Recherches sur la structure des fonctions mesurables.

(Recherches sur les généralisations de la notion de dérivée)[1].

Par

A. Khintchine (Moscou).

Préface.

La définition de la mesure posée par M. Lebesgue a soulevé un grand nombre de problèmes nouveaux dont une grande partie est encore loin d'être résolue. En particulier, la structure des fonctions mesurables d'une variable réelle n'était guère étudiée que sous quelques hypothèses restrictives, et la question semble loin d'être épuisée. Pour ce genre de recher-

[1]) La première Partie de ce Mémoire a été publiée (en russe) dans *Recueil Math. de Moscou* (t. XXX, 1923). Le résultat principal qu'elle renferme est le théorème suivant: *une fonction mesurable $f(x)$ présente presque partout l'un ou l'autre des deux modes de croissance: ou bien elle est asymptotiquement dérivable, ou bien elle a les deux nombres dérivés supérieurs $+\infty$, les deux inférieurs $-\infty$* (chaque de ces quatre circonstances ayant lieu par rapport à un ensemble de densité supérieure 1).

Le théorème équivalent a été démontré déjà en 1916 dans le Mémoire de M. Denjoy *„Sur la totalisation"* (Ann. Ec. Norm. Sup. 3e S. t. XXXIII. 1916. p. 209). La proposition de M. Denjoy n'y est énoncée d'ailleurs que pour les fonctions continues, mais elle présente le même degré de généralité, puisque toute fonction mesurable est, en vertu du théorème bien connu de M. Lusin, identique à une fonction continue sauf aux points d'un ensemble de mesure si petite que l'on veut.

Nous rappelons ici la définition de la notion de dérivée asymptotique introduite par M. Khintchine et jouant un rôle fondamental ausi bien dans la première que dans la deuxième Partie de son Mémoire: *la fonction mesurable $f(x)$ sera dite asymptotiquement dérivable au point x_0 lorsqu'il existe un ensemble mesurable E dont x_0 est un point de densité et tel que le quotient $[f(x) - f(x_0)]:(x - x_0)$ tend vers une limite determinée quand x tend vers x_0 en restant toujours dans E. Cette limite est appelée, par définition, la dérivée asymptotique de $f(x)$ au point x_0.*

Cette definition équivaut évidemment à celle de la notion de dérivée aproximative de M. Denjoy; cette notion fondamentale généralisant celle de dérivée ordinaire, a été introduite, comme on sait, par les deux éminents auteurs indépendemment.

(Remarque de la Rédaction).

ches, le théorème connu de M. Egoroff[1]) était une importante découverte qui a ensuite permis à M. Lusin[2]) d'établir deux propriétés fondamentales des fonctions mesurables; la première de ces propriétés consiste en ce qu'une fonction mesurable devient continue quand on néglige un ensemble de valeurs de l'argument de mesure aussi petite que l'on veut; la seconde établit l'existence d'une infinité de fonctions primitives pour toute fonction mesurable. Ces deux propriétés conduisent ensuite à un grand nombre de conséquences importantes.

Le présent travail constitue la seconde Partie d'un Mémoire consacré à des recherches ultérieures sur la structure métrique des fonctions mesurables; dans la première j'ai cherché à déterminer les caractères de croissance de la fonction mesurable la plus générale; une classe de fonctions, dont la mode de croissance est particulièrement simple, a mérité d'être y étudiée à part. La partie présente est un essai sur les généralisations connues de la notion de dérivée et contient une étude des rapports mutuels ayant lieu entre les différents procédés de dérivation.

Dans ce qui suit, les démonstrations de la mesurabilité d'ensembles ou de fonctions en question, sont souvent laissées au lecteur, même sans indication explicite. On sait que les raisonnements de ce genre sont à la fois très longues et très faciles; on a donc le droit de les supprimer; leur rôle, d'après une remarque très élégante de M. Lusin, est le même que celui des calculations purement algébriques dans l'Analyse classique.

[1]) Egoroff, Sur les suites de fonctions mesurables, *Comptes Rendus*, t. 152, p. 244 (1911).

[2]) V. par ex. Lusin, Sur les propriétés des fonctions mesurables, ibi d. t. 154, p. 1688, (1912).

§ 1. Dérivée symétrique.

1. Dans la théorie des séries trigonométriques et dans quelques questions rattachées à la théorie de la dérivée seconde de Schwarz on parvient à considérer le quotient

$$\frac{f(x+h)-f(x-h)}{h};$$

la limite de cette expression pour $h \to 0$ était étudiée comme une généralisation de la notion de dérivée par M. de la Vallée-Poussin [1]) sous le nom de *dérivée symétrique;* il y a des questions où cette généralisation apparaît tout naturellement; on sait, par exemple, que d'après un théorème de M. Fatou [2]) l'intégrale de Poisson d'une fonction sommable sur une circonférence tend vers la valeur de cette fonction lorsque la variable tend vers un point de la circonférence dans lequel la fonction donnée est la dérivée symétrique de son intégrale indéfinie.

On s'assure inmédiatement que la notion en question nous donne effectivement une généralisation de la notion de dérivée. Nous allons donc examiner le domaine d'appliquabilité de cette méthode de dérivation et les propriétés principales de la dérivée symétrique.

2. Il y a des questions dans la théorie des fonctions où l'on peut entièrement négliger les ensembles de mesure nulle; il nous faudra donc partager notre analyse en deux parties: premièrement, nous étudierons les propriétés de la dérivée symétrique qui ont lieu en tout point de l'intervalle considéré; ensuite, nous allons nous occuper des fonctions qui ont une dérivée symétrique presque partout dans l'intervalle, ou l'ensemble, donné.

En premier lieu, il est évident que les règles élémentaires de dérivation ordinaire subsistent pour la dérivée symétrique, au moins pour les fonctions continues. De plus, la formule

$$\frac{f(x+h)-f(x-h)}{2h}=\frac{1}{2}\left\{\frac{f(x+h)-f(x)}{h}+\frac{f(x-h)-f(x)}{-h}\right\}$$

montre que, pour la dérivée symétrique, le théorème de Rolle peut tomber en défaut.

[1]) *Intégrales de Lebesgue etc. Paris*, 1916.

[2]) *Séries trigonométriques et séries de Taylor, Acta Mathematica*, t. 30.

En effet, on voit que la dérivée symétrique est égale à la demi-somme des dérivées à gauche et à droite en tout point où ces deux dérivées existent séparément. Or, il est évidemment aisé à construire une fonction continue $f(x)$ s'annulant aux points 0 et 1, possédant une dérivée positive entre 0 et $\frac{1}{2}$, une dérivée négative entre $\frac{1}{2}$ et 1, et telle que

$$f'\left(\tfrac{1}{2}+0\right)=0,\ f'\left(\tfrac{1}{2}-0\right)=1.$$

Le même exemple montre que la dérivée symétrique supposée existante en tout point de l'intervalle considéré ne doit pas nécessairement passer par toute valeur intermédiaire entre deux de ses valeurs, comme ceci a lieu pour la dérivée ordinaire.

3. Ces considérations, d'un caractère presque banal, ne sont envisagées que pour caractériser la dérivée symétrique d'une manière plus complète. Voici un problème plus intéressant: chercher l'allure d'une fonction continue dont la dérivée symétrique est positive en tout point d'un intervalle; nous allons montrer qu'une telle fonction doit être croissante dans l'intervalle considéré.

En effet, soit $f(x)$ la fonction donnée, x_1 et x_2 deux points quelconques de l'intervalle considéré, et soit $x_1 < x_2$. Il faut démontrer que l'on

$$f(x_1) < f(x_2).$$

Or, si l'on avait $f(x_1) = f(x_2)$, on pourrait assigner entre x_1 et x_2 un point x_3 tel que $f(x_1) \neq f(x_3)$; on remplacerait l'intervalle (x_1, x_2) par l'intervalle (x_1, x_3) ou par l'intervalle (x_1, x_2), selon que

$$f(x_1) > f(x_3), \text{ ou } f(x_3) > f(x_2) = f(x_1).$$

On peut donc supposer $f(x_1) > f(x_2)$, si le théorème est inexact.

Envisageons la droite

$$y=\frac{1}{2}\left\{f(x_1)+f(x_2)\right\}. \tag{1}$$

$f(x)$ étant supposée continue, il existe entre x_1 et x_2 un point c tel que la courbe $y=f(x)$ passe au-dessous de la droite (1)[1]) dans l'intervalle (c, x_2) et que, au contraire, dans l'intervalle (x_1, c) il y a des points, aussi voisins de c que l'on veut, où la courbe passe au-dessus de cette droite; au point c les deux lignes se confondent. La dérivée symétrique de $f(x)$ ne peut pas être positive au point c, le quotient

$$\frac{f(x+h)-f(x-h)}{2h}$$

devenant négatif pour des valeurs aussi petites que l'on veut de h. La contradiction cherchée étant acquise, on peut affirmer la proposition suivante:

Théorème. *Une fonction continue est nécessairement croissante dans un intervalle où sa dérivée symétrique est positive.*

On démontre de la même manière la proposition analogue pour le cas d'une dérivée symétrique partout négative.

4. Passons au cas d'une dérivée symétrique s'annulant en tout point d'un intervalle (a, b).

Soient c et d deux points de cet intervalle tels que

$$f(c) \neq f(d).$$

Soit pour fixer les idées

$$c < d, f(c) < f(d).$$

Considérons la fonction

$$\varphi(x) = f(x) - \varepsilon(x-c),$$

ε étant un nombre positif quelconque plus petit que $\frac{f(d)-f(c)}{d-c}$.

On trouve en premier lieu

$$\varphi(d) < \varphi(c).$$

D'autre part, la dérivée symétrique de $\varphi(x)$ est visiblement négative en tout point de (a, b); les deux résultats obtenus étant incompatibles en vertu du numéro précédent, on trouve que, pour la dérivée symétrique, subsiste un théorème important du calcul différentiel, à savoir

Théorème. *Deux fonctions continues dont les dérivées symétriques (supposées finies) coincident en tout point d'un intervalle, ne diffèrent que par une constante dans cet intervalle*[2]).

[1]) Tout en pouvant se confondre avec elle en quelques points.

[2]) M. Lusin me communique que ce théorème était connu à M. Sierpiński.

5. En allant maintenant étudier les fonctions qui n'ont de dérivée symétrique que presque partout dans un intervalle, ou dans un ensemble donné, nous commencerons par établir une proposition qui rendra inutiles toutes les recherches ultérieures. En effet, cette proposition nous montrera que, sous la condition de négliger les ensembles de mesure nulle, la notion de dérivée symétrique ne généralise point la notion ordinaire de dérivée; d'une manière plus précise, l'existence d'une dérivée symétrique presque partout dans un intervalle (ou dans un ensemble) entraîne l'existence d'une dérivée ordinaire presque partout dans l'intervalle (resp. ensemble) considéré. Il est donc inutile d'entreprendre une étude spéciale pour la dérivée symétrique.

Pour le voir, il nous suffira d'établir le théorème suivant (qui est d'ailleurs d'un caractère un peu plus général):

Théorème. *$f(x)$ étant une fonction mesurable, elle admet la dérivée unique et finie $f'(x)$ en presque tout point où*

$$\lim_{h \to 0} \sup \frac{f(x+h)-f(x-h)}{2h} < +\infty. \tag{1}$$

D é m o n s t r a t i o n: soit E l'ensemble de valeurs x où l'inégalité (1) ait lieu tandisqu'il n'y ait pas de dérivée $f'(x)$. Supposons, par impossible, que

$$\text{mes}_e E > 0.$$

Désignons, pour tout nombre naturel n, par E_n l'ensemble de points x de E tel que

$$0 < h < \frac{1}{n} \quad \text{entraîne} \quad f(x+h) - f(x-h) < 2hn. \tag{2}$$

On a évidemment en vertu de (2)

$$E = \sum_{n=1}^{\infty} E_n$$

et, parsuite, il existe un nombre n_o, tel que

$$\text{mes}_e E_{n_o} < 0. \tag{3}$$

Posons: $\varphi(x) = f(x) - n_o \cdot x$. En tenant compte de (2), on voit que pour chaque x de E_{n_o} et pour tout $0 < h < \frac{1}{n_o}$, on a:

$$\varphi(x+h)-\varphi(x-h)<0. \tag{4}$$

Soit y un point de densité extérieure de l'ensemble E_{n_0}. Je dis qu'on a

$$\Lambda_d\varphi(y)<0, \quad \Lambda_g\varphi(y)<0.^{1)} \tag{5}$$

En effet, y étant un point de densité extérieure de E_{n_0}, il existe un nombre $k<\frac{1}{n_0}$ tel que, pour tout y',

$$y<y'<y+k\leqslant y+\frac{1}{n_0} \text{ entraîne: } \operatorname{mes}_e E_{n_0}(y,y')>\frac{3}{4}(y'-y), \tag{6}$$

$E_{n_0}(y,y')$ désignant la partie de E_{n_0} contenue dans l'intervalle (y,y').

Je vais démontrer que, pour chaque y', (6) implique encore:

$$\varphi(y)<\varphi(y'). \tag{7}$$

Supposons, dans ce but, que (7) n'ait pas lieu, c.-à.-d. que

$$\varphi(y)\geqslant\varphi(y'). \tag{7bis}$$

Soit A l'ensemble de tous les points x de l'intervalle (y,y') tels que

$$\varphi(x)\leqslant\varphi(y), \tag{8}$$

et B celui d'autres points du même intervalle. On a donc pour tout point x de B:

$$\varphi(x)>\varphi(y)\geqslant\varphi(y'). \tag{9}$$

Désignons ensuite par $\overline{A}$ l'ensemble de tous les points $\overline{x}=\frac{y+x}{2}$, où x parcourt les points de A, et par $\overline{B}$ celui de tous les points $\overline{x}=\frac{x+y'}{2}$, y' appartenant à B.

La fonction $f(x)$, et parsuite la fonction $\varphi(x)$, étant mesurable, les ensembles A et B sont aussi mesurables. La mesure de l'un du moins d'eux doit être $\geqslant\frac{y'-y}{2}$, et, par conséquent, celle de l'un au moins des ensembles $\overline{A}$ et $\overline{B}$ est $\geqslant\frac{y'-y}{4}$, Donc, ces ensembles étant mesurables, en même temps que A et B, l'un au moins d'eux admet, en vertu de (6), des points communs avec E_{n_0}. Supposons, pour fixer l'idée, que c'est $\overline{A}$ qui admet un point commun, soit x_0, avec E_{n_0}.

Le point $x_0+(x_0-y)$ appartient donc à A et, en vertu de (8), on a:

$$\varphi[x_0+(x_0-y)]\geqslant\varphi(y)=\varphi[x_0-(x_0-y)].$$

Or, c'est évidemment contradictoire avec (4), x_0 appartenant à E_{n_0} et (x_0-y) étant inférieur à $(y'-y)\leqslant k<\frac{1}{n_0}$.

Parsuite, l'inégalité (7^{bis}) est aussi contradictoire, et c'est (7) qui a lieu. Il s'ensuit aussitôt qu'on a en y: $\Lambda_d\varphi(y)<0$. On démontre de la même façon la seconde des relations (5).

Ces relations étant valables ainsi en chaque point de densité extérieure de E_{n_0}, donc en presque tout point de E_{n_0}, il s'en suit, en vertu du théorème connu de M^me^ Young[1]), qu'en presque tous les points de $E_{n_0}\subset E$ existe la dérivée finie et unique de $\varphi(x)$, et, par suite, encore celle de $f(x)$.

Nous aboutissons, par conséquent, à une contradiction avec la définition de l'ensemble E, ce qui justifie notre énoncé.

§ 2. La dérivée moyenne de M. Borel.

1. C'est pour illustrer quelques phénomènes physiques que M. Borel[2]) a introduit une généralisation de la notion de dérivée qui paraît présenter un grand intérêt mathématique. Nous dirons que la fonction $f(x)$ a une *dérivée moyenne* (terminologie de M. Borel) an point x lorsque le quotient

$$\frac{1}{h}\int_0^h \frac{f(x+t)-f(x)}{t}\,dt$$

tend vers une limite déterminée, h tendant vers zéro; cette limite est, par définition, la valeur de la dérivée moyenne. On parle d'une dérivée moyenne *à droite*, resp. *à gauche*, lorsque la limite en question existe, h tendant vers zéro par des valeurs

1) $\Lambda_d\varphi$, $\Lambda_g\varphi$ désignent les nombres dérivés supérieurs, resp. *à droite* et *à gauche*, de la fonction φ.

1) *Voir:* G. C. Young. C. R. Note de 13 mars 1916. Aussi: *Fund. Math.* 1924. vol. V. pp. 98—104.

2) Borel, *Modèles arithmétiques etc., Comptes Rendus de l'Ac. de Paris,* t. 154, pp. 1148—1150.

positives, resp. négatives. Enfin, on appelle dérivée moyenne *symétrique* de la fonction $f(x)$ au point x la limite, si elle existe, de l'expression

$$\frac{1}{b}\int_0^b \frac{f(x+t)-f(x-t)}{2t}\,dt,$$

b tendant encore vers zéro. Dans tous les cas l'intégrale est définie comme $\lim\limits_{\varepsilon\to 0}\int\limits_\varepsilon^b$.

L'intérêt que présente cette généralisation est fondé sur ce que la méthode de dérivation envisagée par M. Borel offre une grande analogie avec la méthode de sommation des séries divergentes par les moyennes arithmétiques — méthode dont l'importance capitale pour l'analyse mathématique a été découverte par les travaux récents de MM. Borel, Fejér, Hardy et d'autres savants.

2. En particulier, on est amené d'une façon assez naturelle à l'étude de la dérivée moyenne symétrique (étude qui paraît présenter de difficultés essentielles) par la théorie de la dérivée seconde de Schwarz.

Soit $F(x)$ une fonction absolument continue dans un intervalle déterminé. On a

$$\frac{d}{db}\left\{\frac{F(x+b)+F(x-b)-2F(x)}{b}\right\}=$$
$$=\frac{F'(x+b)-F'(x-b)}{b}-\frac{F(x+b)+F(x-b)-2F(x)}{b^2},$$

en supposant $F'(x+b)$ et $F'(x-b)$ existantes, ce qui aura lieu presque partout; en intégrant les deux membres de cette identité entre les limites p et q $(0<p<q)$, on trouve:

$$\frac{F(x+q)+F(x-q)-2F(x)}{q}-\frac{F(x+p)+F(x-p)-2F(x)}{p}=$$
$$=\int_p^q \frac{F'(x+b)-F'(x-b)}{b}\,db-$$
$$-\int_p^q \frac{F(x+b)+F(x-b)-2F(x)}{b^2}\,db,$$

la première intégrale du second membre étant prise au sens de M. Lebesgue. Or, faisons tendre p vers zéro. En supposant que la fonction $F(x)$ possède une dérivée seconde de Schwarz au point x, le deuxième quotient du premier membre tendra vers zéro, et la fonction à intégrer dans la dernière intégrale sera bornée, de sorte que nous aurons à la limite

$$\frac{F(x+q)+F(x-q)-2F(x)}{q}=\int_0^q \frac{F'(x+b)-F'(x-b)}{b}\,db-$$
$$-\int_0^q \frac{F(x+b)+F(x-b)-2F(x)}{b^2}\,db,$$

où $\int\limits_0^q$ signifie pour la première intégrale $\lim\limits_{p\to 0}\int\limits_p^q$, l'existence de cette limite étant démontrée justement par ce qui précède. Multiplions les deux membres par $\frac{1}{q}$ et faisons tendre q vers zéro. Le premier membre tendra vers la valeur de la dérivée seconde de Schwarz de la fonction $F(x)$ au point x, valeur que nous désignerons par $S(x)$. Il sera de même pour la dernière intégrale du second membre; on s'en assure par une application du premier théorème de la moyenne. Donc,

$$S(x)=\lim_{q\to 0}\frac{1}{q}\int_0^q \frac{F'(x+b)-F'(x-b)}{b}\,db-S(x),$$

d'où

$$\lim_{q\to 0}\frac{1}{q}\int_0^q \frac{F'(x+b)-F'(x-b)}{2b}\,db=S(x),$$

c'est-à-dire: *une fonction sommable a une dérivée moyenne symétrique en tout point où son intégrale indéfinie a une dérivée seconde de Schwarz, et les valeurs des deux dérivées coïncident.*

3. La dérivée moyenne possède la propriété nécessaire de toute généralisation: une simple application du premier théorème de la moyenne montre en effet qu'en un point où la fonction donnée est dérivable au sens ordinaire elle admet aussi une dérivée moyenne, les valeurs des deux dérivées étant égales.

Il est évident que la règle de dérivation d'une somme ou différence subsiste pour la dérivée moyenne. Il n'est pas de même de la dérivation d'un produit. Deux fonctions $f_1(x)$ et $f_2(x)$ peuvent admettre des dérivées moyennes en un point x sans qu'il soit de même pour leur produit. En effet, n étant un entier positif plus grand que *1*, partageons l'intervalle $\left(\frac{1}{n}, \frac{1}{n-1}\right)$ en un nombre pair de parties (inégales) de manière que l'intégrale de la fonction $x^{-\frac{3}{4}}$ ait la même valeur pour toutes ces parties et que cette valeur soit plus petite que $\frac{1}{n^2}$, ce qui est évidemment possible. Soit $f(x)$ la fonction égale à $x^{\frac{1}{4}}$ en valeur absolue et de signes contraires dans deux intervalles voisins de subdivision. En posant $f(0)=0$, $f(x)$ sera définie dans l'intervalle $(0 \leqslant x \leqslant 1)$. Pour h situé entre $\frac{1}{n}$ et $\frac{1}{n-1}$, on a

$$\left|\frac{1}{h}\int_0^h \frac{f(t)-f(0)}{t}\,dt\right| \leqslant n\cdot\frac{1}{n^2}=\frac{1}{n};$$

la dérivée moyenne (à droite) de $f(x)$ est donc nulle pour $x=0$.

Or, on a

$$[f(x)]^2 = x^{\frac{1}{2}},$$

donc

$$\frac{1}{h}\int_0^h \frac{[f(t)]^2-[f(0)]^2}{t}\,dt = \frac{2}{\sqrt{h}},$$

ce qui devient infini pour $h \to 0$.

La fonction $f(x)$ est discontinue; mais en la modifiant légèrement on parvient à une fonction continue conservant le caractère qui nous intéresse.

4. Une fonction continue dont la dérivée moyenne s'annule en tout point d'un intervalle est nécessairement constante dans cet intervalle. En effet, lorsque la dérivée moyenne d'une fonction continue $f(x)$ s'annule au point x, il existe une suite dénombrable de points

$$x_1, x_2, \ldots, x_n, \ldots$$

telle que

$$x_i > x_{i+1}$$

et

$$\lim_{i\to\infty} x_i = x,$$

pour laquelle

$$\lim_{i\to\infty} \frac{f(x_i)-f(x)}{x_i - x} = 0,$$

sinon le quotient

$$\frac{f(x+h)-f(x)}{h}$$

garderait pour h positif et assez petit un signe constant, en restant toujours en valeur absolue plus grand qu'un nombre positif fixe, de sorte que la dérivée moyenne de $f(x)$ ne pourrait pas s'annuler en x.

Mais l'existence, pour tout point x de l'intervalle considéré, d'une telle suite de points entraîne, comme on sait[1]), la constance de $f(x)$ dans l'intervalle en question.

On en conclut que deux fonctions continues dont les dérivées moyennes sont égales en tout point d'un certain intervalle ne diffèrent dans cet intervalle que par une constante.

5. Le théorème de Rolle subsiste pour la dérivée moyenne, au moins pour les fonctions continues. En effet, on voit immédiatement qu'en un point où la fonction donnée a un maximum, sa dérivée moyenne à droite ne peut pas devenir positive, tandis que celle à gauche doit être soit positive, soit nulle; une circonstance analogue a lieu pour le cas d'un minimum.

6. Maintenant nous allons voir que la méthode de M. Borel reste encore une généralisation essentielle de la dérivation ordinaire même sous la condition de négliger les ensembles de mesure nulle (ce qui n'a pas lieu pour la dérivée symétrique, comme nous l'avons vu à la fin du paragraphe précédent). A ce but, il suffit évidemment de construire une fonction continue possédant

[1]) C'est un théorème connu; voir p. ex. Denjoy, *Sur les nombres dérivés des fonctions continues, Journ. des Mathématiques.* 1915.

une dérivée moyenne en tout point d'un ensemble de mesure positive et dépourvue de dérivée ordinaire presque partout dans cet ensemble.

Pour obtenir une telle fonction, envisageons dans l'intervalle $(0, 1)$ un ensemble parfait non dense P de mesure non nulle et tel qu'en désignant par

$$\delta_1, \delta_2, \delta_3, \ldots, \delta_n, \ldots$$

les intervalles contigus de P (et aussi les longuers respectives de ces intervalles) la série

$$\sqrt{\delta_1} + \sqrt{\delta_2} + \sqrt{\delta_3} + \cdots + \sqrt{\delta_n} + \cdots \tag{1}$$

soit convergente.

Il est aisé à construire une fonction continue $f(x)$ s'annulant en tout point de P et dont tous les quatre nombres dérivés soient infinis presque partout dans P, de sorte que $f(x)$ n'ait de dérivée ordinaire en aucun point de P, à l'exception d'un ensemble de mesure nulle au plus. Nous allons nous assurer que $f(x)$ admet presque partout dans P une dérivée moyenne nulle.

A ce but, soit u_n l'intervalle de longueur

$$u_n = \delta_n + 2\sqrt{\delta_n}$$

dont le centre coïncide avec celui de l'intervalle δ_n. La série

$$u_1 + u_2 + u_3 + \cdots + u_n + \cdots$$

étant convergente en vertu de la convergence de la série (1), on sait que l'ensemble-limite complet des intervalles u_n $(n = 1, 2, \ldots)$ est de mesure nulle. Soit x un point de P n'appartenant pas à cet ensemble.

On a

$$\int_x^{x+b} \left| \frac{f(\alpha)}{\alpha - x} \right| d\alpha \leqslant \sum{}' \int_{\delta_n} \left| \frac{f(\alpha)}{\alpha - x} \right| d\alpha,$$

où la somme du second membre s'étend à tous les intervalles δ_n qui appartiennent, soit totalement, soit en partie, à l'intervalle $(x, x+b)$.

Or, pour $|b|$ suffisamment petit, x reste en dehors des intervalles u_n correspondants à ces δ_n; on a donc, α faisant partie de δ_n,

$$|\alpha - x| > \sqrt{\delta_n};$$

d'autre part, $f(x)$ étant continue, on a, pour $|b|$ suffisamment petit,

$$|f(\alpha)| < 1$$

pour

$$|\alpha - x| < |b|,$$

donc

$$\int_{\delta_n} \left| \frac{f(\alpha)}{\alpha - x} \right| d\alpha < \frac{\delta_n}{\sqrt{\delta_n}} = \sqrt{\delta_n},$$

d'où

$$\left| \int_x^{x+b} \left| \frac{f(\alpha)}{\alpha - x} \right| d\alpha \right| < \sum{}' \sqrt{\delta_n},$$

la somme du second membre étant encore étendue aux mêmes intervalles que dans le cas précédent.

Soit à présent $F(x)$ la fonction égale à zéro dans P et à $\frac{1}{\sqrt{\delta_n}}$ dans δ_n. L'intégrale de cette fonction prise entre les extrémités de δ_n est égale à $\sqrt{\delta_n}$, donc la convergence de la série (1) nous montre que $F(x)$ est sommable dans l'intervalle $(0, 1)$. Soit $\Phi(x)$ son intégrale indéfinie; lorsque $x + b$ fait partie de P, on a

$$|\Phi(x+b) - \Phi(x)| = \sum{}' \sqrt{\delta_n}.$$

On a donc aussi

$$\left| \int_x^{x+b} \left| \frac{f(x)}{\alpha - x} \right| d\alpha \right| < \left| \Phi(x+b') - \Phi(x) \right|$$

où $b' = b$ dans le cas où $x + b$ est un point de P; dans le cas contraire, $x + b'$ signifie l'extrémité la plus éloignée de x de l'intervalle contigu dont $x + b$ fait partie.

Par suite,

$$\left|\frac{1}{b}\int_{x}^{x+b}\frac{f(\alpha)}{\alpha-x}\,d\alpha\right|<\left|\frac{\Phi(x+b')-\Phi(x)}{b}\right|=\frac{b'}{b}\cdot\frac{\Phi(x+b')-\Phi(x)}{b'}.$$

Supposons que x soit un point principal de P et que la fonction sommable $F(x)$ soit en x la dérivée de son intégrale définie (ces deux conditions sont remplies simultanément presque partout dans P). Alors on a

$$\lim_{b\to 0}\frac{b'}{b}=1,\ \lim_{b\to 0}\frac{\Phi(x+b')-\Phi(x)}{b'}=F(x)=0,$$

et par suite

$$\lim_{b\to 0}\frac{1}{b}\int_{x}^{x+b}\frac{f(\alpha)}{\alpha-x}\,d\alpha=0,$$

ce qui démontre l'affirmation [1]).

Enfin, la méthode connue de condensation des singularités nous donnera aisément, en partant de la fonction construite $f(x)$, une fonction continue dont la dérivée moyenne s'annule presque partout dans l'intervalle $(0,1)$ et qui est dépourvue de dérivée ordinaire presque partout dans le même intervalle.

7. On voit donc que la dérivée moyenne offre une généralisation essentielle de la méthode ordinaire de dérivation même pour les théories où l'on néglige entièrement les ensembles de mesure nulle. Nous allons maintenant comparer de la même manière la méthode de M. Borel avec la dérivation asymptotique dont la définition a été donnée dans la première partie de notre mémoire. L'intérêt et la difficulté de ce problème découlent surtout de ce fait remarquable que la dérivée moyenne est définie par une expression analytique tandis que la définition de la dérivée asymptotique est d'un caractère purement fonctionnel et n'implique aucune formule.

Montrons en premier lieu que, sous la condition de négliger les ensembles de mesure nulle, le méthode de M. Borel ne peut

[1]) Il faut remarquer que l'idée principale de cette démonstration a été souvent employée par M. Borel, surtout dans quelques questions de la théorie des fonctions d'une variable complexe.

s'appliquer qu'aux mêmes fonctions continues qui tombent dans le champ d'application de la méthode asymptotique. D'une façon plus précise, nous allons démontrer la proposition suivante:

Théorème. *Une fonction continue admettant une dérivée moyenne presque partout dans un ensemble mesurable E de mesure non nulle, est asymprotiquement dérivable presque partout dans E.*

Soit, en effet, M l'ensemble des points de E où la fonction en question $f(x)$ n'a pas de dérivée asymptotique, et soit par impossible mes $M>0$. Alors un théorème de la première partie de ce travail (n^0 8, § 2) nous assure que $f(x)$ remplit presque partout dans M la condition désignée par $\overline{A}$. Si x_0 signifie un point de M où cette condition est remplie, nous allons voir que $f(x)$ est dépourvue de dérivée moyenne en x_0, ce qui démontrera évidemment l'affirmation de notre théorème.

D'après l'énoncé de la condition $\overline{A}$, quelque petit que soit $\varepsilon>0$ et quelque grand que soit $n>0$, il existe à droite de x_0 un intervalle (a,b) $(a<b)$, tel que

$\overline{1})$ $a-x_0<b-x_0<\varepsilon$,

$\overline{2})$ $b-a>(1-\varepsilon)(b-x_0)$

$\overline{3})$ $f(y)-f(x_0)>n(y-x_0)$ pour $a<y<b$.

Si $f(x)$ avait une dérivée moyenne $f'_B(x)$ en x_0, on aurait évidemment:

$$\frac{1}{a-x_0}\int_{x_0}^{a}\frac{f(\alpha)-f(x_0)}{\alpha-x_0}\,d\alpha=f'_B(x_0)+\eta(a),$$

avec $\eta(a)\to 0$ pour $a\to x_0$, donc pour $\varepsilon\to 0$.

On trouverait donc

$$\frac{1}{b-x_0}\int_{x_0}^{b}\frac{f(\alpha)-f(x_0)}{\alpha-x_0}\,d\alpha=\frac{1}{b-x_0}\left\{\int_{x_0}^{a}+\int_{a}^{b}\frac{f(\alpha)-f(x_0)}{\alpha-x_0}\,d\alpha\right\}=$$

$$=\frac{1}{b-x_0}(a-x_0)\{f'_B(x_0)+\eta(a)\}+\frac{1}{b-x_0}(b-a)(n+c),$$

avec $c<0$.

Pour $\varepsilon \to 0$, le premier terme du second membre est infiniment petit, tandis que le second devient arbitrairement grand avec n (en vertu de $\overline{2}$)). Or, en vertu de $\overline{1}$), $b - x_0$ est infiniment petit avec ε. On voit donc que l'existence de $f'_B(x_0)$ est impossible.

8. Le raisonnement du n^0 précédent ne démontre que l'existence de la dérivée asymptotique sous les conditions du théorème. Pour le problème de l'égalité des deux dérivées il ne donne rien, grâce à sa structure apagogique. C'est un problème plus délicat qui doit être traité à part. Nous commencerons par démontrer la proposition suivante:

Lemme. *Si la fonction mesurable $f(x)$ a une dérivée asymptotique nulle en tout point d'un ensemble mesurable E de mesure positive, il existe un ensemble parfait P de mesure aussi voisine de celle de E que l'on veut, contenu dans E et tel que, x_0 étant un point quelconque de P, le quotient*

$$\frac{f(x) - f(x_0)}{x - x_0}$$

tend vers zéro lorsque x tend vers x_0 en restant toujours dans P

En effet, soit x un point de E, et $(x - \delta, x + \delta)$ un intervalle ayant x pour centre. Soit ε le plus petit des nombres qui satisfont à la condition suivante: quel que soit $h > 0$, pourvu que $|h| < \delta$, l'inégalité

$$|f(\alpha) - f(x)| \leqslant \varepsilon |\alpha - x|$$

est remplie pour des points α de l'intervalle $(x, x+h)$ qui forment dans cet intervalle un ensemble dont la densité moyenne n'est pas inférieure à $\frac{3}{4}$ (il est évident qu'un tel nombre ε doit nécessairement exister). ε dépend de x et de δ et, la dérivée asymptotique de $f(x)$ s'annulant par hypothèse en x, ε tend vers zéro avec δ, x restant fixe.

D'après le théorème fondamental de M. Egoroff, on peut affirmer l'existence d'un ensemble parfait P contenu dans E, de mesure aussi voisine de mes E que l'on veut et tel que ε tend vers zéro avec δ uniformément par rapport aux différents points x de P. Or, nous allons voir que cet ensemble P satisfait aux conditions énoncées dans notre lemme.

En effet, soit x_0 un point quelconque de P, et $\sigma > 0$ quelconque. Soit $\varepsilon < \sigma$ pour $\delta < k$, quel que soit le point x de P. Enfin, soit x un point de P tel que

$$|x - x_0| < k;$$

je dis qu'on a

$$|f(x) - f(x_0)| < \sigma |x - x_0|,$$

ce qui démontre évidemment le lemme. Pour s'en assurer, soit E_1 l'ensemble des point y de l'intervalle (x_0, x) pour lesquels

$$|f(y) - f(x)| < \sigma |y - x|;$$

soit de même E_2 l'ensemble des points z du même intervalle, pour lesquels

$$|f(z) - f(x_0)| < \sigma |z - x_0|.$$

D'après la définition de ε et de k, la densité de chacun des ensembles E_1 et E_2 dans l'intervalle (x_0, x) est au moins égale à $\frac{3}{4}$; ces deux ensembles ont donc des points communs. Soit y_0 l'un quelconque de ces points, on trouve

$$|f(x) - f(x_0)| \leqq |f(x) - f(y_0)| + |f(y_0) - f(x_0)| <$$
$$< \sigma \{ |x - y_0| + |y_0 - x_0| \} = \sigma |x - x_0|,$$

ce qu'il fallait démontrer.

9. Supposons maintenant que la dérivée asymptotique de la fonction continue $f(x)$ s'annule en tout point d'un ensemble mesurable E de mesure non nulle et que la dérivée moyenne $f'_B(x)$ de $f(x)$ existe presque partout dans E. En vertu du lemme que nous venons de démontrer, il existe un ensemble parfait $\overline{P}$ contenu dans E, de mesure aussi voisine de mes E que l'on veut et tel que, x_0 étant un point quelconque de $\overline{P}$, le quotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

tend vers zéro, lorsque x tend vers x_0 en restant toujours dans $\overline{P}$.

Par hypothèse, la dérivée moyenne de $f(x)$ existe presque partout dans $\overline{P}$; en vertu du théorème fondamental de M. Egoroff, il existe un ensemble parfait P, contenu dans $\overline{P}$, de mesure aussi voisine de mes $\overline{P}$ que l'on veut et tel que 1^0 $f'_B(x)$ est bornée dans P et 2^0 le quotient

$$\frac{1}{b}\int_0^b \frac{f(x+t)-f(x)}{t}\,dt$$

tend vers $f'_B(x)$ *uniformément* dans P pour $b \to 0$. Soit

$$|f'_B(x)| < M > 0$$

dans P; il existe alors un $H > 0$ tel que pour $|b| < H$

$$\left|\frac{1}{b}\int_0^b \frac{f(x+t)-f(x)}{t}\,dt\right| < 2M,$$

x étant un point quelconque de P.

Soit A un point principal de P et $\pi(b)$ la partie de P enfermée dans l'intervalle $(A, A+b)$.

Désignons par $\delta_1, \delta_2, \ldots, \delta_n, \ldots$ les intervalles contigus de $\pi(b)$ par rapport à l'intervalle $(A, A+b)$.

Nous aurons

$$\int_A^{A+b} \frac{f(\alpha)-f(A)}{\alpha-A}\,d\alpha = \int_{\pi(b)} \frac{f(\alpha)-f(A)}{\alpha-A}\,d\alpha + \sum_{i=1}^{\infty}\int_{\delta_i} \frac{f(\alpha)-f(A)}{\alpha-A}\,d\alpha, \quad (2)$$

sous la condition que l'intégrale $\int_{\pi(b)}$ existe étant définie comme $\lim\limits_{\varepsilon\to 0}\int_{[(A+\varepsilon, A+b), \pi(b)]}$, et que la série du second membre converge absolument.

Or, l'existence de l'intégrale en question est une conséquence évidente du fait que le rapport

$$\frac{f(\alpha)-f(A)}{\alpha-A}$$

tend vers zéro avec $|\alpha - A|$, α restant dans $\pi(b)$.

D'une manière plus précise, une application simple du premier théorème de la moyenne nous donne

$$\frac{1}{b}\int_{\pi(b)} \frac{f(\alpha)-f(A)}{\alpha-A}\,d\alpha \to 0 \text{ pour } b \to 0. \qquad (3)$$

Soient a_n et b_n les extrémités de l'intervalle δ_n et soit, pour fixer les idées,

$$|A - a_n| < |A - b_n|.$$

On a

$$\int_{\delta_n} \frac{f(\alpha)-f(A)}{\alpha-A}\,d\alpha = \int_{\delta_n} \frac{f(\alpha)-f(a_n)}{\alpha-A}\,d\alpha + \int_{\delta_n} \frac{f(a_n)-f(A)}{\alpha-A}\,d\alpha.$$

Or,

$$\left|\int_{\delta_n} \frac{f(a_n)-f(A)}{\alpha-A}\,d\alpha\right| = \left|f(a_n)-f(A)\right| \cdot \left|\int_{\delta_n} \frac{d\alpha}{\alpha-A}\right| \leqq$$

$$\leqq \left|\frac{f(a_n)-f(A)}{a_n-A}\right| \cdot \delta_n < \delta_n$$

pour $|b|$ suffisamment petit, puisque le premier facteur tend vers zéro avec b.

D'autre part,

$$\int_{\delta_n} \frac{f(\alpha)-f(a_n)}{\alpha-A}\,d\alpha = \int_{\delta_n} \frac{f(\alpha)-f(a_n)}{\alpha-a_n} \cdot \frac{\alpha-a_n}{\alpha-A}\,d\alpha.$$

La fonction $\dfrac{\alpha-a_n}{\alpha-A}$ étant monotone dans δ_n, on peut appliquer le second théorème de la moyenne, ce qui donne

$$\left|\int_{a_n+\varepsilon}^{b_n} \frac{f(\alpha)-f(a_n)}{\alpha-a_n} \cdot \frac{\alpha-a_n}{\alpha-A}\,d\alpha\right| = \frac{b_n-a_n}{b_n-A}\left|\int_{\xi}^{b_n} \frac{f(\alpha)-f(a_n)}{\alpha-a_n}\,d\alpha\right| \leqq$$

$$\leqq \frac{\delta_n}{|a_n-A|}\left\{\left|\int_{a_n}^{b_n} \frac{f(\alpha)-f(a_n)}{\alpha-a_n}\,d\alpha\right| + \left|\int_{a_n}^{\xi} \frac{f(\alpha)-f(a_n)}{\alpha-a_n}\right|\right\} <$$

$$< 4M\frac{\delta_n^2}{|a_n-A|} = \left\{\frac{4M\delta_n}{|a_n-A|}\right\}\delta_n,$$

où

$$|a_n - A| < |a_n + \varepsilon - A| < |\xi - A| < |b_n - A|,$$

et ε est un nombre arbitrairement petit.

Or, A étant par hypothèse un point principal de $\pi(h)$, on a, pour $|h|$ assez petit,

$$\frac{4M\delta_n}{|\alpha_n - A|} < 1,$$

donc

$$\left|\int_{\delta_n} \frac{f(\alpha)-f(\alpha_n)}{\alpha - A}\right| < \delta_n,$$

et par suite

$$\left|\int_{\delta_n} \frac{f(\alpha)-f(A)}{\alpha - A}\right| < 2\delta_n,$$

d'où découle la convergence absolue de la série dans le second membre de (2); en même temps nous avons

$$\left|\sum_{i=1}^{\infty}\int_{\delta_i} \frac{f(\alpha)-f(A)}{\alpha - A}\,d\alpha\right| < 2\sum_{i=1}^{\infty}\delta_i,$$

donc

$$\frac{1}{h}\sum_{i=1}^{\infty}\int_{\delta_i} \frac{f(\alpha)-f(A)}{\alpha - A}\,d\alpha \to 0 \text{ pour } h\to 0.$$

En vertu de (2) et (3) on voit que

$$\frac{1}{h}\int_{A}^{A+h} \frac{f(\alpha)-f(A)}{\alpha - A}\,d\alpha \to 0 \text{ pour } h\to 0;$$

la dérivée moyenne de $f(x)$ est donc nulle en tout point principal de P, donc presque partout dans E, puisque $\operatorname{mes} E - \operatorname{mes} P$ est arbitrairement petit.

10. Finalement, supposons qu'on sait que la fonction continue $f(x)$ admet une dérivée moyenne presque partout dans un ensemble mesurable M de mesure non nulle. D'après le théorème du n^0 7, la dérivée asymptotique de $f(x)$ existe presque partout dans M; cette dérivée asymptotique est une fonction mesurable que nous désignerons par $f^{[1]}(x)$; en posant par définition $f^{[1]}(x)=0$ en tout point où $f(x)$ n'a pas de dérivée asymptotique, $f^{[1]}(x)$ sera une fonction mesurable définie dans un intervalle total.

D'après un théorème connu de M. Lusin, il existe une fonction continue $F(x)$ telle que

$$F'(x) = f^{[1]}(x)$$

presque partout dans l'intervalle considéré.

Il est évident que la fonction $F(x)-f(x)$ a une dérivée asymptotique nulle presque partout dans M. De même, la dérivée moyenne de cette fonction existe presque partout dans M. Or, la dérivée moyenne de $F(x)$ est égale à $f^{[1]}(x)$ presque partout dans M. D'autre part, la dérivée moyenne de $F(x)-f(x)$ s'annule en presque tout point de M, d'après le résultat du n^0 9; donc, celle de $f(x)$ est égale à $f^{[1]}(x)$ presque partout dans M; nous sommes ainsi amenés à la proposition suivante, plus complète que celle du n^0 7:

Théorème. *Si la fonction continue $f(x)$ admet une dérivée moyenne en tout point d'un ensemble mesurable, elle admet aussi une dérivée asymptotique presque partout dans le même ensemble, et les valeurs de ces deux dérivées sont égales presque partout dans l'ensemble donné.*

11. On voit donc que sous la condition de négliger les ensembles de mesure nulle la méthode asymptotique a un champ d'application qui n'est pas moins étendu que celui de la méthode de M. Borel. En vérité, il est plus vaste, et c'est ce que nous allons montrer à présent; à ce but, il nous faudra définir une fonction continue asymptotiquement dérivable dans un ensemble de mesure positive et dépourvue de dérivée moyenne presque partout dans le même ensemble.

En premier lieu, nous allons construire l'ensemble en question. Soit en général n_k le plus petit entier positif satisfaisant à l'inégalité

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n_k} > 2^{k+1}.$$

On partagera l'intervalle $(0, 1)$ en $2n_1$ parties égales et on désignera les points de subdivision, en comptant de gauche à droite, par

$$0, c_1, c_2, \ldots c_{2n_1-1}, 1.$$

Soit δ_i l'intervalle de centre c_i et de longueur $\dfrac{1}{8n_1}$, on aura

(en désignant par δ_i la longueur de l'intervalle de même dénomination)

$$\sum_{i=1}^{2n_1-1}\delta_i<\frac{1}{4}.$$

Les intervalles $\delta_i\,(i=1,2,\ldots,2n_1-1)$ seront dits les *intervalles du premier groupe.*

Ensuite, on partagera en $2n_2$ parties égales chaque intervalle qui sépare deux δ_i voisins, l'intervalle entre 0 et δ_1 et l'intervalle entre δ_{2n_1-1} et 1, et on désignera les points de subdivision, en comptant toujours de gauche à droite, par

$$c_{2n_1},\ c_{2n_1+1},\ c_{2n_1+2},\ldots$$

On construira autour de chacun de ces points un intervalle de longueur $\frac{1}{32n_1n_2}$ ayant pour centre le point donné, et on appellera ces intervalles les *intervalles du second groupe,* en les désignant respectivement par

$$\delta_{2n_1},\ \delta_{2n_1+1},\ \delta_{2n_1+2},\ldots$$

La somme des longueurs des intervalles du second groupe est évidemment plus petite que

$$\frac{1}{32n_1n_2}(2n_1\,.\,2n_2)=\frac{1}{8}.$$

Enfin, on continuera ce procédé infiniment.

Pour la longueur des intervalles du k^e groupe on choisira le nombre

$$\frac{1}{2^{2k+1}\,n_1n_2\ldots n_k},$$

de sorte que la somme de leurs longueurs sera plus petite que $\frac{1}{2^{k+1}}$.

La somme totale des longueurs des intervalles de tous les groupes sera par suite moindre que

$$\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots=\frac{1}{2}.$$

C'est l'ensemble parfait complémentaire par rapport à ce système d'intervalles qui nous servira pour la définition de la fonction cherchée; nous le désignerons par P et nous aurons évidemment

$$\text{mes}\,P>\frac{1}{2}.$$

Envisageons la série

$$\sum_{i=1}^{\infty}\frac{\delta_i}{|x-c_i|};\qquad(4)$$

nous allons voir que cette série diverge pour tout point x de l'ensemble P.

En effet, quand on commence la $k^{\text{ième}}$ opération du procédé indiqué plus haut, x est enfermé dans l'intervalle qui sépare deux intervalles voisins du $k-1^{\text{ième}}$ groupe. Pour fixer les idées, on supposera que c'est l'extrémité droite de cet intervalle qui est la plus éloignée de x. Par la $k^{\text{ième}}$ opération on partage l'intervalle qui contient x en $2n_k$ parties égales dont au moins n_k tombent par suite à droite de x. Ce sont des intervalles du $k^{\text{ième}}$ groupe que nous désignerons, en comptant de gauche à droite, par

$$\delta_{m+1},\ \delta_{m+2},\ldots,\ \delta_{m+p}\ (p\geqq n_k).$$

Toutes les différences

$$c_{m+1}-x,\quad c_{m+i}-c_{m+i-1}\quad(i=1,2,\ldots,p)$$

sont plus petites que la quantité

$$\frac{1}{2^k\,n_1n_2\ldots n_k};$$

d'autre part, on a

$$\delta_{m+i}=\frac{1}{2^{2k+1}\,n_1n_2\ldots n_k}\quad(i=1,2,\ldots p);$$

donc,

$$\sum_{i=1}^{p}\frac{\delta_{m+i}}{|x-c_{m+i}|}>\frac{1}{2^{2k+1}\,n_1n_2\ldots n_k}\left\{2^kn_1n_2\ldots n_k+\frac{2^kn_1n_2\ldots n_k}{2}+\right.$$

$$\left.+\frac{2^kn_1n_2\ldots n_k}{3}+\cdots+\frac{2^kn_1n_2\ldots n_k}{p}\right\}=$$

$$= \frac{1}{2^{k+1}}\left\{1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{p}\right\} \geqq$$

$$\geqq \frac{1}{2^{k+1}}\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n_k}\right) > 1,$$

d'après la définition de n_k.

Or, m est aussi grand que l'on veut avec k. Donc, la série (4) diverge au point x, ce qu'il fallait démontrer[1]).

12. A présent, je dis *qu'il existe une suite de nombres positifs*

$$\sigma_1, \sigma_2, \ldots, \sigma_k, \ldots$$

telle que $\sigma_k \to 0$ *pour* $k \to 0$ *et que la série*

$$\sum_{l=1}^{\infty} \frac{\sigma_l \delta_l}{|x - c_l|} \tag{5}$$

diverge presque partout dans P.

En effet, soit en général p_k le plus petit entier positif tel que

$$\text{mes}\, P - \text{mes}\, E_k < \frac{1}{2^k},$$

E_k étant l'ensemble des points de P pour lesquels on a

$$\sum_{l=p_1+p_2+\cdots+p_{k-1}+1}^{l=p_1+p_2+\cdots+p_k} \frac{\delta_l}{|x-c_l|} > 2^{k-1}.$$

On posera

$$\sigma_1 = \sigma_2 = \cdots = \sigma_{p_1} = 1,$$

$$\sigma_{p_1+1} = \sigma_{p_1+2} = \cdots = \sigma_{p_1+p_2} = \frac{1}{2},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\sigma_{p_1+p_2+\cdots+p_{k-1}+1} =$$

$$= \sigma_{p_1+p_2+\cdots+p_{k-1}+2} = \cdots = \sigma_{p_1+p_2+\cdots+p_{k-1}+p_k} = \frac{1}{2^{k-1}},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

[1]) Je dois remarquer que des méthodes de construction de séries divergentes analogues à celle du texte étaient souvent employées par M. Lusin.

Pour un point x de P appartenant à E_k on aura par suite

$$\sum_{l=p_1+p_2+\cdots+p_{k-1}+1}^{p_1+p_2+\cdots+p_k} \frac{\sigma_l \delta_l}{|x-c_l|} > 1;$$

la série (5) divergera donc pour tout point de P faisant partie d'une infinité d'ensembles E_k $(k = 1, 2, 3, \ldots)$. Or, les mesures de ces ensembles sont choisies de sorte que cette condition soit remplie presque partout dans P; ce qui démontre l'affirmation.

13. Nous allons maintenant définir la fonction continue $f(\alpha)$ par les conditions suivantes:

1) $f(x) = 0$ dans P;
2) $f(x) = \sigma_n$ pour $x = c_n$;
3) $f(x)$ est linéaire dans chacune des deux moitiés de l'intervalle δ_n.

Soit x un point intérieur[1]) de P pour lequel la série (5) devient divergente. Cette série est la somme de deux séries se rapportant respectivement aux intervalles contigus situés à droite et à gauche de x, dont une au moins est divergente; pour fixer les idées supposons que ce soit celle qui est relative aux intervalles contigus situés à droite de x. Nous allons démontrer que $f(x)$ ne peut pas admettre en x une dérivée moyenne à droite; d'une manière plus précise, nous prouverons que l'intégrale

$$\int_0^b \frac{f(x+\alpha)}{\alpha}\, d\alpha$$

n'a pas de sens pour $b > 0$.

En effet, soient a_n et b_n $(a_n < b_n)$ les extrémités de δ_n; on aura

$$\int_0^b \frac{f(x+\alpha)}{\alpha}\, d\alpha = \sum_n{}' \int_{a_n - x}^{b_n - x} \frac{f(x+\alpha)}{\alpha}\, d\alpha, \tag{6}$$

où la somme du second membre s'étend aux intervalles contigus (ou parties de tels intervalles) compris dans l'intervalle $(x, x+b)$, les deux membres pouvant être à la fois infinis.

[1]) C'est-à-dire un point qui n'est pas une extrémité d'un intervalle contigu.

Or,

$$\int_{a_n-x}^{b_n-x} \frac{f(x+\alpha)}{\alpha}\,d\alpha > \frac{1}{b_n-x}\int_{a_n-x}^{b_n-x} f(x+\alpha)\,d\alpha = \frac{\sigma_n\,\delta_n}{2\,(b_n-x)} =$$

$$= \frac{\sigma_n\,\delta_n}{2\,(c_n-x)} - \frac{\sigma_n\,\delta_n}{2\,(b_n-x)}\cdot\frac{\delta_n}{2\,(c_n-x)}.$$

Donc, si la série du second membre de (6) était convergente, il en serait de même de la série

$$\sum \frac{\sigma_n\,\delta_n}{2\,(b_n-x)},$$

donc aussi, en vertu de l'inégalité évidente

$$\frac{\delta_n}{2\,(c_n-x)} < 1,$$

de la série

$$\sum \frac{\delta_n\,\sigma_n}{2\,(c_n-x)},$$

ce qui est contraire à l'hypothèse.

Donc, le second membre de (6) est infini. Il en est par suite de même du premier membre. On s'assure ainsi que $f(x)$ est dépourvue de dérivée moyenne en x. Or, presque tout point de P remplit les hypothèses faites sur x. D'autre part, puisque $f(x)$ s'annule et tout point de P, sa dérivée asymptotique est nulle en tout point principal de P, donc presque partout dans P. On est ainsi amené à établir la proposition suivante: *il existe une fonction continue asymptotiquement dérivable en tout point d'un ensemble mesurable de mesure non nulle et dépourvue de dérivée moyenne presque partout dans le même ensemble.*

L'étendue de la méthode asymptotique est donc plus grande que celle de la méthode de M. Borel. Nous allons maintenant étudier quelques propriétés importantes de la dérivation asymptotique.

§ 3. La dérivation asymptotique.

1. La notion de dérivée asymptotique nous donne, comme on l'a vu, une généralisation de la notion de dérivée dont le champ d'application est plus vaste que ceux des autres généralisations connues. D'autre part, la définition de cette notion paraît aussi simple et naturelle que possible; son introduction devient, pour des raisons de caractère logique, inévitable dès qu'on a aperçu le rôle fondamental que joue la notion de *densité* dans la théorie métrique des ensembles et des fonctions. L'utilité pratique de la dérivation asymptotique correspond entièrement à ces arguments théoriques; il y a des théories dans l'Analyse où cette notion peut rendre de grandes services; telle est, par exemple, la théorie des séries trigonométriques de Riemann, ainsi que l'a montré M. Lusin [1]).

Mais c'est surtout la théorie de l'intégration que M. Denjoy a développée [2]) qui a rendu inévitable l'introduction de la notion de dérivée asymptotique. En effet, considérons une fonction continue $F(x)$, égale à zéro en tout point d'un ensemble parfait non dense P de mesure non nulle, possédant une dérivée ordinaire en tout point appartenant à un intervalle contigu de P et dont tous les quatre nombres dérivés sont infinis presque partout dans P (nous avons déjà vu qu'il est aisé à construire une telle fonction). Soit $f(x)$ la fonction égale à zéro dans P et à $F'(x)$ en dehors de P.

On sait que $F(x)$ est l'intégrale indéfinie de $f(x)$ prise au sens de M. Denjoy; et tout ce que nous savons, nous fait croire que c'est justement cette fonction $F(x)$ qui doit être considérée comme la fonction primitive la plus simple et naturelle de $f(x)$.

Or, ce n'est pas une fonction primitive au sens exact du mot, car elle n'a pas de dérivée ordinaire dans un ensemble de mesure positive; d'autre part, la dérivée asymptotique de $F(x)$ est évidemment égale à $f(x)$ presque partout. On voit donc qu'on a tous les raisons de choisir pour intégrale indéfinie de la fonction $f(x)$ (fonction qui a, comme toute fonction mesurable, une infinité

[1]) loc. cit. p. 230.

[2]) Denjoy, *Sur la dérivation et son calcul inverse, Ann. de l'éc. Norm. Sup., 1916;* voir aussi ma note *„Sur une extension de l'intégrale de M. Denjoy", Comptes Rendus* t. 162, p. 287, 1916.

de fonctions primitives au sens exact) une fonction qui n'est qu'asymptotiquement dérivables et, par suite, n'entre point dans la famille des primitives exactes de la fonction à intégrer.

2. Les problèmes qui se rattachent à la dérivation asymptotique, se séparent d'une façon naturelle en deux catégories; on étudie d'abord les fonctions qui sont asymptotiquement dérivables *en tout point* d'un certain intervalle; ensuite, on passe aux circonstances qui ont lieu lorsque la dérivée asymptotique n'existe que *presque partout* dans un intervalle ou un ensemble.

Nous avons vu que la méthode asymptotique est la plus puissante des généralisations considérées. Il paraît donc intéressant de savoir combien cette méthode s'éloigne de la méthode ordinaire. Pour être plus précis, on peut poser le problème de deux manières différentes.

En premier lieu, on suppose que la fonction mesurable donnée soit asymptotiquement dérivable en tout point d'un certain intervalle et on demande si l'on peut affirmer qu'elle admet une dérivée ordinaire pour des points formant un ensemble de caractère détermine.

La question posée de cette façon a été résolue par M. Denjoy[1]) qui a démontré que, sous les conditions que nous venons d'énoncer l'intervalle donné contient un ensemble partout dense d'intervalles, dans chacun desquels la fonction en question admet presque partout une dérivée ordinaire.

On peut poser le problème d'une manière différente: on sait que la fonction mesurable $f(x)$ est en tout point d'un certain intervalle la dérivée asymptotique d'une autre fonction mesurable, soit $F(x)$; on demande la condition à ajouter, nécessaire et suffisante pour qu'on puisse affirmer que $f(x)$ soit la dérivée ordinaire de $F(x)$ en tout point de l'intervalle donné.

C'est ce problème que nous allons maintenant traiter.

3. Soit d'abord $F(x)$ une fonction mesurable, non décroissante dans l'intervalle (a, b); supposons que la dérivée asymptotique de $F(x)$ en un certain point x de cet intervalle ait une valeur positive K; nous allons démontrer que $F(x)$ admet au point x une dérivée ordinaire égale à K.

[1]) loc. cit.

En effet, dans le cas contraire admettons pour fixer les idées que $F(x)$ n'ait pas de *dérivée à droite* égale à k au point x. Il existe alors un nombre positif ε inférieur à $\frac{k}{2}$ et satisfaisant à la condition suivante: quelque petit que soit $\delta > 0$, il existe un nombre h remplissant les conditions

$$0 < h < \frac{\delta}{2}, \quad \left| \frac{F(x+h) - F(x)}{h} - k \right| > 2\varepsilon. \tag{7}$$

En supposant ces conditions remplies, admettons en premier lieu que

$$\frac{F(x+h) - F(x)}{h} < k.$$

Alors je dis que pour un point ξ quelconque de l'intervalle $[x + (1 - \frac{k-\varepsilon}{\varepsilon}) h, x+h]$ on aura

$$\left| \frac{F(\xi) - F(x)}{\xi - x} - k \right| < \varepsilon.$$

En effet, soient A et B les points représentant dans un système rectangulaire respectivement les coordonnées $[x, F(x)]$ et $[x+h, F(x+h)]$. Le coëfficient angulaire de la droite AB est par hypothèse plus petit que $k - 2\varepsilon$; on menera les droites AC et AD de coëfficients angulaires respectivement égaux à k et $k - \varepsilon$; enfin, menons par le point B une parallèle à l'axe des x, et soit E le point où cette droite rencontre la droite AD. On s'assure par un calcul élémentaire que la longueur du segment BE est au moins égale à

$$h \frac{\varepsilon}{k - \varepsilon};$$

or, $F(x)$ étant, par hypothèse, non-décroisante, son image entre les abscisses des points E et B est certainement situé au-dessous de BE, donc à fortiori au-dessous de AD, ce qui démontre évidemment notre affirmation.

On démontre de la même manière que dans le cas où

$$\frac{F(x+h) - F(x)}{h} > k,$$

on a

$$\frac{F(\xi)-F(x)}{\xi-x}<k+\varepsilon$$

pour tout point ξ de l'intervalle $[x+b,\ x+(1+\frac{\varepsilon}{k+\varepsilon})b]$.

On voit donc qu'il existe, dans tous les cas, un intervalle aussi petit que l'on veut δ_1 ayant le point x pour extrémité gauche et contenant à son intérieur un autre intervalle δ_2, dont un point arbitraire ξ satisfait à la condition

$$\left|\frac{F(\xi)-F(x)}{\xi-x}-k\right|>\varepsilon;$$

de plus, *le rapport des longueurs des intervalles δ_2 et δ_1 surpasse un nombre positif qui est indépendant de la longueur de δ_1.*

Ceci nons apprend que $F(x)$ ne saurait admettre une dérivée asymptotique égale à k au point x; nous avons donc démontré la proposition suivante:

Lemme. *Pour une fonction non-décroissante, l'existence d'une dérivée asymptotique positive en un point entraîne l'existence d'une dérivée ordinaire au même point, les deux dérivées étant naturellement égales.*

Il est manifeste que la proposition subsiste pour une fonction non-croissante et une dérivée négative.

4. Supposons maintenant que la fonction mesurable $f(x)$ soit en tout point d'un certain intervalle la dérivée asymptotique d'une autre fonction mesurable $F(x)$. Nous dirons pour abréger que $f(x)$ est *majorée* par la fonction $\varphi(x)$, lorsqu'on a

$$\varphi(x)>f(x)$$

en tout point de l'intervalle considéré.

Supposons que $f(x)$ puisse, dans l'intervalle considéré, être majorée par une fonction $\varphi(x)$ qui, en tout point de cet intervalle, est la dérivée ordinaire d'une fonction continue $\Phi(x)$; la fonction $\varphi(x)-f(x)$ sera, pour tous les points de l'intervalle, la dérivée asymptotique de la fonction $\Phi(x)-F(x)$; or, $\varphi(x)-f(x)$ est positive; il en résulte (nous allons le montrer plus loin, au n^0 14) que $\Phi(x)-F(x)$ est non-décroissante; en appliquant le lemme du n^0 précédent, on voit que $\varphi(x)-f(x)$ est la dérivée ordinaire de $\Phi(x)-F(x)$ en tout point de l'intervalle considéré; donc on a aussi

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

pour toutes les valeurs considérées de x.

La conclusion réciproque est d'un caractère trivial; en effet, toute dérivée exacte $f(x)$ est majorée par la dérivée exacte $f(x)+1$. Nous sommes donc amenés à la proposition suivante qui offre une résolution complète du problème posé:

Théorème[1]. *Pour qu'une fonction $f(x)$ qui est, en tout point d'un certain intervalle, la dérivée asymptotique d'une autre fonction soit, en tout point de l'intervalle considéré, la dérivée ordinaire de cette-même fonction, il faut et il suffit que $f(x)$ puisse être majorée par une dérivée exacte,* c'est-à-dire par une fonction qui est, en tout point de l'intervalle donné, la dérivée ordinaire d'une autre fonction.

En particulier, les conditions du théorème sont remplies pour une dérivée asymptotique bornée (même unilatéralement).

5. Le second problème que nous allons nous poser sera la démonstration du théorème de Rolle pour la dérivée asymptotique.

Dans le cas où fonction primitive est continue, les méthodes classiques offrent aisément cette démonstration. Mais la notion de dérivée asymptotique étant étroitement liée avec la théorie de la mesure de M. Lebesgue, il semble intéressant de traiter le cas d'une primitive mesurable quelconque. C'est pourquoi nous allons entreprendre une étude assez laborieuse afin de parvenir à la proposition générale suivante:

Théorème (de Rolle). *Soit $f(x)$ une fonction mesurable dans l'intervalle* $[0,1]$, *asymptotiquement dérivable en tout point de cet intervalle, et soit* $f(0)=f(1)=0$. *Alors il existe entre* 0 *et* 1 *un point c où la dérivée asymptotique de $f(x)$ est nulle*[2].

Nous allons donner pour ce théorème une preuve directe, en indiquant un procédé de construction qui conduira sans aucune ambiguité au point cherché.

[1]) J'ai indiqué ce théorème sans démonstration dans ma note *„Sur la dérivation asymptotique"*, *Comptes Rendus* t. 164, p. 142, 1917.

[2]) *voir* la note précédente.

Faisons quelques remarques préliminaires.

On peut supposer que, dans tout intervalle intérieur au domaine $[0,1]$, la fonction $f(x)$ soit non-bornée; ou bien, lorsqu'elle est bornée, qu'elle ne prenne ses valeurs extrêmes qu'aux extrémités de l'intervalle au plus; en effet, dans le cas contraire la démonstration devient triviale, se réduisant encore au raisonnement classique. Pour la même raison, il est permis de supposer que les racines de l'équation $f(x)=a$ forment, quel que soit a, un ensemble de mesure nulle au plus dans le domaine $[0,1]$.

Nous désignerons par $M(x_1, x_2)$ la borne supérieure des valeurs de $f(x)$ dans l'intervalle $[x_1, x_2]$, en écrivant

$$M(x_1, x_2)=+\infty$$

dans le cas où $f(x)$ n'est pas bornée supérieurement dans l'intervalle $[x_1\, x_2]$. Soient $E(f<A)$, $E(f=A)$ et $E(f>A)$ les ensembles des points où la valeur de $f(x)$ est respectivement $<$, $=$ ou $>$ que la quantité constante A. Nous désignons par $m(x_1, x_2, A)$ la densité moyenne de l'ensemble $E(f>A)$ dans l'intervalle $[x_1, x_2]$. Pour x_1 et A fixes, la quantité $m(x_1, x, A)$ devient une fonction de x; cette fonction est continue sauf, peut-être, pour $x=x_1$; nous désignerons par $\mu(x_1, x_2, A)$ sa borne supérieure pour $x_1<x\leqslant x_2$[1].

On a, quels que soient x_1, x_2 et A,

$$0\leqslant m(x_1, x_2, A)\leqslant 1$$
$$0\leqslant \mu(x_1, x_2, A)\leqslant 1.$$

6. Lemme I. *Pour x_1 et x_2 fixes, $\mu(x_1, x_2, A)$ est une fonction continue de A, sauf, peut-être, pour la valeur $A=f(x_1)$.*

En effet, soient $\varepsilon>0$ quelconque et $A\neq f(x_1)$.

La fonction $f(x)$ admettant par hypothèse une dérivée asymptotique en x_1, il existe un nombre positif δ et un point ξ de l'intervalle $[x_1, x_2]$ tel que, pour tout point x' compris entre x_1 et ξ, la densité moyenne de $E(|f-A|<\delta)$ soit inférieure à ε dans l'intervalle $[x_1, x']$. D'autre part, dés que le point ξ est fixé, on peut affirmer l'existence d'un $\sigma>0$ tel que la mesure de $E(|f-A|<\sigma)$ dans l'intervalle $[x_1, x_2]$ soit plus petite que $|x_1-\xi|\cdot\varepsilon$; c'est une conséquence de l'hypothèse faite sur l'ensemble des racines de l'équation $f(x)=A$.

[1]) on a supposé $x_1<x_2$ pour fixer les idées.

Soit $\triangle A$ un acroissement arbitraire de la quantité A, inférieur en valeur absolue au plus petit des deux nombres δ et σ, et soit η un point arbitraire de l'intervalle $[x_1, x_2]$.

Supposons, en premier lieu, que η soit situé entre x_1 et ξ; la densité moyenne de $E(|f-A|<|\triangle A|)$ dans l'intervalle (x_1, η) est inférieure, ou égale, à celle de $E(|f-A|<\delta)$, donc inférieure à ε. D'autre part, il est évident qu'elle ne peut pas être plus petite que

$$|m(x_1, \eta, A+\triangle A)-m(x_1, \eta, A)|; \tag{8}$$

donc, cette dernière quantité doit être inférieure à ε.

Dans le cas où η est situé entre ξ et x_2 (limites comprises), la quantité (8) ne peut pas encore surpasser la densité moyenne de $E(|f-A|<|\triangle A|)$ dans l'intervalle $[x_1, \eta]$, et celle-ci est au plus égale à la densité moyenne de $E(|f-A|<\sigma)$ dans le même intervalle. Or, cette dernière densité est inférieure à ε; en effet, la mesure de $E(|f-A|<\sigma)$ dans l'intervalle $[x_1, \eta]$ est plus petite que $|x_1-\xi|\cdot\varepsilon$, et la longueur de cet intervalle surpasse $|x_1-\xi|$.

Donc, dans tous les cas, la quantité (8) est plus petite que ε, d'où l'on conclut aisément que l'on doit avoir

$$|\mu(x_1, x_2, A+\triangle A)-\mu(x_1, x_2, A)|<\varepsilon$$

sous la seule condition que $|\triangle A|$ soit suffisamment petit, ce qui démontre le lemme.

7. Lemme II. *On a, x_1 et x_2 étant fixes,*

$$\mu(x_1, x_2, A)\to 0 \quad \textit{pour} \quad A\to M(x_1, x_2),$$

sauf, peut-être, dans le cas où

$$f(x_1)=M(x_1, x_2).$$

En effet, soient ε un nombre positif quelconque et A_1 un nombre arbitraire tel que

$$f(x_1)<A_1<M(x_1, x_2).$$

La fonction $f(x)$ étant par hypothèse dérivable en x_1 au sens asymptotique, il existe entre x_1 et x_2 un point ξ tel que

$$m(x_1, x, A_1)<\varepsilon \quad \text{pour} \quad x_1<x\leqslant\xi.$$

D'autre part, ξ restant fixe, on peut affirmer l'existence d'une quantité A_2 inférieure à $M(x_1, x_2)$ et telle que la mesure de l'ensemble $E(f > A_2)$ dans l'intervalle $[x_1, x_2)$ soit plus petite que $|x_1 - \xi| \cdot \varepsilon$.

Soit x un point quelconque de l'intervalle $[x_1, x_2]$, et soit A un nombre remplissant les inégalités

$$A_1, A_2 < A < M(x_1, x_2).$$

On supposera en premier lieu que x soit situé entre x_1 et ξ; alors on aura

$$m(x_1, x, A) \leqslant m(x_1, x, A_1) < \varepsilon;$$

en second lieu, lorsque x est compris entre ξ et x_2, on trouve

$$m(x_1, x, A) \leqslant m(x_1, x, A_2) < \varepsilon,$$

d'après la définition même de A_2.

On a donc, quel que soit x dans l'intervalle $[x_1, x_2]$,

$$m(x_1, x, A) < \varepsilon,$$

d'où

$$\mu(x_1, x_2, A) < \varepsilon$$

sous la seule condition que A soit assez voisin de $M(x_1, x_2)$ (assez grand dans le cas où $M(x_1, x_2) = +\infty$).

Le lemme se trouve ainsi démontré.

8. Je désignerai par $f^{[1]}(x)$ la dérivée asymptotique de $f(x)$ au point x. En laissant de côté le cas trivial où l'on aurait identiquement $f^{[1]}(x) = 0$, on peut supposer que

$$M(0, 1) > 0,$$

car dans le cas contraire on remplacerait l'étude de $f(x)$ par celle de $-f(x)$.

Pour l'idée de la démonstration que nous avons en vue, la valeur de $f^{[1]}(0)$ est sans importance. C'est seulement pour rendre le raisonnement plus uniforme que nous supposerons que

$$f^{[1]}(0) > 0.$$

Remarquons d'abord que cette hypothèse est sans influence pour la généralité du raisonnement. En effet, la fonction $f(x)$, étant asymptotiquement dérivable, elle prend nécessairement toutes les valeurs intermédiaires entre deux quelconques de ses valeurs [1]; donc, elle doit prendre au moins deux fois toute valeur intermédiaire entre ses bornes supérieure et inférieure dans l'intervalle $[0, 1]$. Soit a un point tel que

$$f^{[1]}(a) \neq 0.$$

D'après ce qui a été remarqué, il existe un point b distinct de a et tel que

$$f(b) = f(a).$$

Si l'on avait

$$f^{[1]}(0) = f^{[1]}(1) = 0,$$

il suffirait de remplacer l'intervalle $[0, 1]$ par l'intervalle $[a, b]$ et la fonction $f(x)$ par la fonction $f(x) - f(a)$; on peut donc supposer tout d'abord qu'une au moins des deux valeurs $f^{[1]}(0)$ et $f^{[1]}(1)$ soit différente de zéro; on supposera pour fixer les idées que ce soit $f^{[1]}(0)$ et que l'on ait

$$f^{[1]}(0) > 0.$$

9. En vertu de la dernière hypothèse, on a

$$\mu(0, 1, 0) = 1;$$

d'autre part, en vertu du lemme II ($n^0\,7$) on sait que

$$\mu(0, 1, A) \to 0 \quad \text{pour} \quad A \to M(0, 1),$$

la fonction $\mu(0, 1, A)$ étant continue pour toutes les valeurs positives de A, en vertu du lemme I ($n^0\,6$). Soit A_1 le plus grand nombre positif satisfaisant à la condition

$$\mu(0, 1, A_1) = \frac{1}{4}. \tag{9}$$

En vertu de $f(0) = 0$ on a, pour x assez petit,

$$m(0, x, A_1) < \frac{1}{4}.$$

[1] *Voir:* Denjoy, *Sur les fonctions dérivées sommables, Bulletin de la Soc. Math. de France*, 1915.

D'autre part, la condition $f(1)=0$ nous donne, pour b suffisamment petit,

$$m(1-b, 1, A_1) < \frac{1}{4};$$

on a aussi, d'après la définition de la fonction μ, comme conséquence de (9)

$$m(0, 1-b, A_1) \leqslant \frac{1}{4}.$$

Les deux dernières inégalités nous donnent

$$m(0, 1, A_1) < \frac{1}{4}. \tag{10}$$

Donc, la fonction continue de x

$$m(0, x, A_1)$$

doit atteindre son maximum $\left(\text{qui est } \frac{1}{4}\right)$ à l'intérieur de l'intervalle $[0,1]$; soit x_1 la plus grande valeur de x telle que

$$m(0, x_1, A_1) = \frac{1}{4}.$$

Nous allons étudier la fonction $f(x)$ au voisinage de x_1. On s'assure d'abord aisément que la relation

$$f(x_1) < A_1$$

est impossible; en effet, le raisonnement qui nous a conduit à (10), nous donnerait dans ce cas

$$m(0, x_1, A_1) < \frac{1}{4},$$

contrairement à la définition de x_1.

Mais si l'on avait

$$f(x_1) > A_1,$$

on pourrait assigner à droite de x_1 un point x tel que

$$m(x_1, x, A_1) > \frac{1}{4},$$

ce qui donnerait, en vertu de

$$m(0, x_1, A_1) = \frac{1}{4},$$

la relation

$$m(0, x, A_1) > \frac{1}{4},$$

ce qui est contraire à (9).

On a donc nécessairement

$$f(x_1) = A_1.$$

Etudions la valeur de $f^{[1]}(x_1)$. Si l'on avait $f^{[1]}(x_1)=0$, le théorème serait démontré. Le cas $f^{[1]}(x_1)>0$ est impossible; on s'en assure par la même voie qui nous a prouvé l'impossibilité de l'hypothèse $f(x_1) > A_1$.

Dans le seul cas qui nous reste à étudier, l'intervalle $[0, x_1]$ est donc assujetti aux conditions suivantes:

1⁰ $f(0)=0$; $f(x_1)=A_1$;

2⁰ $f^{[1]}(0) > 0$; $f^{[1]}(x_1) < 0$;

3⁰ pour $x \leqslant x_1$, $m(0, x, A_1) \leqslant \frac{1}{4}$.

10. D'après les propriétés indiquées de l'intervalle $[0, x_1]$, cet intervalle contient un point a_1 parfaitement déterminé par les conditions suivantes:

1) pour $a_1 < x < x_1$ on a $m(x, x_1, A_1) \geqslant \frac{3}{4}$;

2) quel que soit $\varepsilon > 0$, il existe un nombre positif ξ tel que

$$a_1 - \varepsilon < \xi < a_1 \text{ et } m(x_1, \xi, A_1) < \frac{3}{4}.$$

On a évidemment

$$f(a_1) \leqslant A_1, \; m(a_1, x_1, A_1) = \frac{3}{4}. \tag{11}$$

Nous allons montrer que

$$x_1 - a_1 < \frac{1}{2} x_1.$$

En effet, dans le cas contraire le rapport de la mesure de $E(f > A_1)$ dans l'intervalle $[a_1, x_1]$ à la longueur de cet intervalle serait

$$m(a_1, x_1, A_1) \leqslant \frac{\operatorname{mes}[E(f > A_1), (0, x_1)]}{|a_1 - x_1|} =$$

$$= \frac{\frac{1}{4} x_1}{|a_1 - x_1|} \leqslant \frac{\frac{1}{4} x_1}{\frac{1}{2} x_1} = \frac{1}{2},$$

ce qui est contraire à (11).

Dans tout ce qui suit, nous ne quitterons plus l'intervalle $[a_1, x_1]$.

11. Il résulte de 1^0 et 2^0 que

$$\mu(x_1, a_1, A_1) = 1.$$

En vertu des lemmes I et II, $\mu(x_1, a_1, A)$ est une fonction continue de A pour $A > A_1$, tendant vers zéro pour $A \to M(x_1, a_1)$. Il existe donc un nombre déterminé $A_2 > A_1$ tel que

$$\mu(x_1, a_1, A_2) = \frac{1}{4},$$

$$\mu(x_1, a_1, A) < \frac{1}{4} \quad \text{pour} \quad A > A_2.$$

Comme auparavant, on s'assure aisément que

$$m(x_1, x, A_2) < \frac{1}{4}$$

pour x suffisamment voisin de x_1; d'autre part, la relation

$$f(a_1) \leqslant A_1 < A_2$$

nous donne sans peine

$$m(x_1, a_1, A_2) < \frac{1}{4}.$$

Ces deux relations nous montrent qu'il existe entre x_1 et a_1 des points x satisfaisant à la condition

$$m(x_1, x, A_2) = \frac{1}{4}.$$

Soit x_2 celui de ces points qui est le plus voisin de a_1. Tout analoguement au cas précédent, on s'assure que

$$f(x_2) = A_2, \; f^{[1]}(x_2) \geqslant 0,$$

et dans la dernière relation il ne reste à examiner que le cas de l'inégalité.

D'autre part, x_2 étant situé entre a_1 et x_1, on a en vertu de 1)

$$m(x, x_1, A_1) \geqslant \frac{3}{4} \quad \text{pour} \quad x_2 \leqslant x < x_1.$$

On peut donc énoncer les résultats établis au sujet de l'intervalle $[x_1, x_2]$ de la manière suivante:

1'. $f(x_1) = A_1, \; f(x_2) = A_2 > A_1$;

2'. $f^{[1]}(x_1) < 0, \; f^{[1]}(x_2) > 0$;

3'. Pour x situé entre x_1 et x_2,

$$m(x_1, x, A_1) \geqslant \frac{3}{4}, \; m(x_1, x, A_2) \geqslant \frac{1}{4};$$

4'. $|x_1 - x_2| < \frac{1}{2} \cdot |x_1 - 0|$.

12. Nous continuerons indéfiniment le procédé envisagé. Tout intervalle (x_{n-1}, x_n) fait partie de l'intervalle (x_{n-2}, x_{n-1}) et jouit des quatre propriétés suivantes:

$1^{(n-1)}$. $f(x_{n-1}) = A_{n-1}, \; f(x_n) = A_n$;

$2^{(n-1)}$. $f^{[1]}(x_{n-1}) > 0, \; f^{[1]}(x_n) < 0$;

(ou réciproquement, selon que n est impair ou pair);

$3^{(n-1)}$. On a

$$m(x_{n-1}, x, A_{n-1}) \geqslant \frac{3}{4}, \quad m(x_{n-1}, x, A_n) \leqslant \frac{1}{4}$$

pour x situé entre x_{n-1} et x_n.

$4^{(n-1)}$. $|x_n - x_{n-1}| < \frac{1}{2} |x_{n-1} - x_{n-2}|$.

En vertu de la dernière propriété, les intervalles (x_{n-1}, x_n), $n = 1, 2, 3, \ldots$, $x_0 = 0$, ont un seul point commun que nous désignerons par c. Nous allons maintenant démontrer que

$$f^{[1]}(c) = 0,$$

ce qui achèvera la démonstration du théorème de Rolle.

13. I. En premier lieu, la propriété $3^{(n-1)}$ conduit à la proposition suivante: quelque petit que soit $\varepsilon > 0$ et quelque grand que soit l'entier positif N, il existe un entier positif $n > N$ et un intervalle $[c, x]$ de longueur inférieure à ε, dont c est l'extrémité droite ou gauche à l'arbitraire et tel que la densité moyenne de l'ensemble $E\{A_n < f < A_{n+1}\}$ soit plus grande que $\frac{1}{2}$ dans cet intervalle.

II. Maintenant je dis que, pour n infini, A_n tend vers une limite positive α et que

$$f(c) = \alpha.$$

En effet, en vertu de $A_n > A_{n-1}$, on a toujours

$$\lim_{n \to \infty} A_n = \alpha,$$

α pouvant se réduire à $+\infty$. Nous allons montrer que l'hypothèse $f(c) \neq \alpha$ (donc en particulier l'hypothèse $\alpha = +\infty$) conduit à une contradiction. Admettons cette hypothèse. La fonction $f(x)$ étant asymptotiquement dérivable au point c, la densité moyenne de $E(|f(x) - f(c)| > \varepsilon)$ dans un intervalle ayant c pour extrémité devient plus petite que $\frac{1}{4}$, quel que soit $\varepsilon > 0$, quand l'intervalle est suffisamment petit. En choisissant ε de façon que α soit en dehors de l'intervalle $[f(c) - \varepsilon, f(c) + \varepsilon]$, on obtient une contradiction avec la proposition I.

III. $f^{[1]}(c)$ ne peut pas être positive; sinon, pour un intervalle assez petit $[c, d]$ ayant c pour extrémité gauche on aurait

$$m(c, d, \alpha) > \frac{1}{2},$$

ce qui est en contradiction avec I.

IV. On démontre tout analoguement que l'inégalité $f^{[1]}(c) < 0$ est impossible.

V. On a donc nécessairement

$$f^{[1]}(c) = 0,$$

et le théorème de Rolle est démontré.

14. Il est inutile d'énoncer les conséquences nombreuses de ce résultat. Il suffit de remarquer que le théorème des accrosisements finis (théorème de Lagrange) en est une conséquence immédiate; en particulier, le théorème suivant nous était déjà utile: si la fonction mesurable $f(x)$ est asymptotiquement dérivable en tout point d'un certain intervalle et si cet intervalle contient deux points x_1 et x_2 tels que

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} < 0,$$

il existe entre x_1 et x_2 un point c tel que

$$f^{[1]}(c) < 0.$$

15. Pour le cas des fonctions qui ne sont asymptotiquement dérivables que presque partout dans un certain intervalle, nous allons nous borner à un seul problème: établir le critère caractéristique pour la structure d'une fonction dont la dérivée asymptotique existe presque partout dans un intervalle.

Supposons que la fonction mesurable $f(x)$ admette une dérivée asymptotique $f^{[1]}(x)$ presque partout dans un intervalle $[a, b]$.

Désignons par E_m l'ensemble des points x de l'intervalle $[a, b]$ tels que, pour tout intervalle contenu dans $[a, b]$, ayant une extrémité au point x et la longueur $\leqslant \frac{1}{m}$, la densité moyenne dans cet intervalle des points t satisfaisant à l'inégalité

$$|f(t) - f(x)| \leqslant m\,|t - x|,$$

surpasse le nombre $\frac{3}{4}$.

On a évidemment, pour tout m, $E_m \subset E_{m+1}$ et, si l'on néglige l'ensemble de mesure nulle, on a: $[a, b] = \sum_1^\infty E_m$. Il existe donc, pour chaque $\varepsilon > 0$, un ensemble E_{m_0} ne différant de $[a, b]$

que par un ensemble de mesure $\leqslant \frac{\varepsilon}{2}$. L'ensemble E_{m_0} étant mesurable, il existe également un ensemble parfait $\pi \subset E_{m_0}$ et coïncidant avec l'intervalle $[a, b]$ à un ensemble de mesure $\leqslant \varepsilon$ près.

Soient maintenant $x < y$ deux points de l'ensemble π tels que

$$|x - y| \leqslant \frac{1}{m_0}.$$

Ces deux points appartenant, par conséquent, à l'ensemble E_{m_0}, on peut affirmer, en tenant compte de la définition de E_{m_0}, l'existence d'un point t dans l'intervalle (x, y) satisfaisant, à la fois, à deux inégalités:

$$|f(t) - f(x)| \leqslant m(t - x), \text{ et } |f(t) - f(y)| \leqslant m(y - t),$$

d'où évidemment:

$$(12) \quad |f(y) - f(x)| \leqslant m_0 |y - x|, \text{ lorsque } x, y \varepsilon \pi \text{ et: } |y - x| \leqslant \frac{1}{m_0}.$$

Désignons maintenant par $\overline{f}(x)$ la fonction continue dans $[a, b]$, identique à $f(x)$ aux points de π et linéaire dans les intervalles contigus à π.

En vertu de (12) la fonction $\overline{f}(x)$ ainsi définie est évidemment à variation bornée, et même absolument continue (elle satisfait même à la condition de Lipschitz), et ne diffère de $f(x)$ qu'aux points formant un ensemble de mesure $\leqslant \varepsilon$.

D'une manière inverse, supposons qu'une telle fonction $\overline{f}(x)$, à variation bornée, existe. Soit x_0 un point tel que

1^0 $\overline{f}'(x_0)$ existe; 2^0 $f(x_0) = \overline{f}(x_0)$; 3^0 x_0 est un point principal de l'ensemble $E(f = \overline{f})$. Il est évident que $f^{[1]}(x_0)$ existe et que

$$f^{[1]}(x_0) = \overline{f}'(x_0).$$

Or, on voit d'après le précédent que la mesure de l'ensemble des points tels que x_0 est aussi voisine que l'on veut de la longueur de l'intervalle $[a, b]$. Donc, $f(x)$ est asymptotiquement dérivable presque partout dans l'intervalle $[a, b]$. Nous avons donc démontré la proposition suivante.

Théorème. *Pour qu'une fonction mesurable soit asymptotiquement dérivable presque partout dans un certain intervalle, il faut et il suffit qu'elle coïncide dans cet intervalle avec une fonction continue à variation bornée (absolument continue), à un ensemble de mesure arbitrairement petite près.*

16. On peut donner au critère caractéristique trouvé une autre forme se prêtant d'une façon plus commode à quelques applications et faisant voir quelques rapports entre la théorie de la dérivation asymptotique et la théorie de l'intégrale de M. Denjoy.

Nous dirons que la fonction mesurable $f(x)$ est absolument continue dans un ensemble mesurable E lorsque la somme des oscillations de $f(x)$ dans un système d'intervalles en nombre fini et sans points communs tend uniformément dans tout l'intervalle donné vers zéro avec la somme des longueurs de ces intervalles, *à condition qu'on ne tient compte que des valeurs de $f(x)$ aux points de l'ensemble E.*

On dira que $f(x)$ est, dans un intervalle $[a, b]$, absolument continue *à un ensemble de mesure arbitrairement petite près*, lorsque, quelque petit que soit $\varepsilon > 0$, il existe dans l'intervalle $[a, b]$ un ensemble parfait P tel que

$$\operatorname{mes} P > |b - a| - \varepsilon$$

et que $f(x)$ est absolument continue dans P.

Théorème. *Pour qu'une fonction mesurable soit asymptotiquement dérivable presque partout dans un certain intervalle, il faut et il suffit qu'elle soit dans cet intervalle absolument continue à un ensemble de mesure arbitrairement petite près* [1]).

En effet, une fonction absolument continue à un ensemble de mesure arbitrairement petite près coïncide, à un ensemble de mesure arbitrairement petite près, avec une fonction continue à variation bornée et admet par suite presque partout une dérivée asymptotique, en vertu du théorème démontré dans le numéro précédent. En effet, on s'assure aisément que lorsqu'une fonction $f(x)$ est absolument continue dans un ensemble parfait P, la fonction continue égale à $f(x)$ dans P et linéaire dans les intervalles contigus de P sera à variation bornée (et même absolument continue) dans l'intervalle total [2]).

1) J'ai démontré ce théorème dans ma note *„Sur la dérivation asymptotique"*, *Comptes Rendus* 1917, t. 164, p. 142. La démonstration du texte est beaucoup plus simple.

2) Pour les détails voir ma Note *„Sur le procédé d'intégration de M. Denjoy"*, *Recueil mathématique de la soc. math. de Moscou, 1918* (en russe).

La réciproque est évidemment contenue dans l'énoncé du No. précédent.

Le théorème est donc démontré.

On peut aussi dire que la condition nécessaire et suffisante pour la dérivabilité asymptotique d'une fonction mesurable presque partout dans un certain intervalle consiste en ce que *cet intervalle puisse être partagé en une infinité dénombrable d'ensembles mesurables dont le premier soit de mesure nulle et que $f(x)$ soit absolument continue dans chacun des autres.*

Dans ma note dernièrement citée j'ai démontré que l'absence du premier ensemble (de mesure nulle) est une condition nécessaire et suffisante pour que $f(x)$ soit, par rapport à sa dérivée asymptotique, l'intégrale indéfinie obtenue par la méthode de M. Denjoy.

§ 4. Recherches ultérieures.

1. La notion classique de dérivée est fondée sur la notion de limite. Il en découle que toute généralisation de la notion de limite nous offre une méthode généralisée de dérivation. Et nous avons vu en effet que la méthode de M. Borel et celle de dérivation asymptotique — les plus importantes connues jusqu'aujourd'hui — étaient fondées sur des généralisations simples de la notion de limite; c'est la méthode des moyennes arithmétiques qui est le fondement de la méthode de M. Borel, tandis que pour la dérivation asymptotique la base est le principe de négliger les ensembles de densité nulle dans l'évaluation des limites de fonctions d'un paramètre continu. Dans les deux cas, la généralisation de l'idée de limite est telle qu'on doit l'appliquer *à chaque point séparément*. Or, une dérivée est toujours une *fonction*, et le passage à la limite qui y amène est appliqué à une fonction de deux variables qui tend vers une fonction d'une seule variable lorsque l'autre variable indépendante tend vers une limite déterminée. On a, par exemple, dans le cas classique,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Une conséquence de ce fait est que, outre les *méthodes de sommation* qui généralisent la notion de limite pour un seul point (une seule valeur de x), toute *généralisation de la notion de convergence d'une série de fonctions* peut servir de base pour une généralisation de la notion de dérivée. En particulier, nous allons traiter une telle généralisation, considérée par M. F. Riesz [1]) et appelée „convergence en mesure".

2. On dira que la fonction $f(x)$ est dans un certain intervalle la *dérivée générale* d'une autre fonction $F(x)$ lorsque, quel que soit $\varepsilon > 0$, la mesure de l'ensemble

$$E\left(\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| > \varepsilon\right)$$

dans l'intervalle considéré tend vers zéro avec h; cette définition n'a aucun sens pour un point considéré à part; en effet, en modifiant d'une façon absolument arbitraire les valeurs de $f(x)$ dans un ensemble quelconque de mesure nulle, on obtient évidemment une fonction que l'on est obligé appeler encore une dérivée générale de $F(x)$. On peut dire que la dérivée générale d'une fonction donnée n'est définie qu'à un ensemble arbitraire de mesure nulle près.

Les propriétés bien connues de la convergence en mesure [2]) montrent que toute fonction qui est presque partout dans un intervalle la dérivée ordinaire d'une autre fonction, est aussi une dérivée générale de la même fonction dans l'intervalle considéré.

D'autre part, la dérivée générale est définie d'une manière unique (en négligeant les ensembles de mesure nulle). En effet, si deux fonctions $f_1(x)$ et $f_2(x)$ sont différentes dans un ensemble de mesure non nulle, il existe un $\varepsilon > 0$ tel que

$$\text{mes}\, E(|f_1 - f_2| > \varepsilon) > 0.$$

Les mesures des ensembles

$$E\left|\left(\frac{F(x+h)-F(x)}{h} - f_1(x)\right| > \frac{\varepsilon}{2}\right)$$

[1]) *Comptes Rendus de l'Ac. des Sc. de Paris*, t. 148, p. 1303, 1909.

[2]) Voir par ex. le cours d'analyse de M. De la Vallée Poussin, t. I, troisième édition, 1915.

et

$$E\left(\left|\frac{F(x+h)-F(x)}{h}-f_2(x)\right|>\frac{\varepsilon}{2}\right)$$

ne peuvent donc pas être à la fois infiniment petites avec h.

3. La méthode de dérivation que nous venons de définir est basée sur la théorie de la mesure de M. Lebesgue; il paraît donc intéressant de la comparer avec la méthode asymptotique qui a le même fondement; cette comparaison nous montrera que c'est la méthode nouvelle qui est plus puissante.

Soit E un ensemble mesurable quelconque. Un point x sera dit appartenir à l'ensemble E_h, h étant un nombre quelconque, lorsque le point $x-h$ appartient à l'ensemble E. L'ensemble E_h est donc le même ensemble E sur lequel on a effectué un déplacement caractérisé par le nombre h.

Lemme. *Soit E un ensemble de points formant un système d'intervalles sans points communs* (en nombre fini ou dénombrable), *on a*

$$\lim_{h\to 0} \text{mes}\,(E+E_h)=\text{mes}\,E.$$

En effet, soit ε un nombre positif quelconque. Je désignerai par

$$\delta_1, \delta_2, \ldots, \delta_n, \ldots$$

les intervalles formant l'ensemble E et les longueurs respectives de ces intervalles. Soit N un entier positif tel que

$$\sum_{k=N+1}^{\infty} \delta_k < \frac{\varepsilon}{2},$$

et h un nombre positif inférieur à $\frac{\varepsilon}{2N}$.

L'ensemble $E+E_k$ est contenu dans les trois systèmes d'intervalles suivantes:

1) $\delta_1, \delta_2, \ldots, \delta_n, \ldots$

2) Les N intervalles de longueur h ayant pour extrémités gauches les extrémités droites des intervalles

$$\delta_1, \delta_2, \ldots, \delta_N.$$

3) Les intervalles

$$\delta_{N+1}^{(h)}, \delta_{N+2}^{(h)}, \ldots, \delta_{N+k}^{(h)}, \ldots,$$

où l'intervalle $\delta_l^{(h)}$ est défini par la condition que x appartienne à $\delta_l^{(h)}$ lorsque $x-h$ appartient à δ_l.

Le premier système constitue l'ensemble E.

La longueur totale des intervalles du second système est au plus égale à $N\cdot h<\frac{\varepsilon}{2}$; enfin, celle des intervalles du troisième système est

$$\sum_{l=1}^{\infty} \delta_{N+l} < \frac{\varepsilon}{2};$$

on a donc

$$\text{mes}\,(E+E_h)<\text{mes}\,E+\varepsilon,$$

ce qui démontre le lemme, puisque le cas d'un h négatif peut être traité de la même manière.

Corollaire. *Pour un ensemble parfait P, on a*

$$\text{mes}\,(P, P_h)\to\text{mes}\,P \quad \textit{pour}\ h\to 0.$$

Pour la démonstration, on applique le lemme que nous venons de prouver, aux ensembles complémentaires.

4. Théorème. *Si la fonction mesurable $f(x)$ admet une dérivée asymptotique presque partout dans un certain intervalle, elle admet aussi dans cet intervalle une dérivée générale, et les deux dérivées sont égales à un ensemble arbitraire de mesure nulle près.*

Soient en effet δ et η deux nombres positifs arbitraires. D'après ce que nous savons (§ 3, n^0 15), il existe une fonction continue à variation bornée $\varphi(x)$ telle qu'on a

$$f(x)=\varphi(x)$$

en tout point d'un ensemble parfait P dont la mesure diffère aussi peu que l'on veut, de la longueur de l'intervalle considéré; nous supposerons que cette différence soit inférieure à η. Cette fonction $\varphi(x)$ est dérivable au sens ordinaire presque partout dans l'intervalle donné et admet par suite dans cet intervalle une dérivée générale qui est évidemment égale à $f^{[1]}(x)$ presque partout dans P.

Or, posons

$$\psi(x)=f(x)-\varphi(x).$$

Nous aurons

$$\frac{f(x+h)-f(x)}{h}=\frac{\varphi(x+h)-\varphi(x)}{h}+\frac{\psi(x+h)-\psi(x)}{h};$$

par suite, la mesure de l'ensemble

$$E=E\left(\left|\frac{f(x+h)-f(x)}{h}-f^{[1]}(x)\right|>2\varepsilon\right)$$

ne peut pas surpasser la somme des mesures des deux ensembles

$$E_1=E\left(\left|\frac{\varphi(x+h)-\varphi(h)}{h}-f^{[1]}(x)\right|>\varepsilon\right)$$

et

$$E_2=E\left(\left|\frac{\psi(x+h)-\psi(x)}{h}\right|>\varepsilon\right).$$

Pour $h\to 0$, la limite supérieure de mes E_1 est inférieure ou égale à η, puisque $f^{[1]}(x)$ coïncide avec la dérivée générale de $\varphi(x)$ presque partout dans P. D'autre part, on a

$$\psi(x+h)=\psi(x)=0$$

pour tout point x faisant partie à la fois des deux ensembles P et P_h. D'après le corollaire du n^0 3, il en découle que la limite supérieure de mes E_2 pour $h\to 0$ est inférieure ou égale à η. Celle de mes E est donc au plus égale à 2η, ce qui démontre le théorème, puisque η et ε sont arbitraires et indépendants l'un de l'autre.

5. Nous allons maintenant montrer que la méthode nouvelle est plus puissante que la méthode asymptotique. A ce but, nous définirons une fonction mesurable dont la dérivée générale dans l'intervalle $[0,1]$ sera nulle et qui sera dépourvue de dérivée asymptotique presque partout dans le même intervalle.

1^0. *Définition de la fonction* $f_1(x)$.

On pose

$$f_1(x)=-\frac{1}{2} \text{ pour } 0\leqslant x<\frac{1}{2}$$

et

$$f_1(x)=+\frac{1}{2} \text{ pour } \frac{1}{2}<x\leqslant 1.$$

2^0. *Définition de la fonction* $f_2(x)$.

On partage le premier intervalle où $f_1(x)$ est constante (l'intervalle $[0,\frac{1}{2}]$) en un nombre pair de parties égales de manière que la longueur de chaque partie soit inférieure à $\frac{1}{3}$ de la longueur du plus petit intervalle de subdivision qu'on a construit jusqu'à ce moment (il suffit de prendre quatre parties); soit δ la longueur d'une telle partie. On pose dans ces parties alternativement, en allant de gauche à droite,

$$f_2(x)=-\delta \text{ et } f_2(x)=+\delta.$$

Ensuite, on partage l'autre intervalle où $f_1(x)$ est constante (l'intervalle $[\frac{1}{2},1]$) en un nombre pair de parties égales de manière que chaque partie ait une longueur inférieure à $\frac{1}{4}$ de la longueur du plus petit intervalle de subdivision défini jusqu'ici (il suffit de prendre 18 parties). Soit de nouveau δ la longueur d'une partie obtenue, on posera dans ces parties alternativement, en allant toujours de gauche à droite,

$$f_2(x)=-\delta \text{ et } f_2(x)=+\delta.$$

3^0. *Définition de la fonction* $f_3(x)$.

Le premier (en comptant toujours de gauche à droite) des intervalles où $f_2(x)$ est constante doit être partagé en un nombre pair de parties égales de façon que chaque partie soit plus petite que $\frac{1}{5}$ de la longueur du plus petit intervalle de subdivision qu'on a construit jusqu'à ce moment; on pose dans les parties ainsi obtenues alternativement

$$f_3(x)=-\delta \text{ et } f_3(x)=+\delta,$$

δ désignant de nouveau la longueur d'une quelconque des parties obtenues.

Ensuite, on partage le second (de gauche à droite) intervalle d'invariabilité de $f_2(x)$ en un nombre pair de parties égales de sorte que chaque partie soit plus petite que $\frac{1}{6}$ de la longueur du plus petit intervalle de subdivision obtenu jusqu'ici. On pose alternativement dans les parties nouvelles

$$f_3(x) = -\delta \quad \text{et} \quad f_3(x) = +\delta,$$

δ ayant une signification nouvelle aisée à comprendre

Et ainsi de suite.

n^0. Définition de la fonction $f_n(x)$.

Pour la fonction $f_{n-1}(x)$, le domaine $[0,1]$ est partagé en un certain nombre de parties dans chacune desquelles cette fonction reste constante.

On partage le premier (toujours de gauche à droite) de ces intervalles en un nombre pair de parties égales de sorte que la longueur δ d'une quelconque de ces parties soit inférieure à $\frac{1}{p}$ de la longueur du plus petit intervalle de subdivision construit jusqu'à ce moment, p étant l'entier dont est venu le tour d'après la loi simple, claire par les définitions précédentes. On pose dans ces parties alternativement

$$f_n(x) = -\delta \quad \text{et} \quad f_n(x) = +\delta;$$

on passe ensuite au second intervalle d'invariabilité de $f_{n-1}(x)$.

Et ainsi de suite.

Pour les points de subdivision, on posera

$$f_i(x) = 0 \quad (i = 1, 2, \ldots).$$

On a évidemment

$$|f_{k+1}(x)| \leqslant \frac{1}{2} |f_k(x)| \quad (k = 1, 2, 3, \ldots),$$

donc la série

$$f_1(x) + f_2(x) + \cdots + f_n(x) + \cdots$$

converge absolument et uniformément dans l'intervalle $[0,1]$ et définit par suite une fonction de première classe d'après la classification de M. Baire que nous désignerons par $f(x)$. Cette fonction satisfait, comme nous allons montrer, à toutes les conditions du problème.

6. Soient

$$\delta_1, \delta_2, \ldots, \delta_n, \ldots$$

toutes les longueurs des différents intervalles de subdivision dont on a parlé dans le n^0 5, rangées par ordre de grandeur. On a

évidemment

$$\delta_i < \frac{1}{i} \delta_{i-1} < \frac{1}{i} \quad (i = 2, 3, \ldots).$$

Soit $\varepsilon > 0$ quelconque et t un entier positif plus grand que $\frac{1}{\varepsilon}$. Pour la définition de la fonction $f(x)$, on avait une fois partagé le dernier (de gauche à droite) intervalle de longueur δ_t; soit δ_n $(n > t)$ la longueur des intervalles partiels correspondants, et soit b un nombre tel que

$$|b| < \delta_{n+1}.$$

On définira un entier positif k par les conditions

$$\delta_{k+1} > |b| \geqslant \delta_{k+2},$$

et l'on aura évidemment $k \geqslant n$.

Les intervalles de longueur δ_{k+1} correspondent à une certaine fonction $f_m(x)$, pour la définition de laquelle ils ont été construits.

Désignons, en comptant de gauche à droite, les intervalles où $f_m(x)$ est constante par

$$\Delta_1, \Delta_2, \ldots, \Delta_s. \tag{12}$$

Nous allons négliger les intervalles de longueur δ_{k+1} et δ_{k+2}. Pour définir $f_m(x)$, on a obtenu ces intervalles par la subdivision de deux intervalles de longueur certainement inférieure à

$$\delta_t < \frac{1}{t} < \varepsilon;$$

la somme de leurs longueurs est donc inférieure à 2ε.

Considérons ceux des intervalles (12) qui sont situés à gauche des intervalles de longueur δ_{k+1}; dans chaque de ces intervalles nous négligerons les points dont la distance de l'une des extrémités de l'intervalle en question est inférieure à $|b|$; on néglige de cette manière, dans chaque intervalle Δ_i considéré, deux intervalles de longueur $|b|$; or, on a

$$2|b| < 2\delta_{k+1} < 2\frac{\Delta_i}{k+1} < 2\frac{\Delta_i}{n} < 2\frac{\Delta_i}{n} < 2\varepsilon\,\Delta_i;$$

la somme des longueurs de tous ces intervalles négligés est donc inférieure à

$$2\varepsilon \sum \Delta_l < 2\varepsilon.$$

Pour tout point x conservé dans un des Δ_l du type considéré, on a évidemment

$$f_i(x+h) - f_i(x) = 0 \quad (i = 1, 2, 3, \ldots, m).$$

D'autre part, chacune des deux quantités $|f_{m+1}(x+h)|$ et $|f_{m+1}(x)|$ est au plus égale à la longueur du plus grand intervalle d'invariabilité de la fonction $f_{m+1}(x)$, donc au plus égale à

$$\delta_{k+3} < \frac{\delta_{k+2}}{k+3} < \frac{\delta_{k+2}}{n} < \varepsilon\,\delta_{k+2} \leqslant \varepsilon\,|h|\,[1]);$$

en vertu de

$$|f_{m+l+1}(x)| \leqslant \frac{1}{2}|f_{m+l}(x)|,$$

$$|f_{m+l+1}(x+h)| \leqq \frac{1}{2}|f_{m+l}(x+h)|,$$

on a donc

$$\left|\sum_{l=1}^{\infty} f_{m+l}(x+h)\right| < \varepsilon\,|h|\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = 2\varepsilon\,|h|,$$

$$\left|\sum_{l=1}^{\infty} f_{m+l}(x)\right| < \varepsilon\,|h|\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = 2\varepsilon\,|h|,$$

et par suite

$$\left|f(x+h) - f(x)\right| = \left|\sum_{l=1}^{\infty} f_l(x+h) - f_l(x)\right| < 4\varepsilon\,|h|,$$

d'où

$$\left|\frac{f(x+h) - f(x)}{h}\right| < 4\varepsilon.$$

[1]) Une exception offre le cas où δ_{k+1} serait la longueur du plus petit intervalle où $f_m(x)$ est constante; dans ce cas la longueur des premiers intervalles d'invariance de $f_{m+1}(x)$ serait δ_{k+2}; or, les intervalles de cette longueur ont été négligés tout d'abord.

Considérons maintenant, s'il en existe, ceux des intervalles du système (12) qui sont situés à droite des intervalles de longueur δ_{k+1}. Ces intervalles constituent dans leur ensemble un ou plusieurs intervalles d'invariance de la fonction $f_{m-1}(x)$; nous désignerons ces derniers par

$$d_1, d_2, \ldots, d_r, \tag{13}$$

en comptant toujours de gauche à droite. Dans chacun de ces intervalles, nous négligerons les points dont la distance à une extrémité de l'intervalle est inférieure à $|h|$. Dans un intervalle d_l, la totalité des points exclus forme deux intervalles dont la somme des longueurs est

$$2|h| < 2\delta_{k+1} < 2\frac{d_l}{k+1} < 2\frac{d_l}{n} < 2\frac{d_l}{t} < 2\varepsilon\,d_l;$$

la somme des longueurs des intervalles exclus des différents intervalles d_l est donc inférieure à

$$2\varepsilon \sum \delta_l < 2\varepsilon.$$

Pour un point quelconque x non exclu, on a évidemment

$$f_i(x+h) - f_i(x) = 0 \quad (i = 1, 2, 3, \ldots, m-1).$$

D'autre part, chacune des quantités $|f_m(x+h)|$ et $|f_m(x)|$ est au plus égale à la longueur du plus grand intervalle d'invariance de $f_m(x)$ inférieur au d_l. Cette longueur est

$$\delta_{k+3} < \frac{\delta_{k+2}}{k+3} < \frac{\delta_{k+2}}{n} < \delta_{k+2}\cdot\varepsilon \leqslant \varepsilon\,|h|;$$

on trouve donc analoguement au raisonnement précédent

$$\left|\sum_{l=0}^{\infty} f_{m+l}(x+h)\right| < \varepsilon\,|h|\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = 2\varepsilon\,|h|$$

$$\left|\sum_{l=0}^{\infty} f_{m+l}(x)\right| < \varepsilon\,|h|\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = 2\varepsilon\,|h|,$$

donc

$$\left|\sum_{l=1}^{\infty} f_l(x+h) - f_l(x)\right| = \left|f(x+h) - f(x)\right| < 4\varepsilon\,|h|,$$

et par suite

$$\left|\frac{f(x+b)-f(x)}{b}\right|<4\varepsilon. \tag{14}$$

L'inégalité (14) est remplie pour tous les x sauf les points exclus qui forment un ensemble dont la mesure doit nous occuper à présent. On a d'abord exclu les intervalles de longueur δ_{k+1} et δ_{k+1} et l'on a vu que la somme des longueurs de ces intervalles est inférieure à 2ε. Ensuite on a supprimé de petits intervalles de longueur $|b|$ à l'intérieure des intervalles Δ_l et d_k. La somme des longueurs de ces intervalles a été inférieure à 4ε. La mesure de l'ensemble des points où (14) tombe en défaut, est donc en somme inférieure à 6ε, dans la seule condition que $|b|$ soit suffisamment petit.

Enfin, $\varepsilon>0$ étant quelconque, on en conclut que $f(x)$ a une dérivée générale identiquement nulle dans l'intervalle $[0,1]$.

7. Nous allons maintenant démontrer que la fonction $f(x)$ est dépourvue de dérivée asymptotique presque partout dans l'intervalle $[0,1]$. A ce but, remarquons d'abord que pour tout point x de cet intervalle, à l'exception d'un ensemble de mesure nulle, la suite

$$f_1(x), f_2(x), \dots, f_n(x), \dots \tag{15}$$

contient une infinité de termes soit positifs, soit négatifs, comme on s'assure par un raisonnement tout élémentaire. Appelons intervalles *extrêmes* ceux des intervalles d'invariance de $f_m(x)$ qui ont une extrémité commune avec un intervalle d'invariance de $f_{m-1}(x)$; alors le même raisonnement élémentaire nous conduira à l'affirmation suivante: *pour tout point x de l'intervalle* $[0,1]$ *à l'exception d'un ensemble de mesure nulle, la série* (15) *contient une infinité de termes soit positifs, soit négatifs, le point x tombant pour les fonctions correspondantes dans des intervalles non-extrêmes.*

Soit x_0 l'un de ces points; soit $\delta>0$ quelconque, et soit m un entier positif tel qu'on ait

$$f_m(x_0)<0,$$

x_0 tombant dans un intervalle d'invariance non-extrême, et que la plus grande longueur des intervalles d'invariance $f_m(x)$ soit plus

petite que $\frac{\delta}{2}$. Désignons par δ_k la longueur de l'intervalle qui contient x_0; cet intervalle étant, par hypothèse, non-extrême, la longueur de l'intervalle voisin à droite est encore δ_k; soient a et b les extrémités de cet intervalle, et soit x un point quelconque de l'intervalle $[a,b]$. D'après la définition de $f_m(x)$, on a

$$f_m(x_0)=-\delta_k,\ f_m(x)=+\delta_k,$$

d'où

$$f_m(x)-f_m(x_0)=2\delta_k.$$

D'autre part, on a évidemment

$$f_i(x)-f_i(x_0)=0 \quad \text{pour} \quad i=1,2,\dots,m-1.$$

Enfin, pour $s>1$,

$$|f_{s+1}(x)|\leqslant\frac{1}{4}|f_s(x)|,$$

donc

$$|f_{m+1}(x_0)|\leqslant\frac{1}{4}\delta_k,\ |f_{m+2}(x_0)|\leqslant\frac{1}{4^2}\delta_k,\dots$$

$$|f_{m+1}(x)|\leqslant\frac{1}{4}\delta_k,\ |f_{m+2}(x)|\leqslant\frac{1}{4^2}\delta_k,\dots$$

En somme, on trouve

$$f(x)-f(x_0)=\sum_{i=1}^{\infty}[f_i(x)-f_i(x_0)]\geqslant$$

$$\geqslant[f_m(x)-f_m(x_0)]-\sum_{k=1}^{\infty}|f_{m+k}(x)|-\sum_{k=1}^{\infty}|f_{m+k}(x_0)|\geqslant$$

$$\geqslant[f_m(x)-f_m(x_0)]-2\delta_k\cdot\frac{1}{4}\left(1+\frac{1}{4}+\frac{1}{4^2}+\cdots\right)=$$

$$=2\delta_k-\frac{2}{3}\delta_k>\delta_k.$$

Or, on a

$$x-x_0<2\delta_k,$$

d'où

$$\frac{f(x)-f(x_0)}{x-x_0}>\frac{1}{2}.$$

Le résultat que nous venons d'obtenir peut être énoncé de la manière suivante: quelque petit que soit $\delta > 0$, il existe à droite de x_0 un intervalle $[a, b]$ tel que:

1) $b - x_0 < \delta$,

2) $b - a \geqslant \frac{1}{2}(b - x_0)$

3) $\frac{f(x) - f(x_0)}{x - x_0} > \frac{1}{2}$, x étant un point quelconque de l'intervalle $[a, b]$.

Par une voie tout à fait analogue, en considérant une fonction $f_n(x)$ dont les intervalles d'invariance sont de longueur inférieure à $\frac{\delta}{2}$, et telle que

$$f_n(x_0) > 0,$$

x_0 tombant encore dans un intervalle non-extrême, on parviendrait au résultat suivant: quelque petit que soit $\delta > 0$, il existe à droite de x_0 un intervalle $[c, d]$ tel que:

1) $d - x_0 < \delta$,

2) $d - c \geqslant \frac{1}{2}(d - x_0)$,

3) $\frac{f(x) - f(x_0)}{x - x_0} < -\frac{1}{2}$, x étant un point quelconque de l'intervalle $[c, d]$.

La comparaison des deux résultats nous montre que $f(x)$ n'est pas asymptotiquement dérivable au point x_0, ce qui démontre notre affirmation.

8. La méthode nouvelle est donc essentiellement plus puissante que toutes les méthodes considérées jusqu'ici. Il paraît donc intéressant de savoir s'il existe des fonctions mesurables auxquelles cette méthode ne serait pas applicable.

On sait[1]) qu'une suite de fonctions mesurables qui converge „en mesure" vers une certaine fonction dans un intervalle contient une suite convergente presque partout dans le même intervalle vers la même fonction dans le sens ordinaire.

Ce théorème subsiste pour le cas où le paramètre varie d'une façon continue. En particulier, pour la dérivation on a le

[1]) Voir p. ex. Lalesco, *Introduction à la théorie des équations intégrales*, Paris 1912.

théorème suivant: *lorsque la fonction mesurable $F(x)$ admet dans un intervalle $[a, b]$ une dérivée générale $f(x)$, il existe une suite décroissante de nombres positifs*

$$b_1 > b_2 > \cdots > b_n > \cdots > 0 \tag{16}$$

telle que $b_n \to 0$ pour $n \to \infty$ et que

$$\frac{F(x + b_n) - F(x)}{b_n} \to f(x) \quad \text{pour} \quad n \to \infty$$

presque partout dans l'intervalle $[a, b]$.

Ce théorème peut donner lieu à une généralisation nouvelle de la notion de dérivée; on peut, en effet, considérer $f(x)$ comme une dérivée généralisée de $F(x)$ dans la condition de l'existence d'une suite telle que (16). Il faudrait alors tout d'abord traiter la question de l'unicité de la dérivée ainsi définie.

Ce que nous savons, c'est que cette méthode doit être applicable à toute fonction admettant une dérivée générale.

Or, nous allons construire une fonction $f(x)$ qui est inattaquable par cette méthode; d'une manière plus précise, nous démontrerons la proposition suivante: *quelle que soit la suite infinie*

$$b_1 > b_2 > \cdots > b_n > \cdots > 0, \quad [b_n \to 0 \text{ pour } n \to \infty],$$

le quotient

$$\frac{f(x + b_n) - f(x)}{b_n}$$

est, presque partout dans l'intervalle $[0, 1]$, *dépourvu de limite déterminée pour $n \to \infty$.*

Il en découle évidemment que $f(x)$ ne saurait pas admettre une dérivée générale dans l'intervalle $[0, 1]$.

9. Soit donc

$$f(x) = \frac{a_2(x)}{2^2} + \frac{a_4(x)}{2^4} + \cdots + \frac{a_{2n}(x)}{2^{2n}} + \cdots,$$

où $a_n(x)$ désigne le n-ième chiffre du développement de x en une fraction dyadique infinie, et soit

$$b_1 > b_2 > \cdots > b_n > \cdots > 0, \quad b_n \to 0 \text{ pour } n \to \infty.$$

A tout entier positif i faisons correspondre un autre entier positif défini par les inégalités

$$\frac{1}{2^{2n_i}} > b_i \geqslant \frac{1}{2^{2(n_i+1)}}.$$

Un raisonnement tout élémentaire nous démontrerait aisément les affirmations suivantes.

1) Les relations

$$a_{2n_i}(x) = 0, \ a_{2n_i+1}(x) = 1 \tag{17}$$

sont, presque partout dans l'intervalle $[0, 1]$, simultanément remplies pour une infinité de valeurs de i.

2) Les relations

$$\left.\begin{aligned} a_{2n_k+2}(x) &= 1, \\ a_{2n_k+1}(x) &= 1, \\ a_{2n_k}(x) &= 1, \\ a_{2n_k-1}(x) &= 0 \end{aligned}\right\} \tag{18}$$

sont, presque partout dans le même intervalle, remplies une infinité de fois simultanément.

Soit x un point pour lequel chacun des systèmes de relations (17) et (18) est rempli une infinité de fois (les points tels que x forment un ensemble dont la mesure est égale à l'unité).

Soit, en premier lieu, le système (17) rempli pour l'entier positif i. En vertu de

$$x = \frac{a_1(x)}{2} + \frac{a_2(x)}{2^2} + \cdots + \frac{a_{2n_i-1}(x)}{2^{2n_i-1}} + \frac{a_{2n_i}(x)}{2^{2n_i}} + \frac{a_{2n_i+1}(x)}{2^{2n_i+1}} + \cdots,$$

$$x + b_i = \frac{a_1(x+b_i)}{2} + \frac{a_2(x+b_i)}{2^2} + \cdots + \frac{a_{2n_i+1}(x+b_i)}{2^{2n_i-1}} +$$

$$+ \frac{a_{2n_i}(x+b_i)}{2^{2n_i}} + \frac{a_{2n_i+1}(x+b_i)}{2^{2n_i+1}} + \cdots$$

et de

$$a_{2n_i}(x) = 0,$$

on doit avoir

$$a_m(x+b_i) = a_m(x) \text{ pour } m = 1, 2, \ldots, 2n_i - 1. \tag{19}$$

En effet, dans le cas contraire, en posant

$$S_n(x) = \sum_{k=1}^{n} \frac{a_k(x)}{2^k}, \quad R_n(x) = \sum_{k=n+1}^{\infty} \frac{a_k(x)}{2^k},$$

on aurait

$$S_{2n_i-1}(x+b_i) - S_{2n_i-1}(x) \geqq \frac{1}{2^{2n_i-1}},$$

$$a_{2n_i}(x+b_i) - a_{2n_i}(x) \geqq 0,$$

$$|R_{2n_i}(x+b_i) - R_{2n_i}(x)| < \frac{1}{2^{2n_i}},$$

d'où

$$(x+b_i) - x = b_i \geqq \frac{1}{2^{2n_i}},$$

ce qui est contraire à la définition du nombre n_i.

Si

$$a_{2n_i}(x+b_i) = 1,$$

on a

$$a_{2n_i}(x+b_i) - a_{2n_i}(x) = 1,$$

$$\sum_{k=1}^{\infty} \left| \frac{a_{2(n_i+k)}(x+b_i)}{2^{2(n_i+k)}} - \frac{a_{2(n_i+k)}(x)}{2^{2(n_i+k)}} \right| \leqq$$

$$\leqq \frac{1}{2^{2(n_i+1)}} \left\{ 1 + \frac{1}{2^2} + \frac{1}{2^4} + \cdots \right\} = \frac{4}{3 \cdot 2^{2(n_i+1)}},$$

et par suite

$$f(x+b_i) - f(x) \geqq \frac{1}{2^{2n_i}} - \frac{1}{3 \cdot 2^{2n_i}} > \frac{2}{3} b_i,$$

d'où

$$\frac{f(x+b_i) - f(x)}{b_i} > \frac{2}{3}.$$

Si

$$a_{2n_i}(x+b_i) = 0,$$

on doit avoir

$$\left.\begin{aligned} a_{2n_l+2}(x)&=0,\\ a_{2n_l+2}(x+b_l)&=1,\\ a_{2n_l+1}(x+b_l)&=1.\end{aligned}\right\}. \tag{20}$$

En effet, la dernière de ces relations est une conséquence de

$$a_{2n_l+1}(x)=1,$$

car autrement on aurait

$$x+b_l<x,$$

ce qui est absurde. Par la même raison, la combinaison

$$a_{2n_l+1}(x)=1,\ a_{2n_l+1}(x+b_l)=0$$

est impossible.

D'autre part, si l'on avait

$$a_{2n_l+2}(x+b_l)=a_{2n_l+2}(x),$$

il en suivrait

$$(x+b_l)-x=\sum_{k=1}^{\infty}\frac{a_{2n_l+2+k}(x+b_l)-a_{2n_l+2+k}(x)}{2^{2n_l+2+k}}<$$

$$<\frac{1}{2^{2n_l+2}}\left[\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots\right]=\frac{1}{2^{2n_l+2}},$$

ce qui est encore en contradiction avec la définition de n_l. Les relations (20) sont donc démontrées, et nous avons

$$f(x+b_l)-f(x)=\frac{1}{2^{2n_l+2}}+\sum_{k=1}^{\infty}\left\{\frac{a_{2(n_l+1+k)}(x+b_l)-a_{2(n_l+1+k)}(x)}{2^{2(n_l+1+k)}}\right\}>$$

$$>\frac{1}{2^{2n_l+2}}-\sum_{k=1}^{\infty}\frac{1}{2^{2(n_l+1+k)}}=\frac{1}{3}\,\frac{1}{2^{2n_l+1}},$$

d'où

$$\frac{f(x+b_l)-f(x)}{b_l}>\frac{1}{3\cdot 2^{2n_l+1}}:\frac{1}{2^{2n_l}}=\frac{1}{6}.$$

Donc, dans la condition que les relations (17) soient remplies, on a dans tous les cas

$$\frac{f(x+b_l)-f(x)}{b_l}>\frac{1}{6}.$$

Cette proposition était notre premier but.

10. Soit maintenant k un entier positif tel que les relations (18) soient remplies.

On a d'abord

$$a_m(x+b_k)=a_m(x) \text{ pour } m=1,2,\ldots,2n_k-2.$$

En effet, dans le cas contraire on aurait

$$S_{2n_k-2}(x+b_k)-S_{2n_k-2}(x)\geq\frac{1}{2^{2n_k-2}},$$

ce qui donnerait, en vertu de

$$a_{2n_k-1}(x+b_k)-a_{2n_k-1}(x)\geqq 0$$

et

$$|R_{2n_k-1}(x+b_k)-R_{2n_k-1}(x)|<\frac{1}{2^{2n_k-1}},$$

la relation

$$(x+b_k)-x=b_k>\frac{1}{2^{2n_k-1}},$$

ce qui est en contradiction avec la définition du nombre n_k.

Je dis ensuite que

$$a_{2n_k-1}(x+b_k)=1;$$

en effet, dans le cas contraire on aurait

$$b_k=(x+b_k)-x<\sum_{r=1}^{\infty}\frac{1}{2^{2n_k+r+2}}=\frac{1}{2^{2n_k-2}},$$

contrairement à la définition de n_k.

Enfin, on doit avoir

$$a_{2n_k}(x+b_k)=0,$$

sinon on aurait

$$b_k=x+b_k-x>\frac{1}{2^{2n_k-1}}-\sum_{r=1}^{\infty}\frac{1}{2^{2n_k+r}}=\frac{1}{2^{2n_k}},$$

ce qui contient encore une contradiction avec la définition de n_k.

On a donc

$$a_{2m}(x+b_k)-a_{2m}(x)=0,\ [m=1,2,\ldots,n_k-1]$$
$$a_{2n_k-1}(x+b_k)=1,\ a_{2n_k-1}(x)=0,$$
$$a_{2n_k}(x+b_k)=0,\ a_{2n_k}(x)=1,$$

d'où

$$f(x+b_k)-f(x) < -\frac{1}{2^{2n_k}}+\sum_{r=1}^{\infty}\frac{1}{2^{2(n_k+r)}}=$$

$$=-\frac{1}{2^{2n_k}}+\frac{4}{3}\cdot\frac{1}{2^{2n_k+2}}=-\frac{1}{3\cdot 2^{2n_k}},$$

et par suite

$$\frac{f(x+b_k)-f(x)}{b_k} < -\frac{2}{3},$$

puisque

$$b_k < \frac{1}{2^{2n_k}}.$$

Cette inégalité a lieu chaque fois que k satisfait aux relations (18).

Il y a donc, dans la suite

$$b_1, b_2, \ldots, b_p, \ldots$$

une infinité de nombres satisfaisant à la relation

$$\frac{f(x+b_k)-f(x)}{b_k} - \frac{1}{6}$$

et une infinité de nombres remplissant l'autre condition

$$\frac{f(x+b_k)-f(x)}{b_k} < -\frac{2}{3}$$

On en conclut que la quantité

$$\frac{f(x+b_n)-f(x)}{b_n}$$

ne peut pas tendre vers une limite déterminée, pour $n\to\infty$. Puisque x est un point quelconque d'un ensemble dont la mesure est égale à l'unité, l'affirmation à la fin du n^0 8 est complètement démontrée.

11. Nous avons vu qu'il existe une fonction mesurable sans dérivée générale. Cette fonction est discontinue; la question de l'existence d'une fonction *continue* sans dérivée générale n'est donc pas encore résolue. Or, la réponse est encore affirmative, et c'est ce que nous allons voir à présent.

D'après le théorème fondamental de M. Lusin, il existe une fonction continue $\varphi(x)$ telle qu'on a

$$f(x)=\varphi(x)$$

en tout point d'un ensemble parfait P dont la mesure est aussi voisine que l'on veut de l'unité. Notre affirmation sera démontrée si nous montrerons que pour toute suite infinie

$$b_1, b_2, \ldots, b_n, \ldots \; (b_n > b_{n+1} > 0, \; \lim_{n\to\infty} b_n = 0)$$

le quotient

$$\frac{\varphi(x+b_n)-\varphi(x)}{b_n}$$

est dépourvu de limite déterminée, pour $n\to\infty$, x étant un point quelconque d'un ensemble de mesure non nulle.

Supposons, par impossible, que ce quotient ait une limite déterminée presque partout dans l'intervalle $[0,1]$. Posons

$$\varepsilon_n = \operatorname{mes} P - \operatorname{mes}(P, P_{-b_n}).$$

D'après le corollaire du n^0 3, on a

$$\lim_{n\to\infty}\varepsilon_n = 0.$$

Il existe donc une suite croissante d'entiers positifs

$$n_1, n_2, \ldots, n_k, \ldots,$$

telle que la série

$$\varepsilon_{n_1}+\varepsilon_{n_2}+\cdots+\varepsilon_{n_k}+\cdots$$

soit convergente. D'après un théorème fondamental de la théorie métrique des ensembles, on en conclut que tous les points x de P, sauf au plus un ensemble de mesure nulle, remplissent la condition que dans la suite

$$x+b_{n_1}, x+b_{n_2}, \ldots, x+b_{n_k}, \ldots$$

il n'y a qu'un nombre fini de points non-appartenants à l'ensemble P.

D'après l'hypothèse faite, le quotient

$$\frac{\varphi(x+b_{n_k})-\varphi(x)}{b_{n_k}}$$

tend presque partout vers une limite déterminée pour $k \to \infty$. Puisque $f(x) = \varphi(x)$ dans P, on en conclut que le quotient

$$\frac{f(x+b_{n_k}) - f(x)}{b_{n_k}}$$

tend, pour k infini, vers une limite déterminée presque partout dans P, ce qui est contraire à la propriété de la fonction $f(x)$ démontrée dans le n^0 précédent. Notre affirmation est donc vraie.

Ivanovo — Wosniesiensk,
Octobre, 1919.

Remarque post scriptum. La démonstration du théorème du § 2, n^0 7 (p. 227) de ce Mémoire n'étant pas (comme je viens de remarquer pendant la correction d'épreuves) complète, on le remplacera par la suivante.

Désignons par E l'ensemble des points où $f(x)$ n'est pas dérivable asymptotiquement, mais est dérivable au sens de M. Borel. Supposons, par impossible, que mes $E > 0$.

Soit E_m l'ensemble des point $x \varepsilon E$ tels que, pour chaque t

$$0 < t - x < \frac{1}{m} \text{ entraîne: } \left| \frac{1}{t-x} \int_x^t \frac{f(u)-f(x)}{u-x} du \right| < m. \qquad (1)$$

On a évidemment $E = \sum_1^\infty E_m$, et il existe, par conséquent, tout au moins une valeur de l'indice m, soit m_0, telle que mes $E_{m_0} > 0$.

Soit F un sous-ensemble de E_{m_0}, fermé et de mesure positive; un tel ensemble existe, car tous les ensembles E_m sont évidemment mesurables.

A presque tout point de E, donc, en particulier, à presque tout point x de F, vient correspondre (d'après le théorème fondamental rappelé au début de ce Mémoire [1]) un ensemble M_x ayant dans x un point de densité supérieure 1 et tel que l'on ait

$$\lim \frac{f(y)-f(x)}{y-x} = +\infty,$$

lorsque y tend vers x dans M_x.

[1]) voir la *Remarque de la Rédaction* (p. 212).

Soit x_0 un de tels points de F. On peut évidemment admettre qu'il soit à la fois un point principal de F.

Pour tout couple de nombres positifs ε et N, il existe donc un nombre positif $b < \varepsilon, \frac{1}{m_0}$ tel que

la densité moyenne de F dans l'intervalle $(x_0, x_0 + b)$ est $\geqslant \frac{3}{4}$, (2)

$$\text{mes } G > \frac{3}{4} b, \qquad (3)$$

où G désigne l'ensemble des points y de l'intervalle $(x_0, x_0 + b)$ satisfaisant à l'inégalité

$$f(y) - f(x_0) > N(y - x_0). \qquad (4)$$

En raison de la continuité de $f(x)$ G est un ensemble ouvert; il est donc somme d'une suite d'intervalles ouverts et disjoints $\{\delta_n\}$. Nous allons modifier ce système de la manière suivante: lorsque δ_n ne contient ancun point de F, il sera simplement supprimé; dans le cas contraire, on le remplacera par un intervalle ouvert δ'_n dont l'extrémité gauche coïncide avec celle de δ_n et l'extrémité droite est la borne supérieure de la partie de l'ensemble F contenue dans δ_n.

L'ensemble $\sum_1^\infty \delta'_n$ est évidemment contenu dans $G = \sum_1^\infty \delta_n$, mais ne diffère de celui-ci que par les points n'appartenant pas à F, donc par l'ensemble de points qui, en vertu de (2), est de mesure $\leqslant \frac{1}{4} b$. On a parsuite, en tenant compte de (3):

$$\text{mes} \left(\sum_{n=1}^{\infty} \delta'_n \right) \geqslant \text{mes } G - \frac{1}{4} b > \frac{1}{2} b.$$

Par conséquent, pour un nombre p assez grand, il vient:

$$\text{mes} \left(\sum_1^p \delta'_n \right) > \frac{1}{4} b. \qquad (5)$$

Posons: $\delta'_n = (a_n, b_n)$; on peut évidemment supposer les intervalles δ'_n $(n = 1, 2, \ldots, p)$ rangés de manière que l'on ait toujours: $b_n < a_{n+1} < b_{n+1}$ où $1 \leqslant n < p$. On peut admettre eu outre que $b_0 = x_0$.

Ceci posé, on a en, vertu de (4):

$$\int_{a_n}^{b_n} \frac{f(u)-f(x_0)}{u-x_0}\,du > N\cdot(b_n-a_n),\ \text{d'où, d'aprés (5):}$$

$$\sum_{n=1}^{p}\int_{a_n}^{b_n} \frac{f(u)-f(x_0)}{u-x_0}\,du > \frac{N}{4}\,b \geqslant \frac{N}{4}(b_p-x_0). \tag{6}$$

D'autre part, pour $0 \leqslant n < p$:

$$\begin{aligned}\int_{b_n}^{a_{n+1}} \frac{f(u)-f(x_0)}{u-x_0}\,du &= \int_{b_n}^{a_{n+1}} \frac{f(u)-f(b_n)}{u-x_0}\,du \\ &+ [f(b_n)-f(x_0)]\cdot\int_{b_n}^{a_{n+1}} \frac{du}{u-x_0};\end{aligned} \tag{7}$$

or, d'après (4) on a pour tout $y \varepsilon G$: $f(y)-f(x_0) \geqslant 0$, donc on a aussi pour chaque b_n: $f(b_n)-f(x_0) \geqslant 0$, et il s'ensuit de (7) que:

$$\int_{b_n}^{a_{n+1}} \frac{f(u)-f(x_0)}{u-x_0}\,du \geqslant \int_{b_n}^{a_{n+1}} \frac{f(u)-f(b_n)}{u-x_0}\,du.$$

En appliquant enfin au second membre de cette inégalité le second théorème de la moyenne, on obtient:

$$\begin{aligned}\int_{b_n}^{a_{n+1}} \frac{f(u)-f(x_0)}{u-x_0}\,du &\geqslant \int_{b_n}^{a_{n+1}} \frac{f(u)-f(b_n)}{u-b_n}\,\frac{u-b_n}{u-x_0}\,du \\ &\geqslant \frac{a_{n+1}-b_n}{a_{n+1}-x_0}\int_{\xi}^{a_{n+1}} \frac{f(u)-f(b_n)}{u-b_n}\,du- \\ &\geqslant \frac{a_{n+1}-b_n}{a_{n+1}-x_0}\left[\int_{b_n}^{a_{n+1}} \frac{f(u)-f(b_n)}{u-b_n}\,du - \int_{b_n}^{\xi} \frac{f(u)-f(b_n)}{u-b_n}\,du\right],\end{aligned} \tag{8}$$

où: $b_n \leqslant \xi \leqslant a_{n+1}$.

Or, b_n appartenant à $F \subset E_{m_0}$, on a, en vertu de (1):

$$\left|\int_{b_n}^{a_{n+1}} \frac{f(u)-f(b_n)}{u-b_n}\,du\right| \leqslant (a_{n+1}-b_n)\cdot m_0,$$

$$\left|\int_{b_n}^{\xi} \frac{f(u)-f(b_n)}{u-b_n}\,du\right| \leqslant (\xi-b_n)\,m_0$$

donc, en raison de (8):

$$\begin{aligned}\int_{b_n}^{a_{n+1}} \frac{f(u)-f(x_0)}{u-x_0} &\geqslant -2\,(a_{n+1}-b_n)\cdot m_0\cdot\frac{a_{n+1}-b_n}{a_{n+1}-x_0} \geqslant - \\ &\geqslant -2\,(a_{n+1}-b_n)\cdot m_0,\end{aligned}$$

d'où, en tenant compte de (5):

$$\begin{aligned}\sum_{n=0}^{p-1}\int_{b_n}^{a_{n+1}} \frac{f(u)-f(x_0)}{u-x_0}\,du &\geqslant -2\,m_0\sum_{n=0}^{p-1}(a_{n+1}-b_n) \geqslant \\ &\geqslant -\frac{3}{2}(b_p-x_0)\,m_0.\end{aligned}$$

On déduit de là et de (6):

$$\begin{aligned}\frac{1}{b_p-x_0}\int_{x_0}^{b_p} \frac{f(u)-f(x_0)}{u-x_0}\,du &= \frac{1}{b_p-x_0}\left[\sum_{n=1}^{p}\int_{a_n}^{b_n} \frac{f(u)-f(x_0)}{u-x_0}\,du\right. \\ &\left.+\sum_{n=0}^{p-1}\int_{b_n}^{a_{n+1}} \frac{f(u)-f(x_0)}{u-x_0}\,du\right] \geqslant \frac{1}{4}(N-6\,m_0),\end{aligned}$$

où m_0 est un nombre fixe, N est un nombre positif arbitrairement grand et la différence $b_p-x_0 < b < \varepsilon$ peut être aussi petite que l'on veut.

On a, par conséquent:

$$\limsup_{t\to x_0}\int_{x_0}^{t} \frac{f(u)-f(x_0)}{u-x_0}\,du = +\infty.$$

Nous sommes donc en contradiction avec la supposition que $f(x)$ admette la dérivée moyenne au point $x_0 \varepsilon E$.

Le théorème est ainsi démontré.