JOURNAL

DE

MATHÉMATIQUES

PURES ET APPLIQUÉES

FONDÉ EN 1836 ET PUBLIÉ JUSQU'EN 1874

PAR JOSEPH LIOUVILLE

MAURICE FRÉCHET

Sur la notion de différentielle

Journal de mathématiques pures et appliquées 9e série, tome 16, n° 1-4 (1937), p. 233-250.

<http://www.numdam.org/item?id=JMPA_1937_9_16_1-4_233_0>

Sur la notion de différentielle;

PAR MAURICE **FRÉCHET**.

Résumé. — L'auteur montre que la définition de la différentielle totale de Stolz-Young est équivalente à la définition due à M. Hadamard. Par contre, quand on étend cette dernière aux fonctionnelles elle devient plus générale que celle de l'auteur. Enfin la définition due à M. Paul Lévy, ne vérifiant pas nécessairement le théorème des fonctions composées, est plus générale encore, mais, pour cette même raison, peut-être trop générale.

INTRODUCTION. — M. Hadamard s'est occupé, à plusieurs reprises, de la notion de différentielle, soit dans l'Analyse classique, soit dans l'Analyse fonctionnelle. Ses publications et son enseignement nous ont incité à étudier cette notion et à en donner des extensions successives en 1911 et en 1925. On ne s'étonnera pas de trouver dans un volume dédié à l'illustre mathématicien, un Mémoire ayant pour but de préciser les relations qui existent entre les définitions dues à M. Hadamard (ou qu'on peut déduire des siennes) avec des définitions — équivalentes ou non, mais en tout cas de formes distinctes — proposées par d'autres auteurs.

1. LA DIFFÉRENTIELLE DANS L'ANALYSE CLASSIQUE. — Supposant connue la définition de la différentielle d'une fonction d'une variable, M. Hadamard a proposé (¹) de définir la différentielle totale d'une fonction de plusieurs variables, par exemple, $f(x, y)$, comme étant

(¹) *La notion de différentielle dans l'enseignement* (*Scripta Univ. Ab. Bib. Hierosolymitanarum*, Jérusalem, 1923).

l'expression classique

$$dz = \frac{\partial z}{dx}dx + \frac{\partial z}{\partial y}dy. \tag{1}$$

Mais, pour lui, cette expression n'est qu'un *symbole d'opérations* : « que signifie, dit-il, l'égalité (1) ? Que si x, y et dès lors $z = f(x, y)$ sont exprimés en fonction d'une variable auxiliaire quelconque u, on a, quelles que soient ces expressions,

$$\frac{dz}{du} = \frac{\partial z}{\partial x}\frac{dx}{du} + \frac{\partial z}{\partial y}\frac{dy}{du}. \tag{2}$$

Tel est le sens unique de l'égalité (1). L'égalité (2) ayant lieu quelle que soit la variable indépendante en fonction de laquelle les deux autres variables sont exprimées, on supprime la mention de u. L'avantage précieux de la notation différentielle consiste précisément en la possibilité de ne pas préciser quelle est la variable que l'on considérera comme indépendante. »

Pour pouvoir comparer en toute rigueur cette définition avec d'autres définitions, il sied de bien préciser le sens qu'il faut lui attribuer. On sait depuis longtemps qu'il ne suffit pas qu'on puisse *écrire* la relation (2) pour qu'elle soit exacte. Autrement dit, il ne suffit pas que $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{dx}{du}, \frac{dy}{du}$ existent pour que $\frac{dz}{du}$ existe et vérifie (2). Dès lors, en supposant vraie l'égalité (2), M. Hadamard impose à la fonction $z(x, y)$ une condition supplémentaire. Nous dirons donc : *une fonction $f(x_1, \ldots, x_n)$ admet au point $(a_1, \ldots, a_n)$ une différentielle au sens de M. Hadamard, si f admet en ce point toutes ses dérivées partielles du premier ordre et si,* DE PLUS, *quand $\varphi_1(u), \ldots, \varphi_n(u)$ sont des fonctions de u égales à $a_1, \ldots, a_n$ pour $u = u_0$ et dérivables pour $u = u_0$, alors $f[\varphi_1(u), \ldots, \varphi_n(u)]$ est dérivable pour $u = u_0$, et l'on a*

$$\left[\frac{d}{du}f[\varphi_1(u), \ldots, \varphi_n(u)]\right]_{u=u_0} = f'_{a_1}\left[\frac{d\varphi_1}{du}\right]_{u=u_0} + \ldots + f'_{a_n}\left[\frac{d\varphi_n}{du}\right]_{u=u_0}. \tag{3}$$

La différentielle df sera le symbole d'opération

$$df(x_1, \ldots, x_n) = f'_{a_1}dx_1 + \ldots + f'_{a_n}dx_n. \tag{4}$$

La définition de M. Hadamard met l'accent sur ses propriétés opératoires. D'autres définitions de la différentielle mettent plutôt en évidence sa propriété de constituer une sorte de partie principale de l'accroissement de la fonction [1]. Elles sont et ne restent rigoureuses qu'en donnant au mot partie principale un sens un peu différent de son sens habituel. Il est, en effet, facile de voir que même pour des fonctions $f(x, y)$ très simples, le rapport $\frac{\Delta f}{df}$, c'est-à-dire

$$\frac{f(x+\Delta x, y+\Delta y)-f(x, y)}{f'_x \Delta x+f'_y \Delta y},$$

ne tend pas nécessairement vers l'unité, de quelque façon que Δx et Δy tendent vers zéro. Et même ce rapport n'a pas de sens quand f'_x et f'_y sont nuls au point considéré sans que f soit constant.

D'après Stolz (qui a introduit une nouvelle définition) et W. H. Young (qui est le premier à en avoir montré en détail les avantages) *une fonction $f(x_1, \ldots, x_n)$ est différentiable au point $a_1, \ldots, a_n$ si les dérivées partielles $f'_{a_1}, \ldots, f'_{a_n}$ existent, et si,* DE PLUS, *on a*

$$(5) \qquad \left\{ \begin{aligned} & f(a_1+\Delta x_1, \ldots, a_n+\Delta x_n)-f(a_1, \ldots, a_n) \\ & \quad =\Delta x_1(f'_{a_1}+\varepsilon_1)+\ldots+\Delta x_n(f'_{a_n}+\varepsilon_n) \end{aligned} \right.$$

où $\varepsilon_1, \ldots, \varepsilon_n$ sont infiniment petits avec $|\Delta x_1|+\ldots+|\Delta x_n|$. (Et alors $df=f'_{a_1}\Delta x_1+\ldots+f'_{a_n}\Delta x_n$.)

Nous avons proposé [2] deux autres définitions dont l'une, de forme très différente de celle-ci, est la suivante pour le cas de deux variables.

Une fonction $z=f(x, y)$ est différentiable au point a, b, si la surface $z=f(x, y)$ admet au point $a, b, c=f(a, b)$ un plan tangent non parallèle à Oz. Et si ce plan a pour équation

$$Z-c=p(X-a)+q(Y-b),$$

la différentielle de f est

$$df=p\,dx+q\,dy.$$

---

[1] De telles définitions quand on s'en inspire pour définir les différentielles d'ordre supérieur ont conduit fréquemment à des contradictions ou, au moins des obscurités que M. Hadamard a justement relevées. Nous croyons cependant qu'il ne s'agit là que de défauts d'exposition qu'il est possible d'éviter.

[2] *Voir* par exemple : *Sur la notion de différentielle totale* (*Nouv. Ann. Math.*, t. 12, 1912, p. 385-403 et 433-449).

Nous avons démontré [1] que cette dernière définition est équivalente à celle de Stolz-Young. La seconde définition, qui s'est présentée à nous comme un cas particulier de la définition de la différentielle d'une fonctionnelle à laquelle nous avions été conduit d'abord, et directement, se rapproche beaucoup plus, dans sa forme même, de celle de Stolz-Young. On peut la résumer en disant : la différentielle de $f(x_1, \ldots, x_n)$ est une fonction des accroissements $\Delta x_1, \ldots, \Delta x_n$ qui constitue une représentation *simple et approchée* de l'accroissement correspondant de la fonction. Et nous avons précisé ainsi le sens de cette définition : *une fonction $f(x_1, \ldots, x_n)$ est différentiable au point $a_1, \ldots, a_n$ s'il existe une fonction de $\Delta x_1, \ldots, \Delta x_n$, du premier degré et homogène, soit*

$$A_1 \Delta x_1 + \ldots + A_n \Delta x_n,$$

qui ne diffère de l'accroissement correspondant Δf de la fonction f que par une quantité infiniment petite par rapport à la distance

$$r = \sqrt{\Delta x_1^2 + \ldots + \Delta x_n}$$

du point fixe $(a_1, \ldots, a_n)$ et du point varié

$$a_1 + \Delta x_1, \ \ldots, \ a_n + \Delta x_n.$$

Et alors cette fonction auxiliaire $\Sigma A_i \Delta x_i$ est la différentielle de f correspondant à ces accroissements Δx_i.

La différentielle se traduit donc par les deux égalités [2] :

$$\Delta f - df = o(r) = o\left(\sqrt{\Delta x_1^2 + \ldots + \Delta x_n^2}\right),$$
$$df = A_1 \Delta x_1 + \ldots + A_n \Delta x_n.$$

Il est d'ailleurs facile de voir qu'alors $f'_{a_1}, \ldots, f'_{a_n}$ existent, qu'on a

$$A_1 = f'_{a_1}, \ \ldots, \ A_n = f'_{a_n},$$

et de s'assurer que cette dernière définition est aussi équivalente à celle de Stolz-Young.

---

(1) *Cf.* la note (2) de la page précédente.

(2) Le symbole $o(\varepsilon)$ signifiant ici, comme d'habitude, une quantité infiniment petite par rapport à ε.

Ainsi les trois dernières définitions sont équivalentes entre elles. Nous allons voir que la définition de M. Hadamard leur est aussi équivalente.

En divisant l'équation (5) par Δu et faisant tendre Δu vers zéro, on voit que si $x_1, \ldots, x_n$ sont des fonctions de u dérivables en u en un point u_0 où

$$x_1(u_0) = a_1, \quad \ldots, \quad x_n(u_0) = a_n,$$

l'égalité (3) a lieu : toute fonction différentiable au sens de Stolz-Young est différentiable au sens de M. Hadamard.

La réciproque de cette proposition bien connue est moins évidente : nous allons la démontrer.

Nous prendrons, pour simplifier les notations, le cas de $n = 2$, le cas général se traitant de même.

Nous supposons donc que $f(x, y)$ est différentiable au point (a, b) au sens de M. Hadamard. Il s'agit de montrer que la quantité

$$R = \Delta f(x, y) - f'_a \Delta x - f'_b \Delta y$$

est infiniment petite par rapport à $r = \sqrt{\Delta x^2 + \Delta y^2}$. Dans le cas contraire, il existerait un nombre $\varepsilon > 0$ et une suite de couples de valeurs h_n, k_n de Δx, Δy, tendant vers zéro avec $\frac{1}{n}$ et tels que $\left|\frac{R_n}{r_n}\right| > \varepsilon$, avec

$$R_n = l_n - h_n f'_a - k_n f'_b; \quad l_n = f(a + h_n, b + k_n) - f(a, b), \quad r_n = \sqrt{h_n^2 + k_n^2}.$$

Considérons deux fonctions $x(t)$, $y(t)$ dérivables pour $t = t_0$ et telles que

$$x(t_0) = a, \qquad y(t_0) = b.$$

Par hypothèse, la fonction

$$\varphi(t) = f[x(t), y(t)]$$

sera dérivable pour $t = t_0$ et l'on aura

$$\varphi'(t_0) = f'_a x'_{t_0} + f'_b y'_{t_0}.$$

Supposons qu'on puisse choisir $x(t)$, $y(t)$ de sorte que

$$x(t_n) = a + h_n, \qquad y(t_n) = b + k_n$$

avec $t_n - t_0 = r_n$. Alors on aura

$$x'_{t_0} = \lim_{t_n \to t_0} \frac{h_n}{t_n - t_0}, \qquad y'_{t_0} = \lim_{t_n \to t_0} \frac{k_n}{t_n - t_0},$$

d'où

$$\lim_{n \to \infty} \frac{R_n}{t_n - t_0} = \lim_{n \to \infty} \left\{ \frac{\varphi(t_n) - \varphi(t_0)}{t_n - t_0} - \frac{x(t_n) - a}{t_n - t_0} f'_a - \frac{y(t_n) - b}{t_n - t_0} f'_b \right\}$$
$$= \varphi'_{t_0} - f'_a x'_{t_0} - f'_b y'_{t_0} = 0.$$

Or,

$$\left| \frac{R_n}{t_n - t_0} \right| = \left| \frac{R_n}{r_n} \right| > \varepsilon.$$

On aboutit bien à la contradiction annoncée.

Il reste à établir l'existence de fonctions $x(t)$, $y(t)$ satisfaisant aux conditions indiquées. Il suffit de le faire en remplaçant la suite S des couples h_n, k_n par une suite s convenablement extraite de S. On peut d'abord, puisque h_n et k_n tendent vers zéro, extraire de S une première suite σ, telle que $r_n = \sqrt{h_n^2 + k_n^2}$ converge vers zéro en décroissant constamment et effectivement $\left(\text{la considération de } \frac{R_n}{r_n} \text{ n'a d'ailleurs de sens que si déjà pour S, } r_n \text{ restait constamment } \neq 0\right)$. Ceci fait, puisque le point P_n de coordonnées $\alpha_n = \frac{h_n}{r_n}$, $\beta_n = \frac{k_n}{r_n}$, reste sur le cercle $x^2 + y^2 = 1$, on peut extraire de σ une suite s telle que la suite des points P_n soit convergente, Soit (α, β) le point limite des P_n. La courbe $\Gamma : x = x(t)$, $y = y(t)$ a été assujettie à passer par le point (a, b) et par les points connus $x_n = a + k_n$, $y_n = b + k_n$, pour $t = t_0$, et $t = t_n = t_0 + r_n$. Prenons pour $x(t)$ et $y(t)$ des fonctions du premier degré de t_n à t_{n+1} quel que soit n. Alors Γ sera la ligne polygonale passant par les points (x_n, y_n).

Les fonctions $x(t)$, $y(t)$ sont continues pour $t_0 < t \leqq t_0 + r_1$. Elles sont même, pour $t = t_0$, continues et dérivables. Car, lorsque t est compris entre t_{n+1} et t_n,

$$x(t) - a, \quad y(t) - b, \quad \frac{x(t) - a}{t - t_0}, \quad \frac{y(t) - b}{t - t_0}$$

étant continues et homographiques restent comprises entre leurs valeurs respectives pour $t = t_{n+1}$ et $t = t_n$; or, ces quatre fonctions convergent vers des limites respectives, 0, 0, α, β quand t converge vers t_0 en prenant seulement les valeurs t_n.

On peut d'ailleurs supposer $x(t)$, $y(t)$ définis aussi pour $t < t_0$, en prenant

$$x(t_0 - \theta) - a = -[x(t_0 + \theta) - a], \qquad y(t_0 - \theta) - b = -[y(t_0 + \theta) - b]).$$

Ainsi, DANS L'ANALYSE CLASSIQUE, *la définition opératoire de la différentielle totale due à M. Hadamard est entièrement équivalente aux trois définitions équivalentes rappelées ci-dessus.* Et pourtant ces dernières ont trouvé leurs origines dans des ordres d'idées tout à fait différents de celui dont s'est inspiré M. Hadamard.

D'autres définitions plus générales pourraient être envisagées. On aurait pu penser à donner la définition suivante : on observe comme M. Hadamard, que la différentielle classique de $f(x, y)$ est une fonction déterminée de x, y, dx, dy, qui, lorsque x, y sont des fonctions de u, est la différentielle de la fonction de u qu'est alors $f(x, y)$. Mais, contrairement à ce que fait M. Hadamard, on ne se donne pas la forme de cette différentielle.

On dira alors que $f(x, y)$ est différentiable au point (a, b) s'il existe une fonction $g(x, y, dx, dy)$ — qui sera alors df — telle que pour $u = u_0$

$$\frac{d}{du} f[x(u), y(u)] = \frac{g[x(u), y(u), x'_u du, y'_u du]}{du},$$

pourvu que $x(u_0) = a$, $y(u_0) = b$ et que $x(u)$, $y(u)$ soient dérivables pour $u = u_0$.

Mais comme l'a fait observer M. Paul Lévy, une telle définition n'est pas suffisante, car une fonction différentielle à ce sens peut perdre d'importantes propriétés de la différentielle des fonctions simples et en particulier la propriété (3). Tel est, par exemple, le cas pour la fonction

$$f(x, y) = x\sqrt{\frac{y^2}{x^2+y^2}} \qquad \text{pour } x^2+y^2 \neq 0 \quad \text{avec} \quad f(0, 0) = 0.$$

Cette fonction est différentiable au sens précédent pour le point $a = 0$, $b = 0$, comme on le voit en prenant $g(x, y, dx, dy) \equiv f(dx, dy)$. Elle possède, d'autre part, en ce point, des dérivées partielles $f'_a = 0$, $f'_b = 0$. Mais il est bien clair que la propriété fondamentale (3) ne sera pas vérifiée par cette fonction.

Dans son très intéressant ouvrage sur l'Analyse fonctionnelle, M. Paul Lévy a proposé de corriger ce défaut en adoptant une définition de la différentielle d'une fonctionnelle qui dans le cas particulier d'une fonction numérique de deux variables numériques se réduit à la suivante :

Une fonction $f(x, y)$ est différentiable au point a, b, si:

1° L'expression

$$\rho = \frac{f(a + \lambda \Delta x, b + \lambda \Delta y) - f(a, b)}{\lambda}$$

a, quels que soient Δx, Δy, une limite quand λ tend vers zéro;

2° Cette limite est du premier degré et homogène en Δx, Δy. Alors cette limite est, par définition, la différentielle de f relative aux accroissements Δx, Δy.

On voit facilement qu'alors f'_a et f'_b existent et que la différentielle est égale à $f'_a \Delta x + f'_b \Delta y$.

Il est facile de voir que cette définition relève de la même objection que la précédente. On le voit par l'exemple suivant.

Considérons le cas où $a = b = 0$ et appelons $\mathcal{R}$ la région comprise, dans le plan des x, y, entre deux courbes γ_1, γ_2, chacune symétrique par rapport à l'origine et tangentes à Ox à l'origine : par exemple

$$y = x^3 \quad \text{et} \quad y = 3x^3.$$

Prenons $f(x, y) \equiv 0$ en dehors de la région $\mathcal{R}$ et $f(0, 0) = 0$. Il est clair que dans le cas actuel $f(a + \lambda \Delta x, b + \lambda \Delta y) - f(a, b)$, sans être nul quel que soit λ, est nul pour λ assez petit. Donc $\lim\limits_{\lambda \to 0} \rho$ existe et est nul : f est différentiable à l'origine au sens de M. Paul Lévy et sa différentielle est nulle.

Cependant la formule (3) n'est pas vérifiée nécessairement, comme on le voit en définissant convenablement la valeur de $f(x, y)$ quand x, y est dans la région $\mathcal{R}$. Rien n'empêche, par exemple, tout en maintenant partout la continuité de f, de prendre $f(x, y) = x$ quand le point (x, y) parcourt la courbe $y = 2x^3$ de $\mathcal{R}$. Alors en prenant sur cette courbe $x = u$, $y = 2u^3$, on aura pour $u = 0$,

$$\frac{d}{du} f[x(u), y(u)] = \frac{dx}{du} = 1 \quad \text{et non pas} \quad \frac{df}{du} = 0.$$

Il en résulte que la définition de la différentielle totale due à M. Paul

Lévy, est plus générale que les quatre définitions équivalentes examinées d'abord.

La généralisation naturelle d'un point anguleux d'une courbe plane est un point d'une surface où les tangentes forment un cône ou bien un point d'une ligne double avec deux demi-plans tangents.

La différentielle de M. Paul Lévy n'est donc pas assimilable à une généralisation de la dérivée à droite ou à gauche. Et parmi les surfaces les plus communément étudiées, on en trouvera beaucoup possédant des points singuliers de la nature spécifiée ci-dessus, mais aucune qui soit différentielle au sens de M. Lévy sans l'être au sens de Stolz-Young. En résumé, la plus grande généralité de la différentielle de M. Lévy ne paraît guère susceptible d'applications dans l'Analyse classique. Elle pourrait, par contre, rendre des services dans la théorie des ensembles et particulièrement dans cette théorie des « contingents et paratingents » dont l'utilité a été signalée d'abord par M. Beppo Levi, puis par M. Séveri, mais dont on doit à M. Bouligand et à ses élèves d'avoir entrepris l'étude systématique. Enfin, M. Lévy a bien voulu nous signaler lui-même l'intérêt qu'il y a aussi « à partir d'une définition trop générale, pour montrer qu'elle n'a toutes les propriétés qu'on attend d'elle que si l'on introduit des conditions restrictives ».

2. Analyse fonctionnelle. — C'est M. Volterra qui a eu le premier l'idée d'étendre le champ d'application du Calcul différentiel à l'Analyse fonctionnelle. Les définitions de la dérivée fonctionnelle lui ont permis de traiter un grand nombre d'applications importantes.

Toutefois M. Hadamard a signalé qu'il y aurait grand intérêt à généraliser les définitions de M. Volterra. Celles-ci laissent échapper, en effet (et M. Volterra, lui-même, l'avait le premier indiqué) des cas simples où la généralisation de la notion de différentielle ne devrait pas souffrir de difficultés. Tel est le cas où l'on a à considérer une fonctionnelle linéaire, $U[f]$, c'est-à-dire une fonctionnelle continue et distributive, cas où l'on posera naturellement

$$dU[f] = U[df].$$

M. Hadamard a montré le chemin qui devait conduire vers des définitions satisfaisantes en proposant d'imposer à la différentielle d'une fonctionnelle la condition d'être *linéaire* par rapport à la différentielle de l'argument.

Cette condition ne suffit d'ailleurs pas pour préciser la définition de la différentielle. Nous avons donné en 1911, une définition précise inspirée par l'idée intuitive que la différentielle d'une fonctionnelle doit fournir une représentation simple et approchée de l'accroissement de cette fonctionnelle. Le mot simple se trouvera précisé par la condition de linéarité de M. Hadamard. Le mot approchée signifiera que la différentielle est une sorte de partie principale de l'accroissement de la fonctionnelle, en ce sens que *leur différence* $\Delta U - dU$ *devra être infiniment petite par rapport à l'accroissement* Δf *de l'argument* f. Comme Δf n'est pas un nombre, on pourra préciser que $|\Delta U - dU|$ doit être infiniment petit par rapport à nombre mesurant la « distance » de l'argument fixe f_0 et de l'argument varié $f_0 + \Delta f$. (Selon le champ fonctionnel considéré cette distance pourra avoir telle ou telle expression.)

On voit immédiatement que, d'après cette définition, si U est linéaire, on a bien

$$dU[f] = U[\Delta f],$$

et que si U est quelconque mais différentiable à ce sens, sa différentielle est unique.

Plus tard, M. Paul Lévy a proposé à la page 52 de son livre, une définition plus générale. A son sens, une fonctionnelle $U[f]$ est différentiable pour l'argument f_0 si :

1° Le rapport $\dfrac{U[f_0 + \lambda\Delta f] - U[f_0]}{\lambda}$ a, pour chaque accroissement Δf, une limite $W[\Delta f, f_0]$.

2° Cette limite est une fonctionnelle linéaire de l'accroissement Δf de l'argument. Et c'est, par définition, la différentielle de la fonctionnelle

$$dU[f] = W[\Delta f, f_0].$$

En appliquant au cas où $f(t)$ dépend linéairement de deux paramètres x et y [par exemple, $f(t) = xt + y$], $U[f]$ devient une fonction de x et y, $F(x, y)$, on a $\Delta f = t\Delta x + \Delta y$; W prend la forme

$$\Delta x W[t, f_0] + \Delta y W[t, f_0] = \Delta x A(x_0, y_0) + \Delta y B(x_0, y_0)$$

et l'on retombe bien sur la définition donnée plus haut pour la différentielle d'une fonction de deux variables numériques au sens de M. Lévy.

On a vu que, dans ce cas beaucoup plus simple, la définition de M. Lévy n'offrirait d'avantage que pour l'étude de fonctions très singulières. *Dans le cas plus complexe de l'Analyse fonctionnelle, la définition de M. Lévy* — qui n'autorise pas l'application des formules les plus fondamentales du Calcul différentiel, telles que la formule (3) — *paraît être d'une généralité prématurée alors que l'étude des fonctionnelles les plus simples est encore dans l'enfance.*

Il n'est pourtant pas impossible de concevoir une définition de la différentielle d'une fonctionnelle plus générale que celle que nous avons donnée en 1911 mais moins générale que celle de M. Lévy, et qui vaudrait la peine d'être étudiée en vue de l'application à des espaces fonctionnels plus généraux que ceux où s'applique notre définition.

Reportons-nous, en effet, à la définition de la différentielle d'une fonction de deux variables numériques donnée par M. Hadamard et rappelée plus haut. M. Glivenko, qui a publié un article pour en signaler l'intérêt, a bien voulu nous exprimer, dans un entretien privé, le sentiment qu'il serait souhaitable d'appliquer à l'Analyse fonctionnelle une définition opératoire de la différentielle basée sur les mêmes principes que cette définition donnée par M. Hadamard pour les fonctions de l'Analyse classique.

En précisant cette dernière sous la forme que nous avons indiquée plus haut, il faudrait, pour la généraliser, généraliser d'abord la notion de dérivée partielle. On peut éviter cette nécessité en exprimant la définition primitive de M. Hadamard sous une forme équivalente mais plus apte à la généralisation. A cet effet, il suffit de s'inspirer précisément de la condition de linéarité recommandée par M. Hadamard pour l'Analyse fonctionnelle.

Nous dirons qu'une fonction ordinaire $f(x, y)$ (fonction numérique de deux variables numériques) est différentiable au point a, b au sens de M. Hadamard, s'il existe une expression linéaire en dx et dy, soit $\mathrm{A}\,dx + \mathrm{B}\,dy$, telle que si x, y sont deux fonctions de u égales en u_0 à a et b et dérivables en ce point, $f[x(u), y(u)]$ soit aussi dérivable en u_0 et qu'on ait, pour $u = u_0$,

$$\frac{d}{du} f[x(u), y(u)] = A x'_u + B y'_u.$$

En prenant $u_0 = a$, $x(u) = u$ et $y(u) = b$, on voit que f'_a existe et est égal à A; de même $B = f'_b$. De sorte que cette définition est bien équivalente à celle donnée plus haut.

La généralisation se présente alors d'elle-même :

Une fonctionnelle $U[f]$ *sera dite différentiable pour* $f \equiv f_0$ *au sens de M. Hadamard généralisé, s'il existe une fonctionnelle* $W[df, f_0]$, *linéaire par rapport à* df, *telle que si l'on considère une fonction* $f(t, \lambda)$ *dérivable par rapport à* λ *pour* $\lambda = 0$, *avec* $f(t, 0) = f_0(t)$, *la fonction de* λ, $U[f(t, \lambda)]$ *soit dérivable en* λ *pour* $\lambda = 0$ *et qu'on ait pour* $\lambda = 0$

$$\frac{d}{d\lambda} U[f(t, \lambda)] = W\left[\frac{df}{d\lambda}, f_0\right] \tag{4}$$

ou avec les notations des « variations »

$$\delta U[f] = W[\delta f, f_0]. \tag{4 bis}$$

Cette définition appelle quelques observations. Tout d'abord, la condition de continuité imposée à W suppose qu'on a affaire à un champ fonctionnel déterminé, $\mathcal{C}$, où la limite d'une suite de fonctions a une signification déterminée. (Suivant les cas, on supposera qu'il y a convergence uniforme, en moyenne quadratique, etc.). D'autre part les fonctions $f(t, \lambda)$ devant appartenir nécessairement à $\mathcal{C}$ pour toute valeur de λ, il faudra, en outre, supposer que $f'_\lambda(t, 0)$ appartient aussi à $\mathcal{C}$. Il suffira pour cela de supposer que la convergence de $\frac{f(t, \lambda) - f(t, 0)}{\lambda}$ vers $f'_\lambda(t, 0)$ soit de la nature admise dans $\mathcal{C}$ (1).

Observons qu'il résulte des diverses définitions données qu'une fonctionnelle différentiable au sens de M. Hadamard généralisé est différentiable au sens de M. Lévy et a la même différentielle. D'autre part, il a été déjà démontré ailleurs, et, au surplus, on voit facilement

(1) Nous croyons que l'extension à l'Analyse fonctionnelle ainsi formulée complètement, de la définition de M. Hadamard donnée pour les fonctions ordinaires, n'avait été jusqu'à maintenant ni étudiée, ni même énoncée.

qu'une différentielle à notre sens vérifie la condition (4), c'est-à-dire qu'une fonctionnelle différentiable à notre sens est différentiable au sens de M. Hadamard généralisé et a la même différentielle.

L'exemple donné plus haut dans le cas des fonctions ordinaires montre que la différentielle de M. Lévy est effectivement plus générale que celle de M. Hadamard.

La différentielle au sens de M. Hadamard généralisé est-elle plus générale que la nôtre? Nous avons démontré plus haut que ces deux définitions sont équivalentes quand on les applique à des fonctions ordinaires. Par contre, nous allons voir qu'elles cessent de l'être quand il s'agit de fonctionnelles. Dès lors sera établi que : *la différentielle au sens de M. Hadamard généralisé qui est équivalente à la nôtre dans l'Analyse classique est plus générale dans l'Analyse fonctionnelle.*

Pour établir ce point, nous nous adresserons à un exemple déjà considéré dans son livre par M. Lévy dans un but analogue.

Exemple. — Prenons pour champ fonctionnel $\mathcal{C}$, celui des fonctions $f(t)$ *continues* sur (a, b), f étant considéré comme limite de f_n quand $f_n(t)$ converge *uniformément* vers $f(t)$ sur (a, b). Prenons d'autre part pour fonctionnelle le maximum $M[f]$ de $f(x)$ sur (a, b). C'est une fonctionnelle bien déterminée sur $\mathcal{C}$. On voit facilement qu'elle y est continue. Car, en désignant par (f, g) le maximum de $|f(t) - g(t)|$ sur (a, b) on a :

$$f = g + f - g \leqq M[g] + (f, g),$$

d'où $M[f] - M[g] \leqq (f, g)$ et, par suite, en permutant f, g :

$$|M[f] - M[g]| \leqq (f, g) = M[|f - g|]. \tag{5}$$

Cette fonctionnelle $M[f]$ est-elle différentiable? Si, pour l'argument $f_0(t)$, elle a une différentielle — que nous pourrons dénoter $\partial_{df} M[f]$ — alors, que ce soit au sens de M. Lévy ou au sens de M. Hadamard généralisé, ou au nôtre, cette différentielle devra être égale à la limite, quand $\lambda \to 0$, de

$$r = \frac{M[f_0 + \lambda\varphi] - M[f_0]}{\lambda}$$

quand elle est relative à un accroissement $\varphi(x)$ de f ou à la « variation » $\delta f = \varphi(t)\,\delta\lambda$ de f.

Le maximum de $f_0(t)$ est atteint en au moins un point c de (a, b). Et l'on a

$$\begin{aligned} o &\leqq M[f_0(t) + \lambda\varphi(t) - f_0(c) - \lambda\varphi(c)] \\ &= M[f_0(t) + \lambda\varphi(t)] - M[f_0] - \lambda\varphi(c) = \lambda[r - \varphi(c)]; \end{aligned}$$

d'où, si $\lambda > o$, $r \geqq \varphi(c)$; si $\lambda < o$, $r \geqq \varphi(c)$. Or, si $M[f]$ a une différentielle, r doit avoir une limite unique, cette limite doit donc être à la fois $\geqq \varphi(c)$; $\leqq \varphi(c)$; elle doit donc être égale à $\varphi(c)$. Cette expression est d'ailleurs admissible, puisque, sur le champ $\mathcal{C}$, c'est une fonctionnelle linéaire de $\varphi(t)$.

Mais si la fonction $f_0(t)$ atteint son maximum en au moins deux points distincts c, c_1, on devrait avoir $\varphi(c) = \varphi(c_1)$ quelle que soit la fonction continue $\varphi(x)$, ce qui est impossible.

En résumé, $M[f]$ est une fonctionnelle continue qui, pour un argument $f_0(t)$ atteignant son maximum en plus d'un point, ne saurait avoir de différentielle en aucun des trois sens indiqués.

Voyons si la différentielle existe quand f_0 n'atteint qu'en un point c son maximum. Si elle existe, elle est égale à $\varphi(c)$; il faut donc au moins prouver que la quantité $\rho = r - \varphi(c)$ tend vers zéro avec λ. Or, on a vu que $\lambda[r - \varphi(c)] \geqq o$. Donc

$$\rho\lambda \geqq o. \tag{6}$$

D'autre part, en posant $\operatorname{sgn}\lambda = \dfrac{\lambda}{|\lambda|}$ pour $\lambda \neq o$, on a

$$\rho \operatorname{sgn}\lambda \geqq o \tag{6 bis}$$

et, d'autre part,

$$\rho \operatorname{sgn}\lambda = \frac{M[f_0 + \lambda\varphi] - f_0(c) - \lambda\varphi(c)}{|\lambda|} = \frac{M\{f_0(t) - f_0(c) + \lambda[\varphi(t) - \varphi(c)]\}}{|\lambda|}$$

d'où

$$\rho \operatorname{sgn}\lambda = M[\theta(t, \lambda)] \tag{7}$$

avec

$$\theta(t, \lambda) = \frac{f_0(t) - f_0(c)}{|\lambda|} + (\operatorname{sgn}\lambda)[\varphi(t) - \varphi(c)].$$

Appelons $\mu_1(\eta)$ le maximum de $\theta(t, \lambda)$ sur le segment $(c-\eta, c+\eta)$ et $\mu_2(\eta)$ son maximum sur (a, b) en dehors de l'intervalle $(c-\eta, c+\eta)$. Il est clair que $M[\theta(t, \lambda)]$ est égal soit à $\mu_1(\eta)$; soit à $\mu_2(\eta)$. D'autre part, d'après (6 *bis*) et (7), $M[\theta(t, \lambda)] \geqq 0$. Or, *puisque* $f_0(t)$ *n'atteint son maximum, qu'en un point*, c, $f_0(c)-f_0(t)$ a un minimum positif $m(\eta)$ en dehors de $c-\eta$, $c+\eta$ et $\frac{f_0(t)-f_0(c)}{|\lambda|} \leqq \frac{-m(\eta)}{|\lambda|}$. Soit, d'autre part, A le maximum de $|\varphi(t)-\varphi(c)|$ sur tout (a, b). On aura :

$$\mu_2(\eta) \leqq A - \frac{m(\eta)}{|\lambda|}.$$

On a évidemment $\mu_1(\eta) \leqq m'(\eta)$, en désignant par $m'(\eta)$ le maximum de $|\varphi(t)-\varphi(c)|$ sur $(c-\eta, c+\eta)$. On peut prendre η assez petit pour que $m'(\eta)$ soit plus petit qu'un nombre positif ω donné d'avance η étant ainsi fixé, on voit que, pour $|\lambda|$ assez petit, $\mu_2(\eta)$ sera $\leqq 0$, et par suite $M[\theta(t, \lambda)] < \omega$. On a vu que $M[\theta(t, \lambda)] \geqq 0$. Dès lors, on a bien $\lim_{\lambda \to 0} \rho = 0$.

Finalement, nous venons de démontrer ces deux résultats, prouvés assez brièvement par M. Lévy dans son livre (p. 52):

$M[f]$ n'est pas différentiable pour l'argument f_0 si $f_0(t)$ atteint son maximum en plus d'un point.

$M[f]$ est différentiable au sens de M. Lévy pour l'argument f_0 si $f_0(t)$ atteint son maximum en un seul point c; et pour l'accroissement $\varphi(t)$ de f_0, cette différentielle est égale à $\varphi(c)$. En notation des « variations »

$$\delta M[f] = \delta f(c).$$

Nous allons maintenant prouver deux résultats que nous croyons nouveaux, mais dont l'importance réside surtout dans la conséquence annoncée plus haut en italiques (p. 245).

1° *Le maximum* $M[f(t)]$ *d'une fonction* $f(t)$ *continue sur* (a, b) *est une fonctionnelle différentiable au sens de M. Hadamard généralisé* (définition, p. 243 et 244) *pour tout argument* $f_0(t)$ *n'atteignant son maximum qu'en un point.*

2° *Cette fonctionnelle* $M[f(t)]$ *n'est différentiable à notre sens* (définition, p. 241 et 242) *pour aucun argument* $f_0(t)$.

Pour prouver le premier point, il suffit de prouver que

$$\lim_{\lambda \to 0} \left\{ \frac{M[f(t, \lambda)] - M[f_0(t)]}{\lambda} - f'_\lambda(c, o) \right\} = o,$$

pourvu que le rapport $\frac{f(t, \lambda) - f_0(t)}{\lambda}$, considéré comme fonction de t appartenant au champ $\mathcal{C}$, converge vers $f'_\lambda(t, o)$ au sens admis dans $\mathcal{C}$, c'est-à-dire avec convergence uniforme quand t parcourt (a, b). Or, cette hypothèse permet de ramener au cas précédemment examiné, — où $f(t, \lambda)$ était du premier degré en λ. En posant pour simplifier l'écriture $f'_\lambda(t, o) = \varphi(t)$, on peut écrire

$$\frac{f(t, \lambda) - f_0(t)}{\lambda} - \varphi(t) = \varepsilon(t, \lambda)$$

et, par hypothèse, la quantité

$$\omega(\lambda) = M[|\varepsilon(t, \lambda)|]$$

tend vers zéro avec λ. Or, en vertu de (5), on a

$$(8) \qquad |M[f(t, \lambda)] - M[f_0(t) + \lambda\varphi(t)]| \leqq |\lambda| \omega(\lambda).$$

Dès lors, la quantité

$$\begin{aligned} Q(\lambda) &\equiv \frac{M[f(t, \lambda)] - M[f_0(t)]}{\lambda} - f'_\lambda(c, o) \\ &\equiv \left\{ \frac{M[f(t, \lambda)] - M[f_0(t) + \lambda\varphi(t)]}{\lambda} \right\} \\ &\quad + \left\{ \frac{M[f_0(t) + \lambda\varphi(t)] - M[f_0(t)]}{\lambda} - \varphi(c) \right\} \end{aligned}$$

est la somme de deux termes dont le premier, en vertu de (8), est au plus égal à $\omega(\lambda)$, en valeur absolue, et dont le second, comme nous l'avons prouvé plus haut, converge vers zéro avec λ quand $f_0(t)$ n'atteint son maximum qu'au point c. De sorte que $Q(\lambda)$ tend vers zéro comme il était annoncé.

Quant au second point, il faut montrer que si $\mu = M|\psi|$ où $\psi(t)$ est une fonction continue quelconque, non identiquement nulle, on n'a pas toujours

$$\lim_{\mu \to 0} \left[\frac{M[f_0 + \psi] - M[f_0] - \psi(c)}{\mu} \right] = o.$$

Or, la quantité entre crochets peut s'écrire :

$$\alpha = \frac{M[f_0(t)+\psi(t)-f_0(c)-\psi(c)]}{\mu} = M\left\{\frac{\psi(t)-\psi(c)}{\mu} - \frac{f_0(c)-f_0(t)}{\mu}\right\}.$$

On a évidemment $\alpha \geqq 0$.

L'énoncé 2° résultant de ce qui précède quand $f_0(t)$ atteint son maximum en plus d'un point, il suffit de l'établir dans le cas contraire. Dans ce cas, pour tout nombre $\eta > 0$, le maximum de $\sqrt{f(c)-f(t)}$ pour $|t-c| \leqq \eta$ est un nombre $\xi(\eta) > 0$. Prenons alors pour $\psi(t)$ une fonction continue égale à $\sqrt{f(c)-(t)}$ pour $|t-c| \leqq \eta$ et constante en dehors de l'intervalle $(c-\eta, c+\eta)$. Alors $\psi(c)=0$ et μ est égal à $\xi(\eta)$, et par suite $\neq 0$. Il est clair que $\xi(\eta)$ tend vers zéro avec η. Soit c_1 le point où ce maximum de $\psi(t)$ est atteint : c_1 est dans $(c-\eta, c+\eta)$; on voit qu'on aura :

$$\alpha \geqq \frac{\psi(c_1)-\psi(c)}{\mu} - \frac{f_0(c)-f_0(c_1)}{\mu} = 1 - \frac{f_0(c)-f_0(c_1)}{\xi(\eta)} > 1-\xi(\eta).$$

Quand η tend vers zéro, le dernier membre tend vers un et μ tend vers zéro; α ne peut donc, dans ce cas, tendre vers zéro avec μ.

3. Analyse générale. — L'intérêt de la définition de M. Hadamard n'est pas épuisé par son utilisation en Analyse fonctionnelle. Il est peut-être plus encore dans la possibilité de son extension en Analyse générale.

Dans ce domaine, on peut généraliser la notion de fonctionnelle et considérer des transformations $X = F[x]$ d'un élément abstrait x en un élément abstrait X. Nous avons pu en 1925 (1) étendre notre définition (rappelée plus haut p. 241 et 242) de la différentielle d'une fonctionnelle, définir la différentielle de $F[x]$ quand X et x appartiennent à des espaces « vectoriels abstraits distanciés » et en établir les propriétés les plus importantes.

La définition au sens de M. Hadamard généralisé présente sur notre

(1) *La notion de différentielle dans l'Analyse générale* (*Ann. Ec. Norm.*, t. XLII, 1925, p. 293-323).

définition l'avantage de garder un sens pour des espaces abstraits vectoriels non distanciés où notre définition ne s'applique pas (ou, tout au moins, ne s'applique pas littéralement).

Il reste à voir si elle conserve les propriétés les plus importantes de la différentielle classique en dehors de la propriété essentielle (généralisant le théorème des fonctions composées) qui lui sert de définition. C'est un point sur lequel nous reviendrons ultérieurement.