DERIVATIVES AND FRÉCHET DIFFERENTIALS

MAX A. ZORN

1. **Generalities.** A function $f(x)$, defined on an open set S of a complex Banach space X, with values in a complex Banach space Y, is said to have a Fréchet differential at a point x_0 of S if for $x=x_0$ the following conditions (G), (D), and (P) are satisfied:

(G) The limit $\lim_{\zeta\to 0}[f(x+\zeta h)-f(x)]/\zeta=\delta_x^h f=\delta f(x, h)$ exists for all h in X; (D) this limit is a continuous linear function of h; (P) the Gâteaux differential $\delta f(x, h)$ is a principal part of the increment, that is, $[f(x+h)-f(x)]-\delta f(x, h)=o(\|h\|)$.

We say that $f(x)$ is F-differentiable on S if these conditions hold at every point of S; if the condition (G) is satisfied for every point in S we call the function G-differentiable on S.

The reader will find in [2][1] or [6] a proof to the effect that a function which is G-differentiable on S—or indeed on more general sets—leads to a function $\delta f(x, h)$ which is linear, in the algebraic sense, with respect to h. We may thus replace the condition (D) by the requirement that the Gâteaux differential be continuous with respect to the argument h, which in turn is equivalent to $\delta f(x, h)$ being $O(1)$, $o(1)$ or $O(\|h\|)$ as $\|h\|$ tends to zero.

Our main purpose is to show that (P) is satisfied automatically if (G) and (D) hold on S, giving a new answer to the question: under which conditions is a G-differentiable function F-differentiable?

Previous solutions of this problem have been of two kinds. The first kind operates with topological conditions on the function $f(x)$, like continuity (see [4]), local boundedness (see [2]), or essential continuity (see [6]). The most general characterization theorem of this type seems to be the following: Let $f(x)$ be G-differentiable on the connected open set S, and bounded on a set $V-M$, where V is a nonvoid open subset of S and M is such that the whole space X is not the sum of a countable number of homothetic images $\alpha_n M+a_n$ of M; under these conditions the function $f(x)$ is F-differentiable on S (see [7]).

A solution of the second kind may be abstracted from [2] or [6]: if the higher differentials $\delta^n f(x; h_1, \cdots, h_n)$ are continuous functions of their h-arguments for one value x_0 of x, then $f(x)$ will be F-differentiable on a suitable neighborhood of x_0. The two kinds of characterizations are rather different; the first type refers to the behaviour

Received by the editors September 4, 1945.

[1] Numbers in brackets refer to the references cited at the end of the paper.

of $f(x)$ on an open set; the second is based on the behaviour on a subset which is only "finitely open" (see the definition (2.1)). The condition (D) belongs to the second class, and we look upon it in this manner:

By virtue of (D) there belongs to x a bounded linear transformation on X to Y, whose value for the argument h is $\delta f(x, h)$. The bounded linear transformations on X to Y, under the standard definition of norms, constitute a (complex) Banach space $[X, Y]$. The above linear transformation is thus the value of a function on S to $[X, Y]$, which we denote by $f'(x)$; the name derivative is justified by the formula $\delta f(x, h) = f'(x)h$.

2. **Derivation of the condition** (P). The functions $f(x)$ we deal with are at first assumed to be G-differentiable on a set D, which is finitely open according to the definition:

(2.1) A subset D of the (complex) Banach space X is finitely open if for x in D, $h_1, \cdots, h_n$ in X, the n-uples $(\zeta_1, \cdots, \zeta_n)$ for which

$$x + \zeta_1 h_1 + \cdots + \zeta_n h_n \in D$$

form an open set of the n-dimensional (complex) number space.

Without making use of the topology or metric of X the G-differential $\delta f(x, h)$ and the higher differentials $\delta^n f(x; h_1, \cdots, h_n)$ may be defined; for instance, $\delta^2 f(x; h, k)$ is $\delta_x^k[\delta f(x, h)]$. We shall use the (trivial) observation that the function $f(x)$ is G-differentiable on D if and only if $f(x+\zeta h)$ is a differentiable function of the complex variable ζ.

The topology of the value space is of course being used; and since we want the values of our functions to be in Banach spaces it becomes understandable that we restrict the concept derivative as follows:

(2.2) A function $f'(x)$ on D to the Banach space $[X, Y]$ of all bounded linear transformations on X to Y is called the derivative of $f(x)$ if $\delta f(x, h) = f'(x)h$, for x in D and h in X.

The value of this definition may be gauged by the lemma:

(2.3) *The derivative is G-differentiable on D*;

and the theorem:

(2.4) *If $f(x)$ is G-differentiable on an open set S and possesses a derivative on a nonvoid, finitely open subset D of S it satisfies the condition* (P) *at every point of D*.

We shall arrange the proofs in such a manner that a maximum of information is derived from the behaviour of the function on the finitely open set D alone.

From the theory of the G-differential in [2] or [6] we shall have to use the theorems:

(2.5.1) *The G-differentials* $\delta^n f(x; h_1, \cdots, h_n)$ *exist and they are G-differentiable with respect to x on D, linear and symmetric with respect to the h-arguments on X*;

(2.5.2) *For x fixed in D we get* $f(x+h)=\sum_0^\infty \delta^n f(x; h, \cdots, h)/n! =\sum_0^\infty p_n(x, h)$, *where h comes from a set* H_x *which is defined by the condition that* $|\zeta| \leqq 1$ *implies* $x+\zeta h \in D$.

From the theory of linear operators we borrow (see [5]):

(2.6) *If the bounded operator* $U(\zeta)$ *depends on the complex number* ζ*—which varies in an open set* Δ*—in such a manner that* $U(\zeta)h$ *is differentiable with respect to* ζ *for any h in X, then* $U(\zeta)$ *is differentiable with respect to* ζ, *on* Δ.

Proceeding now to the proof of the lemma (2.3) we note that it suffices to show that for k in X the quantity $f'(x+\zeta k)$ is differentiable with respect to ζ. This will follow from A. E. Taylor's theorem (2.6) if we know that $f'(x+\zeta k)h$ or $\delta f(x+\zeta k, h)$ is differentiable with respect to ζ for any h in X; that, however, amounts just to G-differentiability of $\delta f(x, h)$ with respect to x, which is asserted by (2.5.1). The lemma is thus proved and we may apply the theory of the G-differential to the function $f'(x)$. Its higher differentials will exist, and they will be bounded linear transformations. If $U(x)$ is a G-differentiable function on D to $[X, Y]$, we shall have the equality $[\delta U(x, k)]h = \delta_x^k\{[U(x)]h\}$, for

$$\left\{\lim_{\zeta\to 0} [U(x + \zeta k) - U(x)]/\zeta\right\} h = \lim_{\zeta\to 0} \{[U(x + \zeta k)]h - [U(x)]h\}/\zeta.$$

With the use of this principle and the symmetry of the differentials in their h-arguments one arrives by way of a mathematical induction at the formula:

$$(2.7) \qquad \delta^n f'(x; h_1, \cdots, h_n)h_{n+1} = \delta^{n+1} f(x; h_1, \cdots, h_{n+1}).$$

The left member of (2.7) is continuous with respect to h_{n+1}; the right member is a symmetric function of the h-arguments, so that the differentials turn out to be partially continuous in these arguments. By a theorem of Mazur and Orlicz (see [3, p. 65] and the references given there; compare also [6, Theorem (3.7)]) they will be continuous jointly in their h-arguments. The functions $p_n(x, h) = \delta^n f(x; h, \cdots, h)/n!$ are therefore continuous in h; in the terminology of [2] and [6] we have shown that the "G-powers" $p_n(x, h)$ are "F-powers" of h.

We show now that in a suitable neighborhood of $h=0$ the power series $\sum_0^\infty p_n(x, h)$ converges uniformly towards a function which has $p_1(x, h)$ as its Fréchet differential. By (2.5.2) the sum of this power series coincides with $f(x+h)$ on the set H_x. The proof is only a slight variation of the arrangement in [6]; we may thus content ourselves with a mere sketch.

The set H_x on which the power series converges is of the second category, since $\bigcup_{n=1}^\infty (nH_x)$ is the whole space X. Since the terms are continuous functions of h, a classical principle shows that they are uniformly bounded on a sphere (compare [1, p. 19]). We may thus assume that for a suitable h_0 and positive numbers ρ, M the inequality $\|h-h_0\| \leqq \rho$ implies, for all n, $\|p_n(h)\| \leqq M$ (we drop the argument x). It is easily seen that due to the homogeneity and G-differentiability of the functions $p_n(h)$ the same uniform bound obtains for $\|h\| \leqq \rho$ (compare Theorem (4.1) of [6]).

For $\|h\| \leqq \sigma < \rho$ we find $\|p_n(h)\| \leqq M(\sigma/\rho)^n$; this ensures uniform convergence of $\sum_0^\infty p_n(h)$ towards a function $g(h)$, for $\|h\| \leqq \sigma < \rho$.

We might stop here with an appeal to the theory of power series (see, for instance, [4, p. 11]); or we may prove, as in [6, Theorem (4.3)], the inequality

$$\| [g(h) - g(0)] - p_1(h)\| \leqq \|h\|^2 M/\rho(\rho - \sigma),$$

which yields for x in D, h in H_x, $\|h\| \leqq \sigma < \rho$,

$$\| [f(x + h) - f(x)] - \delta f(x, h)\| \leqq \|h\|^2 M/\rho(\rho - \sigma), \tag{2.8}$$

where the quantities M and ρ depend on x.

At this point we make use, for the first time, of the premise that $f(x)$ be defined on an open set S which contains D. The number ρ may then be taken so small that for $\|h\| < \rho$ the points $x+h$ are contained in S; we do not ask for more if we want the points $x+\zeta h$ to be in S for $|\zeta| \leqq 1$. By virtue of (2.5.2) the power series $\sum_0^\infty p_n(x, h)$ represents $f(x+h)$ in the sphere $\|h\| < \rho$, so that the inequality (2.8) is valid there. A fortiori, the condition (P) holds at every point of D. We have proved the theorem (2.4); let us add the corollary:

(2.9) *If $f(x)$ possesses a derivative on the open set S it is F-differentiable on S.*

Added in proof, January 20, 1946. Professor A. D. Michal informs me that more than ten years ago, in connection with a first draft of his paper *General tensor analysis* (Bull. Amer. Math. Soc. vol. 43 (1937)), he introduced the notion of a derivative as distinguished from a differential.

REFERENCES

1. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.
2. E. Hille, *Topics in the theory of semi-groups*, Colloquium Lectures, 1944.
3. S. Mazur and W. Orlicz, *Grundlegende Eigenschaften der polynomischen Operationen*, Part 1, Studia Mathematica vol. 5 (1934) p. 50.
4. A. E. Taylor, *Analytic functions in general analysis*, Annali della R. Scuola Normale Superiore di Pisa (2) vol. 6 (1937).
5. A. E. Taylor, *Linear operations which depend analytically on a parameter*, Ann. of Math. vol. 39 (1938).
6. M. A. Zorn, *Characterization of analytic functions in Banach spaces*, Ann. of Math. vol. 46 (1945).
7. M. A. Zorn, *Gâteaux differentiability and essential boundedness* (to appear in Duke Math. J.).

UNIVERSITY OF CALIFORNIA AT LOS ANGELES