# THE FRÉCHET DIFFERENTIAL IN NORMED LINEAR SPACES

by

JOHN HILZMAN

A THESIS

submitted to

OREGON STATE COLLEGE

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1955

APPROVED:

[redacted]

Professor of Mathematics

In Charge of Major

[redacted]

Head of Department of Mathematics

[redacted]

Chairman of School of Science Graduate Committee

[redacted]

Dean of Graduate School

Date thesis is presented May 3, 1955

Typed by Margaret Schmidt

TABLE OF CONTENTS

# THE FRÉCHET DIFFERENTIAL IN NORMED LINEAR SPACES

## INTRODUCTION

Since the beginning of the twentieth century, the trend of Mathematics has been toward greater abstraction and generality. Functional analysis, a recent mathematical discipline originated by Vito Volterra and extended by Maurice Fréchet, E. H. Moore and other mathematicians, has come to the forefront embodying the concepts of topological spaces and modern algebra. Functional analysis is a generalization and an extension of much of classical analysis, the underlying characteristic being the passage from the finite to the infinite dimension. The evolution of functional analysis can be traced back through the theory of infinite systems, to integral and integro-differential equations, the calculus of variations and the boundary value problems of mathematical physics.

A typical example of an operation in functional analysis is $y = Px$ where $x$ and $y$ are elements of any nature, for instance, vectors, functions or sets of functions. This is a wide generalization of the functions in classical analysis. To particularize, consider the relation

$$P\left[g(x)\right] = \int_0^1 g(x)dx \tag{1.1}$$

where $g(x)$ is a single-valued real function. $P\left[g(x)\right]$ is called a <u>functional</u> of $g(x)$. The domain is the set of real functions for which the integral is defined, and the range is a subset of the real continuum.

Much of the theory of abstract spaces and the extension of common notions from classical theories to functional analysis is due to Maurice Fréchet. Most of his works have recently been compiled in the publication, Pages Choisies d'Analyse Generale. One of Fréchet's important contributions was the introduction of the idea of the differential to abstract spaces (4, pp.293-323). The complete analysis of this development can be found in his cited works.

The extension of the differential led to new researches. Development of a differential calculus in abstract spaces followed. In most instances the topological spaces considered were normed linear spaces (to be defined later). Several extensions and modifications of the Fréchet differential were made by R. Gateaux, A. D. Michal (11, pp.532-536), E. W. Paxson and D. H. Hyers (8, pp.315-316). A differential calculus of implicit functions in functional analysis was studied by Hildebrandt and Graves (6, pp.127-153).

Although much of the work involving the Fréchet differential has been of a theoretical nature, there have been investigations on the applied level which have led to practical results. The study by Michal of solutions of Volterra and Fredholm integral equations as functionals of resolvent kernels has application to the problem of obtaining approximations with precise error estimates (13, pp.252-258). A most direct application of the Fréchet differential was in the generalization of Newton's method by Kantorovich for the approximate solution of functional equations (9, pp.154-183). Also, on a research project directed by Dr. A. T. Lonseth at Oregon State College, the generalized Newton's method was used to obtain a numerical solution to the Chandrasekhar non-linear integral equation which arises in astrophysics. In a recent paper on gradient mappings (15, pp.5-19), E. H. Rothe utilizes the Fréchet differential in the treatment of integral equations in boundary value problems.

A brief analysis will be presented here of the generalization of the classical differential to normed linear spaces. Several rules for Fréchet differentiation and examples will be given. The bibliography will contain a substantial list of references.

## FUNDAMENTAL DEFINITIONS

Basic to the understanding of differentials and differentiation, in the sense of Fréchet, is the theory of complete normed linear spaces and operations in them. An outline of this theory will be presented below. Proofs for many results will be found in the works cited throughout the text of this paper.

A set of elements $X = \{x\}$ will be called a <u>real linear space</u> whenever it satisfies the following conditions. Here $x_1$, $x_2$ are elements of the set and $a$, $b$ are real numbers.

(1) The <u>sum</u> $x_1 + x_2$ is defined and lies in $X$.

(2) The <u>product</u> $ax$ is defined and lies in $X$.

(3) The elements of the set form a commutative (abelian) group under <u>addition</u>. The identity element of $X$ is denoted by $0$.

(4) Multiplication is associative, i.e., $a(bx) = (ab)x$.

(5) Multiplication is doubly distributive, i.e., $(a+b)x = ax + bx$ and $a(x_1+x_2) = ax_1 + ax_2$.

(6) $1 \cdot x = x$ and $o \cdot x = 0$.

If the numbers $a$ and $b$ were complex, the set $X$ would be called a complex linear space. Only real linear spaces will be considered unless otherwise specified.

If for every element $x$ in $X$ there is defined a non-negative real number $||x||$, called the <u>norm</u> or <u>length</u> of $x$, which satisfies the following conditions:

(7) $||x|| \geq o$, $||x|| = o$ if and only if $x = 0$,

(8) $||x_1 + x_2|| \leq ||x_1|| + ||x_2||$ .

(9) $||ax|| = |a| \cdot ||x||$ ;

then X is called a <u>normed linear space</u> (9, p.4).

In a normed linear space X a sequence $\{x_n\}$ converges to an element x in X if $||x - x_n|| \to o$ as $n \to \infty$. The sequence $\{x_n\}$ is said to be a Cauchy sequence if $\lim ||x_{n+p} - x_n|| \to o$ as $n \to \infty$ for every positive integer p. A normed linear space is called <u>complete</u> if every Cauchy sequence in it has a limit, that is, converges (7, p.5 and 9, p.4). In a complete space the convergence of the series $\sum_{k=1}^{\infty} ||x_k||$ implies the existence of the limit

$$\lim_{n\to\infty} \sum_{k=1}^{n} x_k = \sum_{k=1}^{\infty} x_k ,$$

that is, an absolutely convergent series is convergent.

Several examples of normed linear spaces are the following.

(1) The set of all real numbers with addition and multiplication defined in the usual way is a normed linear space. The norm is

$$||x|| = |x| . \tag{2.1}$$

(2) The set of all complex numbers with addition and multiplication by real (or complex) numbers defined in the usual way is a normed linear space. The norm is

$$||x + iy|| = |x + iy| = (x^2 + y^2)^{\frac{1}{2}} \tag{2.2}$$

(3) The n-dimensional vector space (Euclidean space $R^n$) where

addition is defined as

$$x + y = (\xi_1, \xi_2, \ldots, \xi_n) + (\eta_1, \eta_2, \ldots, \eta_n)$$
$$= (\xi_1 + \eta_1, \xi_2 + \eta_2, \ldots \xi_n + \eta_n)$$

and multiplication by a real (or complex number), $\alpha$, as

$$\alpha x = (\alpha \xi_1, \alpha \xi_2, \ldots \alpha \xi_n)$$

is a normed linear space. The norm in this space may be defined in several ways, as

$$||x|| = \left( \sum_{k=1}^{n} \xi_k^2 \right)^{\frac{1}{2}}, \tag{2.3}$$

$$||x|| = \max_{(i)} |\xi_i| , \tag{2.4}$$

$$||x|| = \sum_{j=1}^{n} |\xi_j| . \tag{2.5}$$

(4) The set C of all continuous real functions defined on the interval (a,b) with addition and multiplication defined in the usual way is a normed linear space. The norm is

$$||x|| = \max_{a \leq s \leq b} |x(s)| . \tag{2.6}$$

(5) The set $L^2$ of all real functions that are square-integrable on the interval (a,b) is a normed linear space. The norm is

$$||x|| = \left( \int_a^b x^2(s)ds \right)^{\frac{1}{2}} \tag{2.7}$$

(6) The set H (after Hilbert) of all real sequences

$x = (x_1, x_2, \ldots)$ such that $\sum_{i=1}^{\infty} x_i^2$ converges is a normed linear space. The norm is

$$(2.8) \qquad ||x|| = \left( \sum_{i=1}^{\infty} x_i^2 \right)^{\frac{1}{2}} .$$

Several of the real spaces just cited have extensions in the complex field with slight alterations.

Let $X$ and $Y$ be two normed linear spaces with $x$ an element of $X$ and $y$ an element of $Y$. A function $y = Px$ which maps $X$ into $Y$ is called an <u>operation.</u> The set $X$ is referred to as the <u>domain</u> of $P$ and $Y$ the <u>range</u> of $P$. The <u>operator</u> $P$ is called a <u>functional</u> if the range is the space $R$ of real numbers. If in the space $X$ the relation $\lim x_n = x$ implies that $\lim Px_n = Px$ in the space $Y$, then $P$ is said to be <u>continuous</u> (16, p.133). The operator $P$ is called <u>additive</u> if

$$(2.9) \qquad P(x_1 + x_2) = Px_1 + Px_2$$

for all $x_1, x_2$ in the domain of $P$. Consequently $P0 = 0$ and $P(-x) = -Px$. The operator $P$ is said to be <u>linear</u> if it is continuous and additive. A linear operator $P$ is homogeneous, that is,

$$(2.10) \qquad P(ax) = a\,P\,x$$

for any real number $a$.

Consider the operation $y = Px$ where $y$ is in $Y$ and $x$ is in $X$. The operator $P$ is <u>bounded</u> if there exists a fixed non-negative number $M$ such that

$$\|Px\| \leq M \|x\| \tag{2.11}$$

for all $x$ in $X$. The smallest such number satisfying this inequality is called the <u>norm</u> or <u>length</u> of the operator $P$ and is denoted by $\|P\|$ (16, p.134). The left side of the inequality refers to the norm as defined in the space $Y$, and on the right the norm refers to the space $X$. The operation $y = Px$ has an <u>inverse</u> $x = P^{-1}y$ if and only if there is a one-to-one correspondence between the domain and the range of the operation. The existence of the inverse follows from the condition that $x_1 \neq x_2$ implies $Px_1 \neq Px_2$ (7, p.26). The inverse of a linear operation is again linear (16, p.162).

A necessary and sufficient condition that the linear operation $y = Px$ have a bounded inverse is the existence of an $m > o$ such that $\|Px\| \geq m\|x\|$ for all $x$ in the domain of $P$. The largest admissible value of $m$ is the reciprocal of the norm of $P^{-1}$ (7, p.26 and 16, p.163). To show this, let $P^{-1}$ exist and be a bounded operation. Then there exist a fixed non-negative number $M$ such that

$$\|P^{-1}y\| \leq M \|y\| , \tag{2.12}$$

or

$$\|x\| \leq M \|Px\| , \tag{2.13}$$

so that

$$\|Px\| \geq \frac{1}{M} \|x\| . \tag{2.14}$$

Thus, if $M$ is the norm of $P^{-1}$, then $\frac{1}{M}$ is the largest value of $m$ for which $||Px|| \geq m||x||$. Conversely, if $||Px|| \geq m||x||$, then $Px = o$ if and only if $x = o$ (16, p.162); and $P^{-1}y$ exists so that

$$||P^{-1}y|| \leq \frac{1}{m}||y|| . \tag{2.15}$$

Consequently, for a linear operator there exist two non-negative real numbers

$$M(P) = \underset{||x|| \neq o}{\text{l.u.b.}} \frac{||Px||}{||x||} \tag{2.16}$$

and

$$m(P) = \underset{||x|| \neq o}{\text{g.l.b.}} \frac{||Px||}{||x||} , \tag{2.17}$$

where $M(P)$ and $m(P)$ are called respectively the <u>upper</u> <u>bound</u> and <u>lower</u> <u>bound</u> of $P$. Their existence implies the inequalities

$$m(P)\ ||Px|| \leq ||Px|| \leq M(P)\ ||Px|| \tag{2.18}$$

for any $x$.

If $P$ and $Q$ are two operators which map space $X$ into space $Y$, the <u>sum</u> $(P + Q)$ is defined by $(P + Q)x = Px + Qx$ for all $x$ in $X$. When $P$ and $Q$ are linear, the norm of $(P + Q)$, $M(P + Q)$, satisfies the inequality

$$M(P + Q) \leq M(P) + M(Q) . \tag{2.19}$$

If $P$ maps $Y$ into $Z$ and $Q$ maps $X$ into $Y$, the <u>product</u> $PQ$ is defined by $(PQ)x = P(Qx)$ for all $x$ in $X$. When $P$ and $Q$ are linear the norm of $(PQ)$, $M(PQ)$, satisfies the inequality

$$M(PQ) \leq M(P)\, M(Q) \; . \tag{2.20}$$

As an instance of a linear operation in a normed linear space, consider the linear operation $y = Px$ which maps $X = R^n$ into $Y = R^m$. The norm is defined as $||x|| = \max_{(i)} |\xi_i|$. For any $x$ in $X$ where $x = (\xi_1, \xi_2, \dots \xi_n) = \xi_1 x_1 + \xi_2 x_2 + \dots + \xi_n x_n$; let $x_1 = (1, o, \dots, o)$, $x_2 = (o, 1, \dots o)$, $\dots$ $x_n = (o, o, \dots 1)$. Also let $y_k = Px_k = (a_{1k}, a_{2k}, \dots a_{mk})$, $k = 1, 2, \dots n$. Then

$$\begin{aligned} y = Px &= P(\xi_1 x_1) + P(\xi_2 x_2) + \dots + P(\xi_n x_n) \; , \\ &= \xi_1 Px_1 + \xi_2 Px_2 + \dots + \xi_n Px_n \; , \\ &= \xi_1 y_1 + \xi_2 y_2 + \dots + \xi_n y_n \; . \end{aligned} \tag{2.21}$$

If $y = (\eta_1, \eta_2, \dots \eta_m)$, then it is seen that

$$\eta_j = \sum_{k=1}^{n} a_{jk}\, \xi_k, \; (j = 1, \dots m) \; . \tag{2.22}$$

Thus the operator $P$ is a linear transformation defined by the <u>matrix</u>

$$A = (a_{ij}) \; . \tag{2.23}$$

The norm of this operator is found as follows.

$$||y|| = \max_{(i)} |\eta_i| = \max_{(i)} \left| \sum_{k=1}^{n} a_{ik}\, \xi_k \right| \leq \max_{(k)} |\xi_k| \cdot \max_{(i)} \sum_{k=1}^{n} |a_{ik}| \; , \tag{2.24}$$

or

$$(2.25) \quad ||Px|| \leq \max_{(k)} |\xi_k| \cdot \max_{(i)} \sum_{k=1}^{n} |a_{ik}| = ||x|| \cdot \max_{(i)} \sum_{k=1}^{n} |a_{ik}| .$$

Thus, in any case,

$$(2.26) \qquad ||P|| \leq \max_{(i)} \sum_{k=1}^{n} |a_{ik}| ,$$

and the exact equation

$$(2.27) \qquad ||P|| = \max_{(i)} \sum_{k=1}^{n} |a_{ik}|$$

can be established (9, pp.8-9).

The operators considered so far were operators of <u>one place</u> which mapped a space $X$ into a space $Y$. The collection of all such operators $L$ which are linear, forms a normed linear space satisfying all conditions defined for these spaces. The space of linear operators from $X$ to $Y$ will be denoted by $(X \rightarrow Y)$. For all $L$ in $(X \rightarrow Y)$, $||L|| = M(L)||x||$.

The <u>cartesian product</u> of the spaces $X$ and $Y$ is the set of all pairs $(x,y)$ where $x$ is in $X$ and $y$ is in $Y$. The cartesian product will be denoted by $X \times Y$ ($X$ cross $Y$). An operation $P$ on $X$ to $Y$ is a subset of $X \times Y$ so that $(x,y)$ is an element of $P$.

A linear operator $B$ which maps $X \times X$ into $Y$ is an operator of <u>two places</u> and is called <u>bilinear</u>. A bilinear operation on $X$ to $Y$ is equivalent to a linear operation on $X$ to $(X \rightarrow Y)$. To show this, consider a linear operation $L = Bx$ mapping $X$ into $(X \rightarrow Y)$. Since $L$ is a linear operator from $X$ to $Y$ ,

$$Bxx' = Lx' = (Bx)\,x' \tag{2.28}$$

is an element of $Y$, and $B$ is defined for all $(x,x')$ in $X \times X$. $B$ is additive with respect to each argument, and hence is equivalent to a bilinear operator from $X$ to $Y$. Conversely, if $Bxx'$ is a bilinear operation from $X$ to $Y$; with $x$ held fixed it represents a linear operation from $X$ to $Y$ so that $Bx$ is an element of $(X \rightarrow Y)$. Thus for $Bx = L$,

$$||B(x,x')|| = ||Lx'|| \leq M(L)\,||x'|| \tag{2.29}$$

(9, pp.155-156).

An example of a bilinear operation is the integral operation

$$y(s) = \int_0^1 \int_0^1 k(s,t,v)\, x_1(t)\, x_2(v)\, dt\, dv\,, \tag{2.30}$$

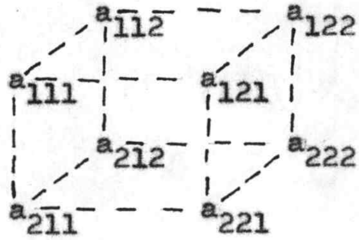
where the space $C$ is the domain and the range of $Bx_1x_2$. Another example is the bilinear operation from $X = R^n$ to $Y = R^m$ with the norm $||x|| = \max\limits_{(i)} |\xi_i|$. The operation $y = Bxx'$ takes the form

$$y_k = \sum_i^n \sum_j^n a_{ijk}\, \xi_i\, \xi'_k\,, \quad (k = 1, \ldots m)\,, \tag{2.31}$$

where $B$ is a three-dimensional matrix $(a_{ijk})$.

In particular, consider the bilinear operation $y = Bxx'$ where $x = (\xi_1, \xi_2)$, $x' = (\xi'_1, \xi'_2)$ and $B$ is the three-dimensional matrix

(2.32)




$$= \left(\begin{array}{cc|cc} a_{111} & a_{112} & a_{121} & a_{122} \\ a_{211} & a_{212} & a_{221} & a_{222} \end{array}\right) .$$

Now

(2.33) $$Bx = \left(\begin{array}{cc|cc} a_{111} & a_{112} & a_{121} & a_{122} \\ a_{211} & a_{212} & a_{221} & a_{222} \end{array}\right)\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} =$$

$$\begin{pmatrix} a_{111}\xi_1 + a_{112}\xi_2 & a_{121}\xi_1 + a_{122}\xi_2 \\ a_{211}\xi_1 + a_{212}\xi_2 & a_{221}\xi_1 + a_{222}\xi_2 \end{pmatrix}$$

so that

(2.34) $$Bxx' = \begin{pmatrix} a_{111}\xi_1\xi_1' + a_{112}\xi_2\xi_1' + a_{121}\xi_1\xi_2' + a_{122}\xi_2\xi_2' \\ a_{211}\xi_1\xi_1' + a_{212}\xi_2\xi_1' + a_{221}\xi_1\xi_2' + a_{222}\xi_2\xi_2' \end{pmatrix} =$$

$$(\eta_1, \eta_2) = y,$$

where

(2.35) $$\eta_k = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ijk}\xi_i\xi_j' , \qquad k = 1, 2.$$

Note that $Bxx' \neq Bx'x$ unless B is symmetric or $x = x'$. In general, for an operation B( )( ) of two places, a permutation on the places gives different results, unless certain conditions on symmetry are satisfied. In what follows, the notation $B^-$ will mean a permutation on the places of the operator B; for $Bx_1x_2$, $Bx_2x_1 = B^-x_1x_2$.

For example, consider the bilinear operations

$$(2.36) \qquad y = Bx_1x_2; \qquad y(s) = \int_a^b \int_a^b k(s,t,v)x_2(t)x_1(v)dtdv \ .$$

If the kernel $k(s,t,v)$ is symmetric with respect to $t$ and $v$, then

$$(2.38) \qquad Bx_1x_2 = B^{-}x_1x_2 \ .$$

If $k(s,t,v)$ is not symmetric, then $Bx_1x_2 \neq B^{-}x_1x_2$. The kernel $k(s,t,v) = \frac{s+tv}{s-tv}$ would satisfy (2.38).

## THE DIFFERENTIAL AND DERIVATIVE

In extending the differential to abstract spaces, it was desirable to ascertain what is essential in the notion of the differential. For this, Frechet turned to J. Hadamard who stated: "The differential of a function is a linear function of the differentials of the variables" (4, p.295). To this as a starting point, Frechet added that the differential of a function ought to be a simple expression, approximate to the increment of the function. By simple he meant the linearity of the expression and by approximate he meant that the differential of the function is the principal part of the increment of the function, that is, the differential and the increment should differ by infinitesimals of higher order.

For instance, the differential of the real function $z = F(x,y)$,

$$dz = \frac{\partial z}{\partial y} dx + \frac{\partial z}{\partial y} dy \, , \tag{3.1}$$

is a linear function of the differentials $dx$ and $dy$. The increment of $z$ is

$$\Delta z = \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y}\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y \, , \tag{3.2}$$

where $\epsilon_1$ and $\epsilon_2$ tend to zero with $\Delta x$ and $\Delta y$. The principal part of $\Delta z$ is defined as $\frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y$ .

The definition of the differential for numerical functionals was the first step in the extension of the differential by Frechet to abstract spaces. The concept of variation of a functional in the calculus of variations played an important role in the formulation of the definition. Consider the functional $F = F(x,y,y')$ which depends on $y(x)$ and $y'(x)$ for a fixed value of $x$. If $y(x)$ is replaced by $y(x) + \epsilon\eta(x)$, where $\epsilon\eta(x) = \delta y$ is called the variation of $y(x)$, the corresponding change in $F$, for a fixed $x$, is

$$\Delta F = F(x,y+\epsilon\eta, y'+\epsilon\eta') - F(x,y,y') \; . \tag{3.3}$$

Expanding the right side of (3.3) in powers of $\epsilon$ and neglecting terms with powers of $\epsilon$ higher than the first, one obtains the first order approximation

$$\delta F = \frac{\partial F}{\partial y}\epsilon\eta + \frac{\partial F}{\partial y'}\epsilon\eta' = \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' \; , \tag{3.4}$$

called the variation of $F$ (5, pp.130-131). This is analogous to the notation of the differential calculus as portrayed by (3.1). Consequently, if the variation, $\delta F(y,\Delta y)$, of a differentiable functional $F(y)$ is a linear functional of the increment $\Delta y = \delta y$, one has the familiar expression

$$F(y+\Delta y) - F(y) = \delta F(y,\Delta y) + h\,\phi(\Delta y) \; , \tag{3.5}$$

where $\phi(\Delta y)$ denotes the distance of $\Delta y$ from the function $\widetilde{\Delta y} = 0$, and $h$ goes to zero with $\phi(\Delta y)$.

Fréchet, after a more rigorous analysis, of course, formulated the following definition of the differential for numerical functionals.

The functional $Px$ allows a differential for a given $x_0$ of the function $x$ if there exists a functional $L\Delta x$, linear with respect to the increment $\Delta x$, such that $|\Delta Px_0 - L\Delta x|$ is infinitely small with respect to the number which measures the difference between the elements $x_0$ and $x_0 + \Delta x_0$. The differential of $Px$ at $x_0$ will be $L\Delta x$ (4, p.296).

Without modification, Fréchet extended this definition to operations on normed linear spaces to normed linear spaces as follows:

An operation $y = Px$, linear or otherwise, mapping space $X$ into space $Y$ is (once) differentiable for a given value of $x$ if there exists a linear operator $L$ an element of $(X \to Y)$ such that

$$||P(x+\Delta x) - Px - L\Delta x|| = o(||\Delta x||) \, , \tag{3.6}$$

where the notation $o(\ )$ means that the function $o(||\Delta x||)$ is infinitely small with respect to $||\Delta x||$.

To particularize, let spaces $X$ and $Y$ each be the space $R$ of real numbers. Then $y = P(x)$ is a real function of a real variable and the definition of the abstract differential coincides with the classical definition of the differential (2, pp.59-60), that is,

$$P(x+\Delta x) - P(x) - P'(x)\Delta x = o(\Delta x) \, . \tag{3.7}$$

For $Px$ differentiable at a given $x$, the linear operator $L$ defined by (3.6) is unique. For suppose the existence of another linear operator $H$, distinct from $L$, which fulfills the same role as $L$ for a given $x$. Let $x' = \dfrac{\Delta x}{||\Delta x||}$ and consider the relation

$$(3.8) \qquad ||L\Delta x - H\Delta x|| = ||\Delta x|| \cdot ||Lx' - Hx'|| .$$

The right-hand side being independent of $||\Delta x||$ goes to zero with $||\Delta x||$. Consequently, $L\Delta x = H\Delta x$.

The linear operation $L\Delta x$ is called the <u>first Fréchet differential</u> of $Px$ for a given $x$ and will be denoted by $d_xP$. The definition of the second differential follows quite naturally.

An operation $y = Px$ is <u>twice differentiable</u> if: 1) $P$ has a first differential defined everywhere in a sphere $S$ about $x$: 2) there exists a <u>bilinear</u> operator $B$ such that for all $(x+\Delta x)$ in $S$

$$(3.9) \qquad ||d_{x+\Delta x}P - d_xP - B\Delta x|| = o(||\Delta x||) .$$

The linear operation $B\Delta x$ is called the <u>second Fréchet differential</u> of $Px$ and will be denoted by $d_x^2 P$.

In general, an operation $y = Px$ is <u>n-differentiable</u> if: 1) $P$ is $(n-1)$-differentiable everywhere in a sphere $S$ about $x$: 2) there exists an <u>n-linear</u> operator $N$ such that

$$(3.10 \qquad ||d^{n-1}_{x+\Delta x} P - d_x^{n-1} P - N\Delta x|| = o(||\Delta x||) .$$

The linear operation $N\Delta x$ is called the <u>$n^{th}$ Fréchet differential</u> of $Px$ and will be denoted by $d_x^n P$.

From the definition of the differential, it follows that if the operation $y = Px$ is differentiable at $x_o$ it is continuous at $x_o$. This is readily discerned by rewriting (3.6) in the form

$$(3.11) \qquad ||P(x_o+\Delta x_o) - Px_o|| \leq M(L)||\Delta x_o|| + o(||\Delta x_o||)$$

and letting $||\Delta x_o|| \to o$ .

Consider again the differentials $d_x P = L\Delta x$, $d_x^2 P = B\Delta x$ and $d_x^n P = N\Delta x$ defined by (3.6), (3.9) and (3.10) respectively. The notion of the <u>derivative</u> is apparent when L, B and N are called respectively the <u>first</u>, <u>second</u> and <u>$n^{th}$ Fréchet derivatives</u> (7, pp.73-74 and 9, p.159). They will be denoted by $D_x P$, $D_x^2 P$, and $D_x^n P$. Thus, in the meaning of differential calculus one has

(3.12) $$d_x P = D_x P\Delta x$$

(3.13) $$d_x^2 P = D_x^2 P\Delta x$$

and

(3.14) $$d_x^n P = D_x^n P\Delta x \ .$$

The notation $D_x P\Delta x$ shall be understood to mean that $D_x P$ is operating on the element $\Delta x$. Sometimes, during the course of a discussion involving Px, the argument of P may become, say, z, then the derivative of Pz will be displayed as $D_z P_z$ .

## RULES FOR DIFFERENTIATION

The analogy implied in this section between the rules for abstract differentiation and those for differentiation of real functions is intended for intuitive purposes only. The analogy will not hold in many instances because of the nature of the elements and the operations on them. However, the rules for real functions are special cases of the rules for normed linear spaces.

THEOREM 1. <u>The Fréchet derivative of a constant operation is null</u>.

PROOF: Let $Px = k$, where $k$ is a constant element. Then

$$P(x+\Delta x) - Px = 0 \ .$$

Set $V\Delta x = 0$ so that

$$||P(x+\Delta x) - Px - V\Delta x|| = ||0|| = o(||\Delta x||).$$

Thus by (3.6) $d_x P = V\Delta x = 0$ and

$$D_x P = 0 \ . \tag{4.1}$$

In the case real functions, $D_x(c) = o$ where $c$ is a constant.

THEOREM 2. <u>The Fréchet derivative of a linear operation is the operator itself</u>.

PROOF: Let $y = Lx$ be a linear operation from $X$ to $Y$. Then by (2.9)

$$L(x+\Delta x) - Lx = L\Delta x \ .$$

By (3.6) it follows immediately that

(4.2) $$D_x L = L \ .$$

Furthermore, since $D_{x+\Delta x}L = D_x L = L$, by theorem (1)

(4.3) $$D_x^2 L = 0 \ .$$

In the case of real functions, $D_x(cx) = c$ and $D_x^2(cx) = o$ .

THEOREM 3. Given that $P$ and $Q$ are differentiable at $x$, the Fréchet derivative of a sum operation $y = (P+Q)x$ is the sum of the derivatives of the operations $Px$ and $Qx$ .

PROOF: Let $y = (P+Q)x$ be a sum operation from $X$ to $Y$. By definition, $(P+Q)x = Px + Qx$ and $(P+Q)(x+\Delta x) = P(x+\Delta x) + Q(x+\Delta x)$. Set $V\Delta x = (d_x P + d_x Q)$ so that

$$\begin{aligned} &||(P+Q)(x+\Delta x) - (P+Q)x - V\Delta x|| \\ &\qquad = ||P(x+\Delta x) - Px - d_x P + Q(x+\Delta x) - Qx - d_x Q|| \\ &\qquad \leq ||P(x+\Delta x) - Px - d_x P|| + ||Q(x+\Delta x) - Qx - d_x Q|| \\ &\qquad = o(||\Delta x||) + o(||\Delta x||) \\ &\qquad = o(||\Delta x||) \ . \end{aligned}$$

Thus by (3.6) $d_x(P+Q) = V\Delta x = d_x P + d_x Q$ and

(4.4) $$D_x(P+Q) = D_x P + D_x Q \ .$$

If $P$ and $Q$ are linear operators, it follows from theorem (2) that

(4.5) $$D_x(P+Q) = P + Q \ .$$

In the case of real functions, $D_x[f(x)+g(x)] = D_xf(x) + D_xg(x)$ .

THEOREM 4. Given that $z = Qy$ and $y = Px$ are differentiable operations at $y$ and $x$ respectively, the Fréchet derivative of the product operation $z = (QP)x$ is $D_yQD_xP$ .

PROOF: By hypothesis, $Q$ is differentiable at $y$ so that

$$||Q(y+\Delta y) - Qy - D_yQ\Delta y|| = o(||\Delta y||) .$$

Set $\Delta y = P(x+\Delta x) - Px$. Then

$$\begin{aligned} ||\Delta y|| &= ||P(x+\Delta x) - Px - D_xP\Delta x + D_xP\Delta x|| \\ &\leq ||P(x+\Delta x) - Px - D_xP\Delta x|| + ||D_xP\Delta x|| \\ &= o(||\Delta x||) + M(D_xP) \cdot ||\Delta x|| . \end{aligned}$$

Thus to infinitesimals of higher order

$$o(||\Delta y||) = o(||\Delta x||) .$$

Now set $V\Delta x = D_yQD_xP\Delta x$ so that

$$\begin{aligned} &||QP(x+\Delta x) - QPx - V\Delta x|| \\ &= ||QP(x+\Delta x) - QPx - D_yQD_xP\Delta x|| \\ &= ||QP(x+\Delta x) - QPx - D_yQ[P(x+\Delta x) - Px] + D_yQ[P(x+\Delta x) - Px - D_xP\Delta x]|| \\ &= ||QP(x+\Delta x) - QPx - D_yQ\Delta y + D_yQ[P(x+\Delta x) - Px - D_xP\Delta x]|| \\ &\leq ||Q(y+\Delta y) - Qy - D_yQ\Delta y|| + ||D_yQ[P(x+\Delta x) - Px - D_xP\Delta x]|| \\ &= o(||\Delta y) + M(D_yQ) \cdot o(||\Delta x||) \\ &= o(||\Delta x||) . \end{aligned}$$

Thus $d_xQP = V\Delta x = D_yQD_xP\Delta x$ and

$$D_xQP = D_yQD_xP . \tag{4.6}$$

This theorem is the analogue of the <u>chain rule</u> for real functions in the ordinary differential calculus, that is, $D_x G[f(x)] = D_y G(y) D_x f(x)$.

If $Q$ is a linear operator, one has the following corollary to theorem (4).

COROLLARY 1. <u>Given that</u> $z = Qy$ <u>and</u> $y = Px$ <u>are differentiable operations at</u> $y$ <u>and</u> $x$ <u>respectively and</u> $Q$ <u>is a linear operator, the Fréchet derivative of the product</u> $z = QPx$ is $QD_xP$.

PROOF: From theorems (2) and (4) it follows readily that

$$D_x QP = QD_x P \;. \tag{4.7}$$

In the case of real functions, $D_x[cf(x)] = cD_x f(x)$.

THEOREM 5. <u>Given a bilinear operation</u> $y = Bx_1x_2$ <u>on</u> $X \times X$ <u>to</u> $Y$, <u>the Fréchet differential is</u> $Bx_1\Delta x_2 + B^- x_2\Delta x_1$ <u>and the Fréchet derivative is</u> $Bx_1(\;) + B^- x_2(\;)$.

PROOF: Since $B$ is linear with respect to its places,

$$B(x_1+\Delta x_1)(x_2+\Delta x_2) = Bx_1x_2 + B\Delta x_1x_2 + Bx_1\Delta x_2 + B\Delta x_1\Delta x_2$$

and

$$B(x_1+\Delta x_1)(x_2+\Delta x_2) - Bx_1x_2 = Bx_1\Delta x_2 + B^- x_2\Delta x_1 + B\Delta x_1\Delta x_2 \;.$$

Set $V\Delta x_1\Delta x_2 = Bx_1\Delta x_2 + B^- x_2\Delta x_1$ so that

$$\begin{aligned} ||B(x_1+\Delta x_1)(x_2+\Delta x_2) - Bx_1x_2 - V\Delta x_1\Delta x_2|| &= ||B\Delta x_1\Delta x_2|| \\ &\leq ||B||\;||\Delta x_1||\;||\Delta x_2|| \\ &= o(||\Delta x_1|| \cdot ||\Delta x_2||) \;. \end{aligned}$$

Thus it follows that

$$d_{x_1x_2}B = Bx_1\Delta x_2 + B^- x_2\Delta x_1 \tag{4.8}$$

and

(4.9) $$D_{x_1x_2}B = Bx_1(\ ) + B^{-}x_2(\ ) .$$

Furthermore, if $x_1 = x_2$ ,

(4.10) $$d_xB = (B+B^{-})\ x\Delta x$$

(4.11) $$D_xB = (B+B^{-})\ x$$

and

(4.12) $$D_x^2\ B = (B+B^{-}) .$$

In the case of real functions, a specialization of (4.8) would be $d(cxy) = cydx+cxdy$ likewise, for $y = x^2$, (4.10) specializes to $d(x^2) = 2xdx$.

THEOREM 6. *Let* $y_1 = Hx$ *and* $y_2 = Qx$ *be two differentiable operations at* x. *If* $y = By_1y_2$ *is a bilinear operation from* Y x Y *to* Y, *then*

$$d_xB_{y_1y_2} = B(Hx)(D_xQ\Delta x) + B^{-}(Qx)(D_xH\Delta x) .$$

PROOF: From (4.9)

(4.13) $$d_{y_1y_2}B = By_1\Delta y_2 + B^{-}y_2\Delta y_1 .$$

By hypothesis

$$\Delta y_1 = H(x+\Delta x) - Hx = D_xH\Delta x + o(\Delta x) ,$$
$$\Delta y_2 = Q(x+\Delta x) - Qx = D_xQ\Delta x + o(\Delta x) .$$

Substituting the last two expressions, into (4.13), it is readily discerned that

$$d_x B_{y_1 y_2} = B(Hx)(D_x Q\Delta x) + B^-(Qx)(D_x Hx) \tag{4.14}$$

as $||\Delta x|| \to o$. Furthermore, if $H$ and $Q$ are linear operators,

$$d_x B_{y_1 y_2} = B(Hx)(Q\Delta x) + B^-(Qx)(H\Delta x) \ . \tag{4.15}$$

In the case of real functions, $d_x[f(x)g(x)] = g(x)\, D_x f(x)dx + f(x)\, D_x g(x)dx$ .

THEOREM 7. <u>Let</u> $y = N\overbrace{x \ldots x}^{n}$ <u>be an n-linear operation on</u> $X$ <u>to</u> $Y$. <u>If</u> $N$ <u>is differentiable at</u> $x$, <u>then</u>

$$d_x N = N_1\overbrace{x \ldots x}^{n-1}\Delta x + N_2\overbrace{x \ldots x}^{n-1}\Delta x + \ldots + N_n\overbrace{x \ldots x}^{n-1}\Delta x \ .$$

(The subscripts on $N$ indicate the places, from left to right, that $\Delta x$ occupies).

PROOF: Since $N$ is linear,

$$N\overbrace{(x+\Delta x)\ldots(x+\Delta x)}^{n} - N\overbrace{x \ldots x}^{n} = N_1\overbrace{x \ldots x}^{n-1}\Delta x + \ldots + N_n\overbrace{x \ldots x}^{n-1}\Delta x + (\text{terms involving } \Delta x \text{ in more than one place}) \ .$$

Set $V\overbrace{\Delta x \ldots \Delta x}^{n} = N_1\overbrace{x \ldots x}^{n-1}\Delta x + \ldots + N_n\overbrace{x \ldots x}^{n-1}\Delta x$ so that

$$||N\overbrace{(x+\Delta x)\ldots(x+\Delta x)}^{n-1} - N\overbrace{x \ldots x}^{n} - V\overbrace{\Delta x \ldots \Delta x}^{n}\,|| \leq |||N||\cdot||\text{ terms involving } \Delta x \text{ in more than one place}|| = o(||\Delta x||) \ .$$

Thus $d_x N = V\overbrace{\Delta x \ldots \Delta x}^{n}$ and

$$d_x N = N_1\overbrace{x \ldots x}^{n-1}\Delta x + N_2\overbrace{x \ldots x}^{n-1}\Delta x + \ldots + N_n\overbrace{x \ldots x}^{n-1}\Delta x \ . \tag{4.16}$$

Consequently, if $N$ is a symmetric operator, (4.17) $$d_x N = nN\overbrace{x \ldots x}^{n-1}\Delta x \ ,$$

and

(4.18) $$D_x N = nN\overbrace{x \ldots x}^{n-1} \ .$$

In the case of real functions, $D_x(x^n) = n\, x^{n-1}$ .

THEOREM 8. Given that $y = Px$ is an operation from $X$ to $Y$, and the inverse operation $x = Qy$ from $Y$ to $X$ exists, then $D_x P = (D_y Q)^{-1}$ .

PROOF: By hypothesis

(4.19) $$Ix = QPx$$

where $I$ is the identity operator. From (4.19) and theorem (4)

$$D_x I = I = D_y Q D_x P \ .$$

If $D_y Q \neq 0$, it follows that

(4.20) $$D_x P = (D_y Q)^{-1}$$

Furthermore, if $P$ is linear then $Q$ is linear, as mentioned earlier, and it follows from theorem (2) that

(4.21) $$P = Q^{-1} \ .$$

as it should be.

EXAMPLES

(1) Consider an operation $y = Px$ mapping $X = R^n$ into $Y = R^m$ where $x = (\xi_1, \dots \xi_n)$ and $\eta_j = F_j(\xi_1, \dots \xi_n)$; $j = 1, \dots m$. The operation will be differentiable at $x$ if the functions $F_j$ are differentiable at $x$. To infinitesimals of higher order

$$(5.1) \qquad \Delta y = \{\Delta \eta_j\}; \quad \Delta \eta_j = \sum_{k=1}^{n} \partial_{\xi_k} F_j d\xi_k (j=1,\dots m; k=1,\dots n) .$$

The Fréchet differential of $Px$ is seen to be represented by the <u>matrix product</u>

$$(5.2) \qquad d_x P = (\mathcal{D}_{\xi_k} F_j)\,(d\,\xi_k) ,$$

and the Fréchet derivative is the transformation matrix

$$(5.3) \qquad D_x P = (\mathcal{D}_{\xi_k} F_j), \ (j=1, \dots m;\ k=1, \dots n),$$

where $\mathcal{D}_{\xi_k} F_j = \dfrac{\partial F_j}{\partial \xi_k}$ .

(2) Consider the bilinear integral operation

$$(5.4) \qquad y(s) = Bx_1x_2 = \int_0^1 \int_0^1 K(s,t,v)\, x_1(t)x_2(v)dvdt$$

on $C$ to $C$. Applying rules (4.8) and (4.9) of theorem 5, the Fréchet differential and derivative of (5.4) are

$$(5.5) \qquad d_{x_1x_2}B = \int_0^1 \int_0^1 K(s,t,v)\,\Delta x_1(t)x_2(v)dvdt + \int_0^1 \int_0^1 K(s,t,v)x_1(t)\Delta x_2(v)dvdt$$

and

$$(5.6)\quad D_{x_1x_2}B = \int_0^1\int_0^1 K(s,t,v)(\ )x_2(v)dvdt + \int_0^1\int_0^1 K(s,t,v)x_1(t)(\ )dvdt\ .$$

(3) Suppose that in example (2) $x_1 = Hx = x^2$ and $x_2 = Qx = x^3$ . Then use of (4.14) theorem 6 gives

$$(5.7)\quad d_xB_{x_1x_2} = 2\int_0^1\int_0^1 K(s,t,v)x(t)\Delta x(t)x^3(v)dvdt + 3\int_0^1\int_0^1 K(s,t,v)x^2(t)x^2(v)\Delta x(v)dvdt\ .$$

(4) Consider the non-linear integral operation

$$(5.8)\qquad y(s) = Px = \int_0^1 K(s,t)[x(t)]^2\,dt$$

from C to C. Now

$$(5.9)\quad P(x+\Delta x) - Px = \int_0^1 K(s,t)[x(t)+\Delta x(t)]^2dt - \int_0^1 K(s,t)[x(t)]^2\,dt$$

$$= 2\int_0^1 K(s,t)x(t)\Delta x(t)dt + \int_0^1 K(s,t)[\Delta x(t)]^2\,dt\ .$$

Using definition (3.6) and setting $L\Delta x = 2\int_0^1 K(s,t)x(t)\Delta x(t)dt$ ,

$$(5.10)\quad ||P(x+\Delta x) - Px - L\Delta x|| = ||\int_0^1 K(s,t)[\Delta x(t)]^2dt\ ||$$

$$\leq M(K)||\Delta x||^2$$

$$= o(||\Delta x||)\ .$$

Thus

$$(5.11)\qquad d_xP = 2\int_0^1 K(s,t)x(t)\Delta x(t)\,dt\ ,$$

$$D_x P = 2 \int_0^1 K(s,t)x(t)(\ )dt \ . \tag{5.12}$$

These results are readily obtainable by treating (5.8) as a <u>product operation</u> on x as follows. Let

$$y = Lz = \int_0^1 K(s,t)z(t)dt \tag{5.13}$$

be a linear operation, and

$$z = Px = [x(t)]^2 \tag{5.14}$$

a non-linear operation. Then (5.8) is equivalent to the product operation

$$y(s) = LPx = \int_0^1 K(s,t)\,[x(t)]^2 dt \ . \tag{5.15}$$

Rule (4.7) of corollary 1 gives immediately

$$LD_x P = 2 \int_0^1 K(s,t)x(t)(\ )dt \ . \tag{5.16}$$

The second Fréchet derivative of (5.8) is seen to be the bilinear operator

$$D_x^2 P = 2 \int_0^1 K(s,t)(\ )(\ )\,dt \ . \tag{5.17}$$

(5) Consider the non-linear integral operation

$$x(s) = 1 + x(s) \int_0^1 \frac{\phi(t)}{s+t} \quad x(t)dt \tag{5.18}$$

in C. This is the Chandrasekhar non-linear integral equation which arises in astrophysics in connection with radiative transfer. The function $\phi(t)$ is determined by the type of scattering being considered. Equation (5.18) may be written

$$Px = Ix - Bxx - 1 . \tag{5.19}$$

Applying rules (4.1), (4.2) and (4.11) of theorems 1, 2 and 5 respectively, the Fréchet derivative of (5.14) becomes

$$D_x P = I - (B + B^{-})x . \tag{5.20}$$

Applying rules (4.3) and (4.12), the second Fréchet derivative becomes

$$D_x^2 P = -(B + B^{-}) . \tag{5.21}$$

These Fréchet derivatives were used in Newton's method to arrive at a numerical solution to (5.18). This was carried out on a research project at Oregon State College.

MEAN VALUE THEOREMS

The following theorem is an abstract analogue to the classical mean value theorem of the differential calculus (9, pp.161-162).

THEOREM 9. Let $Px$ be an operation, linear or otherwise, from $X$ to $Y$. If $P$ is a differentiable operator,

$$||P(x+\Delta x) - Px|| \leq \max_{0 \leq \theta \leq 1} ||D_\xi P_\xi|| \, ||\Delta x||, \quad \xi = (x+\theta\Delta x).$$

PROOF: Set $P(x+\Delta x) - Px = y$. Let there exist in space $Y$ a linear functional $T$ (1, p.55 and 16, pp.144-148) such that

$$Ty = ||y||; \quad ||T|| = 1 . \tag{6.1}$$

Consider the real function of a real variable t,

$$f(t) = T\,P(x+t\Delta x) = T\ P\,\xi' \ . \tag{6.2}$$

Using theorem 4 for a chain of three operations, (6.2) takes the form

$$f'(t) = T\,D_{\xi'}\,P_{(x+t\Delta x)}\Delta x = T\,D_{\xi'}\,P_{\xi'}\,\Delta x \ . \tag{6.3}$$

From $y = P(x+\Delta x) - Px$ and (6.2),

$$Ty = T[P(x+\Delta x) - Px] = f(1) - f(o) \ , \tag{6.4}$$

and applying the law of finite increments for real functions to (6.4)

$$f(1) - f(o) = f'(\theta) = T\,D_\xi P_{(x+\theta\Delta x)}\Delta x = T\,D_\xi\,P_\xi\,\Delta x \ . \tag{6.5}$$

Thus from (6.1), (6.4) and (6.5) it follows that

$$(6.6)\quad ||P(x+\Delta x) - Px|| \leq M(T)\ ||D_\xi P_\xi||\ ||\Delta x|| \leq \max_{o\leq\theta\leq 1} ||D_\xi P_\xi||\ ||\Delta x|| \ .$$

The abstract analogue to Taylor's formula follows quite naturally (9, pp.162-163).

THEOREM 10. *Let* $Px$ *be an operation, linear or otherwise, from* $X$ *to* $Y$. *If* $P$ *is a twice Fréchet differentiable operator,*

$$||P(x+\Delta x) - Px - D_x P\Delta x|| \leq \frac{1}{2} \max_{o\leq\theta\leq 1} ||D^2_\xi P_\xi||\ ||\Delta x||^2, \quad \xi = (x+\theta\Delta x) \ .$$

PROOF: Set $P(x+\Delta x) - Px - D_x P\Delta x = y$. Again select a linear functional $T$ such that

$$(6.7)\qquad Ty = ||y||;\ ||T|| = 1 \ .$$

Consider the real function of a real variable $t$,

$$(6.8)\qquad f(t) = T\,P(x+t\Delta x) = T\,P\,\xi' \ .$$

The first and second Fréchet derivatives of (6.8) are

$$(6.9)\qquad f'(t) = T\,D_{\xi'}\,P_{(x+t\Delta x)}\Delta x = T\,D_{\xi'}\,P_{\xi'}\,\Delta x \ ,$$

$$(6.10)\qquad f''(t) = T\,D^2_{\xi'}\,P_{(x+t\Delta x)}\Delta x\Delta x = T\,D_{\xi'}P_{\xi'}\Delta x\Delta x \ ,$$

where $D^2_{\xi'}\ P_{\xi'}$ is a bilinear operator defined by (3.9) .

From $y = P(x+\Delta x) - Px - D_x P\Delta x$ and (6.8),

$$(6.11)\qquad Ty = T[P(x+\Delta x) - Px - D_x P] = f(1) - f(o) - f'(o) \ ,$$

and applying Taylor's formula for real functions to (6.11)

$$(6.12) \qquad f(1) - f(o) - f'(o) = \frac{1}{2} f''(\theta) = \frac{1}{2} T\, D^2_\xi P_\xi \Delta x \Delta x \,.$$

Thus from (6.7), (6.11) and (6.12) it follows that

$$(6.13) \quad \|P(x+\Delta x) - Px - DxP\Delta x\| \leq \frac{1}{2} M(T) \, \|D^2_\xi P_\xi\| \, \|\Delta x\|^2$$

$$\leq \frac{1}{2} \max_{o \leq \theta \leq 1} \|D^2_\xi P_\xi\| \, \|\Delta x\|^2 \,.$$

Consequently, if $P$ is n-differentiable,

$$(6.14) \quad \|P(x+\Delta x) - Px - D_x P\Delta x - \frac{1}{2!} D^2_x P\Delta x\Delta x - \dots - \frac{1}{(n-1)!} D^{n-1}_x P \overbrace{\Delta x \dots \Delta x}^{n}\|$$

$$\leq \frac{1}{n!} \max_{o \leq \theta \leq 1} \|D^n_\xi P_\xi\| \, \|\Delta x\|; \quad \xi = x + \theta\Delta x \,.$$

Newton's method for approximations of solutions of the real function

$$(6.15) \qquad f(x) = o,$$

takes the form

$$(6.16) \qquad x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})} \,,$$

where $x_1, x_2, \dots, x_n$ are successive approximations to $x$. Kantorovich (9, pp.166-198) has demonstrated that Newton's process can be extended to the operation

$$Px = o, \tag{6.17}$$

where $P$, an operator on $X$ to $Y$, is twice differentiable. If the derivative $D_x P$ has an inverse $(D_x P)^{-1}$, the analogue to (6.16) is

$$x_n = x_{n-1} - (D_x P)^{-1} P x_{n-1} \, . \tag{6.18}$$

The conditions for convergence of (6.18) to the exact solution, and applications of Newton's method can be found in Kantorovich's paper.

BIBLIOGRAPHY

1. Banach, Stefan. Théorie des opérations linéaries. Warsaw, Subweniji funduszu kultury narodowej, 1932. 217p.

2. Courant, Richard. Differential and integral calculus. Vol. 2. New York, Nordemann, 1936. 682p.

3. Courant, Richard and David Hilbert. Methods of mathematical physics. Vol. 1. New York, Interscience, 1953. 561p.

4. Fréchet, Maurice. La notion de différentielle dans l'analyse générale. Annales scientifiques de l'école normale supérieure 42:293-323. 1925.

5. Hildebrand, F. B. Methods of applied mathematics. New York, Prentice-Hall, 1952. 523p.

6. Hildebrandt, T. H. and L. M. Graves. Implicit functions and their differentials in general analysis. Transactions of the American mathematical society 29:127-153. 1927.

7. Hille, Einar. Functional analysis and semi-groups. New York, American mathematical society, 1948. 528p.

8. Hyers, D. H. A generalization of Fréchet's differential. Proceedings of the national academy of sciences 27:315-316. 1941.

9. Kantorovich, L. V. Functional analysis and applied mathematics (Russian). Uspehi matematiceskih nauk 3:89-185. 1948. (Translation in NBS report 1509, Los Angeles, national bureau of standards, 1952. 202p.)

10. Kerner, M. Die differentiale in der Allgemeinen Analysis. Annals of mathematics 34:546-572. 1933.

11. Michal, A. D. General differential geometries and related topics. Bulletin of the American mathematical society 45:529-563. 1939.

12. Michal, A. D. Differential calculus in linear topological spaces. Proceedings of the national academy of sciences 24:340-342. 1938.

13. Michal, A. D. Differential equations in Fréchet differentials occurring in integral equations. Proceedings of the national academy of sciences 31:252-258. 1945.

14. Michal, A. D. The Fréchet differentials of regular power series in normed linear spaces. Duke mathematical journal 13:57-59. 1946.

15. Rothe, E. H. Gradient mappings. Bulletin of the American mathematical society 59:5-19. 1953.

16. Zaanen, Adrian Cornelis. Linear analysis. New York, Interscience, 1953. 600p.

17. Zorn, Max A. Derivatives and Fréchet differentials. Bulletin of the American mathematical society 52:133-137. 1946.