1966

Comparative Definitions of the Derivative

Richard Francis Barry
College of William & Mary - Arts & Sciences

COMPARATIVE DEFINITIONS OF THE DERIVATIVE

A Thesis

Presented to

The Faculty of the Department of Mathematics

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

By

Richard Francis Barry, Jr.

1966

APPROVAL SHEET

This thesis is submitted in partial fulfillment of

the requirements for the degree of

Master of Arts

Richard V. Jamny V.

Author

Approved, May 1966

Thomas L. Reynolds

Thomas L. Reynolds, Ph.D

Benjamin R Cato, Jr.

Benjamin R. Cato, Jr., M.A.

Hugh B. Easler

Hugh B. Easler, M.S.

ACKNOWLEDGMENTS

The writer wishes to express his appreciation to Assistant Professor Hugh B. Easler, under whose guidance this survey was conducted, for his patient guidance, gentle criticism, and the generous giving of his time throughout the survey. The author is also indebted to Professor Thomas L. Reynolds and Associate Professor Benjamin R. Cato, Jr. for their careful reading and helpful criticism of the manuscript.

COMPARATIVE DEFINITIONS

OF THE

DERIVATIVE

TABLE OF CONTENTS

ABSTRACT

The purpose of this paper is to examine some of the different definitions of derivatives found in contemporary mathematical literature, to compare them, and to verify the statements in some cases.

Since the topological derivative is defined as a special limit of a function at a point, all such definitions are essentially the same. A survey of the various mathematical society periodicals, as well as a search of the majority of available mathematical titles, gives several definitions of the abstract algebraic derivative. That of Bourbaki is considered to be the standard.

The topological and algebraic derivatives are the same only under special circumstances. Algebraic derivatives are defined as endomorphisms or homomorphisms in rings, integral domains, and fields, and in polynomial domains over all three. Certain linear mappings behave like the abstract derivative operator and consequently may be so considered.

INTRODUCTION

Differentiation, or the process of finding the derivative of a function, is a fundamental operation of calculus. Leibniz and Newton invented calculus in the period 1665-1675, apparently working independently of each other. Subsequently, one of the most famous controversies among not only mathematicians, but among scholars in general, arose between Newton and Leibniz and their respective supporters as to the priority of discovery. Since the notation for the derivative played a significant role in the dispute [2]*, we shall examine the notations involved.

Inasmuch as Newton's calculus was oriented to his study of motion, he took the time, t, as an independent variable and the dependent variable as x, calling it the *fluent* [29]. The velocity of the *fluent* he called the *fluxion*, denoting it by x, the derivative with respect to t. The higher derivatives were denoted by $\ddot{x}$, $\dddot{x}$, ... The inverse process, integration, Newton symbolized by x' . These notations, originated in 1665, persisted in England throughout the Eighteenth century [7], and can still be found in British mathematical papers [23].

In 1675, Leibniz abbreviated *omnes lineae* to *omn l* and then to *S l*, to mean the sum of lines. Consequently, the familiar integral sign is derived from the first letter of the word *summa*. At the same time he used the symbol d in the denominator of a variable to denote a difference,

---

* Numbers in brackets refer to bibliography.

and then continued with the now classical notations dx, dy, dy/dx, and the integral $\int y\, dy$. "Perhaps no mathematician has seen more clearly than Leibniz the importance of good notation in mathematics" [8]. His symbols have stood the test of nearly three centuries of mathematical progress.

CHAPTER I

DEFINITIONS OF DERIVATIVES OF POLYNOMIALS

1-1: The derivative is most commonly defined in a polynomial domain. Polynomials are expressions of the form

$$\sum_{i=0}^{n} a_{(i)} x^i,$$

where the a's are from some algebraic system, i is an integer, and x is an indeterminate. We shall examine several different definitions of derivatives of polynomials.

1-2: Bourbaki [5] defines a derivative in the following manner: "Let f be a polynomial of the ring (commutative, with unity element) $A [X_1, X_2, \cdots, X_p] = B$. In the ring $A [X_1, \cdots X_p, Y_1, \cdots Y_p]$ of polynomials with 2 p indeterminates X_i, Y_i $(1 \leq i \leq p)$, let us consider the polynomial $f (X_1 + Y_1, X_2 + Y_2, \cdots +, X_p + Y_p)$; this polynomial can be written as a polynomial in Y_i, with coefficients in B, such that its constant term is $f (X_1, X_2, \cdots , X_p)$.

Then if $\Delta f = f (X_1 + Y_1, \cdots, X_p + Y_p) - f (X_1, \cdots , X_p)$, the polynomial Δf (also written $\Delta f (X_1, \cdots, X_p, Y_1, \cdots Y_p)$ is thus a polynomial of $B [Y_1, Y_2, \cdots, Y_n]$ without a constant term.

Definition 1. The derivative of the polynomial f, denoted Df, is the homogeneous part of the first degree of the polynomial Δf, considered as a polynomial in Y_i, with coefficients in

$B = A [X_1, X_2, \cdots, X_p]$.

From this definition, then, one has $Df = \sum_{i=1}^{p} g_i Y_i$,

where $g_1, g_2, \cdots , g_p$ are the elements of B, that is, of the polynomials of the ring $A[x_1, x_2, \cdots, x_p]$.''

Example: Let $f(x) = x^2 + ax + b \qquad a, b \in A$

Consider the polynomial

$$f(x + y) = (x + y)^2 + a(x + y) + b, \text{ then}$$

$$\Delta f = x^2 + 2xy + y^2 + ax + ay + b - (x^2 + ax + b)$$

$$= 2xy + ay + y^2 = (2x + a)y + y^2$$

The coefficient of y is $2x + a$. Therefore $Df = D(x^2 + ax + b) = 2x + a$.

Consider a polynomial in two variables:

Let $f(x,y) = x^2 y + xy^2 + 4$

then $\Delta f = f(x + x_1, y + y_1) - f(x, y)$

$$= (x + x_1)^2 (y + y_1) + (x + x_1)(y + y_1)^2 + 4 - x^2y - xy^2 - 4$$

$$= x^2y + 2xx_1y + x_1^2 y + x^2 y_1 + 2xx_1 y_1 + x_1^2 y_1$$

$$+ xy^2 + 2xyy_1 + xy_1^2 + x_1 y^2 + 2x_1 y y_1$$

$$+ x_1 y_1^2 - x^2y - xy^2$$

$$= (2xy + y^2) x_1 + (x^2 + 2xy) y_1 + (y + y_1) x_1^2$$

$$+ (2x + 2y) x_1 y_1 + (x + x_1) y_1^2$$

The coefficients of x_1 and y_1, in this case, are the partial derivatives of $f(x,y)$ with respect to x and y respectively, and are denoted

$$\frac{\partial f(x,y)}{\partial x} \quad \text{and} \quad \frac{\partial f(x,y)}{\partial y}$$

Their sum is called the total derivative of $f(x,y)$, denoted

$$D f(x,y) = \frac{\partial f(x,y)}{\partial x} + \frac{\partial f(x,y)}{\partial y}$$

1-3: van der Waerden [31] defines the derivative, without mention of a limit, in a manner similar to Bourbaki, but uses a congruence which simplifies the definition:

Given a commutative ring R, let $f(x)$ be a polynomial in R $[x]$. Then form the polynomial $f(x + h)$ in R $[x,h]$,

$$f(x + h) = f(x) + h\, f_1\,(x) + h^2\, f_2\,(x) + \cdots\cdots$$

where

$$f(x + h) \equiv f(x) + h\, f_1\,(x) \pmod{h^2} \qquad (1.1)$$

The derivative $f(x)$ then, is defined to be the coefficient of h in (1.1) above.

Consequently

$$f(x + h) + g\,(x + h) \equiv f(x) + hf'\,(x) + g(x) + hg'\,(x) \pmod{h^2}$$

whence

$$(f + g)' = f' + g' \qquad (1.2)$$

and

$$f(x + h)\, g(x + h) \equiv [f\,(x) + h\, f'\,(x)]\, [g\,(x) + hg'\,(x)] \pmod{h^2}$$

$$\equiv f\,(x)\, g\,(x) + h\, [f'\,(x)\, g(x) + f\,(x)\, g'\,(x)] \pmod{h^2}$$

whence

$$(fg)' = f'g + fg' \qquad (1.3)$$

The above definition is then used to define the derivative of a rational function.

Given polynomials $f(x)$ and $g(x)$ with coefficients in a field F, let

$$s(x) = \frac{f\,(x)}{g\,(x)}$$

then let

$$s(x + h) - s(x) = \frac{f\,(x + h)}{g\,(x + h)} \qquad \frac{f\,(x)}{g\,(x)}$$

$$= \frac{f(x + h)\, g(x) - f(x)\, g\,(x + h)}{g(x)\, g(x + h)} \qquad (1.4)$$

If $h = 0$, the numerator of (1.4) becomes zero, which implies that h is a factor. Dividing both sides of (1.4) by h, one obtains

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

$$= \frac{q(x,h)}{g(x)\ g(x+h)}$$

The expression on the right is, therefore, a rational function of h, which has a particular value when h = 0, since the denominator does not vanish. The derivative of s(x) is defined to be this particular value, ie,

$$s'(x) = \frac{q(x,0)}{g(x)^2}$$

The value q (x,0) is determined by expressing the numerator of (1.4) in terms of ascending powers of h, dividing by h, then setting h= 0, to obtain $\frac{f(x+h)}{h}\ g(x) - f(x)\ \frac{g(x+h)}{h}$

$$= \frac{1}{h}\ [f(x) + hf_1(x) + h^2 f_2(x) + \cdots + h^n f_n(x)]\ g(x)$$

$$- f(x)\ (\frac{1}{h})\ [g(x) + hg_1(x) + h^2 g_2(x) + \cdots + h^n g_n(x)]$$

$$= [\frac{f(x)}{h} + f_1(x) + h f_2(x) + \cdots + h^{n-1} f_n(x)]\ g(x)$$

$$- f(x)\ [\frac{g(x)}{h} + g_1(x) + h\ g_2(x) + \cdots + h^{n-1} g_n(x)]$$

$$= [f_1(x) + h\ f_2(x) + \cdots + h^{n-1}\ f_n(x)]\ g(x)$$

$$- f(x)\ [g_1(x) + hg_2(x) + \cdots + h^{n-1} g_n(x)]$$

and setting h = 0, we get

$$f_1(x)\ g(x) - f(x)\ g_1(x)$$

However, $f_1(x)$ and $g_1(x)$, are the derivatives of f(x) and g(x), as defined in (1.1).

Therefore,

$$s'(x) = \left(\frac{f(x)}{g(x)}\right)' \quad \frac{f'(x)\ g(x) - f(x)\ g'(x)}{g(x)^2}$$

MacDuffee's [21] definition of a derivative coincides with that given by van der Waerden.

1-4: Schreier and Sperner [30] define the derivative of a polynomial in a different fashion;

Let the polynomial $f(x) = (x-a_1)\ (x-a_2) \cdots (x-a_n)$, $a_i \in F$, a field; define $\frac{f(x)}{x-a_i}$ to be the polynomials $g_i\ (x)$. Then the sum of the polynomials $g_i\ (x)$ is defined to be the derivative of $f(x)$, ie,

$$f'\ (x) = \sum_{i=1}^{n} g_i\ \ (x) \tag{1.5}$$

Example: let $f(x) = x^3 - 6x^2 + 11x - 6 = (x-1)\ (x-2)\ (x-3)$,

then $\frac{f(x)}{x-1} + \frac{f(x)}{x-2} + \frac{f(x)}{x-3} = \frac{(x-1)(x-2)\ (x-3)}{x-1} + \frac{(x-1)\ (x-2)\ (x-3)}{x-2} + \frac{(x-1)\ (x-2)\ (x-3)}{x-3}$ so that $f'(x) = 3\ x^2 - 12\ x + 11$,

Consider the derivative of a constant:

If $f(x) = x$, then $f'(x) = \sum_{i=1}^{n} g_i\ \ (x) = \frac{x}{x} = 1$

If $f(x) = x + c$, c a constant, then $f'(x) = \sum_{i=1}^{n} g_i\ \ (x) = \frac{x+c}{x+c} = 1$

By (1.2) we have $f'\ (x + c) = f'\ (x) + f'\ (c)$

Therefore $1 = 1 + f'\ (c)$ implies that $f'\ (c) = 0$

1-5: Another definition of the derivative of a polynomial over a field is found in Dubreil [12]:

Let $f\ (x) \in K\ [x]$, K a field

Denote $F\ (x,y) = \frac{f(x) - f(y)}{x - y}$, $f(y) \in K\ [x,y]$, and

$$f(x) = \sum_{i=1}^{n} a_i x^i \quad \text{and} \quad f(y) = \sum_{i=1}^{n} a_i y^i$$

Then the derivative of $f(x)$ is defined to be $f'(x) = F\ (x,x)$, where x is set equal to y after $f(x)$ and $f(y)$ have been evaluated and the expression simplified. Note that $(x-y)$ will have been cancelled out before we set $x = y$.

Example:

Let $f(x) = 2x^2 + 3x + 1$

$$F(x,y) = \frac{2x^2 + 3x + 1 - 2y^2 - 3y - 1}{x - y}$$

$$= \frac{2(x^2 - y^2) + 3(x - y)}{x - y}$$

$$= 2(x + y) + 3$$

Let $y = x$ and we get

$$F(x,x) = f'(x) = 4x + 3$$

Also, we have

$$[f(x)\, g(x)]' = \frac{f(x)\, g(x) - f(y)\, g(y)}{x - y} \qquad \text{(from 1.6)}$$

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

$$= \frac{f(x)\, [g(x) - g(y)]}{x - y} + \frac{[f(x) - f(y)]\, g(x)}{x - y}$$

(and from 1.6)

$$= f(x)\, g'(x) + f'(x)\, g(x)$$

1-6: Herstein [15] defines the derivative of a polynomial $f(x) \in F[x]$, F a field, to be $f'(x) = na_0x^{n-1} + (n-1)\, a_1x^{n-2} + \cdots + a_{n-1}$, (1.7) where

$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$, $a \in F$, and gives as lemmas:

(1) $[f(x) + g(x)]' = f'(x) + g'(x)$

(2) $[af(x)]' = af'(x)$

(3) $[f(x)\, g(x)]' = f'(x)\, g(x) + f(x)\, g'(x)$

Proof of (3):

Let $f(x) = a_0 + a_1x + \cdots + amx^m = \sum_{i=0}^{m} a_ix^i$

$$g(x) = b_0 + b_1x + \cdots + b_nx^n = \sum_{j=0}^{n} b_jx^j$$

Then $f'(x) = a_1 + 2a_2x + \cdots + ma_mx^{m-1} = \sum_{i=0}^{m} ia_ix^{i-1}$

$$g'(x) = b_1 + 2\, b_2 x + \cdots + n b_n x^{n-1} = \sum_{j=0}^{i=0} j b_i x^{i-1}$$

Therefore, $f(x)\, g(x) = \displaystyle\sum_{i=0}^{m} \sum_{j=0}^{n} a_i b_j x^{i+j}$

and $[f(x)\, g(x)]' = \displaystyle\sum_{i=0}^{m} \sum_{j=0}^{n} (i + j)\, a_i b_j x^{i+j-1}$

$$= \sum_{i=0}^{m} \sum_{j=0}^{n} i a_i b_j x^{i+j-1} + \sum_{i=0}^{m} \sum_{j=0}^{n} j a_i b_j x^{i+j-1}$$

$$= \sum_{i=0}^{m} i a_i x^{i-1} \sum_{j=0}^{n} b_j x^j + \sum_{i=0}^{m} a_i x^i \sum_{j=0}^{n} j b_j x^{j-1}$$

$= f'(x)\, g(x) + f(x)\, g'(x)$

Herstein observes that if the field F is of characteristic $p \neq 0$ (p a prime), the derivative of the polynomial x^p is $p\, x^{p-1} = 0$. Therefore the usual rule that an element, whose derivative is zero, is a constant is not valid in all fields; it is either a constant or a polynomial in x^p. If, however, the characteristic of the field F is 0, $f(x) \in F[x]$, then $f'(x) = 0$ implies $f(x) = a \in F$.

The treatment of the derivative in polynomials by Albert [1] and Jacobson [17] is essentially the same as that of Herstein.

1-7: Barnes [3] considers the integral domain F [x] (F, any field) as an infinite dimensional vector space over F, defines the derivative as in 1.7), and calls it an F-endomorphism of the vector space F [x], where F is the kernel of D, the derivative mapping.

1-8: Birkhoff and MacLane [4], as well as Jacobson [17], refer to the derivative of polynomial rings as the "formal derivative", presumably in contradistinction to the topological derivative. Generally, however, in the literature, the labelling, abstractly, as a derivative, of the mapping of a polynomial into a polynomial is sufficiently meaningful because it coincides with the notion of the derivative operator from analysis. From the preceding it is clear that we can define abstractly the derivative of a polynomial and one would suspect that it might be possible to define abstractly the derivative of functions in general.

1-9: The derivative from calculus is a topological derivative. A standard calculus text (Hart [14] defines the derivative as follows:

"At a given point $x = x_0$, if the ratio of the increment of the function $f(x)$ to the corresponding increment Δx of x approaches a limit as $\Delta x \to 0$, this limit is called the derivative of $f(x)$ with respect to x at $x = x_0$. Or, the derivative of $f(x)$ at $x = x_0$ is the instantaneous rate of change of $f(x)$ with respect to x at $x = x_0$".

Let $y = f(x)$ be a polynomial. The derivative of y with respect to x at $x = x_0$, denoted dy/dx, is given by

$$\lim_{\Delta x \to 0} \frac{f(x) + \Delta x) - f(x)}{x}$$

This derivative is also denoted as y', $f'(x)$, $D_x y$, $Df(x)$.

The result obtained by the above method is the same as that obtained by the abstract methods given in the preceding sections [14].

CHAPTER II

THE ALGEBRAIC DERIVATIVE

2-1: Let M be a module over a ring A, such that for all $a, b \in A$ and $x, y \in M$, the following holds:

i) $a(x + y) = a x + a y$

ii) $(a + b) x = a x + b x$

iii) $a(bx) = (ab) x$

iv) $1 \cdot x = x$

Chevalley [9] defines an algebra E over A to be "a module over A with an associative multiplication which makes E a ring satisfying

v) $a(xy) = (ax) y = x(ay)$"

2-2: The algebraic derivative is defined by Bourbaki [5] as follows:

"Let E be an algebra over a commutative ring A (having a unity element). A derivative, D, of E is defined to be an endomorphism, D, of the A-module E such that

$D(xy) = y D(x) + x D(y)$".

It follows that

$D(x + y) = D x + D y$, and

$D(a x) = a D(x) \qquad a \in A$

2-3: The derivative of a constant may be determined as follows (see [5]):

Let E be an algebra over a commutative ring A, having a unity element e.

Then

$D(e) = D(e^2) = D(e)\,e + e\,D(e) = 2\,D(e)$,

which infers that $D(e) = 0$.

Therefore,

$D(ne) = n(D(e)) = 0$ for all integers n, and $D(ae) = a\,D(e) = 0$ for all $a \in A$.

2-4: Proposition: For each derivation D of an algebra E and all elements b of the center of E, the endomorphism $x \rightarrow b\,D(x)$ denoted by bD of the A-module E, is a derivation of the algebra E. [5]

Proof:

Let $x, y \in E$ and $b \in$ center of E

Then

$b\,D(x + y) = b\,D(x) + b\,D(y)$

and

$$b\,D(xy) = b\,D(x)\,y + b\,x\,D(y)$$
$$= [b\,D(x)]\,y + x\,[b\,D(y)]$$

2-5: Proposition: If D_1 and D_2 are any two derivations of E, the endomorphism $D = D_2D_1 - D_1D_2$ of the A-module E is a derivation of the algebra E. [5]

Proof:

Let $x, y \in E$. Then

$$D(x + y) = (D_2D_1 - D_1D_2)(x + y)$$
$$= (D_2D_1 - D_1D_2)\,x + (D_2D_1 - D_1D_2)\,y$$
$$= D\,x + D\,y$$

and

$$D(xy) = D_2[D_1(x)\,y + x\,D_1(y)] - D_1[D_2(x)\,y + x\,D_2(y)]$$
$$= D_2D_1(x)\,y + D_1(x)\,D_2(y) + D_2(x)\,D_1(y)$$
$$+ x\,D_2D_1(y) - D_1D_2(x)\,y - D_2(x)D_1(y)$$

$$- D_1(x)\ D_2(y) - x\ D_1D_2(y)$$

$$= D_2D_1(x)\ y + x\ D_2D_1(y) - D_1D_2(x)\ y - x\ D_1D_2(y)$$

$$= D_2D_1(xy) - D_1D_2(xy)$$

$$= yD\ (x) + x\ D\ (y)$$

One observes that if E is the algebra of polynomials and D an endomorphism of the algebra E satisfying the properties of the definition in 2-2, then D is the derivative of E.

2-6: Let E and F be two algebras over a commutative ring A, having a unity element. Let ϕ be a function of E into F. Every linear mapping D of E into F such that for all x and y $\in$ E:

$$D\ (xy) = D(x)\ \phi\ (y) + \phi\ (x)\ Dy$$

is called a ϕ derivation of the algebra E into the algebra F. ϕ satisfies the properties:

1) If E is a subalgebra of F, ϕ is the canonical injection of E into F;

2) If F is the ring A and E is the set of functions B (i.e., B into A), then $\phi\ (\ f\ )$ is equal to $f\ (x_0)$ at $x_0 \in$ B.

CHAPTER III

THE TOPOLOGICAL AND ALGEBRAIC DERIVATIVES

3-1: We have seen that the topological derivative of a function is a special limit of the function at a point. The abstract algebraic derivative, on the other hand, is an operator which, without mention of continuity or limit of a function, does, in general, behave like the topological derivative operator. Actually, the topological and algebraic derivatives are the same under special circumstances. In this chapter we examine this relation.

3-2: Definition: Let f be a function defined in an interval $I \subset R$, not reduced to a point, with values in a normed vector space over the real numbers R. f is said to be differentiable at $x_0 \in I \subset R$ if

$$\lim_{\substack{x \to x_0,\ x \neq x_0 \\ x \in I}} \left\{ \frac{1}{x - x_0} (f (x) - f (x_0)) \right\} \qquad \text{exists;}$$

the value of this limit being called the derivative of f at x_0 and being denoted by $f' (x_0)$ or $D f (x_0)$. (Bourbaki [6])

Proposition: If f and g are functions defined in the interval $I \subset R$, not reduced to a point, with values in a normed vector space E over R and if f and g are differentiable at a point $x_0 \in I$, then $f + g$ and $a f$ $(a \in R)$ are differentiable at the same point and

i) $D (f + g) (x_0) = D f (x_0) + D g (x_0)$

ii) $D (af (x_0)) = a D f (x_0)$

Furthermore, if f and g are functions defined in the interval $I \subset R$, not reduced to a point, with values in a normed algebra E over R and if f and g are differentiable at $x_0 \in I$, then f g is differentiable at the same point and

iii) $D(fg)(x_0) = Df(x_0) \cdot g(x_0) + f(x_0) \cdot D g(x_0)$

Proof:

From Defn. 3-2 above, we have

$$Df(x_0) = \lim_{x \to x_0, x \neq x_0} \left\{ \frac{1}{x - x_0} (f(x) - f(x_0)) \right\}$$

and from i) above we get

$$D(f(x_0) + g(x_0)) = \lim_{\substack{x \to x_0 \\ x \neq x_0}} \left\{ \frac{1}{x - x_0} [f(x) + g(x)] - [f(x_0) + g(x_0)] \right\}$$

$$= \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} \left\{ [f(x) - f(x_0)] + [g(x) - g(x_0)) \right\}$$

$$= \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} [f(x) - f(x_0)] + \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} [g(x) - g(x_0)]$$

$$= Df(x_0) + Dg(x_0)$$

From ii) above we have

$$D(af(x_0)) = \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} (af(x) - af(x_0))$$

$$= \lim_{\substack{x \to x_0 \\ x \neq x_0}} a\left(\frac{1}{x - x_0}\right) (f(x) - f(x_0))$$

$$= a \lim_{\substack{x \to x_0 \\ x \neq x_0}} \left(\frac{1}{x - x_0}\right) (f(x) - f(x_0))$$

$$= a\, D f(x_0)$$

From iii) above it follows that

$$D(fg)(x_0) = \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} \left(f(x)\, g(x) - f(x_0)\, g(x_0) \right)$$

$$= \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} [f(x) g(x) - f(x_0)\, g(x) - f(x_0)\, g(x_0) + f(x_0)\, g(x)]$$

$$= \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} [f(x) - f(x_0)]\, g(x) + \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{x - x_0} (f(x_0))[g(x) - g(x_0)]$$

$$= Df(x_0) [\lim_{\substack{x \to x_0 \\ x \neq x_0}} g(x)] + [\lim_{\substack{x \to x_0 \\ x \neq x_0}} f(x_0)]\; Dg(x_0)$$

$$= Df(x_0)\, g(x_0) + f(x_0)\, Dg(x_0)$$

3-3: Let f be a function defined and continuous in the interval $I \subset R$. If the derivative f' exists at $x_0 \in I$ and is itself differentiable at x_0, its derivative is called the second derivative of f at x_0 and is denoted $f''(x_0)$ or $D^2 f(x_0)$. If this second derivative exists at all points of I (which implies that f' exists and is continuous in I), $x \to f''(x)$ is a function denoted f'' or $D^2 f$. By recurrence, we define the n^{th} derivative of f in the same manner and denote it $f^{(n)}$ or $D^n f$; by definition, it has for a value at $x_0 \in I$ the derivative of the function $f^{(n-1)}$ at x_0: this definition therefore supposes the existence of all the derivatives $f^{(k)}$ of order $K \leq n-1$ at x_0 and the differentiability of $f^{(n-1)}$ at x_0. If f is n times differentiable at x_0, we will say $f \in \mathcal{F}^n$, where $\mathcal{F}^n$ is the family of those functions which are n times differentiable. If f is indefinitely differentiable at x_0, we will say $f \in \mathcal{F}^\infty$.

3-4: Consider the set of indefinitely differentiable functions, $f \in \mathcal{F}^{\infty}$. We shall say that, if a function is continuous at a point x_0, it belongs to the class C (of functions continuous at x_0).

Theorem from calculus [14]: "If a function is differentiable at a point, then the function is continuous at that point".

If $f'(x_0)$ exists, then $f \in C$ (and since $f \in \mathcal{F}^{\infty}$, $f'(x_0)$ exists).

Consequently, if $f^k(x_0)$ exists (and this is guaranteed), $f^{k-1} \in C$, and by recursion $f^n \in C$, $n < \infty$. Therefore, an indefinitely differentiable function is indefinitely continuous.

Note: If $F \in \mathcal{F}^k$, $k \geq 1$, a mapping D of F into R satisfying the properties:

1) $D(f + g) = D\,f + D\,g$	$f, g \in F$
2) $D(af) = aDf$	$a \in R$
3) $D(fg) = Df \cdot g(x_0) + f(x_0) \cdot Dg$	$x_0 \in I \subset R$

is a ϕ derivation of F into R.

Consider the set of all functions $f \in \mathcal{F}^{\infty}$ which are indefinitely differentiable and continuous at x_0. We define an algebra [9]:

i) $r\,(f + g) = r\,f + r\,g$	$f, g \in \mathcal{F}^{\infty}$
ii) $(r + s)\,f = r\,f + s\,f$	$r, s \in R$
iii) $r\,(sf) = r\,s\,(f)$	
iv) $1 \cdot f = f$	
v) $r\,(fg) = r\,f\,(g) = f\,(r\,g)$	

The set of functions, $f \in \mathcal{F}^{\infty}$ must fulfill the following properties from the topological point of view:

Property (1): $D\,(rf + sg) = rD(f) + sD(g)$

Property (2): $D\,(fg) = D(f)\,g(x_0) + f(x_0)\,D(g)$

Property (1) implies $D(f) = 0$ if f is a constant function, because if $f(x_0) = 1$ and $g(x_0)$ has an arbitrary value $\neq 0$, then Property (2) gives

$$D(1 \quad g(x_0)) = D(1)\, g(x_0) + 1 \cdot D(x_0) \qquad \text{or}$$

$$D\, g(x_0) = D(1) g(x_0) + Dg(x_0),$$

whence $D(1) \quad g(x_0) = 0$, $\qquad g(x_0) \neq 0$

which implies that $D(1) = 0$

Property (3): $D(f) = 0$, if f is a constant function.

3-5: Definition: Let f be a function defined in the block $B \subset R^n$ ($B = I_1 \times I_2 \times \dots \times I_n$, $I_i \quad R$, for $0 \leq i \leq n$) not reduced to a point, with values in a normed vector space over R. f is partially differentiable at the point $t_0 \in B$ if

$$\lim_{t'_0 \to t_0} \left[\frac{1}{t'_0 - t_0} \left\{ f(t'_0) - f(t_0) \right\} \right] \qquad \text{exists}$$

for $t'_0 \neq t_0$, $t'_0 \in B$; the value of this limit is called the partial derivative of t_0 with respect to the i^{th} variable and is denoted $f'_i(t_0)$ or $D_i f(t_0)$, $t_0 \in B$.

Note: $t'_0 = (t_{1_0}, t_{2_0}, \dots, t_{(i-1)_0}, t_i, t_{(i+1_0}, \dots, t_{n_0})$

$$t!_0 = (t_{1_0} \; t_{2_0}, \dots, t_{(i-1)_0}, t_{i_0}, t_{(i+1)_0}, \dots, t_{n_0})$$

When $B = I$, this definition reduces to the definition in 3-2 above. Moreover, it can be shown that the proposition in 3-2 above for derivatives is also valid for partial derivatives.

By $Df(t_0) = \sum_{i=1}^{n} D_i f(t_0)$ we denote the total derivative of f at $t_0 \in B$.

Thus, each derivative from the topological viewpoint can be expressed as a sum of derivatives of this type.

3-6: Proposition: Let F^k be the set of all functions defined in an interval $I \subset R$ (not reduced to a point), with values in a normed algebra E over R, such that $F^k \in \mathcal{F}^k$, $k \geqslant 1$ in I. The linear mapping D of the algebra F^k into the algebra F^{k-1} such that for all f, g $\in F^k$, $D(fg) = Df \cdot g + f \cdot Dg$, is a derivation of the algebra F^k into the algebra F^{k-1}.

Proof: The proof is immediate from the definition of a Φ derivation.

3-7: Theorem: Consider the set of functions defined in an interval $I \subset R$ (not reduced to a point), which are continuous at a point $x_0 \in I$ and which have values in a normed algebra. A derivation, D, of these functions in the topological sense is the same as that in the algebraic sense if and only if D is an endomorphism of $\mathcal{F}^\infty$.

Proof: Let B R^n $(B = I_1 \times I_2 \times \cdots \times I_n)$ and let F be a normed algebra over R. Let $\mathcal{F}^\infty$ be the set of functions which are indefinitely differentiable at $t_0 \in B$. Let $B_0 \subset B$ be a star-shaped open set about t_0, i.e., an open set such that if $z \in B_0$,

$$t_0 + x\,(z - t_0) \in B_0 \qquad \text{for } 0 \leqslant x \leqslant 1 \tag{25}$$

Then $f_{B_0}(z) = f_{B_0}(t_0) + \int_0^1 \frac{d}{dx} f_{B_0}(t_0 + x\,(z - t_0))\,dx$

$$= f_{B_0}(t_0) + f_{B_0}(t_0 + (z - t_0)) - f_{B_0}(t_0)$$

$$= f_{B_0}(t_0) + f_{B_0}(t_0) + f_{B_0}(z) - f_{B_0}(t_0) - f_{B_0}(t_0)$$

$= f_{B_0}(z)$, which verifies the identity.

So, $f_{B_0}(z) = f_{B_0}(t_0) + \int_0^1 \frac{d}{dx} f_{B_0}(t_0 + x\,(z - t_0))\,dx$

$$= f_{B_0}(t_0) + \sum_{i=1}^{n} (z_i - t_{i_0}) \int_0^1 \frac{\partial f}{\partial t_i}(t_0 + x\,(z - t_0))\,dx$$

Let $\int_0^1 \frac{\partial f}{\partial t_i}(t_0 + x\,(z_i - t_0))\,dx$ be denoted by

$$g_i(t_0) = \frac{\partial f}{\partial t_{i_0}} (t_0) \quad , \text{ where } g_i \quad \in \mathcal{F}^{\infty} \quad \text{and where}$$

z_i = coordinates of z.

Therefore,

$$f_{B_0}(z) = f_{B_0}(t_0) + \sum_{i=1}^{n} (z_i - t_i) \quad g_i(t_0)$$

For D satisfying properties (1), (2) and (3), with t_0 fixed and z variable:

$$Df_{B_0}(z) = 0 + \sum_{i=1}^{n} \left\{ D(z_i - t_{i_0}) \quad g_i(t_0) + (z_i - t_0) \quad Dg_i \quad (t_0) \right\}$$

$$= 0 + \sum_{i=1}^{n} \left\{ [D(z_i - 0] \quad . \quad g_i(t_0) + (z_i - t_0) \quad . \quad 0 \right\}$$

$$= \sum_{i=1}^{n} D(z_i)\, g_i(t_0)$$

$$= \sum_{i=1}^{n} D(z_i) \frac{\partial f(t_0)}{\partial t_{i_0}}$$

which is consistent with the definition of a total derivative of a function defined in a block. The above proof fails if we start from $\mathcal{F} = \mathcal{F}^k$, $k < \infty$, since the functions g_i are not necessarily in the set for which properties (1), (2) and (3) hold, inasmuch as a function is continuously differentiable if and only if its partial derivatives exists and are continuous.

CHAPTER IV

OTHER DERIVATIVES IN CONTEMPORARY LITERATURE

4-1: In this chapter we will survey some definitions of derivatives in general, as found in the contemporary literature. The derivatives of polynomials, which we have already considered, are algebraic in nature, and are endomorphisms of the ring of polynomials, which are indefinitely differentiable.

4-2: Chevalley [10] defines a derivation of R into S, where R is a subfield of a field S, to be a mapping such that:

(1) $D(x + y) = D x + D y \qquad x, y \in R$

(2) $D(xy) = D(x) y + x D(y)$,

where (1) is a homomorphism of the additive group of R into that of S, and

(3) $(u D)(y) = u (Dy), \qquad u \in S$

whence the derivations of R into S form a vector space over S.

This is an algebraic definition, more inclusive than that of Bourbaki, differing in that it is defined as a mapping of a subfield into a field, rather than an endomorphism of a ring R.

Note: (1) and (2) above will hereinafter be referred to as the standard forms for the derivatives of sums and products.

4-3: Kawada [19] defines a derivation D to be an operation obeying (1) and (2) above, where the domain is a commutative ring R and the range M is an R-module.

This is an algebraic definition, more inclusive than that of Bourbaki, and differing only in that it is the mapping of a ring into a ring module rather than of an algebra of a ring into itself.

4-4: Consider an arbitrary collection of elements, which obey the field postulates, as an algebraic field F, of characteristic zero. Ritt [28] defines a derivation to be an operation which replaces every element of F by its derivative, where the derivative operator obeys (1) and (2) above.

This is an algebraic definition. Since it is defined in a field, it is more inclusive than that of Bourbaki, which is defined in a ring.

4-5: Weil [32] defines a derivation D in a field F to be a mapping of F into itself which obeys (1) and (2) above. He further defines a derivation D in F over a subfield K of F to be one such that $Dx = 0$ for every x in K.

This is an algebraic definition, more inclusive than that of Bourbaki. In the abstract field F, the elements $x \in K$ are constants in differentiation.

4-6: The derivation of an arbitrary algebra R over a commutative field F is defined by Jacobson [18] to be a single-valued mapping of R onto itself such that (1) and (2) are obeyed. He notes that $D_1 \pm D_2$ is a derivation, but that D_1D_2 is not, in general, a derivation. However, D^p is a derivation, where F has characteristic $p \neq 0$:

$D\,(xy) = D(x) \cdot y + x\,D(y)$

$D^2(xy) = D^2(x)\,y + 2\,D(x)\,D(y) + x\,D^2(y)$

$D^3(xy) = D^3(x)\,y + 3\,D^2(x)\,D(y) + 3\,D(x)\,D^2(y) + xD^3(y)$

which follows the pattern

$(a + b)^p = a^p + \binom{p}{1} a^{p-1} b + \cdots + \binom{p}{i} a^{p-i} b^i \ldots + b^n,$

where $\binom{p}{i} \equiv 0 \pmod{p}$

Therefore $D^p(xy) = D^p(x)\,y + x\,D^p(y)$

whence D^p is a derivation.

Jacobson's definition is an algebraic one, which is more inclusive than that of Bourbaki. It would be the same as Bourbaki's if a ring were specified rather than a field.

4-7: Zariskie and Samuel [34] define a derivation to be a mapping D of a ring R into a subring S such that (1) and (2) above are obeyed, and add that if R is an integral domain and D a derivation of R in a field F which contains R, that the derivative of the quotient field of F is

$$D\left(\frac{x}{y}\right) = \frac{yD(x) - xD(y)}{y^2} \quad , \quad \frac{x}{y} \in F \, ; \, x, y \in R, \, y \neq 0.$$

This follows from (2) above:

$$D(x) = D\left(\frac{x}{y} \cdot y\right) = D\left(\frac{x}{y}\right) \cdot y + \frac{x}{y} D(y)$$

$$D\left(\frac{x}{y}\right) y = D(x) - \frac{x}{y} D(y)$$

$$D\left(\frac{x}{y}\right) = \frac{y\, D(x) - xD(y)}{y^2}$$

Furthermore,

$$D(x^2) = D(x \cdot x) = D(x) \cdot x + xD(x) = 2xD(x)$$

$$D(x^3) = D(x^2 \cdot x) = D(x^2) \cdot x + x^2 D(x) = 3x^2 D(x)$$

and for every integer $n \geqslant 1$,

$$D(x^n) = n\, x^{n-1} D(x)$$

The derivative defined by Zariskie and Samuel is an algebraic one. It differs from that of Bourbaki only in that it is a mapping of a ring R into a subring of R, rather than an endomorphism of R.

4-8: Finkbeiner (13) states that the derivative of a matrix exists if and only if each element a_{ij} is a differentiable function. Let A and $m \times n$ matrices of differentiable functions. Then

(i) $D(A + B) = D(A) + D(B)$

(ii) $D(AB) = D(A)B + A D(B)$

(iii) $D(A^{-1}) = -A^{-1} D(A) A^{-1}$, if A is a non-singular matrix.

(i) and (ii) follow from (1) and (2) above. Consider (iii):

$0 = D(A^{-1}A) = D(A^{-1}) A + A^{-1} D(A)$, then

$D(A^{-1}) A = -A^{-1} D(A)$

Multiplying both sides on the right by A^{-1}, we get

$D(A^{-1}) A \cdot A^{-1} = -A^{-1} D(A) A^{-1}$

whence,

$D(A^{-1}) = -A^{-1} D(A)A^{-1}$

Since the multiplication of matrices is not in general commutative, Nehring [24] points out that particular attention must be paid to the order of all factors, e.g.,

$D(A^2) = D(A) A + AD(A)$

This definition of the derivative of a matrix is algebraic. It is a special case of Bourbaki's definition, i.e., an endomorphism of the ring of matrices.

4-9: Let F be a set of indefinitely differentiable functions which form an algebra over the field of real numbers and V a differentiable manifold of dimension n, with the set of real-valued functions $F = F(V)$ on V. Nomizu [26] considers a vector field X as a linear mapping of F into the algebra of all real-valued functions on V such that

$X(fg) = Xf \cdot g + f \cdot Xg \qquad f, g \quad F$

It follows that, since X may be considered as a linear mapping, then

$X(f + g) = X f + X g$

Since X fulfills the abstract requirements of a derivative, it may be considered as an algebraic derivative. This differs from Bourbaki's definition in that this is a homomorphism of a subalgebra into the algebra.

4-10: If A_p is the set of analytic functions which are defined at p, and R is the set of real numbers, Cohn [11] defines a mapping L of A_p into R to be a tangent vector if and only if it is linear over R:

$$L\ (af + bg) = a \cdot L\ f + b \cdot Lg \qquad (f,\ g \in A_p,\ a,\ b \in R)$$

and satisfies

$$L\ (fg) = Lf \cdot g(p) + f\ (p) \cdot Lg \qquad (f,\ g \in A_p)$$

Again we have an operator which fulfills the abstract requirements of a derivative and may be likewise considered as an algebraic derivative. The operation, in this case, is a ϕ derivation.

4-11: Let K be a field of characteristic 0 and K $[x_1, \cdots, x_n]$ a ring of integral (i.e., with exponents 0) formal power series in n variables $x_1, \cdots, x_n$ over K. Hochschild [16] defines a mapping D of K $[x_1, \ldots, x_n]$ into itself to be a derivation, if:

(i) $D(a) = 0$, for every $a \in K$

(ii) For any two power series p and q, D obeys (1) and (2) above.

This derivation is an endomorphism of a ring and coincides with the definition given by Bourbaki.

4-12: Leger [20] defines a linear transformation T on the vector space of a Lie algebra L to be a derivation of L if:

$$T\ (x \cdot y) = T\ (x)\ .\ y + x\ .\ T\ (y), \qquad \text{for all } s,\ y \in L.$$

Since T is a linear transformation,

$$T\ (x + y) = T\ (x) + T\ (y)$$

and we are justified in classifying T among the abstract algebraic operators. This is a less inclusive definition than that of Bourbaki, inasmuch as it describes an endomorphism of the vector space of an algebra.

4-13: If R is a ring, not necessarily associative, Patterson [27] defines a derivation of R to be a mapping D of R into itself such that (1) and (2) above are obeyed.

An algebraic derivation, this is a special case of that defined by Bourbaki, differing in that it is an endomorphism of a non-associative ring.

4-14: Let f be a mapping of an open sub-set R of E^n into E^m, where R is the set of real numbers and E^m and E^n are Euclidean spaces. For any point $p \in R$ and any vector v of E^n, Whitney [33] defines the derivative of f at p with respect to v as:

$$D_v f(p) = \lim_{t \quad 0^+} \frac{1}{t} \; (f \; (p + tv \; - f(p) \;)$$

if this exists; this is the vector in E^m.

Clearly the definition of a derivative in the topological sense, this is a special case of that given by Bourbaki.

4-15: In summary, the algebraic derivative is variously defined as homomorphisms or endomorphisms of rings and fields without mention of a limit. In the algebraic sense, any linear mapping of an algebra into an algebra, which obeys the sum and product properties defined in 2-2, may be called a derivative. In order to compare the definitions of the derivative from the algebraic and topological points of view, it is necessary to consider the algebra of functions. In the algebra of functions, if the derivative exists in the algebraic sense, it always exists in the topological sense. On the other hand, the topological derivative exists at points where the required limit exists, that is, differentiation of a function from the topological point of view is a pointwise operation. The definition of the derivative from the topological point of view is therefore more general than the definition from the algebraic point of view, since a derivation of an algebra from the topological point of view is a derivation from the algebraic point of view only if the algebra is that of indefinitely differentiable functions.

BIBLIOGRAPHY

1. Albert, A. Adrian. Modern Higher Algebra. Chicago: University of Chicago Press, 1937.

2. Anthony, H. D. Sir Isaac Newton. London: Abelard-Schumann, Limited, 1960.

3. Barnes, Wilfred E. Introduction to Abstract Algebra. Boston: D. C. Heath and Co., 1963.

4. Birkhoff, Garrett, and Saunders MacLane. A Brief Survey of Modern Algebra. New York: Macmillan Co., 1962.

5. Bourbaki, Nicolas. Éléments de Mathématique, Livre II, Algebre, Chapitre 4. Paris: Hermann, 1958.

6. Bourbaki, Nicolas. Éléments de Mathématique, Livre IV, Fonctions d'une Variable Reelle. Paris: Hermann, 1958.

7. Brewster, Sir David, K. H. Memoirs of the Life, Writings, and Discoveries of Sir Isaac Newton, Vol. 1. Edinburgh: Thomas Constable and Co., 1855.

8. Cajori, Florian. A History of Mathematics. London: Macmillan Co., 1919.

9. Chevalley, Claude. The Construction and Study of Certain Important Algebras. Tokyo: The Mathematical Society of Japan, 1955.

10. Chevalley, Claude. Introduction to the Theory of Algebraic Functions of One Variable. New York: American Mathematical Society, 1951.

11. Cohn, P. M. Lie Groups. Cambridge: Cambridge University Press, 1957.

12. Dubriel, Paul. Algèbre, Vol. I. Paris: Gauthier-Villars, 1954.

13. Finkbeiner, Daniel T., II. Introduction to Matrices and Linear Transformations. San Francisco: W. H. Freeman and Co., 1960.

14. Hart, William L. Calculus. Boston: D. C. Heath and Co., 1955.

15. Herstein, I. N. Topics in Algebra. New York: Blaisdell Publishing Co., 1964.

16. Hochschild, G. "Lie Algebras and Differentiations in Rings of Power Series". American Journal of Mathematics, Vol. 72, 1950.

17. Jacobson, Nathan. Lectures in Abstract Algebra. Vol. III. Princeton: D. Van Nostrand Co., Inc., 1964.

18. Jacobson, Nathan. "Abstract Derivation and Lie Algebras". American Mathematical Society Transactions, Vol. 42, 1937.

19. Kawada, Yukiyosi. "On the Derivations in Number Fields". Annals of Mathematics, Vol. 54, No. 2, 1951.

20. Leger, George F., Jr., "A Note on the Derivations of Lie Algebras". American Mathematical Society Proceedings, Vol. 4, 1953.

21. Macduffee, Cyrus Colton. An Introduction to Abstract Algebra. New York: John Wiley and Sons, Inc., 1940.

22. Miller, Kenneth S. Elements of Modern Abstract Algebra. New York: Harper and Brothers, 1958.

23. Mirsky, L. An Introduction to Linear Algebra. Oxford: The Clarendon Press, 1955.

24. Nehring, Evar D. Linear Algebra and Matrix Theory. New York: John Wiley and Sons, Inc., 1963.

25. Nicholson, H. K., N. E. Steenrod, and D. C. Spencer. Advanced Calculus. Princeton: D. Van Nostrand Co., Inc., 1959.

26. Nomizu, Katsumi. Lie Groups and Differential Geometry. Tokyo: The Herald Printing Co., Ltd., 1956.

27. Patterson, E. M. "On Certain Types of Derivatives". Proceedings of Cambridge Philosophical Society, 54, 1958.

28. Ritt, Joseph Fels. Differential Algebra. New York: American Mathematical Society, 1950.

29. Rosenthal, Arthur. "The History of Calculus". American Mathematics Monthly, Vol. 58, 1951.

30. Schreier, Otto, and Emanuel Sperner. Modern Algebra and Matrix Theory. New York: Chelsea Publishing Co., 1959.

31. van der Waerden, B. L. Modern Algebra. New York: Frederick Ungar Publishing Co., 1948.

32. Weil, André. Foundations of Algebraic Geometry. New York: American Mathematical Society, 1946.

33. Whitney, Hassler. Geometric Integration Theory. Princeton: Princeton University Press, 1957.

34. Zariskie, Oscar, and Pierre Samuel. Commutative Algebra. Princeton: D. Van Nostrand Co., Inc., 1958.

VITA

Richard Francis Barry, Jr.

Born in Quincy, Massachusetts, December 26, 1915. Graduated from Quincy High School, June 1933, B.S., United States Naval Academy, 1939, Armed Forces Staff College, 1953. Retired from United States Navy in the grade of Captain, June 1965.