QUATERNION CALCULUS

by

DAVID JOSEPH EDELBLUTE

B. S., Kansas State University, 1962

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1964

Approved by:

Leonard E. Fuller

Major Professor

TABLE OF CONTENTS

INTRODUCTION 1

THE "ex" FUNCTION 3

GREAT CIRCLES 7

ROTATIONS 9

DIFFERENTIATION 12

SOME SPECIAL FUNCTIONS 20

HIGHER DIFFERENTIATION 22

THE TAYLOR SERIES 25

ACKNOWLEDGMENT 29

REFERENCES 30

INTRODUCTION

In this paper it will be assumed that the reader is familiar with the algebra of quaternions.[1] We shall restrict our consideration to the real quaternions, i.e., all quantities of the form $x = x_0 + ix_1 + jx_2 + kx_3$, where x_0, x_1, x_2, and x_3 are real numbers. The vector space consisting of all real quaternions will be denoted by "Q." The real numbers will be regarded as the subset of Q formed by the condition that $x_1 = x_2 = x_3 = 0$.

By the term "vector" we shall mean a quantity of the form $ix_1 + jx_2 + kx_3$, i.e., the vectors shall be the subset of Q formed by the condition that $x_0 = 0$. In this way we shall retain the original concept that a quaternion is a scalar plus a vector. We shall see that this concept has several notational advantages. The scalar part of x is x_0 and the vector part of x, denoted by x_v, is $ix_1 + jx_2 + kx_3$. Then $x = x_0 + x_v$.

First we shall establish a notation based on this concept. Then we shall see how some of the ideas of calculus can be applied to quaternions. To do this, we shall develop a number of formulas which are extensions of familiar formulas in real variables. In conclusion, we shall show how Taylor's series can be used to describe quaternions which are functions of other quaternions.

Let us define some functions which are basic to our analysis.

---

[1] For a discussion of quaternion algebra the reader is referred to P. G. Tait, An Elementary Treatise on Quaternions.

The norm of x, denoted by N(x) or by $|x|$, will be defined to be

$$|x| = N(x) = (x_0^2 + x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}} .$$

A metric on Q will be denoted by m(x, y) and defined by m(x, y) = N(x-y). The quantity which we have defined as N(x) was called the "tensor" of x in the original literature on quaternions, while the "norm" was originally the square of the tensor. This terminology is no longer convenient, since the term "tensor" currently has another meaning.

Another useful function of x is the conjugate of x, denoted by $\bar{x}$ or by K(x). If $x = x_0 + ix_1 + jx_2 + kx_3$ then

$$\bar{x} = K(x) = x_0 - ix_1 - jx_2 - kx_3 = x_0 - x_v.$$

Throughout this paper multiplication will be denoted by juxtaposition. We shall list here a few properties of multiplication which will be important in our development.

1) $x_v y_v = -x_v \bullet y_v + x_v X y_v$. Here $x_v \bullet y_v$ and $x_v X y_v$ are the familiar dot and cross products of vector algebra.

2) $x\bar{x} = \bar{x}x = |x|^2$

3) $x^{-1} = \bar{x}/|x|^2$

4) $\overline{(x)(y)} = (\bar{y})(\bar{x})$

5) $|xy| = |x||y|$.

Much of the geometric convenience of quaternions is due to Property 1. It could be used to define quaternion multiplication.

Note that $xy = (x_0 + x_v)(y_0 + y_v)$

$$= x_0 y_0 + x_0 y_v + y_0 x_v + x_v y_v$$

$$= x_0 y_0 + x_0 y_v + y_0 x_v - x_v \bullet y_v + x_v X y_v$$

$$= x_0 y_0 - x_v \bullet y_v + x_0 y_v + y_0 x_v + x_v X y_v.$$

This notation links the noncommutativity of quaternion multiplication with the anticommutativity of the cross product. We have now established the criterion for two quaternions to commute under multiplication: Their vector parts must be proportional.

THE "ex" FUNCTION

We shall now develop a notation based on a function which we shall denote as ex(x). It bears a strong resemblance to the familiar exponential function and, as we shall see, actually reduces to e^x if x is a real or complex number.

First let us define

$$ex(x_v) = \cos|x_v| + (x_v/|x_v|) \sin|x_v| .$$

Let $|x_v| = \theta$. Then $x_v/\theta = E = iE_1 + jE_2 + kE_3$, is a unit vector in the direction of x_v, and x can be written as $x = x_0 + x_v = x_0 + E\theta$. Note that θ is a real number, $|E| = 1$, and $\overline{E} = E^{-1} = -E$. Also

$$ex(E\theta) = \cos(\theta) + E \sin(\theta),$$

which resembles the form for e to an imaginary power.

Now consider $|ex(E\theta)|$.

$$|ex(E\theta)| = \left\{(ex(E\theta))\,\overline{(ex(E\theta))}\right\}^{\frac{1}{2}}$$
$$= \left\{(\cos\theta + E\sin\theta)(\cos\theta - E\sin\theta)\right\}^{\frac{1}{2}}$$
$$= (\cos^2\theta - E^2\sin^2\theta)^{\frac{1}{2}} .$$

Note that from property 1), $E^2 = -1$, so

$$|ex(E\theta)| = (\cos^2\theta + \sin^2\theta)^{\frac{1}{2}}$$
$$= (1)^{\frac{1}{2}}$$
$$= 1 .$$

Thus, $ex(x_v)$ is a quaternion of unit norm, whose vector part points in the direction of x_v, for every x_v.

To define $ex(x) = ex(x_0 + x_v)$, let

$$ex(x_0 + x_v) = ex(x_0)ex(x_v),$$

and

$$ex(x_0) = e^{x_0}.$$

Note that for any quaternion y, $y = |y|(y/|y|)$. This implies that y can be written as the product of a real number, $|y|$, and a quaternion of unit norm. The quantity $y/|y|$ will be called the versor of y and denoted by $U(y)$ or $\vec{y}$.

We shall now show that for any quaternion, y, there exists a quaternion x such that y = ex(x). This is similar to the polar representation of complex numbers. To do this, we shall write $x = \ln(r) + E\theta$, where $r = |y|$ and E and θ are defined as above.

Let $E = y_v / |y_v|$ and $\tan\theta = |y_v|/y_0$. Note that

$$(|y_v|^2 + y_0^2)^{\frac{1}{2}} = (y_1^2 + y_2^2 + y_3^2 + y_0^2)^{\frac{1}{2}} = |y| = r.$$

Then since $\tan\theta = |y_v|/y_0$, $\sin\theta = |y_v|/r$ and $\cos\theta = y_0/r$.

Note that $|y_v| = (r)\sin\theta$, so $E = y_v/((r)\sin\theta)$, or $y_v = rE\sin\theta$. Also, since $\cos\theta = y_0/r$, $y_0 = (r)\cos\theta$. Thus we have shown that

$$\begin{aligned} y_0 + y_v &= (r)\cos\theta + rE\sin\theta \\ &= r(\cos\theta + E\sin\theta) \\ &= e^{\ln(r)}(\mathrm{ex}(E\theta)) \\ &= (\mathrm{ex}(\ln(r)))(\mathrm{ex}(E\theta)) \\ &= \mathrm{ex}(\ln(r) + E\theta) \\ &= \mathrm{ex}(x). \end{aligned}$$

This completes the proof.

We can now see that the complex case arises whenever the direction of E is fixed. Whenever this happens, we are restricted to the two dimensional subspace of Q spanned by 1 and E. In this space the basis vectors have the familiar multiplication properties of complex variables.

To illustrate this, let $x = (r)(ex(E\theta))$ and $y = (s)(ex(F\phi))$. Then

$$xy = (r(ex(E\theta)))(s(ex(F\phi)))$$
$$= (rs)(ex(E\theta))(ex(F\phi))$$
$$= (rs)(\cos\theta + E\sin\theta)(\cos\phi + F\sin\phi)$$
$$= (rs)(\cos\theta\cos\phi - E\bullet F\sin\theta\sin\phi + E\sin\theta\cos\phi$$
$$+ F\cos\theta\sin\phi + EXF\sin\theta\sin\phi).$$

This reduces to the familiar form in complex variables when $E = F$, for then $E\bullet F = E\bullet E = |E|^2 = 1$ and $EXF = EXE = 0$.

$$xy = (rs)(\cos\theta\cos\phi - \sin\theta\sin\phi + E\sin\theta\cos\phi + E\cos\theta\sin\phi)$$
$$= (rs)(\cos(\theta + \phi) + E\sin(\theta + \phi))$$
$$= (rs)(ex(E(\theta + \phi))).$$

In this notation it is easy to examine rational powers of a quaternion. If n is an integer and $x = r(ex(E\theta))$, then an obvious induction based on the above formula gives $x^n = r^n(ex(En\theta))$.

Example[1]: Given $w = s(\cos\phi + F\sin\phi)$, find x such that $x^n = w$. Let $x = r(ex(E\theta))$.

Case 1: $\sin\phi \neq 0$. Choose $E = F$. $r^n = s$, so $r = s^{1/n}$. $\mathrm{Cos}(n\theta =$ $\cos\phi$ and $\sin(n\theta) = \sin\phi$, so $\theta = (\phi + 2\pi m)/n$, for $m = 0, 1, 2, \ldots, n-1$.

---

[1] L. Brand, Vector and Tensor Analysis, pp. 412-413

Case 2: $\sin\phi = 0$, i.e., the direction of F is arbitrary.

a) If $w > 0$, $\phi = 2\pi m$, and $\theta = 2\pi m/n$, for $m = 0, 1, 2, \ldots, n-1$. If $n = 2$, we get two roots, $\pm\sqrt{w}$. If $n > 2$, some values of θ ($\neq 0, \pi$) give non-real roots with E in an arbitrary direction.

b) If $w < 0$, $\phi = \pi$, $\theta = (2m+1)\pi/n$, for $m = 0, 1, 2, \ldots, n-1$. In every case some values of θ ($\neq \pi$) give non-real roots with E in an arbitrary direction.

GREAT CIRCLES

Although the lack of a simple multiplication rule detracts from the value of the function ex(x), this function can still be useful for many purposes. One example of this is in spherical geometry. For this purpose, let a and b be vectors of unit norm. Recall that for any quaternions x, y, and z, $xy = z$ implies that $|x||y| = |z|$.

We can visualize unit vectors a and b as vectors which terminate at points A and B, respectively, on a sphere of unit radius, center at the origin.

A directed great circle arc is thus defined from A to B. We shall be concerned with adding such arcs "vectorially." Note that if A and B are permitted to move on the great circle connecting them, while the angle (a, b) remains fixed, we still have, for our purposes, the same arc. To add two great circle arcs, then, we first find a point of intersection of the two great circles. Then we shift the arcs along their great circles until the terminal point of the first coincides with the initial point of the second at that point. The result, or "sum" is the great circle arc from the initial point of the first to the terminal point of the second.

We shall now show that great circle arcs under addition correspond to versors, or unit quaternions, under multiplication.

The product ba^{-1} may now be regarded as an operator which will rotate the vector a into the position b. We shall now show that ba^{-1} is of the form $\cos\theta + E\sin\theta$, where θ is the angle between a and b, and E is perpendicular to the plane of a and b so that a, b, and E form a dextral set.

Recall that for two vectors p and q the product is $pq = -p \cdot q + pXq$. Also, since a is a vector of unit norm, $a^{-1} = -a$. Thus $ba^{-1} = b(-a) = (-1)ba = (-1)(-b \cdot a) + (-1)(bXa) = (-1)(-\cos\theta) + (-1)(-aXb) = \cos\theta + aXb = \cos\theta + E\sin\theta$.

Note that if A and B are permitted to move on a great circle whose plane is perpendicular to E while the angle (a, b) is fixed the product ba^{-1} remains unchanged.

Thus for every unit quaternion $ex(E\theta)$ there is a corresponding great circle arc of a unit sphere centered at the origin, and conversely. This arc lies on a great circle whose plane is perpendicular to E and is directed in a right-handed direction with respect to E.

Let A, B, and C denote points of a unit sphere centered at 0. Let $0A \rightarrow a$, $0B \rightarrow b$, and $0C \rightarrow c$. Then arc (AB) = ba^{-1} and arc (BC) = cb^{-1}. Arc (AC) = $(cb^{-1})(ba^{-1}) = ca^{-1}$.

At this point we need to note some special cases. Let q be a unit quaternion. If $q = 1$ ($\theta = 0$) or $q = -1$ ($\theta = \pi$), the direction of E is arbitrary. Any point on the sphere may represent $q = 1$ and any great semicircle may represent $q = -1$. If $q = E$ ($\theta = \pi/2$) then q represents a quadrental arc whose plane is perpendicular to E.

Let p, q, r, and s be unit quaternions. Then we have corresponding arcs; arc(p), arc(q), arc(r), and arc(s). From the foregoing argument, arc(p) + arc(q) = arc(qp), or in general, arc(p) + arc(q) + arc(r) + arc(s) = arc(srqp).

If arc(p), arc(q), arc(r) and arc(s) form the sides of a closed spherical polygon, then we say

$$\mathrm{arc}(p) + \mathrm{arc}(q) + \mathrm{arc}(r) + \mathrm{arc}(s) = 0.$$

Hence, arc(srqp) = 0, or srqp = 1.

ROTATIONS

Another useful product of quaternions is $r' = qrq^{-1}$. We shall consider $q(\ )q^{-1}$ to be an operator, acting on r to product r'. From the elementary properties of multiplication it follows readily that $q(\ )q^{-1}$ is distributive with respect to both addition and multiplication of quaternions, i.e., if r and s are any two quaternions, then $q(r + s)q^{-1} = q(r)q^{-1} + q(s)q^{-1}$ and $q(rs)q^{-1} = (q(r)q^{-1})\ (q(s)q^{-1})$.

Let $q = |q|(\cos\theta + E\sin\theta)$. We shall show that this operator has three characteristics:

1) $|r| = |r'|$

2) $r_0 = r'_0$

3) r_v is rotated about E through an angle of 2θ in a right-handed direction to obtain r'_v.

To establish property 1), consider $|r'| = |qrq^{-1}| = |q||r||q^{-1}| = |q||r||q|^{-1} = |q||q|^{-1}|r| = |r|$.

To establish 2) and 3), note that

$$\begin{aligned} r'_0 + r'_v &= q(r_0 + r_v)q^{-1} \\ &= qr_0q^{-1} + qr_vq^{-1} \\ &= qq^{-1}r_0 + qr_vq^{-1} \\ &= r_0 + qr_vq^{-1}. \end{aligned}$$

Note that we can, without loss of generality, assume that $|r| = |q| = 1$. Let $r = \cos\phi + F\sin\phi$. We now need to consider $qr_vq^{-1} = q(F\sin\phi)q^{-1} = \sin\phi(qFq^{-1})$.

Let us use $F = F_{\parallel} + F_{\perp}$, where $F_{\parallel}$ is the component of F parallel to E and $F_{\perp}$ is the component of F perpendicular to E. Note that $F_{\perp}$ lies in the plane of E and F. Further, from page 3, since $F_{\parallel}$ is parallel to the vector part of q, $F_{\parallel}$ and q commute.

$$
\begin{aligned}
qFq^{-1} &= q(F_{\parallel} + F_{\perp})q^{-1} \\
&= qF_{\parallel}q^{-1} + qF_{\perp}q^{-1} \\
&= qq^{-1}F_{\parallel} + qF_{\perp}q^{-1} \\
&= F_{\parallel} + qF_{\perp}q^{-1}.
\end{aligned}
$$

We now need only to consider $qF_{\perp}q^{-1}$.

$$
\begin{aligned}
qF_{\perp}q^{-1} &= (\cos\theta + E\sin\theta)F_{\perp}(\cos\theta - E\sin\theta) \\
&= (\cos\theta + E\sin\theta)(F_{\perp}\cos\theta - F_{\perp}E\sin\theta).
\end{aligned}
$$

Note that $F_{\perp} \cdot E = 0$.

$$
\begin{aligned}
qF_{\perp}q^{-1} &= (\cos\theta + E\sin\theta)(F_{\perp}\cos\theta - (F_{\perp} X E)\sin\theta) \\
&= F_{\perp}\cos^2\theta + (E X F_{\perp})\sin\theta\cos\theta - (F_{\perp} X E)\sin\theta\cos\theta \\
&\quad -E(F_{\perp} X E)\sin^2\theta.
\end{aligned}
$$

Note that $F_{\perp} X E = -E X F_{\perp}$. Also, $E(F_{\perp} X E) = -E\cdot(F_{\perp} X E) + E X(F_{\perp} X E)$. But $F_{\perp} X E$ is perpendicular to E, so $E\cdot(F_{\perp} X E) = 0$.

$$
\begin{aligned}
E X(F_{\perp} X E) &= F_{\perp}(E\cdot E) - E(F_{\perp}\cdot E) \\
&= F_{\perp}.
\end{aligned}
$$

$$
\begin{aligned}
\text{So } qF_{\perp}q^{-1} &= F_{\perp}\cos^2\theta + (E X F_{\perp})2\sin\theta\cos\theta - F_{\perp}\sin^2\theta \\
&= F_{\perp}(\cos^2\theta - \sin^2\theta) + (E X F_{\perp})2\sin\theta\cos\theta \\
&= F_{\perp}(\cos 2\theta) + (E X F_{\perp})\sin 2\theta.
\end{aligned}
$$

Thus if $q = ex(E\theta)$, $q(\ )q^{-1}$ can be regarded as a rotation operator whose only effect is to rotate r_v about E through an angle of 2θ.

Successive rotations, then are easily handled. Given two rotations, $p(\ )p^{-1}$ and $q(\ )q^{-1}$, if we perform them in that order the result is

$$q(p(\ )p^{-1})q^{-1} = qp(\ )p^{-1}q^{-1} = qp(\ )(qp)^{-1}.$$

DIFFERENTIATION

If a quaternion, y if a function of a scalar, t, the operation of differentiation presents no difficulty.

$$\frac{dy}{dt} = \frac{dy}{dt}0 + \frac{idy}{dt}1 + \frac{jdy}{dt}2 + \frac{kdy}{dt}3\ .$$

If, however, we have a quaternion, y, which is a function of another quaternion, x, the problem of differentiation cannot be so easily resolved. The familiar concept of differential coefficients or derived functions has been found to be, in general, inapplicable to quaternions. As we shall see, this is because of the non-commutativity of quaternions under multiplication. In the case of scalar functions, the most general linear functions are of the form ax + b. In quaternion functions a linear form may be considerably more complex.

We need, then, a definition of differentiation which is applicable to any normed vector space and which will reduce to the customary definition whenever it is applicable. Such a definition was made by Hamilton:

'''<u>Simultaneous Differentials</u> are <u>Limits</u> <u>of Equimultiples</u> <u>of Simultaneous and Decreasing Differences</u>.'

'And conversely, whenever any <u>simultaneous differences</u>, of any system of variables, all <u>tend</u> to <u>vanish together</u>, according to any <u>law</u>, or system of laws; then, if any <u>equimultiples</u> of those decreasing differences all <u>tend</u> together to any system of finite <u>limits</u>, those <u>Limits</u> are <u>said</u> to be <u>Simultaneous Differentials</u> of the related <u>Variables</u> of the <u>System</u>; and are <u>denoted</u>, as such, by prefixing the <u>letter</u> d, as a <u>characteristic of differentiation</u>, to the <u>Symbol</u> of each such <u>variable</u>.'''[1]

More symbolically, let x, y, and z be related variables in any normed vector spaces. Let Δx, Δy, and Δz denote arbitrary increments in these variables. Then let the sums $x + \Delta x$, $y + \Delta y$, and $z + \Delta z$, be related in the same way that x, y, and z originally were. Thus if Δx, Δy, and Δz occur simultaneously, a relation between them is prescribed. Now consider $dx = n \Delta x$, $dy = n\Delta y$, and $dz = n\Delta z$, where n is a suitably chosen scalar. If $|\Delta x|$, $|\Delta y|$, and $|\Delta z|$ tend toward zero, we may expect dx, dy, and dz to do the same. But in many cases, we can make n tend toward infinity in such a way that dx, dy, and dz approach a system of finite limits. If this happens, dx, dy, and dz are said to be simultaneous differentials.

Now that the operation of differentiation has been defined and explained, let us illustrate it with an example. Let x be a quaternion and $y = x^2$. Let us find dy.

---

[1] W. R. Hamilton, <u>Elements of Quaternions</u>, p. 431.

$$y = x^2$$

$$y + \Delta y = (x + \Delta x)^2$$

$$\Delta y = (x + \Delta x)^2 - x^2$$

$$= x^2 + x\Delta x + (\Delta x)x + (\Delta x)^2 - x^2$$

$$= x(\Delta x) + (\Delta x)x + (\Delta x)^2$$

$$n\Delta y = nx(\Delta x) + n(\Delta x)x + n(\Delta x)^2$$

$$= x(n\Delta x) + (n\Delta x)x + \frac{(n\Delta x)^2}{n} .$$

We shall adjust n so that $(n\Delta x) = dx = \text{constant} \neq 0$.

$$n\Delta y = x(dx) + (dx)x + \frac{(dx)^2}{n} .$$

Now we shall let $\Delta x \rightarrow 0$ and $n \rightarrow \infty$. In the limit, $n\Delta y \rightarrow dy$, so

$$dy = x(dx) + (dx)x.$$

Note that the concept of infinitesimals plays no role in this definition; dy and dx are finite quantities and need not be small. Even when they are large, there can be no additional terms involving $(dx)^2$, for this would imply that for a constant $c = dx$, the term c/n did not tend to zero as n tended to infinity.

Out definition can be stated formally as

$$dy = \lim_{\substack{\Delta x \rightarrow 0 \\ n \rightarrow \infty \\ n\Delta x = \text{const.} \neq 0}} \left\{ n(y(x + \Delta x) - y(x)) \right\}$$

We can simplify this definition by introducing $dx = n\Delta x$ = constant.

Then $\Delta x = dx/n$, and

$$dy = \lim_{n\to\infty} n(y(x + \frac{dx}{n}) - y(x))$$

This is the definition which is normally given.

Starting then with a function $y(x)$ we arrive at a function $dy = y'(x, dx)$. We shall now develop two properties of $y'(x, dx)$ which may be considered fundamental properties independent of the form of y.

First we shall show that $y'(x, dx)$ is homogeneous of first order with respect to dx. Let s be an arbitrary scalar.

$$\begin{aligned} y'(x, sdx) &= \lim_{n\to\infty} n(y(x + \frac{sdx}{n}) - y(x)) \\ &= (s)\lim_{n\to\infty}(n/s)(y(x + \frac{sdx}{n}) - y(x)) \\ &= (s)\lim_{n\to\infty}(n/s)(y(x + dx/(n/s)) - y(x)) \\ &= (s)\lim_{(n/s)\to\infty}(n/s)(y(x + dx/(n/s)) - y(x)) \\ &= (s)y'(x, dx) \end{aligned}$$

Now we shall show that $y'(x, dx)$ is distributive with respect to dx, that is, if $dx = dr + ds$, then $y'(x, dr + ds) = y'(x, dr) + y'(x, ds)$.

$$y'(x, dr + ds) = \lim_{n\to\infty} n(y(x + \frac{dr + ds}{n}) - y(x))$$

$$= \lim_{n\to\infty}\left\{n\left(y\left(x + \frac{dr}{n} + \frac{ds}{n}\right) - y\left(x + \frac{ds}{n}\right) + y\left(x + \frac{ds}{n}\right)\right.\right.$$

$$\left.\left. -y(x)\right)\right\}$$

$$= \lim_{n\to\infty}\left\{n\left(y\left(x + \frac{dr}{n} + \frac{ds}{n}\right) - y\left(x + \frac{ds}{n}\right)\right)\right\}$$

$$\lim_{n\to\infty}\left\{n\left(y\left(x + \frac{ds}{n}\right) - y(x)\right)\right\}$$

$$= y'(x, dr) + y'(x, ds)$$

In the real field the most general function, y', which can satisfy both of these conditions is dy = (u) (dx), where u is an element of the field. Thus the differential always takes the form of a real number times dx. This may be regarded as an unusual property of the real numbers.

In the complex field the problem is not quite so simple. If $y = y_0 + Ey_1$, $E^2 = -1$, $x = x_0 + Ex_1$, and y is a function of x, then dy =

$$\frac{\partial y_0}{\partial x_0} dx_0 + \frac{\partial y_0}{\partial x_1} dx_1 + E\left\{\frac{\partial y_1}{\partial x_0} dx_0 + \frac{\partial y_1}{\partial x_1} dx_1\right\}.$$

If we assume that there exists a function f = u + Ev such that dy = (f) (dx), then $dy = (u + Ev)(dx_0 + Edx_1)$

$$= (u)(dx_0) - (v)(dx_1) + E\left\{(v)(dx_0) + (u)(dx_1)\right\}.$$

Thus we see that the desired f exists only if $\frac{\partial y_0}{\partial x_0} = \frac{\partial y_1}{\partial x_1}$ and $\frac{\partial y_1}{\partial x_0} = -\frac{\partial y_0}{\partial x_1}$.

So if we restrict our consideration to functions satisfying the Cauchy-Riemann equations, differentiation in the complex field resembles differentiation in the real numbers.

If, however, we attempt to do the same thing with quaternions, we find that the class of functions which are satisfactory seems too restricted to be useful.

To see this, consider the quaternion function $y = x^2$.

$$\begin{aligned} dy &= x(dx) + (dx)x \\ &= \left\{x + (dx)x(dx)^{-1}\right\} dx \end{aligned}$$

There does not exist a function of y alone which can be multiplied by dx on the right to produce dy. We encounter the same problem if we look for a right multiplier.

Thus, if we desire something which we can call a derivative, we must consider the derivative to be an operator. If $y = x^2$, then $y'(x,\ ) = x(\ ) + (\ )x$.

Let us develop a few more rules of differentiation. Note that it follows from the definition that $d(x) = dx$. First let $y = uv$, where u and v are functions of x.

$$dy = \lim_{n\to\infty}\left\{n\left(u\left(x + \frac{dx}{n}\right)v\left(x + \frac{dx}{n}\right) - u(x)v(x)\right)\right\}$$

$$= \lim_{n\to\infty}\left\{n\left(u\left(x + \frac{dx}{n}\right)v\left(x + \frac{dx}{n}\right) - u(x)v\left(x + \frac{dx}{n}\right)\right.\right.$$

$$+ u(x)v(x + \frac{dx}{n}) - u(x)v(x) \Big) \Big\}$$

$$= \lim_{n\to\infty} \Big\{ n\big(u(x + \frac{dx}{n})v(x + \frac{dx}{n}) - u(x)v(x + \frac{dx}{n})\big) \Big\}$$

$$+ \lim_{n\to\infty} \Big\{ n\big(u(x)v(x + \frac{dx}{n}) - u(x)v(x)\big) \Big\}$$

$$= (du)\,(v) + (u)\,(dv)$$

Example: Find $d(x^3)$.

$$d(x^3) = d(x^2 x)$$

$$= d(x^2)x + x^2\, dx$$

$$= \big((dx)x + x(dx) \big)x + x^2 dx$$

$$= (dx)x^2 + x(dx)x + x^2 dx$$

Example: Find $d(x^{-1})$.

$$(x)(x^{-1}) = 1$$

$$(x)d(x^{-1}) + (dx)\,(x^{-1}) = 0$$

$$(x)d(x^{-1}) = -(dx)(x^{-1})$$

$$d(x^{-1}) = -(x^{-1})(dx)(x^{-1})$$

Example: Find $d(x^{1/2})$

$$x^{\frac{1}{2}} x^{\frac{1}{2}} = x$$

1) $$x^{\frac{1}{2}} d(x^{\frac{1}{2}}) + d(x^{\frac{1}{2}})x^{\frac{1}{2}} = dx$$

$$d(x^{\frac{1}{2}})\overline{(x^{\frac{1}{2}})} + (x^{-\frac{1}{2}})d(x^{\frac{1}{2}})\,|x^{\frac{1}{2}}|^2 = x^{-\frac{1}{2}}(dx)\overline{(x^{\frac{1}{2}})}$$

Note that $x^{-\frac{1}{2}} = \overline{(x^{\frac{1}{2}})} / |x^{\frac{1}{2}}|^2$, so $x^{-\frac{1}{2}} |x^{\frac{1}{2}}|^2 = \overline{(x^{\frac{1}{2}})}$.

2) $$d(x^{\frac{1}{2}})\overline{(x^{\frac{1}{2}})} + \overline{(x^{\frac{1}{2}})}d(x^{\frac{1}{2}}) = x^{-\frac{1}{2}}\,dx\overline{(x^{\frac{1}{2}})}$$

Adding 1) and 2),

$$x^{\frac{1}{2}}d(x^{\frac{1}{2}}) + \overline{(x^{\frac{1}{2}})}d(x^{\frac{1}{2}}) + d(x^{\frac{1}{2}})x^{\frac{1}{2}} + d(x^{\frac{1}{2}})\overline{(x^{\frac{1}{2}})} = dx + x^{-\frac{1}{2}}\,dx\overline{(x^{\frac{1}{2}})}.$$

$$(x^{\frac{1}{2}} + \overline{x^{\frac{1}{2}}})d(x^{\frac{1}{2}}) + d(x^{\frac{1}{2}})(x^{\frac{1}{2}} + \overline{x^{\frac{1}{2}}}) = dx + x^{-\frac{1}{2}}\,dx\overline{(x^{\frac{1}{2}})}$$

$$2(x^{\frac{1}{2}})_0 d(x^{\frac{1}{2}}) + d(x^{\frac{1}{2}})2(x^{\frac{1}{2}})_0 = dx + x^{-\frac{1}{2}}\,dx\overline{(x^{\frac{1}{2}})}$$

$$4(x^{\frac{1}{2}})_0 d(x^{\frac{1}{2}}) = dx + x^{-\frac{1}{2}}\,dx\overline{(x^{\frac{1}{2}})}$$

$$d(x^{\frac{1}{2}}) = \frac{dx + x^{-\frac{1}{2}}(dx)\overline{x^{\frac{1}{2}}}}{4(x^{\frac{1}{2}})_0}$$

The reader may have noted that the manipulations in the last example were a little bit difficult. This illustrates the previous remark that linear equations in Q are not as simple as linear equations in the real numbers. We easily found dx in terms of a linear equation in $d(x^{\frac{1}{2}})$, but to get our answer we needed to solve for $d(x^{\frac{1}{2}})$ as a linear function in dx.

Since dy is a linear function in dx, the solution of linear equations is a very important topic in the theory of quaternion differentiation. Hamilton developed this subject extensively. However, it is beyond the scope of this paper. We shall content ourselves by noting that the solution of linear equations can be regarded as a matrix inversion problem. This can be seen by writing dy and dx as a pair of column vectors.

$$dy = \begin{bmatrix} dy_0 \\ dy_1 \\ dy_2 \\ dy_3 \end{bmatrix}, \quad dx = \begin{bmatrix} dx_0 \\ dx_1 \\ dx_2 \\ dx_3 \end{bmatrix}$$

Then
$$\begin{bmatrix} dy_0 \\ dy_1 \\ dy_2 \\ dy_3 \end{bmatrix} = \begin{bmatrix} \frac{\partial y_0}{\partial x_0} & \frac{\partial y_0}{\partial x_1} & \frac{\partial y_0}{\partial x_2} & \frac{\partial y_0}{\partial x_3} \\ \frac{\partial y_1}{\partial x_0} & \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_0} & \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \\ \frac{\partial y_3}{\partial x_0} & \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} & \frac{\partial y_3}{\partial x_3} \end{bmatrix} \begin{bmatrix} dx_0 \\ dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = D \begin{bmatrix} dx_0 \\ dx_1 \\ dx_2 \\ dx_3 \end{bmatrix}$$

Thus, if the matrix D^{-1} can be found, we can obtain dx as a function of dy.

SOME SPECIAL FUNCTIONS

To further illustrate the rules of differentiation, let us consider some functions which are fundamental to the algebra of quaternions: $\bar{x}$, x_0, x_v, $|x|$, and $\vec{x}$.

The differentials of the first three can be found quite readily, for let f(x) be any function which is homogeneous and distributive, i.e., $f(x + y) = f(x) + f(y)$ and, if s is a scalar, $f(sx) = sf(x)$.

$$df(x) = \lim_{n\to\infty} \left\{ n\left(f\left(x + \frac{dx}{n}\right) - f(x)\right) \right\}$$

$$= \lim_{n\to\infty} \left\{ n\left(f(x) + f\left(\frac{dx}{n}\right) - f(x)\right) \right\}$$

$$= \lim_{n\to\infty} \{n(f(dx/n))\}$$

$$= \lim_{n\to\infty} \{n(1/n)(f(dx))\}$$

$$= \lim_{n\to\infty} \{(f(dx))\}$$

$$df(x) = f(dx)$$

This gives us the formulas,

$$d(\bar{x}) = \overline{(dx)},$$

$$d(x_0) = (dx)_0,$$

and

$$d(x_v) = (dx)_v.$$

To find $d(|x|)$ we shall first find $d(|x|^2)$. Then we shall use the previous result for the differential of a square root.

$$|x|^2 = \bar{x}x$$

$$d(|x|^2) = (\bar{x})(dx) + d(\bar{x})(x)$$

$$= (\bar{x})(dx) + (\overline{dx})(x)$$

$$= (\bar{x})(dx) + \overline{((\bar{x})(dx))}$$

$$= 2((\bar{x})(dx))_0$$

From the formula for $d(q^{\frac{1}{2}})$, $d(|x|) = \dfrac{d(|x|^2) + |x|^{-1}d(|x|^2)\overline{|x|}}{4|x|_0}$

Note that $|x|$ is a scalar, $|x| = |x|_0$.

$$d(|x|) = \frac{d(|x|^2) + d(|x|^2)}{4|x|}$$

$$= \left\{2(\,(\bar{x})(dx)\,)_0 + 2(\,(\bar{x})(dx)\,)_0\right\}/(4|x|)$$

$$= \frac{(\,(\bar{x})(dx)\,)_0}{|x|}$$

This is often given in the more easily remembered form,

$$\frac{d(|x|)}{|x|} = (x^{-1}dx)_0$$

To obtain this, note that $((\bar{x})(dx))_0/|x|^2 = ((\bar{x}/|x|^2)(dx))_0 = (x^{-1}dx)_0.$

Now we can find $d(\vec{x})$.

$$x = |x|\vec{x}$$

$$dx = d(|x|)(\vec{x}) + |x|d(\vec{x})$$

$$|x|d(\vec{x}) = dx - d(|x|)\vec{x}$$

$$d(\vec{x})(\vec{x}^{-1}) = dx(x^{-1}) - d(|x|)/|x|$$

$$= dx(x^{-1}) - (\,(x^{-1})(dx)\,)_0$$

$$= (dx(x^{-1})\,)_v$$

$$d(\vec{x}) = (dx(x^{-1})\,)_v\vec{x}$$

HIGHER DIFFERENTIATION

In order to arrive at an understanding of what is meant by higher differentiation we need to consider the differential of a function of two variables. Let z be a function of two quaternions, x and y.

$$dz = \lim_{n\to\infty} n\left\{z(x + \frac{dx}{n},\ y + \frac{dy}{n}) - z(x, y)\right\}$$

$$= \lim_{n\to\infty} n\left\{z(x + \frac{dx}{n},\ y + \frac{dy}{n}) - z(x, y + \frac{dy}{n})\right.$$

$$\left. + z(x, y + \frac{dy}{n}) - z(x, y)\right\}$$

$$= \lim_{n\to\infty} n\left\{z(x + \frac{dx}{n},\ y + \frac{dy}{n}) - z(x, y + \frac{dy}{n})\right\}$$

$$+ \lim_{n\to\infty} n\left\{z(x,\ y + \frac{dy}{n}) - z(x, y)\right\}$$

$$= z'_x(x, y, dx) + z'_y(x, y, dy)$$

$$= d_x z + d_y z.$$

Here $z'_x(x, y, dx)$ and $d_x z$ denote the result which will be obtained if we assume that x is the variable and y is a constant. Note that if y is, in fact, a constant then $dy = 0$. Since $z'_y(x, y, dy)$ is homogeneous of first order in dy, $z'_y(x, y, 0) = 0$.

This reasoning leads us easily to an understanding of what form higher differentials must take. We start with a function f(x). Differentiating once, we obtain df(x, dx). To differentiate this, we shall simply treat df(x, y) as a function of two "independent" variables. The independence is modified, however, by the identification of y with dx.

1) $$d^2 f = d_x(df(x, y, dx)) + d_y(df(x, y, dy))$$

We shall use d^2x to denote dy and think of d^2x as the second differential of x.

$$d^2 f = d_x(df(x, dx, dx)) + d_{dx}(df(x, dx, d^2x))$$

Note that the first term on the right is no longer necessarily homogeneous of the first degree in dx. However, the second term must be homogeneous of the first degree in d^2x. This will be implied when we write the above form simply as $d^2f(x, dx, d^2x)$.

Higher differentials will be treated in the same way.

We shall now establish some notation which will be essential in the development of Taylor's series.

Note that d^2x was considered the differential of dx. If we assume that dx is a constant, this implies that

$$0 = d^2x = d^3x = \cdots = d^nx = \cdots .$$

We may expect that we could then differentiate df(x, dx) as if x were the only variable. Equation 1) shows that this will give $d^2f(x, dx, 0)$. Similarly, a third differentiation, still made on the assumption that x is the only variable involved, will give $d^3f(s, dx, 0, 0)$. For notational convenience, then, we shall simply write $d^2f(x, dx)$ and $d^3f(x, dx)$ with the understanding that dx is being treated as a constant.

Example: If $y = x^2$, find d^2y and d^3y.

$$dy = x(dx) + (dx)x$$

$$d^2y = (dx)^2 + x(d^2x) + (d^2x)x + (dx)^2$$

$$= x(d^2x) + 2(dx)^2 + (d^2x)x$$

$$d^3y = x(d^3x) + (dx)(d^2x) + 2(dx)(d^2x) + 2(d^2x)(dx)$$

$$+ (d^3x)x + (d^2x)(dx)$$

$$= x(d^3x) + 3(dx)(d^2x) + 3(d^2x)(dx) + (d^3x)x$$

Here x, dx, d^2x, and d^3x are three independent variables. However, if dx is considered a constant, $d^2x = d^3x = 0$. Then

$$d^2y = 2(dx)^2$$

$$d^3y = 0$$

THE TAYLOR SERIES

Let a and b be two quaternion constants and s a scalar. If f is a quaternion function of the quantity (a + sb), then f is a function of s. Consider df/ds.

$$\frac{df}{ds} = \lim_{\Delta s \to 0} \frac{f(a + sb + \Delta sb) - f(a + sb)}{\Delta s}$$

$$= \lim_{\Delta s \to 0} \left(\frac{1}{\Delta s}\right)\left(f(a + sb + \frac{b}{1/\Delta s}) - f(a + sb)\right)$$

$$= \lim_{(1/\Delta s) \to \infty} \left(\frac{1}{s}\right)\left(f(a + sb + \frac{b}{1/\Delta s}) - f(a + sb)\right)$$

$$= df(a + sb, b)$$

If this is evaluated at $s = 0$, $df/ds = df(a, b)$. Also

$$\frac{d^2f}{ds^2} = \frac{d}{ds}(f(a + sb, b))$$

$$= d^2f(a + sb, b)$$

$$\frac{d^nf}{ds^n} = d^nf(a + sb, b)$$

Now consider a quaternion function $f(x + sdx)$, where we shall treat x and dx as constants. We shall think of f as a function of s. We can write the Taylor series of f, subject, of course, to questions of convergence which are beyond the scope of this paper. We shall expand the series about $s = 0$.

$$f(x + sdx) = f(x) + sdf + \frac{s^2d^2f}{2!} + \cdots + \frac{s^nd^nf}{n!} + \cdots .$$

Note that although x and dx are considered constant, they are arbitrary.

Example: If $df = x(dx) + (dx)x$, and $f(x)$ is known for some $x = a$, find $f(x)$ for any $x \neq a$.

We shall assume that dx is a constant.

$$d^2f(x, dx) = (dx)(dx) + (dx)(dx) = 2(dx)^2$$

$$d^3f(x, dx) = 0$$

$$f(a + sdx) = f(a) + s(a(dx) + (dx)a) + s^2 2(dx)^2$$

$$= f(a) + a(sdx) + (sdx)a + (sdx)^2$$

Let $c = f(a) - a^2$.

$$f(a + sdx) = a^2 + a(sdx) + (sdx)a + (sdx)^2 + c$$

$$= (a + sdx)^2 + c$$

If $x = a + sdx$, then $f(x) = x^2 + c$.

Example: Let $df = x(dx)x + x^2(dx)$ and let $f(x)$ be given for some $x = a$. Find $f(b)$ for a given $b \neq a$.

Again assume that dx is a constant.

$$d^2f = (dx)^2x + x(dx)^2 + x(dx)^2 + (dx)x(dx)$$

$$d^2f = (dx)^2x + 2x(dx)^2 + (dx)x(dx)$$

$$d^3f = 4(dx)^3$$

$$d^4f = 0$$

$$f(a + sdx) = f(a) + s(a(dx)a + a^2dx)$$
$$+ \tfrac{1}{2}s^2 \left\{(dx)^2a + 2a(dx)^2 + (dx)a(dx)\right\}$$
$$+ \frac{s^3}{6} \left\{4(dx)^3\right\}$$

$$= f(a) + a(sdx)a + a^2(sdx) + \tfrac{1}{2}(sdx)^2a$$
$$+ a(sdx)^2 + \tfrac{1}{2}(sdx)a(sdx) + \frac{2}{3}(sdx)^3$$

In this example no simple expression, such as was obtained in the first example, is apparent. However, this expression will enable us to evaluate $f(b)$ for any b in Q. This can be accomplished simply by letting $sdx = b - a$. When this value is inserted into the series, $f(b)$ is obtained.

The foregoing material suggests two questions:

1. What form does a theory of integration take?

2. To what extent can this theory be extended to an arbitrary vector space, and how?

These questions remain, at present, unanswered. We shall confine ourselves to the unsatisfactory observation that they would be interesting areas of further study.

ACKNOWLEDGMENT

The writer wishes to express his deepest thanks to Professor Leonard E. Fuller for his patience and guidance in the writing of this paper. This has been a fascinating and very educational experience.

REFERENCES

Brand, Louis. Vector and Tensor Analysis. New York: John Wiley & Sons, Inc., 1948.

Hamilton, W. R., Elements of Quaternions, second edition. London: C. Jasper Joly, 1899-1901.

Kelland, P. and Tait, P. G. Introduction to Quaternions. London: Macmillan and Company, 1882.

McAulay, A. Utility of Quaternions in Physics. London: Macmillan & Company, 1893.

Macfarlane, Alexander. Vector Analysis and Quaternions. New York: John Wiley & Sons, Inc., London: Chapman and Hall, Limited, 1906.

Tait, P. G. An Elementary Treatise on Quaternions. Cambridge at the University Press, 1873.

QUATERNION CALCULUS

by

DAVID JOSEPH EDELBLUTE

B. S., Kansas State University, 1962

AN ABSTRACT OF A REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1964

Although quaternions were never widely accepted as an analytic tool for applied fields, ideas which arose from the study of quaternions have had a profound effect on subsequent mathematical development. The purpose of this paper was to investigate the use of the ideas of calculus when dealing with quaternions, and to see what ideas about calculus might arise from this investigation.

Before the discussion of calculus was begun, it was felt advisable to establish a notation which was well adapted for representing quaternions and which exposed some of their geometric properties. The examination of these geometric properties was further aided by a discussion of their application in describing spherical geometry and rotations.

It was found that the usual definition of differentiation, or of differential coefficients, was not useful in the theory of quaternion functions. The definition was modified by Hamilton in what at first seems to be a rather curious way. Using Hamilton's definition a differential calculus of quaternions was developed. This calculus is unusual in that there is (usually) no function which can be regarded as a derivative, and the idea of infinitesimals plays no role.

In conclusion, it was shown that, in one sense, Taylor's series can be used to describe quaternions which are functions of other quaternions.