singularités. Cette valeur de $\mathcal{F}(\mathrm{I})$ sur l'ensemble des singularités se définit par le procédé qui, dans le cas d'une fonction d'intervalle absolument continue, permet de définir la valeur de la fonction sur un ensemble quelconque. Mais ce procédé,

---

[1] Il convient de se rappeler qu'un intervalle I a été dit « la somme d'une infinité d'autres » lorsque tout point intérieur à I est point intérieur ou point frontière d'un des intervalles partiels.

qui suppose seulement $\mathcal{F}(\mathrm{I})$ continue au sens large et non nécessairement positive comme il a été supposé, ne réussit ici à donner une valeur déterminée que pour l'ensemble des singularités, — très largement indéterminé, je le rappelle, — à moins qu'il n'y ait pas de singularités.

42. D'après ce qu'on a vu sur les nombres dérivés de la fonction des singularités $s(\mathrm{I})$, il semble bien que l'ensemble des singularités doive être cherché dans l'ensemble des points où $s(\mathrm{I})$ n'a pas une dérivée nulle. Mais, tout d'abord, de tels points existent-ils certainement ? Non, si l'on prend la définition particulière de la dérivée en P à l'aide d'intervalles de dimensions toutes égales entre elles et de centre P. Par exemple, la fonction $s(\mathrm{I})$, dans le cas d'un seul point singulier A en lequel les sauts ont une somme nulle, a une dérivée au sens étroit, nulle en tout point. Au contraire, avec la définition générale des nombres dérivés, nous allons voir que, *si la fonction des singularités $s(\mathrm{I})$ est différente de zéro dans un intervalle I_0, il y a dans I_0 des points où l'un de ses nombres dérivés est infini et du signe de $s(\mathrm{I}_0)$.*

Dans I_0 on peut trouver une suite σ_0 d'intervalles non empiétants dont la somme des mesures soit ε_0 au plus et telle que $s(\sigma_0)=s(\mathrm{I}_0)$. Alors, supposant $s(\mathrm{I}_0)$ positif, pour l'un au moins de ces intervalles le rapport incrémentiel de s atteint ou surpasse $\frac{s(\mathrm{I}_0)}{\varepsilon_0}$; soit I_1 un tel intervalle. On peut, dans I_1, trouver une suite σ_1 d'intervalles non empiétants dont la somme des mesures soit ε_1 au plus et telle que $s(\sigma_1)=s(\mathrm{I}_1)$. Alors, pour l'un au moins de ces intervalles, le rapport incrémentiel atteint ou surpasse $\frac{s(\mathrm{I}_1)}{\varepsilon_1}$; soit I_2 un tel intervalle. En continuant ainsi, on aura des intervalles I_0, I_1, I_2, ayant au moins un point P en commun, et, en ce point, le plus grand nombre dérivé est au moins égal à la limite supérieure de $\frac{s(\mathrm{I}_k)}{\varepsilon_k}$ qui sera infinie pour un choix convenable des ε_k.

L'ensemble des singularités d'une fonction $f(\mathrm{I})$ est l'ensemble des points où la variation totale de $f(\mathrm{I})$ a l'un de ses nombres dérivés infini (¹).

(¹) Cet ensemble n'est pas nécessairement l'ensemble de tous les points communs à une suite de familles d'intervalles contenues chacune dans la précédente ; mais, bien

Pour la simplicité du langage, supposons f non négative, donc identique à sa variation totale. Soit $D'(P)$ une fonction égale à la dérivée de f là où elle existe et à zéro ailleurs. Soient $\Sigma_1, \Sigma_2, \ldots$ des suites d'intervalles, dont la mesure tend vers zéro ; soit $\Sigma_p(I)$ la somme des valeurs de f dans les parties $\Sigma_{p,I}$ de Σ_p intérieures à I :

$$\alpha(I) = f(I) - \int_I D'(P)\, dP - \Sigma_p(I)$$

n'a jamais de valeurs limites négatives pour p infini, car, pour p infini, $\int_{\Sigma_{p\,I}} D'(P)\, dP$ a une limite nulle et

$$(I) - \int_{I-\Sigma_p I} D'(P)\, dP - \Sigma_p(I)$$

n'est jamais négatif.

Pour prouver le théorème, il suffit de montrer que $\alpha(I)$ a une limite nulle pour p infini quand les Σ_p enferment l'ensemble des points où l'un des nombres dérivés de f est infini. Or, dans ce cas, $f(I) - \Sigma_p(I)$ non négatif a partout des nombres dérivés finis; donc il est absolument continu. Il en est de même de $\alpha(I)$, et, la dérivée de $\alpha(I)$ étant presque partout négative ou nulle, $\alpha(I)$ est négatif ou nul ; le théorème est démontré.

43. Notre théorème ne généralise pas celui de M. de la Vallée-Poussin, d'après lequel l'ensemble des singularités est, dans le cas d'une variable, l'ensemble des points en lequel l'un, arbitrairement choisi à l'avance, partout le même, des nombres dérivés est infini. Cette proposition n'est plus vraie avec notre définition des nombres dérivés ; on le verra facilement dans le cas d'une variable ; mais les

---

entendu, il n'y a pas là de contradiction avec la définition primitive de l'ensemble des singularités.

On prouvera aussi, si l'on veut, qu'on peut prendre pour ensemble des singularités l'ensemble commun à une suite de domaines Σ_p dont les mesures tendent vers zéro et tels que Σ_p contiennent les points où l'un des nombres dérivés de $f(I)$ surpasse p.

Dans l'énoncé du texte, il s'agit, bien entendu, de l'ensemble des points où l'un au moins des nombres dérivés est infini, et non de l'ensemble des points où l'un, nommé à l'avance, des nombres dérivés est infini.

exemples sont plus simples encore pour deux variables : la fonction d'intervalle additive, à variation bornée et continue au sens large, égale dans tout intervalle à la longueur de la droite $x + y = 0$ contenue dans ce rectangle, a son plus petit nombre dérivé nul en tout point.

Au reste, cette fonction est à elle-même sa propre fonction des singularités, et il est évident qu'une telle fonction a en tout point l'un de ses nombres dérivés égal à zéro, puisqu'on peut la considérer comme attachée à un ensemble de mesure nulle.

Notons encore ce fait que prouvent ces exemples : une fonction d'intervalle additive et continue qui n'est pas partout nulle n'a pas nécessairement ses deux nombres dérivés différents de zéro et de même signe en quelque point.

Ceci explique pourquoi nous avons été parfois obligés de suivre une voie assez détournée pour atteindre notre but.

J'énonce encore une proposition qui vient d'être utilisée et qui n'est qu'une généralisation approchée de la proposition correspondante dans le cas d'une seule variable : *Une fonction d'intervalle additive, à variation bornée, mais non absolument continue, n'a pas ses deux nombres dérivés partout finis* (¹).

En d'autres termes : *Une fonction d'intervalle additive, à variation bornée, et dont les nombres dérivés sont partout finis, est l'intégrale indéfinie de ses nombres dérivés*. Cette proposition est la généralisation approchée de la suivante : si une fonction $f(x)$ a l'un de ses nombres dérivés partout fini, la condition nécessaire et suffisante pour qu'il soit sommable est que f soit à variation bornée ; f est alors l'intégrale indéfinie de son nombre dérivé. Cet énoncé même n'est évidemment plus exact avec notre langage actuel.

Il me semble vraisemblable qu'on obtiendrait des généralisations exactes, c'est-à-dire contenant les énoncés relatifs au cas d'une variable comme cas particuliers, du théorème précédent comme de celui de M. de la Vallée-Poussin et de tous ceux qui nous ont paru en

(¹) Cette proposition est en relation évidente avec celle citée dans la Note précédente. Il paraît probable qu'on peut prendre pour ensemble des singularités l'ensemble des points où $f(I)$ a l'un de ses nombres dérivés infini.

défaut quand on applique leurs énoncés au cas de plusieurs variables, en se bornant à la considération des fonctions d'intervalle additives, continues au sens large, et en attachant à chaque point quatre nombres dérivés par le procédé que je vais indiquer dans le cas de deux variables x, y. Soit un point $P(x_0, y_0)$; considérons un carré de côtés parallèles aux axes et de centre P; divisons-le en deux rectangles par la droite $x = x_0$. Les rectangles analogues au rectangle de droite obtenu nous conduiront à la notion des deux nombres dérivés à droite, et pareillement on définira les nombres dérivés à gauche.

Il paraît plus naturel de partager le carré en quatre carrés partiels par les droites $x = x_0$, $y = y_0$ et d'attacher ainsi à P huit nombres dérivés. Mais cette façon de procéder, qui est en somme celle de M. Vitali, présente une différence qui me paraît essentielle avec celle que je propose et qui explique, je crois, l'impossibilité évidente de généraliser certains énoncés avec cette seconde façon de procéder. Tandis que le premier procédé attache à deux points P et Q quatre demi-plans placés de telle façon que le demi-plan de droite de l'un des points contient le demi-plan de droite de l'autre point, tandis que le demi-plan de gauche de ce second point contient le demi-plan de gauche du premier, il n'y a aucune relation de situation analogue entre les quadrants que le second procédé attache à P et Q.

On peut noter d'ailleurs que tous les énoncés qu'il s'agit de généraliser peuvent l'être, comme je l'ai dit, par le procédé des chaînes en adoptant les définitions provisoires des nombres dérivés qui nous ont servi quand nous nous sommes occupés de ce procédé. Or, on attache ainsi, à chaque point P, 4 nombres dérivés, 2 à 2 relatifs respectivement aux parties de droite et de gauche en lesquelles le point P partage la courbe d'Hilbert. Et ces parties de droite et de gauche pour deux points P et Q jouissent des mêmes relations de situation que les demi-plans dont il a été question plus haut.

44. En terminant cette étude trop longue des fonctions d'intervalle additives et à variation bornée les plus générales, je fais remarquer qu'une étude analogue peut être faite pour les fonctions d'ensemble mesurable ou de domaine quarrable additives et à variation bornée.

D'ailleurs, se donner une fonction de domaine, c'est se donner une fonction d'intervalle; il n'y a donc rien à changer à nos considérations; mais je dois dire qu'il n'est nullement certain qu'il existe, dans le cas de plusieurs dimensions, des fonctions définies pour tous les domaines quarrables additives et à variations bornées qui ne soient pas absolument continues. Il est même certain qu'il n'en existe pas quand on prend la propriété d'additivité au sens large. Il en existe, au contraire, si l'on prend cette propriété d'additivité au sens étroit et si l'on se limite aux domaines simplement connexes [1].

Se donner une fonction d'ensemble mesurable, c'est se donner une fonction d'intervalle, mais pas nécessairement additive, parce que certaines parties de frontières d'intervalles peuvent fournir des valeurs non nulles dans la fonction d'ensemble, d'où une fonction des sauts analogue à S, mais attachée cette fois aux points singuliers et variétés singulières elles-mêmes, et non plus aux demi-plans, quadrants, etc., relatifs à ces points. Cette fonction retranchée, il reste une fonction d'intervalle additive et à variation bornée, — laquelle me semble devoir être toujours continue au sens large, mais peu importe, — à partir de laquelle on raisonnera comme précédemment.

VI. — Les dérivées partielles des intégrales indéfinies.

45. Nous venons d'étudier une opération qui, dans le cas des intégrales indéfinies des fonctions continues, se réduit à la dérivation prise une fois par rapport à chacune des k variables.

Nous avons défini cette opération d'un seul coup et non pas grâce à k opérations successives; cela est avantageux en un sens, mais a aussi quelques inconvénients : par exemple celui de ne pas définir des nombres analogues aux dérivées d'ordre moindre que k qu'on rencontre au cours des k déviations successives. Nous allons donc rechercher si l'on peut dériver l'intégrale indéfinie, exprimée à l'aide des coordonnées, par rapport à l'une des variables. Comme nous savons

[1] Pour en construire des exemples, on pourra imiter le procédé qui m'a servi pour la question analogue relative au problème de la mesure des domaines et qui est exposé dans les travaux cités au Chapitre I.

que la dérivation, par rapport à x_1 est presque partout l'inverse de l'intégration par rapport à x_1, cela revient à mettre notre intégrale indéfinie sous la forme d'une intégrale indéfinie en x_1; la question est intimement liée à celle de l'expression d'une intégrale multiple à l'aide d'intégrales simples, question dont je vais tout d'abord m'occuper [1].

Soient x l'une des coordonnées, y l'ensemble des autres, de sorte que y est un point ou une valeur de l'espace à K — 1 coordonnées; on peut d'ailleurs supposer aussi que x et y sont deux ensembles complémentaires de coordonnées.

Soit E un ensemble mesurable; $E(x_0)$ et $E(y_0)$ seront les sections de E par les variétés $x = x_0$ et $y = y_0$. $E(x_0)$ et $E(y_0)$ ne sont pas nécessairement des ensembles mesurables par rapport respectivement aux coordonnées y et x, et $m_y E(x_0)$, $m_x E(y_0)$ ne sont peut-être pas des fonctions mesurables de x_0 et y_0.

Enfermons E dans un ensemble I d'intervalles non empiétants. Pour un intervalle i on a évidemment

$$m(i) = \int m_y [i(x)]\, dx,$$

$m_y[i(x)]$ désignant la mesure dans l'espace de coordonnées y; je dis que la formule analogue est vraie pour I. Il suffit, en effet, de le démontrer pour I borné; alors tous les $I(x)$ sont bornés; si d'ailleurs I est la somme des intervalles $i_1, i_2, \ldots$, la formule $m(I) = \int m_y [I(x)]\, dx$ résulte de la possibilité d'intégrer terme à terme la série à termes positifs $m_y[I(x)] = m_y[i_1(x)] + m_y[i_2(x)] + \ldots$. Prenons maintenant pour l'ensemble I successivement les ensembles $I_1, I_2, \ldots$ tels que $m(E_p)$ tende vers $m(E)$ et que I_p soit contenu dans I_{p-1}. Alors les quantités $m_y[I_p(x)]$ vont en décroissant et, par suite, ont une limite déterminée $\lambda(x)$ positive ou nulle, et, d'après le même théorème sur

---

[1] J'ai traité (**3**), le cas des fonctions bornées. M. Fubini (**1**) a traité le cas général. C'est la méthode qui résulte de ces deux travaux que j'utilise ici, légèrement modifiée. On pourrait simplifier plus encore en se bornant à la considération de celles des fonctions mesurables qui ne conduisent à considérer que des fonctions mesurables dans le raisonnement, ce qui, pratiquement, serait tout à fait suffisant.

l'intégration des séries ou des suites, on a

$$m(\mathrm{E}) = \lim_{p=\infty} m(\mathrm{I}_p) = \int \lambda(x)\, dx.$$

Cela suppose E borné; E est alors compris dans un certain intervalle δ par rapport auquel on prend les complémentaires. En opérant de même sur le complémentaire C(E), on trouve

$$m[\mathrm{C}(\mathrm{E})] = \int \mu(x)\, dx,$$

$\mu(x)$ étant la quantité analogue à $\lambda(x)$. De là il résulte qu'on a

$$m(\delta) = \int [\lambda(x) + \mu(x)]\, dx$$

et par suite on a, presque partout,

$$\lambda(x) + \mu(x) = m_y[\delta(x)].$$

Or $\lambda(x)$ n'est jamais inférieur à la mesure extérieure dans l'espace y de $\mathrm{E}(x)$, ce qui s'écrit $\lambda(x) \geqq m_{e,y}[\mathrm{E}(x)]$. De même on a

$$m_y[\delta(x)] - \mu(x) \leqq m_{i,y}[\mathrm{E}(x)];$$

donc, d'après ce qui précède, on a presque partout

$$m_{e,y}[\mathrm{E}(x)] \leqq m_{i,y}[\mathrm{E}(x)].$$

Mais le signe = étant seul admissible, on voit que presque partout $\mathrm{E}(x)$, en tant qu'ensemble de points de l'espace y, est mesurable. De plus, sa mesure, qui est une fonction de x, est mesurable dans l'ensemble des points où elle existe, puisque dans cet ensemble elle est égale à $\lambda(x)$.

Donc on a

$$m(\mathrm{E}) = \int m_y[\mathrm{E}(x)]\, dx,$$

$m_y[\mathrm{E}(x)]$ existant partout, sauf pour des valeurs de x formant un ensemble de mesure nulle; c'est à l'ensemble des points où $m_y[\mathrm{E}(x)]$ existe que l'intégrale doit être étendue. Ou bien il faut attribuer à

$m_y[E(x)]$ une valeur quelconque, $m_{e,y}[E(x)]$ ou zéro par exemple, aux points où $m_y[E(x)]$ n'existe pas.

Si donc on appelle f la fonction égale à 1 aux points de l'ensemble E et nulle ailleurs, on a

$$m(E) = \int \left[\int f(x, y)\, dy \right] dx,$$

l'intégrale en x ne devant être étendue qu'à l'ensemble des points où l'intégrale en y existe. Si, pour les valeurs de x où cette intégrale n'existe pas, on fait aussi $f = 0$, alors les intégrales existent toujours. Or, opérer ainsi, c'est restreindre E de façon que ses sections par $x = x_0$ deviennent mesurables, ce que nous réalisons donc en négligeant seulement un ensemble de points de mesure nulle.

Ainsi, en supprimant de E un certain ensemble de points de mesure nulle, ce qui revient à modifier f aux points d'un ensemble de mesure nulle, on a

$$\int f(P)\, dP = \int \left[\int f(x, y)\, dy \right] dx.$$

Ceci n'est jusqu'ici démontré que pour les ensembles E bornés; mais l'extension au cas général est immédiate.

46. Soit maintenant $f(P)$ une fonction sommable quelconque; supposons-la positive. Soit E_n l'ensemble mesurable des points où l'on a $f > n$. En supprimant de E_r un certain ensemble de points de mesure nulle, on arrive à $\mathcal{E}_r$ tel que $\mathcal{E}_r(x)$ soit mesurable en y. Faisons cela pour tous les E_r correspondant aux r rationnels rangés dans un certain ordre; cela revient à poser $f = 0$ aux points d'un certain ensemble de mesure nulle. Comme après cela l'ensemble des points où l'on a $f > i$ est l'ensemble somme de ceux pour lesquels on a $f < r$ (r rationnel et plus grand que i), la section de cet ensemble par une variété $x = x_0$ est mesurable, et il en est évidemment de même de la section de l'ensemble des points où $\alpha \leqq f < \beta$.

Nous désignerons par $f_{n,\varepsilon}$ la fonction égale à α dans cet ensemble et nulle ailleurs, en supposant $\alpha = n\varepsilon$, $\beta = (n+1)\varepsilon$. On a alors

$$f = \lim_{\varepsilon=0,\, n\varepsilon=\infty} \left(\sum^{i=n} f_{i,\varepsilon} \right).$$

Peu importe la façon dont $\varepsilon > 0$ tend vers zéro et $n\varepsilon$ vers ∞, que ce soit simultanément ou successivement le second membre croît avec $\frac{1}{\varepsilon}$ et avec $n\varepsilon$; donc on peut intégrer terme à terme et l'on a

$$\int f(\mathrm{P})\,d\mathrm{P} = \lim_{\varepsilon=0,\,n\varepsilon=\infty}\left[\sum_{i=0}^{i=n}\int f_{i,\varepsilon}(\mathrm{P})\,d\mathrm{P}\right].$$

Transformons la somme du second membre; on peut, d'après ce qui précède, l'écrire

$$\int\left[\int\sum_{i=0}^{i=n} f_{i,\varepsilon}(x,y)\,dy\right]dx = \int a(x)\,dx;$$

$a(x)$ tend en croissant vers $\int f(x,y)\,dy$, quantité finie ou non. Si elle n'était pas finie presque partout, pour $n\varepsilon$ assez grand et ε assez petit, $a(x)$ surpasserait tel nombre que l'on voudrait dans un ensemble en x de mesure supérieure à un certain nombre fixe assez petit; donc $\int a(x)\,dx$ serait aussi grande qu'on voudrait, ce qui est en contradiction avec une égalité précédente.

Ainsi $\int f(x,y)\,dy$ a presque partout une valeur finie; négligeons l'ensemble des valeurs de x où elle est infinie; dans nos intégrations on voit que, $a(x)$ tendant en croissant vers $\int f(x,y)\,dy$, $\int a(x)\,dx$ tend vers $\int\left[\int f(x,y)\,dy\right]dx$, et l'on a donc

$$\int f(\mathrm{P})\,d\mathrm{P} = \int\left[\int f(x,y)\,dy\right]dx.$$

Cette égalité est démontrée pour la fonction f transformée, mais il en résulte qu'elle est vraie pour la fonction f primitive si l'on néglige dans l'intégration des ensembles convenables; f peut d'ailleurs évidemment être de signe quelconque; donc : *Une fonction sommable* $f(\mathrm{P})$ *étant donnée dans un ensemble* E, *il existe dans* E *un ensemble* $\mathcal{E}$ *de même mesure que* E *et tel qu'on ait*

$$\int_{\mathrm{E}} f(\mathrm{P})\,d\mathrm{P} = \int_{\mathcal{E}} f(\mathrm{P})\,d\mathrm{P} = \int\left[\int_{\mathcal{E}(x)} f(x,y)\,dy\right]dx.$$

En d'autres termes : *à condition de négliger un ensemble de mesure nulle convenable, une intégrale multiple peut se calculer par des intégrales simples successives, comme dans le cas de l'intégration des fonctions continues.*

Le raisonnement est basé sur la possibilité d'intégrer les suites croissantes terme à terme, et, d'après ce qui a été dit à ce sujet, nous pouvons compléter le résultat en remarquant que, *si l'on suppose seulement $f(\mathrm{P})$ mesurable, on pourra affirmer qu'elle est sommable toutes les fois que* $\int \left[\int_{\mathcal{E}(x)} f(x, y)\, dy \right] dx$ *aura une valeur finie.*

47. Il est facile maintenant d'étudier la différentiation par rapport à chacune des variables d'une intégrale indéfinie. Mais il est nécessaire ici de préciser ce qu'on entend par intégrale indéfinie.

Il ne peut s'agir que de l'intégrale indéfinie exprimée à l'aide des k variables; précédemment, nous entendions par là la fonction $\mathfrak{F}(x_1, x_2, \ldots x_k)$ égale à l'intégrale définie de la fonction donnée f, prise dans l'intervalle dont $(0, 0, \ldots, 0)$, $(x_1, x_2, \ldots, x_k)$ sont deux sommets opposés, augmentée de k fonctions quelconques $f_1, \ldots f_k$ des k groupes formés par $k-1$ des variables. Bien entendu, si nous laissions ces fonctions quelconques, toute dérivation serait impossible; aussi nous prendrons la suite $f_1, \ldots, f_k$ identiquement nulle, ou tout au moins telle que toutes les dérivations de $\mathfrak{F}$ possibles pour $f_1 \equiv f_2 \equiv \ldots \equiv f_k \equiv 0$ restent encore possibles et conservent les mêmes relations. Il en sera ainsi, en particulier, si les $f_1, \ldots, f_k$ ont des dérivées partielles de tous ordres ou encore si la somme $f_1 + f_2 + \ldots + f_k$ peut être remplacée par la somme $\mathrm{C} + \Sigma \mathrm{A}_{\alpha, \beta, \ldots, \lambda}$ où C est une constante et $\mathrm{A}_{\alpha, \beta, \ldots, \lambda}$ une intégrale définie d'une fonction des seules variables $\alpha, \beta, \ldots, \lambda$, prise dans un intervalle de sommets opposés $(0, \ldots, 0)$, $(\alpha, \beta_1, \ldots, \lambda)$; quant à $\alpha, \beta_1, \ldots, \lambda$, c'est un groupe quelconque formé de certaines des variables $x_1, x_2, \ldots, x_k$ en nombre $k-1$ au plus. Ces généralisations seront évidentes quand le cas de $f_1 \equiv f_2 \equiv \ldots \equiv f_k \equiv 0$ aura été étudié. Je ferai donc dorénavant l'hypothèse $f_1 \equiv f_2 \equiv \ldots \equiv f_k \equiv 0$, de façon à éviter toute complication accessoire.

Quant à la définition des dérivées, puisqu'il s'agit ici de fonctions de points, nous revenons à la définition ordinaire, qui s'exprime par l'égalité

$$\frac{\partial \mathcal{F}}{\partial x_1} = \mathcal{F}'_{x_1} = \lim_{h=0} \frac{\mathcal{F}(x_1 + h, x_2, \ldots, x_k) - \mathcal{F}(x_1, x_2, \ldots, x_k)}{h}.$$

Nos résultats resteraient évidemment vrais si l'on remplaçait cette définition par celle qui a été considérée au Chapitre IV pour le cas d'une variable.

48. Nous avons, f désignant la fonction modifiée aux points d'un ensemble de mesure nulle,

$$\mathcal{F}(x_1, x_2, \ldots, x_k) = \int_0^{x_1} \int_0^{x_2} \cdots \int_0^{x_k} f(x_1, x_2, \ldots, x_k)\, dx_k \ldots dx_2\, dx_1 :$$

donc, pour chaque valeur de $(x_2, \ldots, x_k)$, il y a au plus un ensemble de mesure nulle de valeurs de x_1 telles qu'on n'ait pas

$$\frac{\partial \mathcal{F}(x_1, \ldots, x_k)}{\partial x_1} = \int_0^{x_2} \cdots \int_0^{x_k} f(x_1, \ldots, x_k)\, dx_k \ldots dx_2.$$

L'ensemble $\mathcal{E}$ de ces valeurs exceptionnelles de $(x_1, \ldots, x_k)$ est évidemment mesurable, car la fonction $\mathcal{F}'_{x_1}$ est mesurable et par suite a une valeur déterminée finie pour les points d'un ensemble E, et $\mathcal{E}$ est l'ensemble des points extérieurs à E et de ceux où les deux fonctions mesurables $\mathcal{F}'_{x_1}$ et $\int_0^{x_2} \cdots \int_0^{x_k} f(x_1, \ldots, x_k)\, dx_k, \ldots, dx_2$ diffèrent. Les sections $\mathcal{E}(x_2, \ldots, x_k)$ étant de mesure nulle, $\mathcal{E}$ est de mesure nulle.

En négligeant les points de cet ensemble $\mathcal{E}$ il reste un ensemble sur lequel $\mathcal{F}$ est partout dérivable par rapport à x_1, sa dérivée étant finie et l'égalité précédente étant vraie.

Exactement de la même façon, on verra qu'en négligeant un nouvel ensemble de mesure nulle $\mathcal{F}''_{x_1 x_2}$ existera partout et sera égale à l'intégrale de f prise par rapport aux autres variables. Pour prévenir toute confusion, je précise que, si A est l'ensemble restant, ce qui est démontré, c'est qu'en tout point P de A la dérivée $\mathcal{F}'_{x_1}$ existe, qu'on tienne compte pour le calcul de cette dérivée en P des points extérieurs à A ou non, et qu'en P la dérivée $\mathcal{F}''_{x_1, x_2}$ existe quand, pour calculer cette dérivée, on ne tient compte que des points de A, ce qui veut dire que l'on considère $\mathfrak{F}''_{x_1,x_2}$ comme la limite de

$$\frac{\mathfrak{F}'_{x_1}(x_1, x_2+l, x_3, \ldots, x_k) - \mathfrak{F}'_{x_1}(x_1, x_2, x_3, \ldots, x_k)}{l},$$

quand l tend vers zéro par valeur telle que le point $(x_1, x_2+l, x_3 \ldots, x_k)$ soit point de A comme $(x_1, x_2, x_3, \ldots, x_k)$.

En continuant ainsi, on voit qu'en négligeant un certain ensemble de mesure nulle la dérivation successive de $\mathfrak{F}$ par rapport à chacune des variables x_1, x_2, ..., x_k, prise dans l'ordre de ces variables, est possible sur l'ensemble restant. Mais, un résultat analogue étant vrai quel que soit celui qu'on a choisi des k' ordres dans lesquels on peut ranger x_1, x_2, ..., x_k, on pourra énoncer un résultat analogue relatif cette fois à la dérivation dans un ordre quelconque. Notons enfin que nos dérivées d'ordre α s'expriment par $k-\alpha$ intégrales simples successives, dont le résultat, étant équivalent à une intégrale multiple d'ordre $k-\alpha$, est indépendant de l'ordre des intégrations. Donc le résultat de nos dérivations est indépendant de l'ordre de ces dérivations.

Nous résumerons nos résultats en disant :

Une fonction sommable f étant donnée [1], *en négligeant un ensemble de points de mesure nulle, il reste un ensemble sur lequel l'intégrale indéfinie $\mathfrak{F}$ de f, exprimée à l'aide des k variables, est dérivable une fois par rapport à chaque variable. Ces 2^k-1 dérivées sont finies et indépendantes de l'ordre des dérivations. On passe de l'une à l'autre, à f ou à $\mathfrak{F}$, par les mêmes dérivations et intégrations que dans le cas où f est continue*, en ayant soin, par conséquent, de ne pas utiliser deux dérivations ou deux intégrations par rapport à la même variable.

49. Les dérivées partielles, dont l'existence vient ainsi d'être démontrée, auraient pu être obtenues par la méthode des Chapitres précédents.

---

(1) Si f n'est définie que pour les points d'un ensemble, pour la définition de son intégrale indéfinie on lui attribue la valeur zéro là où elle n'était pas primitivement définie.

Soient f la fonction sommable considérée, $\mathcal{F}(\mathcal{E})$ son intégrale indéfinie. Partageons les coordonnées en deux groupes x et y, et soit $\mathcal{F}(x, y)$ l'intégrale indéfinie exprimée à l'aide des coordonnées. $\mathcal{F}(x, y)$ est donc, à des quantités près qui n'interviendront pas dans nos calculs, l'intégrale définie de f dans l'intervalle dont $(0, 0)$, (x, y) sont deux sommets opposés; supposons $\mathcal{F}(x, y)$ exactement égale à cette intégrale définie, pour simplifier. Il est évident que, y étant constant, $\mathcal{F}(x, y)$ est absolument continue en x; donc c'est l'intégrale indéfinie prise par rapport à x d'une certaine fonction $\varphi(x, y)$, et l'on a

$$\mathcal{F}(x, y) = \int_0^x \varphi(x, y)\, dx.$$

Mais on sait qu'on a aussi

$$\mathcal{F}(x, y) = \int_0^x \left[\int_0^y f(x, y)\, dy\right] dx;$$

donc on a presque partout $\varphi = \int_0^y f(x, y)\, dy$.

Or, d'après les propriétés du Chapitre IV, φ est presque partout la dérivée de l'intégrale indéfinie en x, ce qui veut dire que, si l'on s'occupe du point (x_0, y_0), on considère une famille régulière d'ensembles e de l'espace x, dont x_0 fait partie, et dont le diamètre tend vers zéro, et l'on prend la limite du rapport $\frac{\mathcal{F}_{y_0}(e)}{m_x(e)}$, $\mathcal{F}_{y_0}(e)$ étant la fonction d'ensemble de l'espace x qui se déduit de la fonction de point $\mathcal{F}(x, y_0)$.

Si $\mathcal{E}$ est l'ensemble des points de l'intervalle $(0, 0)$, (x, y) dans l'espace x, y, tels que leur x fasse partie de e, cela revient à prendre le rapport $\frac{\mathcal{F}(\mathcal{E})}{m(\mathcal{E})} m_y(0, y_0)$; $m_y(0, y_0)$ étant la mesure dans l'espace y de l'intervalle $(0, y_0)$.

La considération directe de ce rapport nous permettrait de définir la dérivée ou les nombres dérivés de $\mathcal{F}(\mathcal{E})$ le long de l'intervalle $0 \leqq y \leqq y_0$ de la variété $x = x_0$. Et relativement à ces nombres, on pourrait reprendre toute une étude analogue à celle du Chapitre IV.

De la même façon on définirait la dérivée de $\mathcal{F}(\mathcal{E})$ le long d'une partie quelconque d'une variété linéaire de l'espace (x, y); si, par exemple, on se borne au cas d'une partie qui devient un intervalle dans un certain changement de coordonnées, il suffit de prouver que la définition conduit à un résultat indépendant, presque partout, du changement de coordonnées particulier choisi, ce qui se fera sans peine si l'on n'oublie pas qu'avec les nouvelles variables ξ_i, l'intervalle $0 \leqq \xi_i \leqq 1$ $(i=1, 2, \ldots, k)$ n'a plus nécessairement pour mesure l'unité.

On pourrait même définir la dérivée de $\mathcal{F}(\mathcal{E})$ le long de certains ensembles plus généraux, et cela pourrait être utilisé pour les changements de variables dans les intégrales; mais je laisse cela entièrement de côté.

Il suffit de retenir qu'on peut définir la dérivée de $\mathcal{F}(\mathcal{E})$ le long d'un intervalle quelconque de l'une des variétés parallèles aux variétés coordonnées. Cette dérivée, qui existe presque partout, est exprimée par l'intégrale de f prise dans l'intervalle et considérée comme intégrale α-uple, si l'intervalle est porté par une variété à α dimensions.

Ce résultat est évidemment en relation simple avec le précédent relatif à $k-\alpha$ dérivations successives; mais il me semble que ces résultats sont différents, et, par exemple, on ne peut nullement affirmer que, là où toutes les dérivations successives sont possibles jusqu'à l'ordre $k-\alpha$, la dérivation le long de l'intervalle correspondant à α dimensions est possible. Mais on peut affirmer qu'*en laissant de côté les points d'un ensemble de mesure nulle convenablement choisi, il reste un ensemble* E *tel que, d'une part, les dérivations successives par rapport aux k variables de l'intégrale indéfinie $\mathcal{F}(x_1, x_2, \ldots, x_k)$ d'une fonction sommable f soient possibles et conduisent aux nombres qu'on obtient par l'intégration de f par rapport aux variables restées constantes dans les dérivations et que, d'autre part, la dérivée de l'intégrale indéfinie $\mathcal{F}(\mathcal{E})$ de f existe sur chaque intervalle parallèle à une variété coordonnée et dont les sommets font partie de* E, *cette dérivée étant l'intégrale de f prise dans cet intervalle par rapport aux coordonnées variables.*

50. Je vais démontrer que *sur cet ensemble* [1] *l'intégrale indéfinie*

---

[1] Il suffirait que sur cet ensemble les dérivées citées de $\mathcal{F}(\mathrm{E})$ existent, auquel cas les dérivées premières de $\mathcal{F}(x_1, x_2, \ldots, x_n)$ existent nécessairement.

$\mathfrak{F}(x_1, x_2, \ldots, x_k)$ *admet une différentielle totale première*

$$d\mathfrak{F} = \frac{\partial \mathfrak{F}}{\partial x_1} dx_1 + \frac{\partial \mathfrak{F}}{\partial x_2} dx_2 + \ldots + \frac{\partial \mathfrak{F}}{\partial x_k} dx_k,$$

c'est-à-dire que la dérivée de $\mathfrak{F}$ *prise dans la direction de coefficients directeurs* $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ *existe et est égale à l'expression*

$$\alpha_1 \frac{\partial \mathfrak{F}}{\partial x_1} + \alpha_2 \frac{\partial \mathfrak{F}}{\partial x_2} + \ldots + \alpha_k \frac{\partial \mathfrak{F}}{\partial x_k}.$$

Supposons qu'il n'y ait que trois variables x, y, z; occupons-nous du point x_0, y_0, z_0 de E, et soient α, β, γ les cosinus directeurs de la direction que l'on considère. Soit un point de E dans cette direction; ses coordonnées sont $x_0 + \alpha l$, $y_0 + \beta l$, $z_0 + \gamma l$, et l'on veut calculer la limite de

$$\frac{\mathfrak{F}(x_0 + \alpha l, y_0 + \beta l, z_0 + \gamma l) - \mathfrak{F}(x_0, y_0, z_0)}{l}.$$

Désignons, pour simplifier, par $F\begin{pmatrix} a & b & c \\ a_1 & b_1 & c_1 \end{pmatrix}$ la valeur de la fonction d'intervalle qu'on déduit de $\mathfrak{F}$ dans l'intervalle dont a, b, c et a_1, b_1, c_1 sont deux sommets opposés [2]. Le rapport à étudier peut s'écrire

$$\begin{aligned}
&\frac{1}{l}\left[F\begin{pmatrix} x_0+\alpha l & y_0+\beta l & z_0+\gamma l \\ 0 & 0 & 0 \end{pmatrix} - F\begin{pmatrix} x_0 & y_0 & z_0 \\ 0 & 0 & 0 \end{pmatrix}\right] \\
&= l^2\left[\alpha\beta\gamma \frac{1}{\alpha l \beta l \gamma l} F\begin{pmatrix} x_0+\alpha l & y_0+\beta l & z_0+\gamma l \\ x_0 & y_0 & z_0 \end{pmatrix}\right] \\
&\quad + l\left[\beta\gamma \frac{1}{\beta l \gamma l} F\begin{pmatrix} x_0 & y_0+\beta l & z_0+\gamma l \\ 0 & y_0 & z_0 \end{pmatrix}\right. \\
&\quad\quad + \alpha\gamma \frac{1}{\alpha l \gamma l} F\begin{pmatrix} x_0+\alpha l & y_0 & z_0+\gamma l \\ x_0 & 0 & z_0 \end{pmatrix} \\
&\quad\quad \left. + \alpha\beta \frac{1}{\alpha l \beta l} F\begin{pmatrix} x_0+\alpha l & y_0+\beta l & z_0 \\ x_0 & y_0 & 0 \end{pmatrix}\right] \\
&\quad + \left[\alpha \frac{1}{\alpha l} F\begin{pmatrix} x_0+\alpha l & y_0 & z_0 \\ x_0 & 0 & 0 \end{pmatrix} + \beta \frac{1}{\beta l} F\begin{pmatrix} x_0 & y_0+\beta l & z_0 \\ 0 & y_0 & 0 \end{pmatrix}\right. \\
&\quad\quad\quad \left. + \gamma \frac{1}{\gamma l} F\begin{pmatrix} x_0 & y_0 & z_0+\gamma l \\ 0 & 0 & z_0 \end{pmatrix}\right].
\end{aligned}$$

---

(1) *Voir* Chap. II.

La quantité placée dans la première accolade du second membre tend vers la dérivée de $\mathcal{F}(\mathcal{E})$ au point x_0, y_0, z_0, multipliée par $\alpha\beta\gamma$; les quantités placées dans la seconde accolade tendent vers $\beta\gamma$, $\alpha\gamma$, $\alpha\beta$ multipliées respectivement par les dérivées de $\mathcal{F}(\mathcal{E})$ suivant les segments (o, x_0), (o, y_0), (o, z_0) des parallèles aux axes de coordonnées menées par x_0, y_0, z_0. On verrait de même les limites des termes de la troisième accolade; mais il est peut-être plus simple de l'écrire

$$\frac{\alpha}{\alpha l}[\mathcal{F}(x_0+\alpha l, y_0, z_0) - \mathcal{F}(x_0, y_0, z_0)] + \frac{\beta}{\beta l}[\mathcal{F}(x_0, y_0+\beta l, z_0) - \mathcal{F}(x_0, y_0, z_0)] + \frac{\gamma}{\gamma l}[\mathcal{F}(x_0, y_0, z_0+\gamma l) - \mathcal{F}(x_0, y_0, z_0)],$$

et le théorème énoncé résulte immédiatement de cela.

Deux remarques sont cependant utiles : nous avons admis qu'il existait des points de E situés sur la parallèle à la direction considérée et aussi voisins qu'on veut de $P_0(x_0, y_0, z_0)$; cela n'est pas nécessairement vrai. De P_0 il pourra partir un ensemble de mesure nulle de directions sur lesquelles cela n'a pas lieu. Ce qui signifie que les droites exceptionnelles issues de P_0 couvrent un ensemble de points dont la mesure à k dimensions est nulle. Sur ces directions exceptionnelles, variables avec P_0, il n'y a pas lieu de parler de la dérivée en P_0 de $\mathcal{F}(x_1, x_2, \ldots, x_k)$; mais on ne sera jamais conduit à une contradiction en attribuant à cette direction la valeur de la dérivée que fournit la considération de la différentielle totale.

Une autre remarque moins importante est la suivante : le raisonnement précédent s'appuie sur la définition de la dérivée qui, avec les notations précédemment utilisées, se déduit de $\frac{\mathcal{F}(\mathcal{E})}{m_x(e)}$ et non pas sur la définition qui se déduit de $\frac{\mathcal{F}(\mathcal{E})}{m(\mathcal{E})} m_y(o, y_0)$; ces deux définitions sont bien d'accord pour $m_y(o, y_0) \neq o$; mais, pour $m_y(o, y_0) = o$, il faut convenir de donner à $\frac{m(\mathcal{E})}{m_y(o, y_0)}$ la valeur $m_x(e)$.

Peut-on de même définir une différentielle seconde ? Évidemment non, car cela n'est déjà plus possible lorsque la fonction intégrée f, étant continue, n'a pas de dérivées partielles. Cela revient à dire qu'étant données deux directions issues de P_0, on ne peut en général pas parler de la dérivée partielle seconde de $\mathcal{F}(x_1, x_2, \ldots, x_k)$ suivant l'ensemble de ces deux directions, ni, quand cette dérivée existe, la calculer à l'aide des formules du changement de coordonnées.

Cela est possible cependant, on le vérifiera par un raisonnement analogue au précédent, toutes les fois que ces formules du changement de variables ne conduisent pas à prendre deux fois la dérivée par rapport à la même variable. Par exemple, si l'on a 4 variables x, y, z, t, on pourra appliquer les formules du changement de variables pour avoir la dérivée suivant l'ensemble des deux directions

$$x+y=0, \quad z=0, \quad t=0; \quad x=0, \quad y=0, \quad z+t=0.$$

Je crois devoir ici mettre en garde contre une faute très grossière, mais si tentante que, pour ma part, je l'ai commise plusieurs fois au cours de mes réflexions. Effectuer sur $\mathcal{F}(x_1, x_2, \ldots, x_k)$ un changement de coordonnées qui l'exprime en $\xi_1, \xi_2, \ldots, \xi_k$, ce n'est pas du tout former l'intégrale indéfinie de f exprimée à l'aide des variables ξ_1, $\xi_2, \ldots, \xi_k$, et par suite mettre en doute l'existence de dérivées telles que $\frac{\partial^2 \mathcal{F}(x_1, x_2, \ldots, x_k)}{\partial \xi_1 \partial \xi_2}$, ce n'est pas du tout nier l'existence presque partout d'une dérivée en ξ_1, ξ_2 de l'intégrale indéfinie de f exprimée à l'aide des coordonnées $\xi_1, \xi_2, \ldots, \xi_k$. Le gros avantage qu'il y a à considérer une intégrale indéfinie comme fonction d'ensemble, c'est précisément qu'elle est ainsi indépendante des coordonnées choisies.

51. Ceci va nous conduire à examiner la question suivante : Pour les points de l'ensemble, a-t-on

$$\frac{\partial \mathcal{F}[x_1(\xi_1, \ldots), x_2(\xi_1, \ldots), \ldots]}{\partial \xi_2} = \frac{\partial \mathcal{F}}{\partial x_1}\frac{\partial x_1}{\partial \xi_1} + \frac{\partial \mathcal{F}}{\partial x_2}\frac{\partial x_2}{\partial \xi_1} + \ldots + \frac{\partial \mathcal{F}}{\partial x_k}\frac{\partial x_k}{\partial \xi_1};$$

en d'autres termes, la différentielle totale $d\mathcal{F}$ peut-elle être utilisée pour effectuer dans $\mathcal{F}$ un changement de variables ?

Reprenons le cas de trois variables x, y, z; pour ce cas, nous avons défini la dérivée dans la direction α, β, γ issue de x_0, y_0, z_0 par la considération des points $x_0 + \alpha l$, $y_0 + \beta l$, $z_0 + \gamma l$. Pour qu'on puisse effectuer le changement de variables comme il est dit plus haut, il est évidemment nécessaire et suffisant que les résultats concernant l'existence de $d\mathfrak{F}$ subsistent quand on définit la dérivée, dans la direction α, β, γ issue de x_0, y_0, z_0, par la considération des points de l'ensemble $x_0 + \alpha' l, y_0 + \beta' l, z_0 + \gamma' l$; α', β', γ' étant trois cosinus directeurs associés qui tendent vers α, β, γ quand l tend vers zéro (¹). Si nous reprenons le raisonnement fait plus haut, nous voyons qu'il subsiste pourvu que α, β, γ soient tous trois différents de zéro; mais si l'un d'eux, α, est nul, le raisonnement ne subsiste sans changement que si le coefficient correspondant α' est nul aussi. Si α et β sont tous deux nuls, alors α' et β' doivent l'être aussi.

Ainsi *le calcul des dérivées premières* $\dfrac{\partial \mathfrak{F}[x_1(\xi_1, \ldots), x_2(\xi_1, \ldots), \ldots]}{\partial \xi_i}$ *peut être fait à l'aide de la différentielle totale si aucune des dérivées* $\dfrac{\partial x_i}{\partial \xi_j}$ $(i, j = 1, 2, \ldots, k)$ *n'est nulle*, ou encore si $\dfrac{\partial x_i}{\partial \xi_1}$ par exemple étant nulle pour les valeurs $\xi_1^0, \xi_2^0, \ldots, \xi_k^0$ des ξ, $\dfrac{\partial x_i}{\partial \xi_1}$ est nulle quel que soit ξ_1 pour les valeurs $\xi_2^0, \ldots, \xi_k^0$ des autres ξ.

Il faut remarquer que, dans tout ce qui précède, le fait que $\mathfrak{F}$ est une intégrale indéfinie n'est guère intervenu; nous avons prouvé en somme que le fait pour $\mathfrak{F}(x_1, x_2, \ldots, x_k)$ de correspondre à une fonction d'intervalle F ayant une dérivée en tout point, quand on a négligé un certain ensemble de mesure nulle, et le long de tout intervalle des variétés parallèles aux variétés coordonnées qui a pour sommets opposés deux points de l'ensemble conservé, entraîne pour $\mathfrak{F}$ l'existence d'une différentielle totale sur cet ensemble, laquelle, sous les réserves indiquées, peut être utilisée dans les changements de variables. D'ailleurs, en reprenant les raisonnements en sens inverse, on verra que réciproquement l'existence d'une différentielle totale, au sens indiqué, pour $\mathfrak{F}$ entraîne l'existence des dérivées de F. Il n'y a donc dans tout cela que des transformations d'énoncés.

Tenons compte maintenant du fait que $\mathfrak{F}$ est l'intégrale indéfinie d'une fonction sommable f, et reprenons le raisonnement en supposant, par exemple, α seul nul. Les termes qu'il y a lieu d'examiner à nouveau sont les trois qui contiennent α' en dénominateurs, car ces

---

(¹) De sorte que maintenant $\mathfrak{F}$ admet une dérivée suivant chaque direction issue de x_0, y_0, z_0.

termes contiennent des rapports qui ne sont plus relatifs à des intervalles réguliers et dont par suite on ne connaît pas la limite. Le premier peut s'écrire

$$\beta'\gamma' l \frac{1}{\beta' l \gamma' l} \mathrm{F}\begin{pmatrix} x_0+\alpha' l & y_0+\beta' l & z_0+\gamma' l \\ x_0 & y_0 & z_0 \end{pmatrix}.$$

Supposons qu'on s'occupe des points d'un certain intervalle, par exemple $0 \leqq x_i \leqq 1$, et soit F_1 la fonction d'intervalle en y, z qui est l'intégrale indéfinie de $\int_0^1 |f(x, y, z)|\, dx$. Le terme considéré est inférieur en module à $l \frac{1}{\beta' l \gamma' l} \mathrm{F}_1\begin{pmatrix} y_0+\beta' l, & z_0+\gamma' l \\ y_0 & z_0 \end{pmatrix}$, c'est-à-dire au porduit par l d'une quantité qui tend vers la dérivée de F_1 au point (y_0, z_0) si elle existe. Par suite, presque partout ce terme a une limite nulle.

Les deux autres termes sont de forme analogue ; l'un d'eux s'écrit

$$\beta' \frac{1}{\beta' l} \mathrm{F}\begin{pmatrix} x_0+\alpha' l & y_0+\beta' l & z_0 \\ x_0 & y_0 & 0 \end{pmatrix}.$$

Je dis que la limite de ce terme est nulle presque partout. En effet, si E désigne l'ensemble des points où la limite est supérieure en module à ε, ou bien E est de mesure nulle, ou bien on peut trouver un nombre fini de valeurs de x_0 tels que la somme des mesures en y des projections des $\mathrm{E}(x_0)$ sur l'axe des y surpasse tel nombre qu'on veut M, $\mathrm{E}(x_0)$ étant la section par $x = x_0$ de E. Partageons l'axe des x en un nombre fini de parties dans chacune desquelles se trouve un seul nombre x_0. Soient (a, b) une de ces parties, x_0 le nombre qu'elle contient.

A chaque point de $\mathrm{E}(x_0)$ on peut attacher une infinité de parallélépipèdes dont $x_0, y_0, 0$; $x_0+\alpha' l, y_0+\beta' l, z_0$ sont deux sommets opposés tels que $\alpha' l$ et $\beta' l$ tendent vers zéro et tels qu'on ait

$$\left| \frac{1}{\beta l} \mathrm{F}\begin{pmatrix} x_0+\alpha' l & y_0+\beta' l & z_0 \\ x_0 & y_0 & 0 \end{pmatrix} \right| > \frac{\varepsilon}{2}.$$

Ne conservons que ceux de ces parallélépipèdes pour lesquels on a $a < x_0 + \alpha' l < b$. Leurs projections sur l'axe des y sont des intervalles couvrant l'ensemble $\mathrm{E}_y(x_0)$, projection de $\mathrm{E}(x_0)$. Appliquant le théorème de M. Vitali, nous concluons qu'il y a une infinité dénombrable de ces intervalles non empiétants qui couvrent tout $E_y(x_0)$; ils correspondent à des parallélépipèdes non empiétants, et l'on voit que la somme $\Sigma|F(I)|$, étendue à tous ces parallélépipèdes, surpasse

$$\frac{\varepsilon}{2} m_y [E_y(x_0)].$$

En opérant de même sur chaque (a, b), on est conduit à une somme totale $\Sigma|F(I)|$, surpassant $\frac{\varepsilon}{2}M$, et, comme ε peut être pris fixe et M arbitrairement grand, la proposition est démontrée (1).

Ainsi, à condition peut-être de négliger un nouvel ensemble de points de mesure nulle, le cas où l'un des coefficients directeurs est nul fournira le même énoncé que le cas tout d'abord étudié. Il n'y aura pas de difficultés nouvelles si plusieurs de ces coefficients sont nuls, de sorte que nous pouvons conclure : *Si l'on fait abstraction des points d'un ensemble de mesure nulle convenablement choisi, l'intégrale indéfinie* $\mathcal{F}(x_1, x_2, \ldots, x_k)$ *d'une fonction sommable f admet sur l'ensemble restant* E *une différentielle totale première permettant de calculer les nouvelles dérivées premières de $\mathcal{F}$ après un changement de variables quelconque.*

Bien entendu, pour le calcul de la dérivée première de $\mathcal{F}$ en un point P de E et dans une direction PX, on peut évidemment se servir des rapports $\frac{\mathcal{F}(P)-\mathcal{F}(Q)}{PQ}$ quand Q tend vers P, de façon que PQ tende vers PX, Q appartenant à E ou non ; car, si Q n'appartient pas à E, on peut

---

(1) Il me paraît très curieux que, dans cette démonstration, on n'ait pas tenu compte de $\alpha = \lim \alpha' = 0$; bien au contraire, $\alpha' l$ pourrait être infiniment grand vis-à-vis de $\beta' l$ sans que la proposition cesse d'être exacte. Ce résultat serait d'ailleurs évident, s'il était prouvé que les nombres dérivés d'une intégrale indéfinie, définis à l'aide d'intervalles quelconques, et non plus nécessairement réguliers, c'est-à-dire définis comme le fait M. Vitali, sont presque partout finis.

Si l'on suppose $\alpha' < \beta'$ et $\alpha' < \gamma'$, on voit facilement que les termes que nous venons de démontrer être presque partout nuls à la limite sont partout nuls à la limite si f est bornée. Peut-être en est-il encore de même pour f sommable quelconque, de sorte qu'on pourrait démontrer relativement aux intégrales indéfinies une propriété analogue à celle que Riemann démontre pour les fonctions obtenues en intégrant deux fois une série trigonométrique à termes tendant vers zéro (théorème II du paragraphe VIII du Mémoire de Riemann). Sans pouvoir en préciser l'énoncé, il me paraît d'ailleurs certain qu'un théorème de cette nature peut être prouvé en ce qui concerne les intégrales indéfinies.

trouver Q_1 appartenant à E tel qu'on change aussi peu qu'on veut le rapport précédent et la direction PQ, en remplaçant Q par Q_1.

Il faudrait essayer maintenant d'étendre les résultats précédents au cas des fonctions $\mathcal{F}(x_1, x_2, \ldots, x_k)$, qui sont seulement à variation bornée. Je n'aborde ici ni cette question ni les suivantes. L'existence presque partout d'une différentielle totale entraîne-t-elle l'existence de dérivées partielles d'ordre supérieur? Sous quelles conditions peut-on affirmer qu'une fonction $\mathcal{F}(x, y)$, qui, en tant que fonction de x seul et de y seul, est une intégrale indéfinie, est une intégrale indéfinie en xy?

VII. — Quelques théorèmes généraux concernant l'intégration.

52. Je terminerai en donnant les théorèmes généraux relatifs à l'approximation ou au calcul exact des intégrales $\int f(\mathrm{P})\,\varphi(\mathrm{P})\,d\mathrm{P}$. Les théorèmes sur l'approximation résultent des remarques suivantes :

(*a*) *Si l'on a, quel que soit* i,

$$b_i \geqq 0, \qquad m \leqq a_i \leqq \mathrm{M},$$

on a

$$m \Sigma b_i \leqq \Sigma a_i b_i \leqq \mathrm{M} \Sigma b_i.$$

Si l'on suppose que i varie de o à n et si l'on utilise l'identité d'Abel avant d'appliquer les inégalités précédentes, on voit que :

(*b*) *Si les* a_i *vont en croissant et sont positifs, et si l'on a*

$$m \leqq (b_0 + b_1 + \ldots + b_i) \leqq \mathrm{M},$$

on a

$$m a_n \leqq \Sigma a_i b_i \leqq \mathrm{M} a_n.$$

Enfin on a la proposition exprimée par l'inégalité

$$(\Sigma a_i b_i)^2 \leqq \Sigma a_i^2 \Sigma b_i^2.$$

53. L'intégrale de $f\varphi$ pouvant de bien des manières être considérée comme la limite d'une suite $\Sigma a_i b_i$, ces remarques pourront être appliquées de diverses façons pour les résultats que j'ai en vue; on peut

opérer comme il suit : Soient $d_0, d_1, d_2, \ldots, d_n$ des nombres dont nous ferons augmenter le nombre indéfiniment de façon que la différence maximum de deux d'entre eux tende vers zéro et qu'ils finissent par couvrir de d_0 à d_n au moins les intervalles de variation de f, de φ et de $f\varphi$. Soit $E_{\alpha\beta}$ l'ensemble des points pour lesquels on a à la fois $d_\alpha \leqq f < d_{\alpha+1}$, $d_\beta \leqq \varphi < d_{\beta+1}$. L'intégrale est la limite de $\sum\limits_{\alpha,\beta} d_\alpha d_\beta m(E_{\alpha\beta})$. Pour appliquer (a) on fera jouer aux d_α le rôle des a_i et aux $d_\beta m(E_{\alpha\beta})$ le rôle des b_i; on voit ainsi que : *Si φ n'est jamais négatif et si l'on a constamment $m \leqq f \leqq M$, on a aussi*

$$m\int_E \varphi(P)\,dP \leqq \int_E f(P)\,\varphi(P)\,dP \leqq M\int_E \varphi(P)\,dP.$$

C'est *le premier théorème de la moyenne*, qui s'obtient de bien des manières et qu'on a déjà utilisé précédemment.

De même, en faisant jouer à $d_\alpha\sqrt{m(E_{\alpha\beta})}$ le rôle de a_i et à $d_\beta\sqrt{m(E_{\alpha\beta})}$ le rôle de b_i, on est conduit à l'*inégalité de Schwartz* en appliquant la proposition (c). Donc, *si f^2 et φ^2 sont sommables, $f\varphi$ l'est aussi et l'on a*

$$\left[\int_E f(P)\,\varphi(P)\,dP\right]^2 \leqq \int_E f^2(P)\,dP \int_E \varphi^2(P)\,dP.$$

54. J'insisterai davantage sur *le second théorème de la moyenne* qui va résulter de (b). Mais, pour appliquer (b), il faut mettre quelque ordre dans les termes de la somme $\Sigma a_i b_i$ sur laquelle on opère. Ici on fera cela soit en effectuant dans $\sum\limits_{\alpha\,\beta} d_\alpha d_\beta m(E_{\alpha\,\beta})$ les sommations en d_β de façon à obtenir $\sum\limits_\alpha d_\alpha D_\alpha$, soit, ce qui revient au même, comme il suit.

Supposons φ bornée; soient l et L ses limites inférieure et supérieure, et prenons $d_0 = l < d_1 < d_2 < \ldots < d_n = L$. Désignons par $E(d_{i+1})$ l'ensemble des points où l'on a $\varphi < d_{i+1}$ (¹) et par $\mathcal{E}_i$ la différence $E(d_{i+1}) - E(d_i)$. On a

$$\int_E f(P)\,\varphi(P)\,dP = \sum_{i=0}^{i=n}\int_{\mathcal{E}_i} f(P)\,\varphi(P)\,dP = \sum_{i=0}^{i=n} d_i \int_{\mathcal{E}_i} f(P)\,dP,$$

(¹) On peut remplacer le signe $<$ par le signe $\leqq$; d'où des modifications analogues dans l'énoncé suivant.

la dernière égalité n'étant qu'approchée. Donc on a à peu près

$$\int_E f(P)[\varphi(P)-l]\,dP = \sum_{i=0}^{i=n}(d_i - l)\int_{\mathcal{E}_i} f(P)\,dP.$$

Si l'on pose $a_i = d_i - l$, $b_i = \int_{\mathcal{E}_i} f(P)\,dP$, on voit, en appliquant la proposition (b), que : *Si l'on désigne par l et L les limites inférieure et supérieure de $\varphi(P)$ dans E, et par $E(\xi)$, $F(\xi)$ les deux ensembles complémentaires par rapport à E aux points desquels on a respectivement*

$$l \leqq \varphi(P) < \xi, \qquad \xi \leqq \varphi(P) \leqq L,$$

l'intégrale $\int_E f(P)\varphi(P)\,dP$ *est comprise entre les limites inférieure et supérieure de*

$$L\int_{E(\xi)} f(P)\,dP + l\int_{F(\xi)} f(P)\,dP,$$

ξ variant de l à L.

Si, comme cela a fréquemment lieu dans la pratique, l'égalité $f(P) = \xi$ ne peut être réalisée pour chaque valeur de ξ que pour un ensemble de points P de mesure nulle, il existe alors un nombre ξ entre l et L tel qu'on ait exactement

$$\int_E f(P)\varphi(P)\,dP = L\int_{E(\xi)} f(P)\,dP + l\int_{F(\xi)} f(P)\,dP.$$

Pour le cas général on peut donner un énoncé analogue. $m[E(\xi)]$ est en effet une fonction non décroissante. Posons $t = m[E(\xi)] + \xi$. A une valeur de t ne peut correspondre qu'une valeur de ξ au plus; si à t ne correspond aucune valeur et que t soit cependant compris entre $m[E(l)] + l$ et $m[E(L)] + L$, c'est qu'on peut trouver ξ_0 tel qu'on ait

$$\{m[E(\xi)] + \xi\}_{(\xi=\xi_0)} < t \leqq \{m[E(\xi)] + \xi\}_{(\xi=\xi_0+0)};$$

car $m[E(\xi)]$ est évidemment continue à gauche.

Alors définissons un ensemble $e(t)$ de façon que $e(t)$ contienne $e(t')$ si l'on a $t > t'$, que $m[e(t)]$ soit une fonction continue de t et que, si t correspond à ξ, $e(t)$ soit identique à $E(\xi)$. Cela est évidemment toujours possible.

Or il est évident qu'on a toujours

$$\int_{E} f(P)\varphi(P)\,dP = L\int_{e(t)} f(P)\,dP + l\int_{E-e(t)} f(P)\,dP$$

pour une certaine valeur de t. Pour énoncer cela d'une façon générale, convenons de dire qu'on a *une famille continue d'ensembles* $a(\lambda)$ si la mesure de l'ensemble des points de $a(\lambda)$ ne faisant pas partie de $a(\lambda')$ et des points de $a(\lambda')$ ne faisant pas partie de $a(\lambda)$ tend vers zéro avec $\lambda-\lambda'$.

Soit $a(\lambda)$ une famille continue d'ensembles compris dans l'ensemble E *définie pour λ variant de* 0 *à* 1 *et comprenant les ensembles de points en lesquels on a $\varphi(P)<\xi$, quel que soit ξ compris entre les limites inférieure et supérieure l et* L *de $\varphi(P)$ et les ensembles $\varphi(P)\leqq\xi$ qui diffèrent des précédents de plus d'un ensemble de mesure nulle. Soit $b(\lambda)$ la famille complémentaire de $a(\lambda)$ par rapport à l'ensemble d'intégration* E. On a

$$\int_{E} f(P)\varphi(P)\,dP = L\int_{a(\theta)} f(P)\,dP + l\int_{b(\theta)} f(P)\,dP,$$

pour une valeur de θ comprise entre 0 *et* 1, $0\leqq\theta\leqq 1$.

On peut remarquer que, jusqu'ici, il n'a été fait sur f et φ aucune autre hypothèse que celle-ci : f est sommable, φ est mesurable et bornée. Supposons maintenant φ monotone par rapport à chacune des k coordonnées, ce qui veut dire qu'il existe des nombres θ_i égaux à ± 1 tels que $\varphi(\theta_1 x_1, \theta_2 x_2, \ldots, \theta_k x_k)$ soit une fonction croissante de chaque coordonnée. Alors $E(\xi)$ jouit de la propriété suivante : Si $x_1, \ldots, x_k$ en fait partie, tout point $x'_1, x'_2, \ldots, x'_k$ de E satisfaisant aux k inégalités $\theta_i x_i > \theta_i x'_i$ en fait aussi partie. Si l'on considère tous les ensembles A jouissant de cette propriété, il y a parmi les A un A_0 qui peut jouer le rôle de $a(\theta)$ [1], $B = E - A_0$ jouant le rôle de $b(\theta)$.

On retrouve ainsi en particulier l'énoncé classique [2] pour le cas d'une seule variable, énoncé qu'on pourrait étendre au cas général ;

---

[1] Car l'ensemble des points où $\varphi=\xi$ ne contenant qu'un point au plus sur chaque parallèle aux axes de coordonnées est de mesure nulle.

[2] On trouvera une démonstration de cet énoncé, valable pour le cas d'une variable, dans un article de M. Hobson (**2**).

mais je préfère donner une forme du théorème de la moyenne qui paraît moins connue, bien qu'elle soit aussi immédiate et qu'elle soit très commode pour les applications.

55. Pour éviter des complications accessoires je supposerai que E soit un intervalle ; soit O un point de E; prenons ce point pour origine des coordonnées, et ainsi E est divisé en 2^k intervalles au plus dont O est sommet. Considérons celui pour lequel toutes les coordonnées sont positives, soit $\mathcal{C}$. $\varphi(P)$ étant supposée à variation bornée, nous pouvons poser

$$\varphi(P) = \varphi(O) + \psi(P) - \chi(P),$$

$\psi(P)$ et $\chi(P)$ étant *croissantes* avec les coordonnées et nulles en O. Appelons *domaine simple contenant* O un domaine tel que, si P en fait partie, les segments parallèles aux axes coordonnés et compris entre P et les variétés coordonnées en fassent aussi partie ; un tel domaine sera limité dans le quadrant positif par une frontière qu'on pourra définir par une relation de la forme $x_i = F(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$, F étant une fonction non croissante des $k-1$ coordonnées dont elle dépend. On pourrait dire aussi qu'un domaine simple est celui tel que toute parallèle aux axes de coordonnées n'ait au plus qu'un segment commun avec le domaine ; mais cette définition est inutilement large; elle devient équivalente à la précédente si l'on ajoute que le segment en question doit contenir un point de l'une des variétés coordonnées. Ce sont ces domaines simples qui vont jouer pour ψ et pour χ le rôle des ensembles A considérés précédemment. On a donc, pour ψ par exemple, $\overline{u}$ désignant la limite supérieure de u,

$$\left| \int_{\mathcal{C}} f(P)\,\psi(P)\,dP \right| \leqq \overline{\left| \int_{A} f(P)\,dP \right|} \times \overline{\psi(P)},$$

A étant un domaine simple quelconque contenu dans $\mathcal{C}$ et contenant O.

Il en résulte de suite

$$\left| \int_{\mathcal{C}} f(P)\varphi(P)\,dP - \varphi(O)\int_{\mathcal{C}} f(P)\,dP \right| \leqq V \overline{\left| \int_{A} f(P)\,dP \right|}.$$

V désignant la variation totale de φ dans $\mathcal{C}$. Appliquant cela à chaque partie de E, on arrive à la forme suivante : O *étant un point de l'intervalle d'intégration* E, *on désigne par* M *la limite supérieure de la valeur absolue de l'intégrale* $\int f(\mathrm{P})\,d\mathrm{P}$ *étendue à un domaine simple quelconque contenant* O *et contenu dans* E. *Si* V *est la variation totale de* $\varphi(\mathrm{P})$ *dans* E, *on a*

$$\left|\int_{\mathrm{E}} f(\mathrm{P})\varphi(\mathrm{P})\,d\mathrm{P} - \varphi(\mathrm{O})\int_{\mathrm{E}} f(\mathrm{P})\,d\mathrm{P}\right| \leqq \mathrm{MV}.$$

56. Relativement au calcul exact de certaines intégrales $\int f\varphi\,d\mathrm{P}$, il faut citer les deux théorèmes concernant l'intégration par parties et l'intégration par substitution. Je me bornerai aux faits les plus immédiats, presque aussi simples que les résultats classiques.

Soient $\mathrm{F}(x_1, x_2, \ldots, x_k)$, $\Phi(x_1, x_2, \ldots, x_k)$ deux intégrales indéfinies exprimées à l'aide des k coordonnées [1], et soit E un ensemble mesurable. Sauf aux points d'un ensemble e de mesure nulle, F et Φ peuvent être dérivées en tout point de E une fois par rapport à chaque variable, et l'on a identiquement, aux points de $\mathrm{E} - e$,

$$\begin{aligned}&[\mathrm{F}(x_1, x_2, \ldots, x_k)\,\Phi(x_1, x_2, \ldots, x_k)]_{x_1, x_2, \ldots, x_k}\\ &\quad = \Sigma \mathrm{F}_{\alpha_1, \alpha_2, \ldots, \alpha_k}(x_1, x_2, \ldots, x_k)\,\Phi_{\beta_1, \beta_2, \ldots, \beta_k}(x_1, x_2, \ldots, x_k);\end{aligned}$$

dans cette formule, $\alpha_1, \alpha_2, \ldots, \alpha_k$ est une combinaison quelconque des $2k$ indices $x_1, x_2, \ldots, x_k$, o, o, ..., o; $\beta_1, \beta_2, \ldots, \beta_k$ est la combinaison complémentaire; le signe Σ représente une somme étendue à toutes les combinaisons $\alpha_1, \alpha_2, \ldots, \alpha_k$ différentes; après suppression des zéros, les indices indiquent des dérivations.

Si f et φ sont les valeurs de F et Φ en un point d'un ensemble $\mathcal{E}$, U et V leurs variations totales dans $\mathcal{E}$, la variation totale de $\mathrm{F}\Phi$ dans $\mathcal{E}$ est au plus $(f + \mathrm{U})(\varphi + \mathrm{V})$, et de là il résulte de suite que $\mathrm{F}\Phi$ est absolument continue puisqu'on a supposé que F et Φ l'étaient. Il résulte de là que les deux membres de l'identité précédente sont sommables

[1] Il y a lieu de tenir compte ici du fait que les fonctions additives qui figurent dans la définition générale des intégrales indéfinies doivent être prises convenablement dérivables ; on pourra en particulier ne pas prendre de fonctions additives.

dans $E - e$; donc on a

$$[F(x_1, \ldots, x_k)\,\Phi(x_1, \ldots, x_k)]_E$$
$$= \int \cdots \int_{E-e} [\Sigma F_{\alpha_1,\ldots,\alpha_k}(x_1, \ldots, x_k)\,\Phi_{\beta_1,\ldots,\beta_k}(x_1, \ldots, x_k)]\,dx_1 \ldots dx_k.$$

Dans le second membre on pourra remplacer $E - e$ par E si l'on convient de remplacer par zéro la somme Σ là où elle n'a pas de valeur déterminée et finie. Cette formule légitime l'*intégration par parties* (1).

On peut l'utiliser de la manière suivante : on ne peut intervertir les signes $\int \cdots \int$ et Σ, car on ne sait pas si chaque terme $F_{\alpha_1,\ldots,\alpha_k}$, $\Phi_{\beta_1,\ldots,\beta_k}$ est sommable. Mais supposons qu'un de ces termes soit sommable, ce qui sera par exemple toujours le cas pour les termes où tous les $\alpha_1, \ldots, \alpha_k$ ou tous les $\beta_1, \ldots, \beta_k$ sont nuls. On pourra faire sortir ce terme T du signe Σ et poser $\Sigma = \Sigma' + T$; dès lors on aura

$$\int_E T(P)\,dP = (F\Phi)_E - \int_{E-e} \Sigma'(P)\,dP;$$

c'est la forme même de la formule classique d'intégration par parties.

57. Pour légitimer l'*intégration par substitution*, considérons un changement de variables univoque et continu dans les deux sens, pour éviter toute complication inutile, remplaçant les $x_1, x_2, \ldots, x_k$ par les $y_1, y_2, \ldots, y_k$ ou inversement. Par définition même cette transformation fait correspondre à un domaine D_x un domaine D_y; je supposerai que, lorsque D_y est quarrable, D_x l'est aussi, de sorte que l'aire de D_x sera une fonction du domaine D_y; je supposerai cette fonction absolument continue, de sorte qu'on aura

$$m_x(D_x) = \int \cdots \int_{D_y} \varphi(y_1, \ldots, y_k)\,dy_1 \ldots dy_k,$$

(1) M. Léonida Tonelli a donné une formule d'intégration par parties quelque peu différente (**13**).

φ étant une fonction sommable non négative, presque partout égale à la dérivée de $m_x(D_x)$.

Soient E_x, E_y deux ensembles correspondants; supposons E_x mesurable; E_y l'est à cause de la continuité des formules de transformation. Enfermons E_y dans des domaines quarrables A_y que nous diminuerons progressivement, de façon que $m_y(A_y)$ tende vers $m_y(E_y)$. A_x correspond à A_y. Si l'on désigne par χ une fonction égale à φ aux points de A_x et égale à zéro ailleurs, et si l'on suppose E_y et A_y dans un intervalle I_y, on a

$$m_x(A_x) = \int_{I_y} \chi(P)\, dP;$$

d'où, en remarquant que $\chi(P)$ tend presque partout en décroissant vers la fonction $X(P)$ égale à φ aux points de E_y et égale à zéro ailleurs,

$$m_x(E_x) \leqq \int_{I_y} X(P)\, dP,$$

F_x étant le complémentaire de E_x par rapport à I_y; on a de même

$$m_x(F_x) \leqq \int_{I_y} [\varphi(P) - X(P)]\, dP.$$

Mais

$$m_x(E_x) + m_x(F_x) = \int_{I_y} \varphi(P)\, dP;$$

donc ce sont les signes $=$ qui conviennent dans les inégalités précédentes.

On voit en particulier qu'à un E_y de mesure nulle correspond un E_x de mesure nulle; la réciproque n'est peut-être pas vraie, mais, si à E_x de mesure nulle correspond E_y, φ est presque partout nulle sur E_y.

Ceci étant, soit d'abord $f(x_1, \ldots, x_k)$ sommable et bornée; alors, en désignant par $f_1'(y_1, \ldots, y_k)$ ce que devient cette fonction après la substitution, $f'\varphi$ est sommable, et l'on peut comparer les deux fonctions d'ensembles

$$\int \cdots \int_{E_x} f(x_1, \ldots, x_k)\, dx_1 \ldots dx_k,$$

$$\int \cdots \int_{E_y} f'(y_1, \ldots, y_k)\, \varphi(y_1, \ldots, y_k)\, dy_1 \ldots dy_k.$$

Puisque à E_y mesurable correspond E_x mesurable, on peut considérer les deux intégrales comme des fonctions de l'ensemble mesurable E_y. Puisque à E_y de mesure nulle correspond E_x de mesure nulle, ce sont des fonctions absolument continues, et, pour prouver leur égalité, il suffit de montrer qu'elles ont la même dérivée presque partout. La dérivée de la première en un point P_x, ou P_y, sera $f(P_x)$ multiplié par la dérivée en P_x de $m_x(D_x)$ considérée comme fonction de D_y; donc ce sera $f'(P_y)\varphi(P_y)$, sauf peut-être aux points P_y où φ n'est pas la dérivée de son intégrale indéfinie, lesquels forment un ensemble de mesure nulle e_y, et, sauf peut-être aux points d'un ensemble $\mathcal{E}_y$, correspond à l'ensemble $\mathcal{E}_x$ de mesure nulle pour lequel f n'est pas la dérivée de son intégrale indéfinie. Mais on sait presque en tout point de $\mathcal{E}_y$, que, $\varphi(P_y)$ est nul; donc, sauf aux points P_y formant un ensemble de mesure nulle, la dérivée est $f'(P_y)\varphi(P_y)$. Donc on a

$$\begin{aligned}\int\cdots\int_{E_x} f(x_1, \ldots, x_k)\,dx_1\ldots dx_k \\ =\int\cdots\int_{E} f(x_1, \ldots, x_k)\,\varphi(y_1, \ldots, y_k)\,dy_1\ldots dy_k.\end{aligned}$$

C'est la formule d'intégration par substitution; elle n'est démontrée que pour $f(P_x)$ bornée; mais il suffit de prouver que l'hypothèse $f(P_x)$ sommable entraîne cette conséquence $f'(P_y)\varphi(P_y)$ sommable pour qu'elle soit prouvée pour le cas général.

Or désignons par $f_n(P_x), f'_n(P_y)$ les fonctions égales à $f(P_x), f'(P_y)$ quand le module de celles-ci ne dépasse pas n et égales à zéro ailleurs. On a

$$\int_{E_x} |f_n(P_x)|\,d(P_x) = \int_{E_y} |f'_n(P_y)\varphi(P_y)|\,d(P_y);$$

donc, quand n augmente indéfiniment, l'intégrale du second membre n'augmente pas indéfiniment et $f'(P_y)\varphi(P_y)$ est bien sommable.

Je n'aborderai pas ici la question très intéressante consistant à rechercher comment on peut reconnaître, sur les formules du changement de coordonnées, si l'on est ou non dans le cas considéré ici. Une telle recherche trouvera mieux sa place dans une étude sur la quadrature des surfaces. Je me borne à des remarques immédiates. Dans le cas d'une seule variable, le changement de variable le plus général répondant à la question est évidemment défini par la formule

$$x = a \pm \int_{y_0}^{y} \varphi(y)\, dy,$$

$\varphi(y)$ étant une fonction non négative dont l'intégrale indéfinie est constamment croissante. Cette dernière restriction n'est d'ailleurs pas indispensable pour l'exactitude de la formule d'intégration par substitution.

Pour le cas de plusieurs variables, un raisonnement classique, que l'on trouvera, par exemple, dans le Tome I du *Cours d'Analyse* de M. Jordan ou dans le Tome I des *Leçons sur les théories générales de l'Analyse* de M. Baire, prouve qu'on est dans le cas où la formule d'intégration par substitution est démontrée si le changement de variables univoque dans les deux sens est défini à l'aide de formules

$$x_i = \psi_i(y_1, y_2, \ldots, y_k),$$

dans lesquelles les ψ_i sont continues, ont des dérivées premières continues et telles que le déterminant fonctionnel des ψ_i ne s'annule pas. Ce déterminant fonctionnel est, en valeur absolue, égal à $\varphi(y_1, y_2, \ldots, y_k)$.