Commentationes Mathematicae Universitatis Carolinae

Zuhair M. Nashed

Higher order differentiability of nonlinear operators on normed spaces. I.

Commentationes Mathematicae Universitatis Carolinae, Vol. 10 (1969), No. 3, 509--533

Persistent URL: `http://dml.cz/dmlcz/105248`

Commentationes Mathematicae Universitatis Carolinae
10, 3 (1969)

HIGHER ORDER DIFFERENTIABILITY OF NONLINEAR OPERATORS ON NORMED SPACES I [x)]

M.Z. NASHED, Beirut

1. Introduction. This paper is concerned with several notions of higher order differentiability in the theory of nonlinear operators on real normed spaces. We broaden the concepts of higher order differentials of Gâteaux and Fréchet and introduce new variants of such differentials. We also study higher order strong differentials, Hadamard, Peano and Taylor variations and differentials. Various implication relationships among these variations and differentials are obtained. Sufficient conditions for the existence of these differentials are also established.

A differential of order m may be defined in two ways. We may define it directly without reference to lower order differentials, or we may define it inductively assuming the existence of differentials of order less than m. These two approaches may lead to different notions. We shall distinguish between eight notions of Fréchet (Gâteaux) differentials of order m and establish continuity implications of these notions.

----

x) Communicated at the seminar on Nonlinear Functional Analysis at the Mathematical Institute of Charles University on June 22,1968.

We remark that for the differentials considered in this paper, the existence of a differential of order m implies the existence of all lower order differentials. This property is not possessed by all definitions of differentials (for instance, Riemann differentials and difference-differentials [1]).

For an exposition of calculus in normed spaces, we refer to Dieudonné [2], Liusternik and Sobolev [3], Vainberg [4], Kantorovich and Akilov [5], Michal [6], Nashed [7] and Rall [8], and to the older work of Graves and Hildebrandt [9-11], Fréchet [12,13] , Gâteaux [14-16], Levy [17] and Kerner [18-20].

All the limits in the definitions of differentials in this paper are taken in the sense of the norm, and not in the sense of the weak topology. The latter leads to notions of weak differentials.

2. Multilinear operators and differential forms.

The study of higher order Fréchet and Gâteaux differentials essentially involves the approximation in various senses of the difference $f(x_0+h)-f(x_0)$ by abstract polynomials. This leads to consideration of various notions of continuity of mappings from a subset of a normed linear space to a space of multilinear operators, which we shall discuss in this section.

2.1. Let $E_1, E_2, \dots, E_m$ and Y be normed real linear spaces and let Π denote the product space $E_1 \times E_2 \times \dots \times E_m$ equipped with the usual product topology induced by the norms on E_i , $i = 1, 2, \dots, m$. We write

$x = (x_1, \dots, x_m) \in \Pi$ as $x_1 \dots x_m$ and let $\|x\| = \sup \|x_i\|$. We recall that an operator $L_m : \Pi \to Y$ is called multilinear or m-linear if it is separately additive and homogeneous in each of the variables. For a multilinear operator L_m and for each $x_1, \dots, x_{k-1}$, $x_{k+1}, \dots, x_m$ we designate by S_k the linear operator $x_k \to L_m x_1 x_2 \dots x_m$ of the space E_k into Y. We say that L_m is separately continuous if for each $x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_m$, where $x_k \in E_k$, the operator S_k is continuous for $k = 1, \dots, m$. L_m is uniformly separately continuous (bounded) in x_k if the linear operator S_k is continuous (bounded), uniformly on the set $\|x_1\| = \dots = \|x_{k-1}\| = \|x_{k+1}\| = \dots = \|x_m\| = 1$. Finally by a continuous multilinear operator we mean a multilinear operator which is jointly continuous in all the variables, that is, continuous on the product space Π. We shall need certain implications among these notions and characterizations of continuous multilinear operators which are stated in the following:

Theorem 1. Let L_m be a multilinear operator on Π into Y. Then the following implications hold among the following statements:

$a \Rightarrow b \Rightarrow c \Rightarrow d \Rightarrow e \Rightarrow f$,

$f \Rightarrow a$ if the space Π is complete ;

(a) L_m is continuous on Π.

(b) L_m is continuous at the point zero $\theta_1 \dots \theta_m$.

(c) L_m is bounded on each bounded set of Π .

(d) L_m is bounded on Π , that is there exists a constant $M \geq 0$ such that for all $x_1 \dots x_m \in \Pi$,

(1) $\| L_m x_1 \dots x_m \| \leq M \| x_1 \| \dots \| x_m \|$.;

(e) L_m is uniformly separately continuous (bounded)in x_k for some fixed k .

(f) L_m is separately continuous (bounded).

We omit the proof since the various implications of the theorem are either given in the literature or involve simple manipulations. The most important proposition in the theorem is the equivalence of statements (a) and (f) when the space Π is complete. For an exposition of continuous multilinear operators on normed spaces we refer to Dieudonné [2] and Hille and Phillips [21]. For computational aspects of such operators see Rall [8]. Bourbaki [22] and Dieudonné [23] give an account of the theory of multilinear operators on topological vector spaces, where it may be noted that (f) does not imply (a) even for locally convex topological vector spaces. However, the implication holds under weaker conditions than stated here, for example in the case of a bilinear operator which is separately continuous it suffices to take E_1 to be a metrizable barelled vector space, E_2 a metrizable vector space and Y any locally convex space, so that in particular joint continuity follows if the spaces E_i are Fréchet spaces. In this connection we also note that Bourbaki defines a notion of hypercontinuity (and equihypercontinuity) for bilinear operators which is intermediate between the notions of separately continuous and continuous operators. For normed spaces, this notion is equivalent to any of the statements of Theorem 1 if Π is a Banach space. These notions become important for the theory of higher order differentiability in vector spaces without a norm [24], which require modifications to be examined elsewhere. The general theory of continuous multilinear operators is closely connected with topological tensor products which are surveyed in [23].

Let $\mathcal{L}_m$ denote the vector space of all continuous multilinear operators on the product space Π into Y. The greatest lower bound of all constants M satisfying the inequality (1) is given by

$$\| L_m \| = \sup \{ \| L_m x_1 \dots x_m \| : \| x_i \| = 1,\ i = 1, \dots, m \}$$

and is a bonafide norm on the space $\mathcal{L}_m$. Furthermore the space $\mathcal{L}_m$ equipped with this norm is a Banach space if and only if Y is complete.

The inductive definitions of higher order Gâteaux and Fréchet differentials lead to consideration of the space $\mathcal{L}(E_1, \mathcal{L}(E_2, \dots, \mathcal{L}(E_m, Y), \dots))$, where $\mathcal{L}(X, Y)$ denotes the space of all continuous linear operators on X into Y. There is a canonical isometric isomorphism which identifies the space $\mathcal{L}_m(E_1, \dots, E_m; Y)$ with the space $\mathcal{L}(E_1, \mathcal{L}(E_2, \dots, \mathcal{L}(E_m, Y), \dots))$. In the sequel we shall identify the corresponding elements of these spaces under this linear isometry.

Now let $E_1 = E_2 = \dots = E_m = E$. Let σ be a permutation of the set $\{1, 2, \dots, m\}$ and consider x : $:\{1, 2, \dots, m\} \to \Pi$. σ induces a linear transformation $P_\sigma : \mathcal{L}_m(\Pi; Y) \to \mathcal{L}_m(\Pi; Y)$ defined by $(P_\sigma L_m)x = L_m(x \circ \sigma)$, where $x \in \Pi$. The mean of an m-linear operator is defined by $\overline{L}_m x_1 \dots x_m =$ $= \frac{1}{m!} \sum_{\sigma} (P_\sigma L_m) x$, where the sum is taken over all permutations of $\{1, 2, \dots, m\}$. An m -linear operator is said to be symmetric if $E_1 = \dots = E_m$ and $P_\sigma L_m = L_m$ for every permutation σ. The mean is always symmetric; a symmetric m -linear operator coincides with its mean. Note that $\| L_m \| = \| P_\sigma L_m \|$ for every σ.

This notion of symmetry should be distinguished from the weaker notion of symmetric m -linear operator used in exterior algebra where L_m is said to be symmetric if $\sum_{\sigma} s_\sigma P_\sigma L_m = 0$, where $s_\sigma = 1$ if σ is even and $s_\sigma = -1$ if σ is odd.

2.2. The study of differentials of order m leads to mappings from a subset X of a normed linear space to the space $\mathcal{L}_m (E_1 \times \dots \times E_m ; Y)$. We write $D: X \to \longrightarrow \mathcal{L}_m$ and call $D(x; \dots)$ a formal differential operator. $DF(x; h_1, \dots, h_m)$ is called a formal differential form. Several notions of continuity and directional continuity may be defined for $D(x; \dots)$ depending on the topology used for $\mathcal{L}_m$.

<u>Definition 1</u>. A differential operator is said to be <u>pointwise continuous</u> in x at x_0 if for any $h_1 \dots h_m \in \in \Pi$, $\| D(x; h_1 \dots h_m) - D(x_0; h_1 \dots h_m) \|_Y \to 0$ whenever $\| x - x_0 \| \longrightarrow 0$. This is equivalent to considering the space $\mathcal{L}_m$ in the topology of pointwise convergence.

<u>Definition 2.</u> A differential operator is said to be <u>jointly continuous</u> in the variables $x, h_1, \dots, h_m$ at x_0 if for any $x \in X$, $k_1 \dots k_m \in \Pi$, $\| D(x; h_1 \dots h_m) - D(x_0; k_1 \dots k_m) \|_Y \to 0$ whenever

$\| x - x_0 \| \longrightarrow 0$ and $\| h_i - k_i \| \longrightarrow 0$ for $i = 1, \ldots, m$. This is joint continuity in the classical sense.

Definition 3. A differential operator is said to be continuous in x at x_0 if it is continuous as a transformation from the set X to the space $\mathcal{L}_m$:

$$\| D(x; \ldots) - D(x_0; \ldots) \|_{\mathcal{L}_m} \longrightarrow 0 \quad \text{whenever}$$

$\| x - x_0 \| \longrightarrow 0$. This is the same as considering the space $\mathcal{L}_m$ in the topology of uniform convergence on bounded sets.

We remark that Definitions 1 and 3 are implicit in the work of Kerner [19] and Graves and Hildebrandt [11] respectively.

Notions of directional continuity may be defined similarly. For example,

Definition 2d. A differential operator is jointly directionally continuous at x_0 if for any fixed y,

$$\| D(x_0 + ty; h_1 \ldots h_m) - D(x_0; k_1 \ldots k_m) \|_Y \to 0 \quad \text{whenever}$$

$t \to 0$ and $\| h_i - k_i \| \longrightarrow 0$ for $i = 1, \ldots, m$.

Definitions 1d and 3d will denote the analogs of Definitions 1 and 3 for directional continuity.

Notions of weak continuity, demicontinuity and hemicontinuity of differential operators may be defined analogously and are useful in consideration of weak differentials of higher order as will be shown in [25]. (See Remark 1 p. 35 in Vainberg [4] and the recent work of Kolomý [26,27] and Zizler [28] for first order weak differentials.) The implication relationships among these notions are stated in the following:

<u>Theorem 2</u>. (a) A differential operator $D(x; h_1 \dots$ $\dots h_m)$ is continuous at x_0 if and only if it is pointwise continuous at x_0 uniformly with respect to $h_1 \dots h_m$ on the set $\|h_1\| = \dots = \|h_m\| = 1$.
(b) A differential operator is pointwise continuous at x_0 if and only if it is jointly continuous.
Furthermore, the continuity of $D(x; h_1 \dots h_m)$ at x_0 in the sense of any of these definitions implies the existence af a positive number M and a neighborhood $N(x_0)$ of x_0 such that for all $x \in N(x_0)$,

$$(2) \qquad \|D(x; h_1 \dots h_m)\|_Y \leq M \|h_1\| \dots \|h_m\| .$$

Similar implication relations hold among Definitions 1d, 2d and 3d for directional continuity. The directional continuity of $D(x; h_1 \dots h_m)$ at x_0 in the sense of any these definitions implies the existence for fixed y of a positive number M and a neighborhood $N(0)$ such that for all $\tau \in N(0)$,

$$(3) \qquad \|D(x_0 + \tau y; h_1 \dots h_m)\|_Y \leq M \|h_1\| \dots \|h_m\| .$$

<u>Proof</u>. (a) Let D be continuous at x_0. Then given $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|x - x_0\|$ implies

$$\|D(x; \dots) - D(x_0; \dots)\|_{\mathcal{L}_m(E_1 \times \dots \times E_m; Y)} < \varepsilon \text{ , or}$$

$$\sup\{\|D(x; h_1 \dots h_m) - D(x_0; h_1 \dots h_m)\| : \|h_1\| = \dots \|h_m\| = 1\} < \varepsilon .$$

Thus

$\|x - x_0\| < \delta$ implies $\|D(x; h_1 \dots h_m) - D(x_0; h_1 \dots h_m)\| < \varepsilon$ uniformly with respect to $h_1 \dots h_m$ on the set $\|h_1\| = \dots = \|h_m\| = 1$. Conversely if D is pointwise continuous at x_0 uniformly with respect to $h_1 \dots h_m$ on the set $\|h_1\| = \dots = \|h_m\| = 1$, then for each $\varepsilon > 0$, there exists a $\delta(\varepsilon) > 0$ which is independent of $h_1 \dots h_m$ such that

$\|D(x; h_1 \dots h_m) - D(x_0; h_1 \dots h_m)\| < \varepsilon$ for $\|h_1\| = \dots \|h_m\| = 1$ and $\|x - x_0\| < \delta$. Thus $\|D(x; \dots) - D(x_0; \dots)\| < \varepsilon$.

(b) Let D be pointwise continuous at x_0. Let $x \longrightarrow x_0$ and $k_i \longrightarrow h_i$ for $i = 1, \dots, m$. Then

$$\|D(x; k_1 \dots k_m) - D(x_0; h_1 \dots h_m)\| \leq$$
$$\leq \|D(x; k_1 \dots k_m) - D(x_0; k_1 \dots k_m)\| +$$
$$+ \|D(x_0; k_1 \dots k_m) - D(x_0; h_1 \dots h_m)\| .$$

The first term on the right goes to zero by pointwise continuity of D; the second goes to zero since D is a continuous multilinear operator. The converse is trivial. Inequality (2) follows easily using the triangle inequality and part (d) of Theorem 1.

The proof of the implications among the notions of directional continuities is essentially the same.

2.3. The form $L_m x \dots x$ obtained from a symmetric m-linear operator $L_m x_1 \dots x_m$ by setting $x_1 = \dots = x_m = x$ is called an abstract power and is

denoted by $L_m x^m$. The induced operator is called a power operator of degree m. Similarly $L_m x_1^{m_1} \dots x_k^{m_k}$ where m_i are positive integers and $\sum_{i=1}^{k} m_i \leq m$ is an abbreviation for $L_m x_1 \dots x_1 x_2 \dots x_2 \dots x_k \dots x_k$.
Since a symmetric m-linear operator is distributive and commutative in each of the $x_1 \dots x_m$ it follows that

$$L_m (\lambda_1 x_1 + \dots + \lambda_k x_k)^m =$$

$$= \sum_{m_1 + \dots + m_k = m} \frac{\lambda_1^{m_1} \dots \lambda_k^{m_k} \, m!}{m_1! \dots m_k!} L_m x_1^{m_1} \dots x_k^{m_k} .$$

Hence,

$$L_m x_1 \dots x_m = \frac{1}{m!} \frac{\partial^m}{\partial \lambda_1 \dots \partial \lambda_m} L_m (\lambda_1 x_1 + \dots + \lambda_m x_m)^m ;$$

so that the m-linear operator is recovered from the power operator. We note that if L_m is a power operator of degree m and T is a linear operator on the range of L_m, then $T L_m x^m$ is a power of degree m. A power operator of degree m is continuous if and only if it is bounded on some sphere. If so, then it is bounded on each bounded set and satisfies a Lipschitz condition of order one uniformly on such a sphere. An abstract polynomial of degree m is an operator of the form $Q(x) = \sum_{k=1}^{m} L_k x^k$,

where $L_k \in \mathcal{L}(E^k; Y)$, $k = 1, \dots, m$, are abstract forms. Note that there is no loss of generality in assuming that the multilinear operators L_k are symmetric since the polynomial form and its Fréchet derivatives are unchanged if each L_k is replaced by its mean $\bar{L}_k$. Equivalently, an abstract polynomial of degree m is an operator on the space E with values in Y which for $x, h \in E$

and all scalars λ satisfies

$$Q(x+\lambda h) = \sum_{k=1}^{m} C_k(x,h)\lambda^k$$

where $C_k(x,h)$ are independent of λ and $C_m(x,h) \neq 0$. For properties of polynomials over Banach space we refer to Hille and Phillips [21] and Liusternik and Sobolev [3]. The notion of abstract polynomials in real or complex normed linear spaces has passed through several stages. For historical developments we refer to Fréchet [29], Gâteaux [15], Michal [30], Martin [31], Highberg [32] and Gavurin [33,34].

3. Higher order differentials of Fréchet, Gâteaux, Hadamard and related variants.

3.1. Variations. Let E and Y be normed linear spaces. Let F be a mapping from an open subset X of E into Y. Let $x_0 \in X$ and h be an arbitrary nonzero element in E. Then $x_0 + th \in X$ for $|t| \leq r(x_0; h)$. Let $\tau = \sup\{r : |t| \leq r \Rightarrow x_0 + th \in X\}$. Then $\phi(t) = F(x_0+th)$ is defined for $|t| < \tau$. If ϕ has nth order derivative at $t = 0$, then $\phi^{(n)}(0)$ is called the nth Gâteaux-Levy variation and is denoted by $\delta^n F(x_0; h)$. In this form, the definition is due to Levy [17]. The notions of first and higher order variations passed through many stages of generalizations in the work of Gâteaux [14,16], Levy [17], Graves and Hildebrandt [10,11], Kerner [20] and others. It follows easily that F has a first variation on the set X if and only if for each $x \in X$ and any

$h_1, \ldots, h_n \in E$, the function $F(x + \sum_{i=1}^{n} t_i h_i)$ is partially differentiable with respect to t_k, $k = 1, \ldots, n$, in the set of points $t_1, \ldots, t_n$ for which $x + \sum_{i=1}^{n} t_i h_i \in X$. If F has an nth order variation at x_o, then F has variations of all orders up to n at x_o, and for each positive integer $k < n$ and every $h \in E$, $\delta^n F(x_o; h) = \frac{d^k}{dt^k} \delta^{n-k} F(x_o + th; h)|_{t=0}$. $\delta^m F(x_o; h)$ is homogeneous in h of degree n. Every operator which is homogeneous of degree m has mth order variation at θ and $\delta^m F(\theta; h) = m! F(h)$.

The first variation is not necessarily linear nor continuous in h. If F has a first variation at x_o, then F is directionally continuous at x_o, i.e., $\lim_{t \to 0} \| F(x_o + th) - F(x_o) \| = 0$ for fixed h but it is not necessarily continuous at x_o. A linear variation is called a <u>Levy differential</u>. A continuous linear variation is called a <u>Gâteaux differential</u>. Clearly, every Levy differential at x_o is directionally continuous in h:

$$(4) \qquad \lim_{\tau \to 0} \delta F(x_o; h + \tau k) = \delta F(x_o; h).$$

A Levy differential is a Gâteaux differential if and only if (4) holds uniformly with respect to k on the set $\| k \| = 1$.

The nth variation is not necessarily n-linear nor continuous in h.

Higher order variations can also be defined inductively. If $\delta F(x; h)$ exists in a neighborhood of x_o, then $\delta F(x; \cdot)$ is an element of the space of all homogeneous operators on E into Y. If for fixed h, $\delta F(x; h)$ has a first variation at x_0, then we say F has a second variation at x_0 and denote it by $\delta^2 F(x_0; h, k)$, i.e., $\delta^2 F(x_0; h, k) = \delta_x(\delta F(x_0; h); k) =$

$$= \lim_{t \to 0} \frac{1}{t} \{\delta F(x_0 + tk; h) - \delta F(x_0; h)\} .$$

It then follows that

$$\delta^m F(x_0; h_1, \ldots, h_m) = \frac{\partial^m}{\partial t_1 \ldots \partial t_m} F(x_0 + \sum_{i=1}^{m} t_i h_i)\Big|_{t_1 = \ldots = t_m = 0} .$$

If $h_1 = \ldots = h_m = h$, then $\delta^m F(x_0; h, \ldots, h)$ coincides with $\delta^m F(x_0; h)$ defined above.

In order to enrich the theory of differentiation, additional properties on the nth variation are usually imposed. Rather than follow this approach, we shall enlarge the scope of notions of higher order differentials and relate them to variations. For a thorough study of higher order variations in complex spaces, see Hille and Phillips [21].

If $\delta F(x_0 + t\Delta x; h + s\Delta h)$, where Δx and Δh are elements of E, has a total differential at $t = s = 0$, then we say that $\delta F(x; h)$ has a total variation at (x_0, h), and denote it by $\delta\delta F(x_0, h; \Delta x, \Delta h)$. Clearly if $\delta F(x; h)$ has a total variation at $(x_0; h)$, then

$$\delta\delta F(x_0, h; \Delta x, \Delta h) = \delta^2 F(x_0; h, \Delta x) + \delta_h \delta F(x_0; h, \Delta x).$$

3.2. Notions of mth order Fréchet differentials.

We recall that an operator $F: X \to Y$ is said to be Fréchet differentiable at x_0 if there exists a bounded linear operator $L(x_0;\cdot)$ on E into Y such that for all $h \in E$ with $x_0 + h \in X$,

$$F(x_0 + h) - F(x_0) = L(x_0; h) + R(x_0; h)$$

where

$$\lim_{h \to \theta} \frac{\| R(x_0; h) \|}{\| h \|} = 0 .$$

If F is defined in a neighborhood of x_0, then $L(x_0; h)$ is unique and is called the Fréchet differential of F at x_0, and is denoted by $dF(x_0; h)$. The operator $dF(x_0;\cdot)$ is called the Fréchet derivative and is denoted by $F'(x_0)$. If F is Fréchet differentiable at x_0, then it is continuous at x_0. On the other hand, if F is assumed to be continuous, then the requirement of continuity of $dF(x_0; h)$ in h is redundant.

Let X be an open subset of the product space $\Pi = E_1 \times \ldots \times E_n$. Let $F: X \to Y$. The Fréchet partial differential at $u_1, \ldots, u_n$ of F with respect to x_i is defined in the usual way: there exists a bounded linear operator $L(u_1, \ldots, u_n; \cdot)$ such that for all $h_i \in E_i$ with

$$(u_1, \ldots, u_{i-1}, u_i + h_i, u_{i+1}, \ldots, u_n) \in X,$$

$$F(u_1, \ldots, u_{i-1}, u_i + h_i, \ldots, u_n) - F(u_1 \ldots u_n) = L(u_1 \ldots u_n; h_i) + R(u_1 \ldots u_n; h_i)$$

where

$$\frac{\| R(u_1 \ldots u_n; h_i) \|}{\| h_i \|} \to 0 \text{ as } h_i \to \theta .$$

$L(u_1 \dots u_n; h_i)$ is called the <u>Fréchet partial differential</u> and is denoted by $d_i F(u_1 \dots u_n; h_i)$. F is said to be <u>totally differentiable</u> if it is Fréchet differentiable considered as a mapping on $X \subset E_1 \times \dots \times E_n$ into Y, that is, if there exists an $L(u; h)$, $u = u_i \dots \dots u_n \in X$, $h = h_1 \dots h_n \in \Pi$, which is linear and continuous in h such that

$$\lim_{h \to \theta} \frac{\| F(u_1+h_1, \dots, u_n+h_n) - F(u_1 \dots u_n) - L(u_1 \dots u_n; h_1 \dots h_n) \|}{\| h_1 \| + \dots + \| h_n \|} = 0 .$$

$L(u_1 \dots u_n; h_1 \dots h_n)$ is called the total Fréchet differential of F and is denoted by $dF(u_1 \dots u_n; h_1 \dots h_n)$.

An operator $F: X \subset \Pi \longrightarrow Y$ which is totally differentiable at $u_1 \dots u_n$ is partially differentiable with respect to each variable and its total differential is the sum of the differentials with respect to each of the variables. (Fréchet [12;p.319], Dieudonné [2;p.167].) Now we turn to higher order differentials. Suppose that F has a Fréchet derivative on a subset X of E. Then $F': X \times E \longrightarrow Y$ is continuous and linear on E. We shall give eight notions of second order Fréchet differentials of F. (We remark again that weak differentials are not treated here. See Introduction.)

<u>Definition 4</u>. If $dF(x; \cdot)$, considered as a mapping on X into $\mathcal{L}(E; Y)$ has a Fréchet derivative at x_0, then we denote this derivative by $F''(x_0)$ and call it the <u>second order Fréchet derivative</u> of F at x_0.

Accordingly $F''(x_0)$ is an element of the space $\mathcal{L}(E; \mathcal{L}(E; Y))$ which is isometrically isomorphic to $\mathcal{L}(E \times E; Y)$. We make the identification of these two spaces in the sequel so that $F''(x_0) \in \mathcal{L}(E \times E; Y)$. If $F''(x)$ exists on X, then F'' is a map from X to $\mathcal{L}(E \times E; Y)$.

We call $F''(x_0)hk$, for $h, k \in E$, the second order Fréchet differential of F at x_0 and denote it also by $d^2F(x_0; h, k)$. We then have

$$\lim_{k \to 0} \|k\|^{-1} \{ \|dF(x_0 + k; \cdot) - dF(x_0; \cdot) - - d^2F(x_0; \cdot, k)\|_{\mathcal{L}} \} = 0 ,$$

where $\|\cdot\|_{\mathcal{L}}$ denotes the norm in the space of all continuous linear operators on E into Y.

Definition 5. The operator F is said to have at x_0 a second order pointwise Fréchet differential if there exists a continuous bilinear operator $B(x_0; \cdot, \cdot)$ on E such that for each fixed $h \in E$,

$$(5) \lim_{k \to \theta} \|k\|^{-1} \|dF(x_0 + k; h) - dF(x_0; h) - B(x_0; h, k)\|_Y = 0.$$

We call B the second order pointwise Fréchet derivative and denote it by $\partial^2 F(x_0; \cdot, \cdot)$.

Definition 6. The operator F is said to have at x_0 a second order partial Fréchet differential if $dF(x; h)$, considered as a map from $X \times E$ into Y, has at x_0 a partial Fréchet differential with respect to x. We denote the second order partial Fréchet differential by $d_x^2 F(x_0; h, k)$. Accordingly for each $h \in E$,

$$\lim_{k \to \theta} \|k\|^{-1} \|dF(x_0 + k; h) - dF(x_0; h) - d_x^2 F(x_0; h, k)\|_Y = 0,$$

where $d_x^2 F(x_0; h, k)$ is continuous and linear in k. Note that $d_x^2 F(x_0; h, k)$ is not required to be linear nor continuous in h. However, it turns out that it is automatically linear in h.

Definition 7. The operator F is said to have a second order total Fréchet differential at x_0, if $dF(\cdot,\cdot): X \times E \longrightarrow Y$ has a total Fréchet differential at (x_0, h) for all $h \in E$. We denote this differential by $ddF(x_0, h; \Delta x, \Delta h)$. (Note that we do not abbreviate ddF to d^2F since we reserve the latter notation for the second order Fréchet differential in the sense of Definition 4.)

Definition 8. The operator F is said to have at x_0 a second order strong Fréchet differential if there exists an operator $B(x_0; \cdot\,\cdot) \in \mathcal{L}(E \times E; Y)$ such that for each $\varepsilon > 0$, there exists $r > 0$ where $\| dF(y; \cdot) - dF(x; \cdot) - B(x_0; \cdot, y - x) \| \leq \varepsilon \| y - x \|$ for each pair of elements y, x with $\| y - x_0 \| \leq r$, $\| x - x_0 \| \leq r$. The operator $B(x_0; \cdot\,\cdot)$ is called the second order strong Fréchet derivative at x_0 and is denoted by $d^2F(x_0; \cdot\,\cdot)$.

The notions of strong pointwise differential and strong partial differential may be similarly defined.

Before we study various implications among these differentials, we consider continuity properties of $dF(x; h)$ which are implied by the existence of each of these differentials. It is clear that if F has a second order Fréchet differential at x_0, then $F'(x)$

is continuous at x_0; equivalently the differential $dF(x; h)$ is continuous in x at x_0 in the topology of uniform convergence (Definition 3). The implications for the other differentials are stated in the following:

Theorem 3. (a) If F has a second order pointwise Fréchet differential at x_0, then $F'(x)$ is continuous at x_0 (Definition 3), hence also pointwise continuous at x_0.

(b) If F has a second order partial Fréchet differential at x_0, then $dF(x; h)$ is jointly continuous at (x_0, h) (in the sense of Definition 2).

(c) If F has a second order total Fréchet differential, then $dF(x; h)$ is jointly continuous at (x_0, h).

(d) If F has a second order strong Fréchet differential, then $F'(x)$ satisfies a Lipschitz condition in some neighborhood of x_0.

Proof. (a) Consider the function ϕ defined by $\phi(h,k) = \|k\|^{-1} \|dF(x_0+k; h) - dF(x_0; h) - \partial^2 F(x_0; h, k)\|$ if $k \neq \theta$ and $\phi(h, \theta) = 0$. Since $\partial^2 F(x_0; h, k)$ is the second order pointwise Fréchet differential, $\phi(h,k)$ is continuous in k at $k = \theta$. It is also linear and continuous in h. Therefore there exists $M > 0$ and a neighborhood $N(\theta)$ of $\theta \in E$ such that for any $k \in N(\theta)$, $\phi(h, k) \leq M \|h\|$. (See Theorem 2; in particular equation (2)). Thus for $h \in N(\theta)$, $\|dF(x_0+k; h) - dF(x_0; h) - \partial^2 F(x_0; h, k)\| \leq M \|k\| \|h\|$.

Therefore,

$$\|dF(x_0+k;h)-dF(x_0;h)\|-\|\partial^2F(x_0;h,k)\| \leq M\|k\|\;\|h\|$$

and

$$\frac{\|dF(x_0+k;h)-dF(x_0;h)\|}{\|h\|} \leq (M+\|\partial^2F(x_0;.,.)\|_{\mathcal{L}_2})\|k\| .$$

Consequently,

$$\|dF(x_0+k;\,.\,)-dF(x_0;\,.\,)\|_{\mathcal{L}_1} \leq (M+\|\partial^2F(x_0;\,.\,)\|_{\mathcal{L}_2})\|k\| .$$

Thus

$$k \to \theta \text{ implies } \|dF(x_0+k;\,.\,)-dF(x_0;\,.\,)\| \to 0 .$$

(b), (c), (d) follow easily. We prove (d). By definition there exists $r > 0$ such that

$$\|dF(y;\,.\,)-dF(x;\,.\,)-d^2F(x_0;\,.\,.\,.\,)\| \leq \|y-x\|$$

if y and x are in $N(x_0;r)=\{x \in X: \|x-x_0\| \leq r\}$.
Thus $\|dF(y;\,.\,)-dF(x;\,.\,)\|_{\mathcal{L}_1} \leq \|y-x\|+M\|y-x\|$
where $\|d^2F(x_0;\,.\,.\,)\|_{\mathcal{L}_2} \leq M$.
That is,

$$\|dF(y;\,.\,)-dF(x;\,.\,)\|_{\mathcal{L}_1} \leq (M+1)\|y-x\| ,$$

for all $y, x \in N(x_0;r)$.

We now state implication relationships among these notions of differentials.

Theorem 4. (a) F has a second order Fréchet differential at x_0 if and only if F has a second order pointwise differential at x_0 and (5) holds uniformly with respect to h on the set $\|h\| = 1$. In this case the two differentials are equal.

(b) F has a second order pointwise Fréchet differential at x_0 if and only if F has a second order partial Fréchet differential at x_0 and the latter is jointly continuous in h and k (or continuous in

h and the space E is complete).

(c) If $F(x)$ has a second order pointwise Fréchet differential at x_0, then $dF(x; h)$ is totally differentiable (Definition 7) at (x_0, h) and the total differential is given by $\partial^2 F(x_0; h; \Delta x) + dF(x_0; \Delta h)$.

(d) If F has a second order strong Fréchet differential at x_0, then F has a second order Fréchet differential at x_0.

<u>Proof</u>: (a) and (d) are obvious.

(b) First we show that $d_x dF(x_0; h, k)$, the second order partial Fréchet differential at x_0, is automatically linear in h. Let $h_1, h_2, k \in E$, $\|k\| = 1$, α, β and τ be scalars. Then

$$\|d_x dF(x_0; \alpha h_1 + \beta h_2, k) - \alpha d_x dF(x_0; h_1, k) -$$

$$- \beta d_x dF(x_0; h_2, k)\| =$$

$$= \frac{1}{|\tau|} \|d_x dF(x_0; \alpha h_1 + \beta h_2, \tau k) - \alpha d_x dF(x_0; h_1, \tau k) -$$

$$- \beta d_x dF(x_0; h_2, \tau k)\| \leq$$

$$\leq \frac{1}{|\tau|} \|d_x dF(x_0; \alpha h_1 + \beta h_2, \tau k) - dF(x_0 + \tau k; \alpha h_1 + \beta h_2) +$$

$$+ dF(x_0; \alpha h_1 + \beta h_2)\| +$$

$$+ \frac{|\alpha|}{|\tau|} \|dF(x_0 + \tau k; h_1) - dF(x_0; h_1) - d_x dF(x_0; h_1, \tau k)\| +$$

$$+ \frac{|\beta|}{|\tau|} \|dF(x_0 + \tau k; h_2) - dF(x_0; h_2) - d_x dF(x_0; h_2, \tau k)\| .$$

Each of the three terms on the right side of the above inequality tends to zero as $\tau \to 0$, uniformly with respect to k on the set $\|k\| = 1$. But the first term on the left side of the inequality is independent of τ. Consequently, $d_x dF(x_0; \alpha h_1 + \beta h_2, k) = \alpha d_x dF(x_0; h_1, k) + \beta d_x dF(x_0; h_2, k)$ for k with $\|k\| = 1$ and hence for every k since $d_x dF(x_0; h, k)$ is homogeneous in k.

Thus if $d_x dF(x_0; h, k)$ is jointly continuous in h and k (or if $d_x dF(x_0; h, k)$ is continuous in k and the space E is complete (see Theorem 1)), then $d_x dF(x_0; h, k)$ is the second order pointwise differential at x_0. The converse is trivial.

(c) Suppose F has a second order pointwise Fréchet differential at x_0. For fixed $h \in E$, let

$B(x_0, h; \Delta x, \Delta h) = \partial^2 F(x_0; h, \Delta x) + dF(x_0; \Delta h)$.

Clearly B is linear and jointly continuous in Δx and Δh. We now show that $B(x_0, h; \Delta x, \Delta h)$ is the total Fréchet differential at (x_0, h) of $dF(x_0; h)$.

Let

$$R(\Delta x, \Delta h) = dF(x_0 + \Delta x, h + \Delta h) - dF(x_0; h) - B(x_0, h; \Delta x, \Delta h).$$

Then

$$\frac{\|R(\Delta x, \Delta h)\|}{\|\Delta x\| + \|\Delta h\|} \leq \frac{1}{\|\Delta x\|} \|dF(x_0 + \Delta x; h) - dF(x_0; h) - \partial^2 F(x_0; h, \Delta x)\| + \frac{1}{\|\Delta h\|} \|dF(x_0 + \Delta x; \Delta h) - dF(x_0; \Delta h)\| \leq \frac{1}{\|\Delta x\|} \|dF(x_0 + \Delta x; h) - dF(x_0; h) - \partial^2 F(x_0; h, \Delta x)\| + \|dF(x_0 + \Delta x; \cdot) - dF(x_0; \cdot)\|$$

As Δx and Δh tend to zero, the first term on the right goes to zero by definition of $\partial^2 F(x_0; h, \Delta x)$; the second term goes to zero by part (a) of Theorem 3. This shows that $B(x_0, h; \Delta x, \Delta h)$ is the total Fréchet differential of $dF(x_0; h)$.

(Continued in Part II)

References

[1] M.Z. NASHED: On the representation and differentiability of operators (to appear).

[2] J. DIEUDONNÉ: Foundations of Modern Analysis. Academic Pres, New York, 1960.

[3] L.A. LIUSTERNIK and V.J. SOBOLEV: Elements of Functional Analysis. Ungar, New York, 1961.

[4] M.M. VAINBERG: Variational Methods for the Study of Nonlinear Operators. Holden-Day, San Francisco, 1964.

[5] L.V. KANTOROVICH and G.P. AKHILOV: Functional Analysis in Normed Spaces. Pergamon Press, New York, 1964.

[6] A.D. MICHAL: Le calcul différentiel dans les espaces de Banach, vol.1,2. Gauthier-Villars, Paris, 1958, 1964.

[7] M.Z. NASHED: Some remarks on variations and differentials. Amer.Math.Monthly 73(1966), No.4, Part II(Slaught Memorial Papers), 63-76.

[8] L.B. RALL: Computational Solution of Nonlinear Operator Equations. Wiley, New York, 1969.

[9] L.M. GRAVES: Riemann integration and Taylor's theorem in general analysis. Trans.Amer.Math. Soc., 29(1927), 163-177.

[10] L.M. GRAVES: Topics in functional calculus. Bull. Amer.Math.Soc.41(1935),641-662.

[11] T.H. HILDEBRANDT and L.M. GRAVES: Implicit functions and their differentials in general analysis. Trans.Amer.Math.Soc.29 (1927),127-153.

[12] M. FRÉCHET: La notion de différentielle dans l´analyse générale. Ann.Sci.Ecole Norm.Sup. 42(1952),293-323.

[13] M. FRÉCHET: Sur la notion différentielle. J.Math. Pures Appl.16(1937),233-250.

[14] R. GÂTEAUX: Sur les fonctionnelles continues et les fonctionnelles analytiques. Bull. Soc.Math.France,50(1922),1-21.

[15] R. GÂTEAUX: Sur diverses questions du calcul fonctionnel. Bull.Soc.Math.France 50(1922).

[16] R. GÂTEAUX: Fonctions d^une infinite variable indépendantes. Bull.Soc.Math.France 47 (1919),70-96.

[17] P. LÉVY: Leçons d´analyse fonctionnelle, Gauthier-Villars Paris,1922.

[18] M. KERNER: Die Differentiale in der allgemeinen Analysis, Ann.of Math.34(1933),546-572.

[19] M. KERNER: Zur Theorie der Impliziten Funktional Operationen. Studia Mathematica,Tom III (1931),156-173.

[20] M. KERNER: Sur les variations faibles et fortes d´une fonctionnelle. Annali di Matematica Pura ed Applicata,Ser.4,10(1932),145-164.

[21] E. HILLE and R.S. PHILLIPS: Functional Analysis and Semigroups. Amer.Math.Soc.Coll.Publ., Providence,1957.

[22] N. BOURBAKI: Espaces vectoriels topologiques, V. Herman,Paris,1955.

[23] J. DIEUDONNÉ: Recent developments in the theory of locally convex vector spaces. Bull.Soc.Math. France 91(1963),227-284.

[24] A. FRÖLICHER and W. BUCHER: Calculus in Vector Spaces without Norm. Springer-Verlag,Berlin, 1966.

[25] M.Z. NASHED: Some remarks on higher order weak differentials - to appear.

[26] J. KOLOMÝ: On the differentiability of mappings in functional spaces. Comment.Math.Univ.Carolinae 8(1967),315-329.

[27] J. KOLOMÝ and V. ZIZLER: Remarks on the differentiability of mappings in linear normed spaces. Comment.Math.Univ.Carolinae 8(1967),691-704.

[28] V. ZIZLER: On the differentiability of mappings in Banach spaces. Comment.Math.Univ.Carolinae 8(1967),415-430.

[29] M. FRÉCHET: Les polynomes abstraits. Journal de Mathématiques Pures et Appliquées;ser.9,6.8 (1929),p.71.

[30] A.D. MICHAL and R.S. MARTIN: Some expansions in vector spaces.J.Mathématiques Pures et Appliquées,Ser.9,t-13(1934),p.69.

[31] R.S. MARTIN: Contributions to the Theory of Functionals (Thesis, California Institute of Technology, 1932).

[32] I.E. HIGHBERG: A note on abstract polynomials in complex spaces. J.de Math.Pures et Appl. Ser.9, t-16(1937), 307-314.

[33] M.K. GAVURIN: On k-ple linear operations in Banach spaces, Dokl.Akad.Nauk SSSR 22, No.4(1939), 547-551.

[34] M.K. GAVURIN: Analytic methods for the study of non-linear functional transformations. Leningrad Gos.Univ.Uč.Zap.Ser.Mat.Nauk 19 (1950), 59-154.

Department of Mathematics
American University of Beirut
Beirut, Liban

(Oblatum 17.6.1969)