A Survey of Quaternionic Analysis

Christopher Stover

January 5, 2014

Abstract

The group $\mathbb{H}$ of (Hamilton) quaternions can be described, algebraically, as a four-dimensional associative normed division algebra over the ring $\mathbb{R}$ of real numbers. This paper is meant to be a survey on these numbers. In particular, the paper is broken into four main sections. The quaternions are introduced in section one, and in section two, a bit of preliminary information needed to understand the study of functions of a quaternionic variable is introduced. Section three is meant to give an overview of some of the properties of quaternion-valued functions including (but not limited to) properties that parallel main ideas in complex-valued theory. The fourth and final section of the paper describes in varying degrees of detail several different applications of quaternionic analysis. An appendix of notational standards is included for convenience.

Contents

1 Introduction

First invented in 1843 by Irish mathematician W.R. Hamilton as a tool to mathematically describe the quotient of vectors in 3-space, the group $\mathbb{H}$ of (Hamilton) quaternions can be described, algebraically, as a four-dimensional associative normed division algebra over the ring $\mathbb{R}$ of real numbers. Unlike the field $\mathbb{C}$ of complex numbers, the quaternions fail to be (multiplicatively) commutative. More generally, the quaternions hold the distinction of being the first noncommutative division algebra to be discovered [1], as well as being one of only three finite-dimensional associative division algebras over the reals.

Another distinction possessed by the real quaternions is their apparent two-sided nature. To quote [2],:

> The richness of the theory of functions over the complex field makes it natural to look for a similar theory for the only other non-trivial real associative division algebra, namely the quaternions. Such a theory exists and is quite far-reaching, yet it seems to be little known. It was not developed until nearly a century after Hamiltons discovery of quaternions. Hamilton himself and his principal followers and expositors...only developed the theory of functions of a quaternion variable as far as it could be taken by the general methods of the theory of functions of several real variables....

The author of [2] goes on to indicate that neither Hamilton nor his expositors investigated quaternion-valued functions in a way that fully acknowledged the richness of the structure they possessed. In particular, neither Hamilton nor his expositors investigated "special" properties of quaternion-valued functions that would liken them to the well-studied, property-rich class of analytic complex-valued functions, thereby leaving much to be desired in terms of the study and understanding of quaternionic analysis.

This fact was easily noticed and as such formed the basis for the "second wave" of investigation of properties of quaternion-valued functions of a quaternion-variable. Unfortunately for the field, this second wave came only after a period in which the quaternions were deemed in poor taste by many prominent scientists and mathematicians and, as such, were largely ignored in favor of a handful of blossoming alternatives including vector and tensor analysis. A rebirth of sorts occurred in the early to mid 1930s when Swiss mathematician Rudrig Fueter began to investigate quaternionic analogues of key complex-analytic results, the most significant of which came in the form of an analogous set of Cauchy-Riemann equations for functions $f : \mathbb{H} \to \mathbb{H}$, published in 1935.

In the decades since, the usefulness and depth of the field of quaternionic analysis has led to its emergence as a meaningful research topic. The work done by Fueter and his colleagues in the 30s and 40s served, in a sense, as a mere introduction, and despite what [2] calls a lack of rigor, their work remained the most prominent for well over a decade. Since then, Cartan's work on several complex variables (see [3]) as well as treatises by Cullen and Deavours (see [4] and [5], respectively), have brought the field of quaternionic analysis out of the shadows and have served to introduce its appeal and mystique to entirely new generations of mathematicians.

At its core, this article is meant to be a survey of a topic that can be called, at the very least, "extensive." More specifically, the purpose of this writing is to give a summary of the analysis of quaternion functions of a quaternion variable which manages to be thorough while also being succinct enough to be easily-readable. Contained herein will be an exposition intended to be independent and self-contained while still managing to incorporate enough topical breadth to at least convince the reader of the significance of the topic at hand.

2 Quaternion Arithmetic and Algebra

As mentioned in the introduction, the quaternions $\mathbb{H}$ form a four-dimensional associative normed division algebra over the ring $\mathbb{R}$ which fails to be multiplicatively commutative. Recall that an *associative division algebra* is an algebra D over a field $\mathbb{F}$ which satisfies two algebraic properties:

1. There exists a multiplicative identity $1_D \neq 0$ in D
2. For every nonzero $d \in D$, there exist elements $a_1, a_2 \in D$ for which $a_1 d = 1_D = d a_2$.

When the algebra D is multiplicatively commutative, the condition in (2) simplifies, as $a_1 = a_2$. The division algebra D is said to be a *normed division algebra* in the event that it is a normed vector space, i.e. that the vector space D is part of the pair $(D, \|\cdot\|_D)$, where $\|\cdot\|_D : D \to \mathbb{R}$ is a function satisfying the following three properties:

1. $\|av\|_D = |a|\|v\|_D$ for all $a \in \mathbb{F}, v \in D$
2. $\|v\|_D = 0$ if and only if $v = 0$ in D
3. $\|u + v\|_D \leq \|u\|_D + \|v\|_D$ for all $u, v \in D$.

In addition to being endowed with the algebraic structure described above, the Hamilton quaternions can be viewed as a degree four extension of the reals, namely a set of the form

$$\mathbb{H} = \{a + bi + cj + dk \,|\, a, b, c, d \in \mathbb{R}\}$$

with basis elements $\{1, i, j, k\}$ subject to the multiplicative relations

$$i^2 = j^2 = k^2 = -1,\ ij = k = -ji,\ jk = i = -kj,\ ki = j = -ik.$$

Being a division algebra, $\mathbb{H}$ is endowed with two associative operations, namely addition and multiplication, which are defined as follows: Given elements $q_1 = a + bi + cj + dk$ and $q_2 = \alpha + \beta i + \gamma j + \delta k$ in $\mathbb{H}$, their sum and product are defined as follows:

$$q_1 + q_2 = (a+\alpha) + (b+\beta)i + (c+\gamma)j + (d+\delta)k, \text{ and} \tag{2.0.1}$$

$$\begin{aligned} q_1 q_2 = (a\alpha - b\beta - c\gamma - d\delta) + (a\beta + b\alpha + c\delta - d\gamma)i \\ + (a\gamma - b\delta + c\alpha + d\beta)j + (a\delta + b\gamma - c\beta + d\alpha)k. \end{aligned} \tag{2.0.2}$$

Note that in light of its vector space properties, $\mathbb{H}$ is often written as a collection of 4-tuples of the form $(a, b, c, d) = a + bi + cj + dk$, $a, b, c, d \in \mathbb{R}$. In particular, then, equations (2.0.1) and (2.0.2) have the following alternative but equivalent expressions:

$$\begin{aligned} q_1 + q_2 &= (a+\alpha, b_\beta, c+\gamma, d+\delta) \\ q_1 q_2 &= (a\alpha - b\beta - c\gamma - d\delta, a\beta + b\alpha + c\delta - d\gamma, \\ &\qquad a\gamma - b\delta + c\alpha + d\beta, a\delta + b\gamma - c\beta + d\alpha). \end{aligned}$$

As described above, $\mathbb{H}$ is equipped with a norm $\|\cdot\|_{\mathbb{H}}$, the value of which can be written for any $q = (a, b, c, d) \in \mathbb{H}$ as

$$\|q\|_{\mathbb{H}} = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

When there's no risk for confusion, $\|\cdot\|_{\mathbb{H}}$ will be written $|\cdot|$. Other notation utilized in the study of $\mathbb{H}$ can be found in [2], [6], and others, though the following is highlighted because of its prevalence throughout:

$$\overline{q} = a - bi - cj - dk \tag{2.0.3}$$

$$\begin{aligned} \operatorname{Re} q &= \frac{1}{2}(q + \overline{q}) = a \in \mathbb{R} \\ \operatorname{Im} q &= \frac{1}{2}(q - \overline{q}) = bi + cj + dk. \end{aligned}$$

Note, too, that by the definition given in (2.0.3) of the conjugate $\overline{q}$ of $q \in \mathbb{H}$, the following properties analogous to those known in $\mathbb{C}$ follow trivially:

$$\begin{aligned} \overline{q_1 q_2} &= \overline{q_1} \cdot \overline{q_2} \\ \|q\|_{\mathbb{H}} &= \sqrt{q\overline{q}} \end{aligned}$$

$$q^{-1} = \frac{\overline{q}}{\|q\|_{\mathbb{H}}^2}. \tag{2.0.4}$$

Given a quaternion $q \in \mathbb{H}$, property (2.0.4) shows precisely how to compute its multiplicative inverse $q^{-1} \in \mathbb{H}$. Hence, property (2.0.4) confirms that $\mathbb{H}$ is actually a skew field (i.e, a division algebra), a claim which until now had not been verified.

What stands out up until this point is that, with the exception of multiplicative commutativity, operations on $\mathbb{H}$ mirror those on $\mathbb{C}$ almost exactly. It would be especially nice if this trend continued, though verification of such a continuation is no simple feat. For that reason, the remainder of this section will be dedicated to stating and proving some fundamental results which are arithmetic and/or algebraic in nature, the first of which comes from [7].

Theorem 2.1. Let $q \in \mathbb{H}$ be a quaternion. If $q \in \mathbb{H} \setminus \mathbb{R}$, then q has precisely n nth roots. If $q \in \mathbb{R}$ and $q > 0$, then q has two square roots (namely, $\pm\sqrt{q}$); otherwise, if $q < 0$, then q has infinitely many roots.

Proof. Let $q = a + bi + cj + dk$ and let $\theta \in [0, 2\pi)$. Note that q can be written in "polar form" by letting $r = |q|$ and by noting that $\mathrm{Re}(q) = a = r(a/r)$ and that, for $q \notin \mathbb{R}$,

$$\mathrm{Im}(q) = r\left(\frac{|\mathbf{v}|}{r}\right)\left(\frac{\mathbf{v}}{|\mathbf{v}|}\right) \text{ where } \mathbf{v} = \mathrm{Im}(q) = bi + cj + dk.$$

Combining these results and defining

$$\cos\theta = \frac{a}{r}, \quad \sin\theta = \pm\frac{|\mathbf{v}|}{r}, \quad \text{and} \quad \varepsilon = \pm\frac{\mathbf{v}}{|\mathbf{v}|} \tag{2.0.5}$$

gives an identity for $q \notin \mathbb{R}$ in terms of (r, θ), namely

$$q = r\left(\cos\theta + \varepsilon\sin\theta\right). \tag{2.0.6}$$

By convention, let $\varepsilon = 0$ when $q \in \mathbb{R}$. Note that when $\varepsilon \neq 0$,

$$\varepsilon^2 = \frac{\mathbf{v}^2}{b^2 + c^2 + d^2} = -1,$$

and so raising both sides of (2.0.6) to the power n yields

$$\begin{aligned} q^n &= r^n\left(\cos\theta + \varepsilon\sin\theta\right)^n \\ &= r^n\left(\cos n\theta + \varepsilon\sin n\theta\right) \text{ by de Moivre's Theorem.} \end{aligned}$$

Suppose, then, that there exists some $Q \in \mathbb{H}$ for which $q^n = Q$. It follows that Q can be written in a form similar to (2.0.6), namely

$$Q = R(\cos\phi + \varepsilon\sin\phi)$$

for ε as above, $\phi \in [0, \pi)$, $R = r^n$, $\cos\phi = \cos n\theta$, and $\sin\phi = \sin n\theta$. In this form, it follows that the nth roots of Q come from (2.0.6) under the hypotheses

$$\begin{aligned} r &= R^{1/n}, \text{ where the root is positive, and} \\ \theta &= \frac{\phi}{n} + \frac{2\pi m}{n} \text{ for } m = 1, 2, \ldots, n-1. \end{aligned} \tag{2.0.7}$$

Note that the n values given for θ in (2.0.7) designate all angles $\theta \in [0, 2\pi)$ which satisfy the conditions necessary. Moreover, by considering the distinct cases mentioned in the statement of the theorem—namely, the cases where (a) $Q \in \mathbb{H} \setminus \mathbb{R}$, (b) $Q > 0$ is real and $n = 2$, (c) $Q > 0$ is real and $n > 2$, and (d) $Q < 0$ is real for any n—sufficient restrictions on ε in (2.0.6) and on θ in (2.0.7) emerge so as to verify the conclusions thereof. Hence, the result. $\square$

Remark 2.1. The proof of theorem 2.1 is useful for several reasons. Besides exhibiting an explicit analogue to the polar form of a complex number, it also illustrates a useful notational quirk highlighted in [2], among others. In particular, a quaternion $q = a + bi + cj + dk$ can be written as $q = a + \mathbf{v}$ where $\mathbf{v} = bi + cj + dk$ is thought of as an oriented vector in $\mathbb{R}^3$. Moreover, the group $\mathbb{H}$ can thus be written as a direct sum of the form $\mathbb{H} = \mathbb{R} \oplus P$, where P is an oriented three-dimensional Euclidean vector space corresponding to the "imaginary parts" of elements of $\mathbb{H}$, with multiplication given by

$$\mathbf{ab} = -\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b} \text{ for all } \mathbf{a}, \mathbf{b} \in P.$$

It's easily verified that the arithmetic laws given in (2.0.1) and (2.0.2) hold for quaternions written in this form.

The expression of q given in (2.0.6) serves as a sort of generalization of Euler's identity for complex numbers. This notion can be made more precise by observing that the product of the unit quaternion ε and an arbitrary angle θ can be plugged into the "usual" power series[1] for $\exp(z)$ in much the same way that $i\theta$ was in the proof of Euler's identity in $\mathbb{C}$. Doing so yields a quaternionic version of Euler's identity, namely

$$\exp(\varepsilon\theta) = \cos\theta + \varepsilon \sin\theta. \tag{2.0.8}$$

More generally, consider a quaternion $q = a + \mathbf{v}$. The expression for ε in (2.0.5) yields that $\mathbf{v} = \varepsilon|\mathbf{v}|$, and so

$$\begin{aligned}
\exp(q) &= \exp(a + \mathbf{v}) \text{ by definition of } q \\
&= \exp(a)\exp(\mathbf{v}) \\
&= \exp(a)\exp(\varepsilon|\mathbf{v}|), \text{ since } \mathbf{v} = \varepsilon|\mathbf{v}| \\
&= \exp(a)\left(\cos|\mathbf{v}| + \varepsilon \sin|\mathbf{v}|\right) \text{ by (2.0.8)} \\
&= e^a \left(\cos|\mathbf{v}| + \frac{\mathbf{v}}{|\mathbf{v}|}\sin|\mathbf{v}|\right). \qquad (2.0.9)
\end{aligned}$$

Unsurprisingly, the expression in (2.0.9) yields the ability to define several other algebraic properties on $\mathbb{H}$, one of which is the quaternionic logarithm. Recall from complex analysis that the logarithm of an arbitrary element $z \in \mathbb{C}$ is a multi-valued expression defined by

$$\log(z) = \log|z| + i\arg(z), \tag{2.0.10}$$

where $\log|z|$ denotes the real-valued logarithm and where $\arg(z)$ denotes the angles $\theta + 2n\pi$ corresponding to the polar representation of $z = |z|e^{i(\theta+2n\pi)}$. In almost identical fashion, given a quaternion $q = a + \mathbf{v}$, the relations in (2.0.5) yield that $a = |q|\cos\theta$ and hence

[1] The power series being referenced here is $\exp(z) = \sum_{n=1}^{\infty} z^n/n!$.

that $\theta = \arccos(a/|q|) + 2n\pi$. Therefore, mimicking the form of (2.0.10), it follows that the logarithm formula for a quaternion q is precisely

$$\begin{aligned} \log(q) &= \log|q| + \varepsilon \arg(q) \\ &= \log|q| + \frac{\mathbf{v}}{|\mathbf{v}|}\left(\arccos\left(\frac{a}{|q|}\right) + 2n\pi\right). \end{aligned} \tag{2.0.11}$$

The story about quaternions so far has been the ability to mimic within that particular number system the behavior exhibited by the field of complex numbers—a fact that's highlighted in [8], [9], and others. Unfortunately, the pattern that seems so promising after deriving equations (2.0.9) and (2.0.11) ends with those derivations. This fact is subtle, and is a result of the non-commutativity of $\mathbb{H}$. The following proposition summarizes this.

Proposition 2.1. Let $q, t \in \mathbb{H}$, $r \in \mathbb{R}$. The following hold.

1. $\exp(r + q) = \exp(r)\exp(q)$
2. $(\exp(q))^r = \exp(qr)$
3. $\log(rq) = \log(r) + \log(q)$
4. $\log(q^r) = r\log(q)$
5. $\exp(q + t) \neq \exp(q)\exp(t)$ in general
6. $\log(qt) \neq \log(q) + \log(t)$ in general

It's worth noting that facts 5 and 6 fail because the non-commutativity of $\mathbb{H}$ makes the right hand sides of those expressions ambiguous. The effects of this failure carry over into the algebraic properties of $\mathbb{H}$, as well. To see this carry-over, look no further than the following: For complex numbers $z, \zeta \in \mathbb{C}$, the expression z^ζ can be defined by first writing

$$\begin{aligned} z^\zeta &= \exp\left[\log\left(z^\zeta\right)\right] \\ &= \exp\left[\zeta \log(z)\right], \end{aligned} \tag{2.0.12}$$

and then evaluating the logarithm using the formula in (2.0.10) above. The reason the evaluation in (2.0.12) is unambiguous, however, is because elements in $\mathbb{C}$ commute multiplicatively: That is, $\zeta\log(z) = \log(z)\zeta$, and so no ambiguity is present. Because this behavior doesn't translate into the language of the quaternions, hoping to evaluate an expression like q^t for $q, t \in \mathbb{H}$ becomes a completely different animal—one that requires machinery that's both advanced *and* not unanimously agreed upon. In fact, [8] claims that the identity

$$q^t = \exp(\log(q)t)$$

is valid (where the multiplication $\log(q)t$ denotes multiplication in the quaternion sense), while [10] indicates that such an expression requires use of *at least* trigonometric, inverse trigonometric, and hyperbolic trigonometric identities for the vector parts of q and t. It's important to recognize that, to a large extent, such uncertainty and ambiguity is a microcosm of the study of the quaternions. This theme will prove to be a recurring one throughout the remainder of this writing.

We conclude the introduction to arithmetic and algebraic properties of $\mathbb{H}$ by returning to one of the claims made in the introduction to the section. More precisely, we conclude with the following theorem, which was proven in 1877 by its namesake and which serves to characterize the uniqueness of the Hamilton Quaternions among finite-dimensional associative extensions of $\mathbb{R}$.

Theorem 2.2 (Frobenius Theorem)**.** Up to isomorphism, there are precisely *three*[2] associative division algebras which are finite-dimensional extensions of $\mathbb{R}$, namely $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$. These have dimensions 1, 2, and 4 over $\mathbb{R}$, respectively.

Proof. An excellent proof which is also very accessible can be found in [11]. □

3 Quaternionic Analysis and Function Theory

3.1 Differentiability and Analyticity

As demonstrated in appendix section 1.2, one of the most logical starting points regarding analytic theory for functions defined over a given number field lies in the definition of function differentiability. Indeed, taking derivatives *is* one of the logical starting places for performing calculus-type analysis and, which immediately leads to the necessity of defining the quaternionic derivative. Upon doing so, the complications inherent to $\mathbb{H}$ immediately reemerge, however.

The standard definition of the derivative requires investigation of the limiting behavior of a particular difference quotient. For example, given a function $g : \mathbb{C} \to \mathbb{C}$, the derivative g' is considered by investigating what happens tothe expression

$$h^{-1}\left[g(x+h) - g(x)\right] \text{ as } h \to 0. \tag{3.1.1}$$

As pointed out regarding equation (2.0.12) above, the expression in (3.1.1) is unambiguous because of the multiplicative commutativity of $\mathbb{C}$: In particular, the derivative of g can be well-defined by way of (3.1.1) because $h \in \mathbb{C}$ implies that

$$h^{-1}\left[g(x+h) - g(x)\right] = \left[g(x+h) - g(x)\right] h^{-1}. \tag{3.1.2}$$

Note that for a function $g : \mathbb{H} \to \mathbb{H}$, however, the limiting expression h is an element of $\mathbb{H}$, thereby indicating that the expression in (3.1.2) need not be an equality for arbitrary $h \in \mathbb{H}$.

[2]Closely related, Hurwitz' Theorem shows that sans the condition of associativity, there is a fourth division algebra (namely, the octonians $\mathbb{O}$) which is a finite-dimensional extension of $\mathbb{R}$. $\mathbb{O}$ is a dimension-eight extension of $\mathbb{R}$ which is neither associative nor multiplicatively commutative.

The most straight-forward way to remedy this situation is to consider *two* limits—one for each of the left and right difference quotients—and to define instead one-sided differentiability. To that end, consider the following.

Definition 3.1. Let $f : \mathbb{H} \to \mathbb{H}$ be a quaternion-valued function of a quaternion variable.

1. f is said to be **(quaternion) differentiable on the left** if

$$\lim_{h \to 0} \left[h^{-1}(f(q+h) - f(q)) \right]$$

   exists. In this case, the limit will be denoted by df_ℓ/dq or by f'_ℓ.

2. Similarly, f is said to be **(quaternion) differentiable on the right** if

$$\lim_{h \to 0} \left[(f(q+h) - f(q))h^{-1} \right]$$

   exists. In this case, the limit will be denoted by df_r/dq or by f'_r.

At first glance, this piecewise definition of differentiability seems a valid method to overcome the noncommutativity of quaternion multiplication, provided one is willing to ignore the obvious fact that two times the work will be required to develop any sort of meaningful theory whatsoever. One way around *that* caveat is by assuming that *differentiable* in the quaternion sense always means *left differentiable*. Unfortunately, even that is an insufficient work-around.

Consider the straightforward example of the function $f(q) = q^2$. Note that

$$f(q+h) = (q+h)^2 = (q+h)(q+h) = q(q+h) + h(q+h) = q^2 + qh + hq + h^2,$$

where the order of the multiplication is essential. It follows, then, that

$$\begin{aligned} h^{-1}\left[f(q+h) - f(q)\right] &= h^{-1}\left(q^2 + qh + hq + h^2 - q^2\right) \\ &= h^{-1}\left(qh + hq + h^2\right) \\ &= h^{-1}qh + q + h. \end{aligned} \tag{3.1.3}$$

Because $\mathbb{H}$ is non-commutative, the expression in (3.1.3) *can't* be simplified, from which it follows that

$$\lim_{h \to 0} \left[h^{-1}(f(q+h) - f(q)) \right] \text{ fails to exist}$$

even for this most basic function f. A similar conclusion follows if the definition of right differentiability is considered instead, which begs the question: Exactly how bad *is* the quaternion derivative? Consider the following well-known result, the statement of which is in [2] among others.

Theorem 3.1. Suppose the function f is defined and is quaternion-differentiable on the left throughout a connected open set $U \subset \mathbb{H}$. Then on U, f has the form

$$f(q) = a + qb \text{ for some } a, b \in \mathbb{H}.$$

Before proving this, the following well-known lemma regarding analytic functions of a complex variable is needed.

Lemma 3.1 (Hartog's Theorem)**.** Let $G \subset \mathbb{C}^n$ be an open set and write $z = (z_1, \ldots, z_n)$. Let $f : G \to \mathbb{C}$ be a function such that, for each $k = 1, \ldots, n$ and for fixed $z_1, \ldots, z_{k-1}, z_{k+1}, \ldots, z_n$, the function sending w to $f(z_1, \ldots, z_{k-1}, w, z_{k+1}, \ldots, z_n)$ is analytic on the set

$$\{w \in \mathbb{C} \,|\, (z_1, \ldots, z_{k-1}, w, z_{k+1}, \ldots, z_n) \in G\}.$$

Then f is continuous on G.

Proof. The proof of this theorem can be found in [3] and is omitted for brevity. □

Proof of Theorem 3.1. The proof given in [12] proceeds by way of an intricate use of vector calculus on partial derivatives. For a more straightforward argument, we follow the outline utilized by [2]. The argument makes use differentials, which are discussed briefly in appendix section 1.1.

Without loss of generality, suppose $U \subset \mathbb{H}$ a connected open set which is convex. Given an arbitrary quaternion $q = t + ix + jy + kz$, $t, x, y, z \in \mathbb{R}$, write q as the sum $q = v + jw$ where $v = t + ix$, $w = y - iz$ are elements of $\mathbb{C}$. For $f : \mathbb{H} \to \mathbb{H}$ arbitrary and left differentiable on U, note that the differential df has the form

$$dq \frac{df}{dq} = df,$$

and hence,

$$(dt + idx + jdy + kdz) \frac{df}{dq} = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz. \tag{3.1.4}$$

Equating coefficients in (3.1.4) yields that

$$\begin{aligned} \frac{df}{dq} &= \frac{\partial f}{\partial t} \\ &= \frac{1}{i} \frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial x} \\ &= \frac{1}{j} \frac{\partial f}{\partial y} = -j \frac{\partial f}{\partial y} \\ &= \frac{1}{k} \frac{\partial f}{\partial z} = -k \frac{\partial f}{\partial z}. \end{aligned} \tag{3.1.5}$$

Next, write

$$f(q) = g(v, w) + jh(v, w),$$

where $g, h : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$. Because $v = t + ix$ and $w = y - iz$, each of g and h can be differentiated with respect to t and combined with the equations in (3.1.5) to yield:

$$\begin{aligned} \frac{\partial g}{\partial t} &= -i\frac{\partial g}{\partial x} = \frac{\partial h}{\partial y} = i\frac{\partial h}{\partial z}, \text{ and} \\ \frac{\partial h}{\partial t} &= i\frac{\partial h}{\partial x} = -\frac{\partial g}{\partial y} = i\frac{\partial g}{\partial z}. \end{aligned} \tag{3.1.6}$$

Given that g, h are $\mathbb{C}$-valued functions, the identities given in (3.1.6) can be simplified using the v, $\overline{v}$, w, and $\overline{w}$ derivatives. Such simplification yields

$$\frac{\partial g}{\partial \overline{v}} = \frac{\partial h}{\partial \overline{w}} = \frac{\partial h}{\partial v} = \frac{\partial g}{\partial w} = 0 \tag{3.1.7}$$

$$\frac{\partial g}{\partial v} = \frac{\partial h}{\partial w} \tag{3.1.8}$$

$$\frac{\partial h}{\partial \overline{v}} = -\frac{\partial g}{\partial \overline{w}}. \tag{3.1.9}$$

By way of (3.1.7), the fact that g (resp., h) satisfies $\partial g/\partial \overline{v} = 0$ (resp., $\partial h/\partial \overline{w} = 0$) implies that g (resp., h) satisfies the Cauchy-Riemann equation and hence is an analytic function of v and $\overline{w}$ (resp., of w and $\overline{v}$). Therefore, by lemma 3.1, g and h are continuous $\mathbb{C}$-valued functions and hence have continuous partial derivatives of all orders. Therefore, differentiating (3.1.8) with respect to v yields that

$$\frac{\partial}{\partial v}\left(\frac{\partial g}{\partial v}\right) = \left(\frac{\partial^2 g}{\partial v^2}\right) = \frac{\partial}{\partial v}\left(\frac{\partial h}{\partial w}\right) = \frac{\partial}{\partial w}\left(\frac{\partial h}{\partial v}\right),$$

and by (3.1.7), these expressions are all equal to 0. It follows, then, that because U is assumed to be convex, g (resp., h) must be linear in $v, \overline{w}$ (resp., in $w, \overline{v}$). Thus, for complex constants $\alpha_i, \beta_i, \gamma_i, \delta_i$, $i = 1, 2$, g and h can be written

$$\begin{aligned} g(v, w) &= \alpha_1 + \beta_1 v + \gamma_1 \overline{w} + \delta_1 v \overline{w} \\ h(v, w) &= \alpha_2 + \beta_2 \overline{v} + \gamma_2 w + \delta_2 \overline{v} w. \end{aligned} \tag{3.1.10}$$

Equation (3.1.10) can be simplified utilizing the identities given in (3.1.8) and (3.1.9), thus yielding the relations

$$\beta_1 = \gamma_2, \beta_2 = -\gamma_1, \delta_1 = \delta_2 = 0. \tag{3.1.11}$$

Hence, because $f = g + jh$, (3.1.11) implies that

$$\begin{aligned} f &= g + jh \\ &= (\alpha_1 + \beta_1 v + \gamma_1 \overline{w}) + j(\alpha_2 - \gamma_1 \overline{v} + \beta_1 w) \\ &= \alpha_1 + j\alpha_2 + \beta_1 v + j\beta_1 w + \gamma_1 \overline{w} - j\gamma_1 \overline{v} \\ &= \alpha_1 + j\alpha_2 + (v + jw)(\beta_1 - j\gamma_1) \\ &= a + qb, \end{aligned}$$

where $a = \alpha_1 + j\alpha_2$ and $b = \beta_1 - j\gamma_1$. Hence for convex U, the claim is proven. Moreover, because an arbitrary connected open set U can be covered by a collection of convex sets which intersect pairwise, analysis of the intersections shows that $f = a + qb$ for the same $a, b \in \mathbb{H}$ throughout U, thus proving the result. □

By proving 3.1, it's now known that the only functions that are quaternion-differentiable on any given region $U \subset \mathbb{H}$ are linear, thereby proving that any hope of defining "quaternionic analyticity" by way of differentiation throughout regions is fruitless. As indicated in appendix 1.2, an alternative way to define analyticity of a complex-valued function f is by way of power series: That is, $f : \mathbb{C} \to \mathbb{C}$ is said to be *analytic* on a region $U \subset \mathbb{C}$ if for any $z_0 \in U$, $f(z) = \sum_{n=1}^{\infty} a_n (z - z_0)^n$ and the series converges. It stands to reason, then, that a similar approach may be feasible with regards to quaternion-valued functions $f : \mathbb{H} \to \mathbb{H}$.

To proceed, let $q = t + ix + jy + kz$ be an arbitrary quaternion. Note, then, that each of the components t, x, y, z of q can be rewritten *in terms of* q. In particular, [2] shows that sample equations exist of the form

$$\begin{aligned} t &= \frac{1}{4}\left(q - iqi - jqj - kqk\right) \\ x &= \frac{1}{4i}\left(q - iqi + jqj + kqk\right) \\ y &= \frac{1}{4j}\left(q + iqi - jqj + kqk\right) \\ z &= \frac{1}{4k}\left(q + iqi + jqj - kqk\right). \end{aligned} \tag{3.1.12}$$

The equations in (3.1.12) imply that every real polynomial $p(t, x, y, z) : \mathbb{R}^4 \to \mathbb{R}^4$ is, by definition, a quaternionic polynomial of $q \in \mathbb{H}$, thereby implying that no additional structure regarding so-called analyticity of quaternion-valued functions can be obtained by way of investigating quaternionic power series. This is distinctly different than the case of complex-valued functions of a complex variable: In particular, a polynomial $p(x, y) : \mathbb{C} \to \mathbb{C}$ of x, y requires both z and $\overline{z}$ be present due to the fact that $z, \overline{z}$ are linearly independent functions of x, y [13].

To best summarize these results, consider the following quote, taken from [2]:

> [N]either of the two fundamental definitions of [an analytic] function of a complex variable has interesting consequences when adapted to quaternions; one is too restrictive, the other not restrictive enough. The functions of a quaternion variable which have quaternionic derivatives...in the obvious sense...are just the constant and linear functions (and not all of them); the functions which can be represented by quaternionic power series are just those which can be represented by power series in four real variables.

The biggest, most obvious consequence of these results is that any knowledge regarding analytic-type properties of functions $f : \mathbb{H} \to \mathbb{H}$ will require a new level of machinery unnecessary in the study of real or complex analysis. That machinery is the basis of the next section.

3.2 Regularity

The first sucessful attempt at obtaining analytic-style behavior in functions $f : \mathbb{H} \to \mathbb{H}$ by means outside those discussed in section 3.1 came in 1935, when Swiss mathematician Rudolph Fueter[3] defined his so-called "regular" functions by way of generalized Cauchy-Riemann equations. Despite yielding a vast, fruitful theory with many results analogous to traditional complex analysis (see section 4 below), Fueter's definition of regularity left much to be desired and, as a result, many other alternative definitions of regularity have been proposed in the decades since. The most popular of these alternatives is due to Cullen (see [4], [14]), though others can be found in [15], [16], [17], [18], just to name a few. For the sake of brevity, this section is intended to be an investigation of the theories of Fueter and Cullen, though often through varying perspectives (namely, [2], [5], and [13] for Fueter and [14] for Cullen).

Before defining regularity, some notation must be given. Given an arbitrary quaternion $q = t + ix + jy + kz$, the *Left Cauchy-Riemann-Fueter operator* $\overline{\partial}_\ell$ is defined to be

$$\overline{\partial}_\ell = \frac{\partial}{\partial t} + i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}. \tag{3.2.1}$$

Similarly, the *Right Cauchy-Riemann-Fueter operator* $\overline{\partial}_r$ is defined to be

$$\overline{\partial}_r = \frac{\partial}{\partial t} + \frac{\partial}{\partial x}i + \frac{\partial}{\partial y}j + \frac{\partial}{\partial z}k. \tag{3.2.2}$$

Finally, define the *Cullen differential operator* to be

$$\partial_C = \frac{1}{2}\left(\frac{\partial}{\partial t} + \frac{\mathrm{Im}(q)}{r}\frac{\partial}{\partial r}\right), \tag{3.2.3}$$

where $\mathrm{Im}(q) = ix + jy + kz$ and $r^2 = x^2 + y^2 + z^2$. Together, these operators allow the following notions to be defined.

Definition 3.2. A function $f : \mathbb{H} \to \mathbb{H}$ is **Left Regular in the sense of Fueter** at $q \in \mathbb{H}$ if $\overline{\partial}_\ell\, f(q) = 0$ and is **Right Regular in the sense of Fueter** at q if $f(q)\,\overline{\partial}_r = 0$.

Definition 3.3. A function $f : \mathbb{H} \to \mathbb{H}$ is said to be **Regular in the sense of Cullen** at $q \in \mathbb{H}$ if $\partial_C f(q) = 0$.

[3]Fueter's bibliography can be found in [2], [5], etc.

Note that the definition of Cullen regularity is an immediate simplification over the definition(s) of Fueter regularity in that the standard (and oft inconvenient) left- and right-distinction vanishes. To simplify the analysis of Fueter regularity, the convention of considering only left regularity will be adopted.

Remark 3.1. As noted in [13], the operator $\overline{\partial}_\ell$ is one of a pair of operators, the other of which is given by

$$\partial_\ell = \frac{\partial}{\partial t} - i\frac{\partial}{\partial x} - j\frac{\partial}{\partial y} - k\frac{\partial}{\partial z},$$

which are intended to mimic the Cauchy-Riemann operators $\partial/\partial x \pm i\,(\partial/\partial y)$ defined in traditional complex analysis. Analogously, one can define ∂_r as well. [2] defines ∂_ℓ, $\overline{\partial}_\ell$, ∂_r, and $\overline{\partial}_r$ in terms of an evaluation map $\Gamma : \mathbb{H}^* \to \mathbb{H}$ (where $\mathbb{H}^*$ is the dual of $\mathbb{H}$) and points out that such a definition leads naturally to an additional operator Δ defined so that

$$\begin{aligned} \Delta f &= \frac{\partial^2 f}{\partial t^2} + \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \\ &= \partial_r \overline{\partial}_r f \\ &= \partial_\ell \overline{\partial}_\ell f. \end{aligned}$$

Note that in by defining these operators by way of such a map, [2] actually writes $\partial_\ell/2$ instead of ∂_ℓ, as well as in his definitions of $\overline{\partial}_\ell$, ∂_r, $\overline{\partial}_r$, and hence of Δ. This, of course, has no effect on the definition of Fueter regularity and as such can neglected accordingly. Also note that, as in the case of functions $f : \mathbb{C} \to \mathbb{C}$, a function $g : \mathbb{H} \to \mathbb{H}$ is said to be *harmonic* precisely when $\Delta g = 0$.

Despite having little usage in the literature on functions $f : \mathbb{H} \to \mathbb{H}$, the following definition (found in [14] and generalized in [19]) is given for completeness.

Definition 3.4. For an open set $U \subset \mathbb{H}$, a function f is said to be **Fueter holomorphic** if $\overline{\partial}_\ell \Delta f(q) = 0$ for all $q \in U$ and is said to be **Cullen holomorphic** if $\partial_C \Delta f(q) = 0$ for all $q \in U$.

As mentioned above, the notions of Fueter and Cullen regularity given in definitions 3.2 and 3.3 produce close quaternionic analogues to several key results in complex analysis related to *holomorphic* functions $f : \mathbb{C} \to \mathbb{C}$. It seems valid, then, to question whether these definitions are equivalent or whether either (or both) classes of functions coincide with the class of functions defined in 3.4. As mentioned in [14], the classes of Fueter and Cullen regular functions are both strictly contained in the class of holomorphic quaternionic functions and *do not* coincide with one another. This result is summarized in the following theorem.

Theorem 3.2. Let $R_F(\mathbb{H})$, $R_C(\mathbb{H})$, and $\mathrm{Ho}(\mathbb{H})$ denote the classes of quaternionic valued functions $f : \mathbb{H} \to \mathbb{H}$ of a quaternion variable which are Fueter regular, Cullen regular, and holomorphic, respectively. Then:

1. $R_F(\mathbb{H}) \subset \mathrm{Ho}(\mathbb{H})$
2. $R_C(\mathbb{H}) \subset \mathrm{Ho}(\mathbb{H})$
3. $R_F(\mathbb{H}) \neq R_C(\mathbb{H})$.

Proof. The proof of 3 follows from the fact that the identity function $\mathrm{id}_{\mathbb{H}} : q \mapsto q$ is Cullen regular and not Fueter regular. Indeed, for all $q = t + ix + jy + kz \in \mathbb{H}$,

$$\begin{aligned}
\overline{\partial}_\ell \left(\mathrm{id}_{\mathbb{H}}(q)\right) &= \left(\frac{\partial}{\partial t} + i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\right)\left(\mathrm{id}_{\mathbb{H}}(q)\right) \\
&= \left(\frac{\partial}{\partial t} + i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\right)(t + ix + jy + kz) \\
&= 1 + i(i) + j(j) + k(k) = -2,
\end{aligned}$$

while

$$\partial_C(\mathrm{id}_{\mathbb{H}}(q)) = \frac{1}{2}\left(\frac{\partial}{\partial t} + \frac{\mathrm{Im}(q)}{r}\frac{\partial}{\partial r}\right)\left(\mathrm{id}_{\mathbb{H}}(q)\right), \text{ where } r^2 = x^2 + y^2 + z^2. \tag{3.2.4}$$

The key to reducing equation (3.2.4) is to notice that the polynomial $p(q) = q$ can be written as $p(q) = f(t,r) + g(t,r)\,(\mathrm{Im}(q)/r)$, where $f(t,r) = t$, $g(t,r) = r$, and $\mathrm{Im}(q)/r$ is the radial unit vector lying on the pure-imaginary 2-sphere of radius r in $\mathbb{H}$ [13]. As such, the quantity $\mathrm{Im}(q)/r$ commutes with the operator ∂_C and satisfies $(\mathrm{Im}(q)/r)^2 = -1$, from which it follows that (3.2.4) can be reduced to

$$\begin{aligned}
\partial_C(\mathrm{id}_{\mathbb{H}}(q)) &= \frac{1}{2}\left(\frac{\partial}{\partial t} + \frac{\mathrm{Im}(q)}{r}\frac{\partial}{\partial r}\right)\left(f(t,r) + g(t,r)\frac{\mathrm{Im}(q)}{r}\right) \\
&= \frac{1}{2}\left(\frac{\partial}{\partial t}(f(t)) + \left(\frac{\mathrm{Im}(q)}{r}\right)^2\frac{\partial}{\partial r}(g(r))\right), \text{ where } f(t) = t \text{ and } g(r) = r \\
&= \frac{1}{2}\left(1 + (-1)(1)\right) = 0.
\end{aligned}$$

Hence, $f(q) = q$ is Cullen regular by definition 3.3 but *not* Fuetuer regular by definition 3.2, thereby proving statement 3 of the theorem.

Statement 1 is immediate, since

$$\begin{aligned}
\overline{\partial}_\ell\Delta(f) &= \overline{\partial}_\ell\partial_\ell\overline{\partial}_\ell f \text{ by the above remark} \\
&= \overline{\partial}_\ell\partial_\ell(0) \text{ by definition of } R_F(\mathbb{H})..
\end{aligned}$$

Statement 2 can be proved similarly ∎

As the example in theorem 3.2 proposition 3 indicates, Fueter's definition of regularity suffers from the drawback that it omits a large number of key functions of interest, namely polynomials. On the other hand, in the same way that $f(q) = q$ can be shown to satisfy Cullen's definition of regularity, so, too, do all polynomials, and hence all exponential, trigonometric, logarithmic, and rational functions by way of Taylor series [13]. Still, theorem 3.2 shows that both definitions of regularity satisfy a major necessary condition (namely, quaternionic holomorphicity) to allow one to construct a vast number of useful analogues of traditional complex analysis results, which is the focus of the next section.

4 Analogues of $\mathbb{C}$-valued Theory

Throughout this section, the term "regular" will refer to *left* regular and will be quantified when necessary by inclusion of either Fueter or Cullen beforehand to indicate whose theory is being referenced.

4.1 Analogues for Fueter Regular Functions

Many of the main analogues for Fueter regular functions is included in the literature of [2] (though [5] also centers on Fueter regularity) and will thus be borrowed therefrom. As a result, some additional machinery will be needed in order to construct the arguments from that particular perspective, and so establishing these tools is the first order of business.

Given a quaternion $q = t + ix + jy + kz$, two important differential forms related to q are $dq = dt + i\,dx + j\,dy + k\,dz$ and $v = dt \wedge dx \wedge dy \wedge dz$. As shown in [2], these forms satisfy

$$dq \wedge dq = i\,dy \wedge dz + j\,dz \wedge dx + k\,dx \wedge dy$$

and $v(1, i, j, k) = 1$, respectively. From these, the form Dq can be formed, where

$$Dq = dx \wedge dy \wedge dz - i\,dt \wedge dy \wedge dz - j\,dt \wedge dz \wedge dx - k\,dt \wedge dx \wedge dy.$$

Sudbury describes $Dq(a, b, c)$ as a quaternion which is perpendicular to each of a, b, and c and which has the magnitude equal to the volume of the 3-dimensional parallelepiped whose edges are a, b, and c, so in particular, Dq satisfies the identities $Dq(i, j, k) = 1$ and $Dq(1, e_i, e_j) = -\varepsilon_{ijk} e_k$.[4] Given these definitions, the following proposition gives an alternative formulation of Fueter regularity which serves as a prerequisite for the theorems which will follow.

Proposition 4.1. A function $f : \mathbb{H} \to \mathbb{H}$ is left regular at $q \in \mathbb{H}$ if and only if it is real-differentiable at q and if there exists a quaternion $f'_\ell(q)$ such that

$$d(dq \wedge dq\, f) = Dq\, f'_\ell(q). \tag{4.1.1}$$

[4] Here, the quantities e_i, $i = 0, 1, 2, 3$, denote the bases $1, i, j, k$, respectively, of $\mathbb{H}$. Also, ε_{ijk} is the Levi-Civita epsilon, which is defined to be $+1$ if (i, j, k) is an even permutation, -1 if (i, j, k) is odd, and 0 if any index i, j, k is repeated.

Because right regular functions remain untreated in this section, the quaternion $f'(q)$ will often be written in place of $f'_\ell(q)$ and will be called *the derivative of f at q* whenever there's no danger of confusion. Moreover, because (4.1.1) can be rewritten as

$$dq \wedge dq \wedge df = Dq\, f'(q), \tag{4.1.2}$$

the following results hold:

Lemma 4.1. The derivative $f'(q)$ in equations (4.1.1) and (4.1.2) has the form

$$\begin{aligned} f' &= -2\partial_\ell f \\ &= -\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z}. \end{aligned}$$

Proof. This result follows directly by evaluating (4.1.2) at (i, j, k) and $(1, i, j)$ and then solving the resulting system for f'. The details are omitted for brevity. □

Using lemma 4.1 and applying the Fueter regularity criterion given in definition 3.2 yields the following analogue.

Theorem 4.1 (Cauchy-Riemann-Fueter Equations)**.** A real-differentiable function f is regular at $q \in \mathbb{H}$ if and only if

$$\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} = 0. \tag{4.1.3}$$

Writing the quaternion q as a $q = v + jw$ (where $v = t + ix, w = y - iz \in \mathbb{C}$) and rewriting $f(q)$ as the sum $f(q) = g(v, w) + jh(v, w)$, the equation (4.1.3) can be rewritten as a more familiar pair of equations, namely

$$\frac{\partial g}{\partial \overline{v}} = \frac{\partial h}{\partial \overline{w}} \text{ and } \frac{\partial g}{\partial w} = -\frac{\partial h}{\partial v}.$$

Also, as an immediate consequence, the following analogous result holds.

Corollary 4.1. If $f : \mathbb{H} \to \mathbb{H}$ is regular and twice differentiable, then $\Delta f = 0$ and hence, f is harmonic.

Proof. Let f be regular and twice differentiable[5]. Recognizing that (4.1.3) is simply a long form of the differential equation $\partial_\ell f = 0$ and using the identity $\Delta f = 4\partial_\ell \overline{\partial}_\ell$ yields immediately that $\Delta f = 0$. □

Before stating the theorems outlining many other points of analogy, the following technical lemma is needed.

[5]This hypothesis is superfluous. Why? Unsurprisingly, Fueter regular functions are analogous to analytic complex-valued functions of a complex variable in the sense that they are infinitely differentiable.

Lemma 4.2. Let $f, g : \mathbb{H} \to \mathbb{H}$ be differentiable functions. Then:

$$d(g\, Dq\, f) = dg \wedge Dq\, f - gDq \wedge df. \tag{4.1.4}$$

Proof. This is merely a specific instance of the following result, found for example in [35] and discussed in Appendix 1.1: If α and β are k- and ℓ-forms, respectively, then

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta. \tag{4.1.5}$$

Note, in particular, that $\alpha = g$ and $\beta = Dqf$ are 0- and 3-forms, respectively, and that $g \wedge Dq\, df = g\, Dq \wedge df$. Therefore, substituting this information in (4.1.4) yields precisely (4.1.5). □

Remark 4.1. For future reference, it's worth noting that an alternative formulation of (4.1.4) has the form

$$d(g\, Dq\, f) = \left\{(\overline{\partial}_r g)f + g(\overline{\partial}_\ell f)\right\} v,$$

where $v = dt \wedge dx \wedge dy \wedge dz$ is the volume form in $\mathbb{H}$. This formulation will be used in the proof of the Cauchy-Fueter Integral Formula (for parallelepipeds).

Having proved this technical lemma, immediate dividends are now available. In particular, using equation (4.1.4) with $g = 1$ allows Theorem 4.1 to translated into the language of differential forms: In particular, it says that a function f which is differentiable is Fueter regular at $q \in \mathbb{H}$ if and only if $Dq \wedge df\big|_q = 0$. As mentioned in [2], however, Stokes' formula carries over verbatim into the language of quaternionic functions[6], thereby proving the following:

Proposition 4.2. Let $f : \mathbb{H} \to \mathbb{H}$ be regular and continuously differentiable in a domain $\Omega \subset \mathbb{H}$ which has differentiable boundary $\partial\Omega$. Then

$$\int_{\partial\Omega} Dq\, f = 0.$$

Proof. Let Ω be the region described above and let f be regular at every point $q \in \Omega$. By the remark above, it follows that $Dq \wedge df\big|_q = 0$ for every $q \in \Omega$, and hence that $\int_\Omega Dq \wedge df = 0$. By Stokes' theorem, it follows, then, that

$$\begin{aligned} 0 &= \int_\Omega Dq \wedge df \\ &= \int_{\partial\Omega} Dq\, f, \end{aligned}$$

thereby proving the result. □

[6]See 1.1 for a statement of Stokes' Theorem in the language of differential forms and chains.

Recall that Goursat's Theorem (see Appendix 1.2) for analytic functions $g : \mathbb{C} \to \mathbb{C}$ allows one to integrate along triangles Δ (and hence, around squares K) contained in an open set $U \subset \mathbb{C}$ in which g is holomorphic and guarantees that the integral will always equal zero. One standard proof which avoids reference to continuity of partial derivatives involves dissecting the region Δ (or K) into smaller regions whose diameter tends to zero. Unsurprisingly, this proof technique carries over verbatim in $\mathbb{H}$, where the square K is now a four-dimensional parallelepiped. This result is stated for reference, though the proof is withheld.

Lemma 4.3 (Goursat's Theorem)**.** If f is regular at every point in the 4-parallelepiped $K \subset \mathbb{H}$, then

$$\int_{\partial K} Dq\, f = 0.$$

Now, finally, the machinery has been built large enough so that the true fruit of the labor can be enjoyed.

Theorem 4.2 (The Cauchy-Fueter Integral Formula)**.** If f is regular at every point of the positively-oriented parallelepiped K and q_0 is a point in the interior K° of K, then

$$f(q_0) = \frac{1}{2\pi^2} \int_{\partial K} \frac{(q-q_0)^{-1}}{|q-q_0|^2}\, Dq\, f(q). \tag{4.1.6}$$

Proof. The proof given largely follows the outline given in [2]. For an equivalent statement translated from the language of differential forms into the framework of vector fields and their integrals, see the proof on page 12 of [5].

First, let K be a 4-parallelepiped in $\mathbb{H}$ and let q_0 be an element of its interior K°. Define the function g so that

$$\begin{aligned} g(q) &= -\partial_r \left(\frac{1}{|q-q_0|^2} \right) \quad \text{where } \partial_r \text{ is as in Remark 3.1} \\ &= \frac{(q-q_0)^{-1}}{|q-q_0|^2}. \end{aligned}$$

Clearly, then, g is real differentiable at every point $q \neq q_0$ in K. Moreover, a simple computation verifies that $\overline{\partial}_r g = 0$ except at $q = q_0$. Using the identity outlined in remark 4.1, it follows that $d(g\, Dq\, f) = 0$ for all $q_0 \neq q \in K^\circ$ by regularity of f.

Next, we use the dissection method from Goursat's theorem. In particular, consider replacing K by a smaller parallelepiped $\widehat{K}$ for which $q_0 \in \widehat{K}^\circ \subset K$. Because f has a removable singularity at $q = q_0$, it follows that f is continuous at q_0, whereby it follows that choosing $\widehat{K}$ small enough allows approximation of $f(q)$ by $f(q_0)$. Finally, let S be a 3-sphere with center q_0 and Euclidean volume element dS on which

$$Dq = \frac{(q-q_0)}{|q-q_0|} dS. \tag{4.1.7}$$

Such substitution is possible by way of a change of variables in the definition of Dq. This change of variables results in the following:

$$\begin{aligned}
\int_{\partial K} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq\, f(q) &= \int_{\partial K} g\, Dq\, f(q) \text{ by definition of } g \\
&= \int_S g\, \frac{(q-q_0)}{|q-q_0|} dS f(q_0) \text{ by (4.1.7) and since } f(q) \approx f(q_0) \\
&= f(q_0) \int_S \frac{dS}{|q-q_0|^3} \\
&= f(q_0)\left(2\pi^2\right).
\end{aligned}$$

Dividing both sides by $2\pi^2$ yields the exact statement of the result. □

[2] notes that the integral defined in the statement of the Cauchy-Fueter Integral Formula is real-analytic, which thus proves that every function which is regular in a region $\Omega \subset \mathbb{H}$ is automatically real analytic in Ω. As an immediate extension of the above work, the author then goes on to reprove the previous results (namely, the theorems of Stokes and Goursat, as well as the above integral formula) in the case of rectifiable chains in $\mathbb{H}$. The heart of that line of exploration is the statement of Cauchy's Theorem for rectifiable contours in $\mathbb{H}$ which, for all intents and purposes, *must* be shown here. On the other hand, note that the goal is to escape with as little redundancy as possible and so the reinvestigation of previous results in this new, more general context is largely omitted.

First, some terminology. If $C : [0,1]^3 \to \mathbb{H}$ is continuous, then C is said to be *rectifiable* if the variation of $Dq(C_1, C_2, C_3)$ is bounded where C_i indicates indexed partitions over $[0,1]_i$ and constant partitions over $[0,1]_j$ for $j \neq i$, $i,j = 1,2,3$. In this case, C is called a *rectifiable 3-cell.* If C is a rectifiable 3-cell and $f, g : C([0,1]^3) \to \mathbb{H}$ are quaternion-valued functions, then the form $f\, Dq\, g$ is said to be *integrable* over C if the maximum over all $\widehat{C}$ of the sums $f(\widehat{C})Dq(C_1, C_2, C_3)g(\widehat{C})$ has a limit in the sense of Riemann-Stieltjes, where $\widehat{C}$ denotes indexed subpartitions of the partitions of $[0,1]_{i,j,k}$, $i,j,k = 1,2,3$. If it exists, this limit is called the *integral* and is denoted

$$\int_C f\, Dq\, g.$$

The above notions can be generalized. More precisely, a *rectifiable chain* or *rectifiable contour* is a chain comprised of rectifiable 3-cells.[7] The integral of $f\, Dq\, g$ over a rectifiable contour is defined analogously for f, g continuous and for C a rectifiable chain. As mentioned above, this framework allows for the generalization of Stokes' theorem, Goursat's Theorem, and the Cauchy-Fueter Integral Formula (upon definition of a generalized idea of "winding number" which comes from the theory of singular homology of chain complexes), but the main purpose for introducing this framework is the following.

[7] For a brief discussion on cells and chains, see appendix 1.1.

Theorem 4.3 (Cauchy's Theorem for Rectifiable Contours in $\mathbb{H}$)**.** Suppose f is regular in a region $\Omega \subset \mathbb{H}$ and that C is a rectifiable 3-chain which is "homologous to 0"[8] in the singular homology of Ω. Then

$$\int_C Dq\, f = 0. \tag{4.1.8}$$

Sketch of Proof. In the case where Ω is contractible, use Poincaré's Lemma and Stokes' Theorem to get the result. Otherwise, consider C to be the boundary $\partial\widehat{C}$ of a 4-chain $\widehat{C}$ in Ω and dissect $\widehat{C}$ as the sum of a collection $\{\widehat{C}_n\}$ of rectifiable 4-cells lying inside open balls $U_n \subset \Omega$. Doing so yields that the integral of $Dq\, f$ along C is the sum over all n of the corresponding integrals of $Dq\, f$ along $\partial\widehat{C}_n$, each of which can be resolved using the contractible case. □

As mentioned in the text before theorem 4.3, a version of the Cauchy-Fueter Integral Formula for rectifiable contours follows directly from the theorem itself. As it turns out (and parallel to the development of theory in complex analysis), several key results follow as corollaries from the Integral Formula, establishment of which serves as the conclusion of the section on Fueter regular functions. For completeness, the newer, more general version of the Integral Formula is stated as a theorem. Suppose the generalized wrapping number n has been rigorously defined.

Theorem 4.4 (Cauchy-Fueter Integral Formula for Rectifiable Contours)**.** Let f be regular on an open set $\Omega \subset \mathbb{H}$, let $q_0 \in \Omega$, and let C be a rectifiable 3-chain which is homologous to a differentiable 3-chain whose image is ∂B for some ball $B \subset \Omega$. Then

$$\frac{1}{2\pi^2}\int_C \frac{(q-q_0)^{-1}}{|q-q_0|^2}\, Dq\, f(q) = n f(q_0),$$

where n is the wrapping number of C about q_0.

The following results are adapted from exposition in [20]. Note that the framework of this particular literature consists of functions $f : \mathbb{C} \to \mathbb{C}$ and their properties, so only the outlines of these results come therefrom.

Corollary 4.2 (Lioville's Theorem)**.** A quaternionic-valued function $f : \mathbb{H} \to \mathbb{H}$ which is Fueter regular for all values of q and is bounded is a constant function.

Proof. Let $q_1, q_2 \in \mathbb{H}$ and suppose that $f : \mathbb{H} \to \mathbb{H}$ is a Fueter regular function which satisfies $|f(q)| \leq M$ for all q. By theorem 4.4,

$$f(q_1) - f(q_2) = \frac{1}{2\pi^2 n}\int_C \frac{(q-q_1)^{-1}}{|q-q_1|^2}\, Dq\, f(q) - \frac{1}{2\pi^2 n}\int_C \frac{(q-q_2)^{-1}}{|q-q_2|^2}\, Dq\, f(q), \tag{4.1.9}$$

[8]Here, consider this a higher-dimensional analogue of "contractible to a point".

where C is a rectifiable contour including both q_1 and q_2 in its interior. Rewriting (4.1.9) yields

$$f(q_1) - f(q_2) = \frac{1}{2\pi^2 n} \int_C \left[\frac{(q-q_1)^{-1}}{|q-q_1|^2} - \frac{(q-q_2)^{-1}}{|q-q_2|^2} \right] Dq\, f(q).$$

If C is the 3-sphere centered at the origin with radius R such that $R > \max\{|q_1|, |q_2|\}$, then

$$\begin{aligned} |f(q_1) - f(q_2)| &= \left| \frac{1}{2\pi^2 n} \int_C \left[\frac{(q-q_1)^{-1}}{|q-q_1|^2} - \frac{(q-q_2)^{-1}}{|q-q_2|^2} \right] Dq\, f(q) \right| \\ &\leq \int_C \left| \frac{(q-q_1)^{-1}}{|R-q_1|^2} - \frac{(q-q_2)^{-1}}{|R-q_2|^2} Dq\, f(q) \right| \\ &\leq kv(C)M \left[\frac{(q-q_1)^{-1}}{|R-q_1|^2} - \frac{(q-q_2)^{-1}}{|R-q_2|^2} \right], \end{aligned} \tag{4.1.10}$$

where k is some constant, $v(C)$ is the volume of C in $\mathbb{R}^4$ and M is the bounds for f chosen initially. As $R \to \infty$, the quantity in (4.1.10) tends to 0 for q in the ball $C = B(R, \mathbf{0})$, thereby showing that $f(q_1) = f(q_2)$. Because this holds for arbitrary $q_1, q_2 \in \mathbb{H}$, $f(q)$ is constant. □

Corollary 4.3 (Maximum Modulus Theorem)**.** If f is Fueter regular and if $|f(q)| \leq M$ on some rectifiable 3-chain C, then $|f(q)| < M$ at all interior points $q \in C^\circ$. If, therefore, $f(q) = M$ for some $q \in C^\circ$, then $f(q)$ is constant and is identically M throughout C.

Sketch of Proof. Suppose that at an interior point $q_0 \in C^\circ$, $|f|$ has a value at least equal to its value everywhere else. Choosing a 3-sphere Γ lying entirely in C°, the claim can be proven by substituting polar coordinates forms of f, $(q - q_0)$ into the equation in theorem 4.4 and evaluating the integral about Γ of the real part of this integrand. Doing so proves that $f(q) = f(q_0)$ on Γ, from which the claim follows. □

This concludes the discussion on Fueter regular functions. It's worth noting that in [2], Sudbury (albeit briefly) investigates the theory of convergent Taylor and Laurent series expansions of Fueter regular functions, proving that statements similar to those in complex analysis hold for such functions. This theory is omitted here for brevity, though as Deavours points out in [5], Fueter himself proved parallels for quaternions in [21]. The reader is directed there for a more thorough treatment.

4.2 Analogues for Cullen Regular Functions

Cullen's definition of regularity yields an expansive, fruitful array of similarities between quaternion-valued functions of a quaternion variable and complex-valued functions of one complex variable in much the same way that Fueter's definition of regularity yielded an abundance of analogues to complex function theory. In this section, some of the analogues resulting from Cullen's definition are explored. In order to preserve brevity and avoid redundancy, results identical to those in section 4.1 will not be restated *except* in instances in which the author feels them enlightening. Of course, this is entirely dependent upon preference, and so the reader is directed to read the literature in [4] and [14], as well as other cited resources mentioned throughout. This particular exposition will be based largely on [14], which is somewhat more general than the original work done by Cullen in [4].

As pointed out on page 280 in [14], functions which satisfy Cullen's definition of (left) regularity are "closely related to a class of functions of the reduced quaternionic variable" $t + ix + jy$, and so the generalization given below will stem naturally from that perspective. Before advancing, some notation must be introduced. First, let S denote the unit sphere of purely imaginary Hamilton quaternions, namely

$$S = \{q = ix + jy + kz \,:\, x^2 + y^2 + z^2 = 1\}.$$

Note that any element $I \in S$ naturally satisfies $I^2 = -1$ and as a result, such I will be called *imaginary units*. The importance of these imaginary units lies partially in the fact that, for any nonreal element $q \in \mathbb{H}$, there exist unique elements $x, y \in \mathbb{R}$ for which $y > 0$ and $q = x + yI$ for some $I \in S$.

Consider the following definitions related to regular functions and the imaginary unit sphere.

Definition 4.1. Let Ω be a domain in $\mathbb{H}$ and let $f : \Omega \to \mathbb{H}$ be real differentiable. Then f is said to be **(left) regular** if, for every $I \in S$, the restriction f_I of f to the line $L_I = \mathbb{R} + \mathbb{R}I$ passing through the origin and containing $1, I$ is holomorphic on $\Omega \cap L_I$.

Definition 4.2. Let $\Omega \subset \mathbb{H}$ be a region and let $f : \Omega \to \mathbb{H}$ be a real differentiable function. For any element $I \in S$ and any quaternion $q = x + yI \in \Omega$, $x, y \in \mathbb{R}$, the **I-derivative of f at q**, denoted $\partial_I f(q)$, is defined to be

$$\partial_I f(x + yI) = \frac{1}{2}\left(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y}\right) f_I(x + yI).$$

Definition 4.3. Let Ω be a domain in $\mathbb{H}$ and let $f : \Omega \to \mathbb{H}$ be a regular function. Then **the Cullen derivative** of f, denoted $\partial_C f$, is defined as follows:

$$\partial_C f(q) = \begin{cases} \partial_I f(q) & \text{if } q = x + yI, y \neq 0 \\ \dfrac{\partial f}{\partial x} & \text{if } q = x \in \mathbb{R} \end{cases}$$

Remark 4.2. Note that combining definitions 4.1 and 4.2 imply that a function $f : \Omega \to \mathbb{H}$ is regular precisely when $\overline{\partial}_I f(q) = 0$ on $\Omega \cap L_I$. Also note that, while notationally different, definition 4.3 is equivalent to the equation of the Cullen differential operator given in equation (3.2.4); this newer notation is given to simplify the transition to the results that follow.

From remark 4.2, it's easily shown that monomials of the form $q^n a$ are regular for $a \in \mathbb{H}$. Moreover, because definition 4.2 immediately confirms that the collection of regular functions is closed under addition, polynomials of the form $f(q) = q^n a_n + q^{n-1}a_{n-1} + \cdots + qa_1 + q_0$, $a_i \in \mathbb{H}$ for all i, are regular. Finally, if the space of regular functions is considered with the "natural uniform convergence" [14] via convergence on compact subsets, it holds that quaternionic power series $\sum_{n=0}^{\infty} q^n a_n$ can be defined and considered for the first time.

As per usual, given a quaternionic power series $\sum_{n=0}^{\infty} q^n a_n$, there exists a number $R \in [0, \infty]$ called *the radius of convergence* for which the series converges absolutely for every q satisfying $|q| < R$, converges uniformly for every q satisfying $|q| \leq \rho < R$, and diverges for every q such that $|q| > R$. Because many of the same techniques applied to complex power series carry over verbatim in this new context, many of the classical theorems regarding complex power series have exact analogues in the sense of quaternionic power series. Consider, for example, Abel's theorem:

Proposition 4.3 (Abel's Limit Theorem)**.** Let $a = \sum_{n=0}^{\infty} a_n$ be a sequence of quaternions and let $G_a(q) = \sum_{n=0}^{\infty} q^n a_n$ be the power series with coefficients a. If a converges, then $G(q)$ converges to $G(1)$ as q tends to 1 in such a way that $|1-q|/(1-|q|)$ remains bounded.

Proof. The proof given here follows the general framework laid by Ahlfors in [22].

Let $\varepsilon > 0$. Without loss of generality, suppose that $a = 0$; this can be ensured by altering the constant a_0. Consider the partial sum $\sigma_n = \sum_{k=0}^{n} a_k$ of a employ summation by parts:

$$\begin{aligned}\sigma_n(q) &= a_0 + qa_1 + \cdots q^n a_n \\ &= \sigma_0 + (\sigma_1 - \sigma_0)q + \cdots + (\sigma_n - \sigma_{n-1})q^n \\ &= (1-q)\sigma_0 + (q - q^2)\sigma_1 + \cdots + (q^{n-1} - q^n)\sigma_{n-1} + q^n\sigma_n \\ &= (1-q)(\sigma_0 + q\sigma_1 + \cdots + q^{n-1}\sigma_{n-1}) + q^n\sigma_n. \qquad (4.2.1)\end{aligned}$$

Because a converges, $q^n\sigma_n$ tends to 0 as n tends to ∞. Hence, substituting the expression in (4.2.1) into the definition of G_a yields that

$$G_a(q) = (1-q)\sum_{n=0}^{\infty} q^n \sigma_n. \qquad (4.2.2)$$

Suppose that the ratio $|1-q|/(1-|q|)$ is bounded by some real number M. In particular, then, $|1-q| \leq M(1-|q|)$. Because σ_n converges as $n \to \infty$, it follows that $\sigma_n \to 0$ as $n \to \infty$. This implies that, for N large, $|\sigma_n| < \varepsilon$ for all $n \geq N$. Note, then, that $\sum_{n=N}^{\infty} q^n\sigma_n < \varepsilon \sum_{n=N}^{\infty} |q|^n = \varepsilon|q|^N/(1-|q|) < \varepsilon/(1-|q|)$, because N large implies $|q| < 1$.

Using this information in addition to (4.2.2) yields that

$$
\begin{aligned}
|G_a(q)| &= \left|(1-q)\sum_{n=0}^{\infty} q^n \sigma_n\right| \\
&= |1-q|\left|\sum_{n=0}^{N-1} q^n \sigma_n\right| + |1-q|\left|\sum_{n=N}^{\infty} q^n \sigma_n\right| \\
&\le |1-q|\left|\sum_{n=0}^{N-1} q^n \sigma_n\right| + M\varepsilon. \qquad (4.2.3)
\end{aligned}
$$

As stated in the problem, q is assumed to approach 1, from which it follows that the first term on the right hand side of (4.2.3) can be made arbitrarily close to 0. Hence, $G_a(q)$ tends to 0 as n tends to ∞, which by (4.2.2) is precisely the value of $G_a(1)$. Thus, the claim holds. □

Note that many of the familiar properties of complex-valued power series carry over to quaternionic-valued power series because of the structure given to the latter. For example, given the definitions of I- and Cullen- derivatives ∂_I and ∂_C in 3.4 and 4.2, respectively, it's easily verified that any regular function f satisfies $\partial_C(\overline{\partial}_I f) = 0 = \overline{\partial}_I(\partial_C f)$, thereby proving that the Cullen-derivative of f is again regular. Moreover, because convergence of quaternion-valued power series was assumed to come from convergence on all compact subsets of its domain, uniform convergence is immediate, which thus implies (see [23], for example) that such a power series can be (Cullen-)differentiated term-by-term. Hence,

$$\partial_C\left(\sum_{n=0}^{\infty} q^n a_n\right) = \sum_{n=1}^{\infty} q^{n-1} n a_n,$$

where the "new" derivative power series has the same radius of convergence R as the original. In particular, then, power series are infinitely continuously differentiable.

The bulk of the remaining theory for Cullen regular functions centers on uniformly convergent power series and requires the statement of the following two technical lemmas. Though necessary and nontrivial, the proofs of these lemmas fail to be particularly insightful. For this reason, they are omitted, though the interested reader is directed to [14], where the proofs are given in their entireties. For the remainder of this section, the notation B refers to the ball centered at $\mathbf{0} \in \mathbb{H}$ with radius $R > 0$, and the term *orthogonal* is used to denote two imaginary units $I, J \in S$ which satisfy $IJ = 0$, where multiplication here is the usual quaternion multiplication.

Lemma 4.4 (The Splitting Lemma)**.** Let $f : B \to \mathbb{H}$ be regular. Then for every $I \in S$ and every J orthogonal to I, there exist holomorphic functions $F, G : B \cap L_I \to L_I$ for which

$$f_I(q) = F(q) + G(q)J \text{ for all } q = x + yI \in B \cap L_I.$$

Lemma 4.5. Let $f : B \to \mathbb{H}$ be regular. Then for all $n \in \mathbb{N}$, the nth order Cullen derivative $\partial_C^n f : B \to \mathbb{H}$ is again regular and can be rewritten equivalently as

$$\partial_C^n f(x+yI) = \frac{\partial^n f}{\partial x^n}(x+yI)\,.$$

With the statement of these two lemmas, the main, fundamental results of this section can be investigated. The first of these results is the following classification of Cullen regular functions on B.

Theorem 4.5. A function $f : B \to \mathbb{H}$ is regular *if and only if* it has a series expansion of the form

$$f(q) = \sum_{n=0}^{\infty} q^n \frac{1}{n!}\frac{\partial^n f}{\partial x^n}(0) \tag{4.2.4}$$

which converges on B. Hence, regularity of f implies that f is infinitely continuously differentiable on B.

Proof of Theorem. As remarked above, quaternionic power series are infinitely continuously differentiable on their domains, and so the last statement of the theorem is an immediate consequence of the first statement.

Let $I \in S$, let $L_I = \mathbb{R}+\mathbb{R}I$, and let $\Delta_I \subset L_I$ denote the disc centered at $\mathbf{0} \in \mathbb{H}$ with radius a, $0 < a < R$. Using the splitting lemma, there exist functions F and G which are holomorphic in $B \cap L_I$ for which $f(q) = F(q)+G(q)J$ for every $J \in S$ orthogonal to I. Viewing L_I as the complex plane and using the traditional Cauchy Integral Formula from complex analysis, for any $q \in \Delta_I$ and any $\zeta \neq q$ in $B \cap L_I$,

$$f_I(q) = \frac{1}{2\pi I}\int_{\partial\Delta_I} \frac{F(\zeta)}{\zeta - q}\,d\zeta + \left(\frac{1}{2\pi I}\int_{\partial\Delta_I} \frac{G(\zeta)}{\zeta - q}\right) J. \tag{4.2.5}$$

By way of machinery in classical complex analysis, each of the integrands in (4.2.5) have convergent power series[9] in Δ_I, and so (4.2.5) can be transformed into

$$\begin{aligned} f_I(z) &= \sum_{n=0}^{\infty} q^n \left(\int_{\partial\Delta_I} \frac{F(\zeta)}{\zeta^{n+1}}\,d\zeta\right) + \left[\sum_{n=0}^{\infty} q^n \left(\int_{\partial\Delta_I} \frac{G(\zeta)}{\zeta^{n+1}}\,d\zeta\right)\right] J \\ &= \sum_{n=0}^{\infty} q^n \frac{1}{n!}\frac{\partial^n F}{\partial q^n}(0) + \sum_{n=0}^{\infty} q^n \frac{1}{n!}\frac{\partial^n G}{\partial q^n}(0)J, \text{ where } \partial = \partial_C \\ &= \sum_{n=0}^{\infty} q^n \frac{1}{n!}\frac{\partial^n (F+GJ)}{\partial q^n}(0) \text{ by linearity.} \end{aligned}$$

[9] For example, $\frac{F(\zeta)}{\zeta-q} = \frac{1}{1-q/\zeta}\frac{F(\zeta)}{\zeta} = \sum_{n=0}^{\infty}\left(\frac{q}{\zeta}\right)^n \frac{F(\zeta)}{\zeta} = \sum_{n=0}^{\infty} q^n \frac{F(\zeta)}{\zeta^{n+1}}$.

Simplifying thus yields

$$f_I(z) = \sum_{n=0}^{\infty} q^n \frac{1}{n!}\frac{\partial^n f}{\partial q^n}(0), \text{ as } f = F + GJ$$
$$= \sum_{n=0}^{\infty} q^n \frac{1}{n!}\frac{\partial^n f}{\partial x^n}(0) \text{ by lemma 4.5.}$$

Hence, $f_I(z) = \sum_{n=0}^{\infty} q^n a_n$, where the coefficients $a_n = (n!)^{-1}\frac{\partial^n f}{\partial x^n}$ is independent of I, whereby it follows that f_I has such a representation for every $I \in S$. □

In much the same way that representing complex-valued holomorphic functions as convergent power series is the key to proving several well-established, fundamental results about that class of functions, it follows that the previous theorem holds the key to such things in the case of regular functions $f : B \to \mathbb{H}$.

Theorem 4.6 (The Identity Theorem)**.** Let $f : B \to \mathbb{H}$ be regular and let $Z_f = \{q \in B : f(q) = 0\}$ be the zero set of f. If, for some $I \in S$, $L_I \cap Z_f$ has an accumulation point, then $f \equiv 0$ on B.

Proof. Using the splitting lemma, reduce f on $L_I \cap B$ to the form

$$f(x + yI) = F(x + yI) + G(x + yI)J,$$

and note that $q \in Z_f$ precisely when q is in the zero sets of both F and G. Moreover, an accumulation point of $L_I \cap Z_f$ is an accumulation point of both $L_I \cap Z_F$ and $L_I \cap Z_G$, where Z_F, Z_G denote the zero sets of F, G, respectively. Hence, using the identity theorem from complex analysis, both F and G are identically zero on $L_I \cap B$, and hence, $(\partial^n f/\partial x^n)(0) = 0$ for all $n \in \mathbb{N}$. These partial derivatives (up to multiplication by units) are precisely the values of elements a_n shown in theorem 4.5 to be the coefficients of the power series representation of f on $L_I \cap B$, however, and so it follows that f is identically zero there. □

An immediate corollary of Theorem 4.6 is the analogue of a well-known fact from complex analysis: If f and g are two regular functions on B and if $f = g$ on a subset Ω of $L_I \cap B$ which contains an accumulation point, then $f \equiv g$ on $L_I \cap B$. This follows directly, as the function $f - g$ would then be zero on a set with accumulation point. Before proceeding to other main results, consider the following lemma.

Lemma 4.6 (Mean Value Theorem)**.** Every regular function $f : B \to \mathbb{H}$ induces a map $f_I : L_I \cap B \to \mathbb{H}$ which has the mean value property for each $I \in S$.

Proof. Write $f_I(x+yI) = F(x+yI)+G(x+yI)J$ as in the splitting lemma. For all elements $a \in L_I \cap B$, let Δ denote the (open) disc centered at a of radius $r > 0$. Then for all numbers r satisfying $\overline{\Delta} \subset L_I \cap B$, it's easily verified by way of the standard Cauchy Integral Formula that $(2\pi)^{-1} \int_0^{2\pi} f_I(a + re^{I\theta})\, d\theta = f_I(a)$. □

As is the case in complex analysis, the mean value theorem (MVT) is an invaluable tool which can be used to prove many important results. In particular, the MVT plays a fundamental role in the proofs of the maximum modulus theorem, which is given here along with a sketch of its proof.

Theorem 4.7 (Maximum Modulus Theorem)**.** Let $f : B \to \mathbb{H}$ be regular. If $|f|$ has a relative maximum at a point $a \in B$, then f is constant on B.

Sketch of Proof. Suppose $f(a) \neq 0$ is a regular function and assume, without loss of generality, that $f(a) > 0$ and that $a = x_0 + y_0 I$ is an element of S[10]. For r sufficiently small, define $M(r)$ to be $M(r) = \sup_{\theta \in \mathbb{R}} \{|f(a + re^{I\theta})\}$. If it's assumed that $|f|$ has a relative maximum at $a \in B$, then $f(a) \geq M(r)$, and by the MVT, $f_I(a) = M(r)$. Considering $q = x + yI$, define $g(q) = \operatorname{Re}(f_I(a) - f_I(q))$ and note that for sufficiently small values of $r = |q - a|$, $g(q) \geq 0$. Apply the MVT to g to express $g(a)$ as $g(a) = (2\pi)^{-1} \int_0^{2\pi} g(a + r\exp(I\theta))\, d\theta$ and note that $g(a) = 0$ by definition. It follows then that g is identically zero in a closed disc, whereby it follows that $f_I(z) = f_I(a)$ in said disc. The conclusion follows immediately from the identity theorem. ☐

Another fundamental result that follows directly from the Splitting Lemma is Cauchy's Integral Formula which itself is the building block for a large number of analogous results. Before verifying this result, let $q \in B$ and define for convenience the purely imaginary unit I_q to be

$$I_q = \begin{cases} \dfrac{\operatorname{Im}(q)}{|\operatorname{Im}(q)|} & \text{if } \operatorname{Im}(q) \neq 0 \\ \text{arbitrary } u \in S & \text{otherwise} \end{cases} .$$

Under this notation, the expression $(\zeta - q)^{-1}\, d\zeta$ is well-defined for all $\zeta \in L_{I_q}$, $\zeta \neq q$, a fact that leads to the following.

Theorem 4.8 (Cauchy's Integral Formula)**.** Let $f : B \to \mathbb{H}$ be a regular function. For any $q \in B$, define for $r > 0$ $\overline{\Delta_q(0,r)} = \{x + yI_q \,|\, x^2 + y^2 \leq r^2\}$. Then for $\zeta \in L_{I_q} \cap B$ and r such that $q \in \overline{\Delta_q(0,r)} \subset B$,

$$f(q) = \frac{1}{2\pi I_q} \int_{\partial\Delta_q(0,r)} \frac{d\zeta}{\zeta - q} f(\zeta).$$

Remark 4.3. Note that the constant multiple in front of the integral in Theorem 4.8 is different than the one presented in section 4.1, Theorem 4.4. Both of these are also different than the one deduced by Deavours in [5]. The domain and type of integration used plays a fundamental role in this fact and, despite the apparent differences, the integrands of these results absorb and make up for any such topical disparities.

[10] The former can be achieved by multiplying f by a nonzero quaternion, the latter by considering I to be the normalized imaginary part of $x_0 + y_0 I$ and considering f_I.

Proof of Theorem. The result follows directly from the Splitting Lemma decomposition of f_{I_q} into F and G, along with the fact that both F and G can be expressed via the standard complex analysis version of Cauchy's Integral Formula. More precisely,

$$\begin{aligned}
f(q) &= F(q) + G(q)J \\
&= \frac{1}{2\pi I_q}\int_{\partial\Delta_q(0,r)} \frac{F(\zeta)}{\zeta - q}\,d\zeta + \left(\frac{1}{2\pi I_q}\int_{\partial\Delta_q(0,r)} \frac{G(\zeta)}{\zeta - q}\,d\zeta\right) J, \text{ by Cauchy's Integral Formula} \\
&= \frac{1}{2\pi I_q}\int_{\partial\Delta_q(0,r)} \frac{d\zeta}{\zeta - q}\,(F(\zeta) + G(\zeta)J), \text{ by linearity} \\
&= \frac{1}{2\pi I_q}\int_{\partial\Delta_q(0,r)} \frac{d\zeta}{\zeta - q}\, f_{I_q}(\zeta), \text{ by definition of } f \text{ for } \zeta \in L_{I_q} \\
&= \frac{1}{2\pi I_q}\int_{\partial\Delta_q(0,r)} \frac{d\zeta)}{\zeta - q}\, f(\zeta).
\end{aligned}$$

Hence, the claim is proved. □

As is often the case, the Cauchy Estimates are an immediate corollary of the Cauchy Integral Formula.

Corollary 4.4 (Cauchy Estimates)**.** Let $f : B(0,R) \to \mathbb{H}$ be a regular function. Let $r < R$, $I \in S$, and $\partial\Delta_I(0,r) = \{x + yI \,:\, x^2 + y^2 = r^2\}$. If $M_I = \max\{|f(q)|\,|q \in \partial\Delta_I(0,r)\}$ and if $M = \inf_{I\in S}\{M_I\}$, then

$$\frac{1}{n!}\left|\frac{\partial^n f}{\partial x^n}(0)\right| \leq M r^{-n},\; n \geq 0.$$

Proof. Because f is regular, f can be written as a convergent power series of the form in (4.2.4), and because each of the partial derivatives of f is again regular, the Cauchy Integral Formula[11] indicates that the coefficients a_n in the power series representation of f can be expressed as

$$\begin{aligned}
a_n &= \frac{1}{n!}\frac{\partial^n f}{\partial x^n}(0) \\
&= \frac{1}{2\pi I}\int_{\partial\Delta_I(0,r)} \frac{d\zeta}{\zeta^{n+1}}\, f(\zeta).
\end{aligned}$$

In particular, then, for $I \in S$, it follows that

$$\begin{aligned}
\left|\frac{1}{n!}\frac{\partial^n f}{\partial x^n}(0)\right| &= \left|\frac{1}{2\pi I}\int_{\partial\Delta_I(0,r)} \frac{d\zeta}{\zeta^{n+1}}\, f(\zeta)\right| \\
&\leq \frac{1}{2\pi}\int_{\partial\Delta_I(0,r)} \frac{|f(\zeta)|}{|\zeta|^{n+1}}\, d\zeta. \qquad (4.2.6)
\end{aligned}$$

[11]To get the exact expression given, repeated instances of differentiation within the form given in Theorem 4.8 is required. The details are omitted for brevity.

Of course, for r small and for ζ in the region $\partial\Delta_I(0,r)$, $|\zeta|^{-(n+1)} \leq r^{-(n+1)}$ and $|f(\zeta)| \leq M_I$. Hence, the expression in (4.2.6) indicates that

$$\frac{1}{n!}\left|\frac{\partial^n f}{\partial x^n}(0)\right| \leq \frac{1}{2\pi}\int_{\partial\Delta_I(0,r)} \frac{|f(\zeta)|}{r^{n+1}}\,d\zeta \leq \frac{1}{2\pi}\int_{\partial\Delta_I(0,r)} \frac{M_I}{r^{n+1}}\,d\zeta = \left(\frac{1}{2\pi}\frac{M_I}{r^{n+1}}\right)(2\pi r).$$

Hence, reducing and taking the infimum over all $I \in S$ yields the expression in the statement. □

Cauchy's Estimates immediately yield the following corollary.

Corollary 4.5 (Liouville's Theorem)**.** If $f : \mathbb{H} \to \mathbb{H}$ is a function which is regular on all of $\mathbb{H}$ and if there exists an $M \geq 0$ for which $|f(q)| \leq M$ for all $q \in \mathbb{H}$, then f is constant.

Proof. Applying Cauchy's Estimates, it follows that for all $r \in \mathbb{R}$,

$$\frac{1}{n!}\left|\frac{\partial^n f}{\partial x^n}(0)\right| \leq \frac{M}{r^n}.$$

Taking $r \to \infty$, it follows that $(\partial^n f/\partial x^n)(0) = 0$ for all n and hence that $f \equiv f(0)$. □

As the authors of [14] go on to show, Cullen regular functions satisfy Morera's theorem, upon which they shift gears and discuss a variety of geometric properties of Cullen regular functions and of their related power series in $\mathbb{H}$. It's only appropriate, then, that gears here are shifted as well, though the focus will remain on the topic at hand rather than the plethora of interesting, useful consequences of Cullen regularity. In particular, the goal for the remainder of this section is to establish some of the results in [24], most notably, the Open Mapping Theorem for quaternion valued functions of a quaternion variable. Though stated in generally the same framework as given above, some additional machinery is needed before the theorem itself can be derived.

As above, let S denote the sphere of purely imaginary quaternion units. Moreover, for an arbitrary function $f : \Omega \to \mathbb{H}$ defined on a region $\Omega \subseteq \mathbb{H}$, the notation Z_f will be used in the same way it was in Theorem 4.6, that is, to denote the zero set $Z_f = \{q \in \Omega : f(q) = 0\}$ of f. Besides the properties needed for the investigation done here, many independently interesting properties of the sets Z_f are investigated in [24] and its included references. Before stating and proving the Open Mapping Theorem, it's necessary to derive the so-called "Minimum Modulus Theorem," an undertaking which itself requires considerably more effort than in the standard complex analysis sense. First, consider the following definitions.

Definition 4.4. Let $B = B(\mathbf{0}, R)$ for $R > 0$ and let f, g be regular functions on B with convergent power series representations of the form

$$f(q) = \sum_{n=0}^{\infty} q^n a_n \text{ and } g(q) = \sum_{n=0}^{\infty} q^n b_n.$$

For each $q \in B$, define the **regular product of f and g** to be the regular function $f * g : B \to \mathbb{H}$ defined by

$$f * g(q) = \sum_{n=0}^{\infty} q^n \left(\sum_{k=0}^{n} a_k b_{k-n} \right).$$

Definition 4.5. Given B, f as above, **regular conjugate** $f^c : B \to \mathbb{H}$ is defined to be the regular function whose convergent power series has the form

$$f^c(q) = \sum_{n=0}^{\infty} q^n \overline{a_n}.$$

Definition 4.6. Under the same hypotheses as above, the **symmetrization of f** f^s is defined to be the function

$$\begin{aligned} f^s(q) &= f * f^c \\ &= f^c * f \\ &= \sum_{n=0}^{\infty} q^n \sum_{k=0}^{n} a_k \overline{a_{n-k}}, \end{aligned}$$

where the coefficients $\sum_{k=0}^{n} a_k \overline{a_{n-k}}$ are real for each k.

It certainly isn't immediately obvious why definitions 4.4, 4.5, and 4.6 are pertinent to the establishment of the Open Mapping Theorem. As it turns out, the proof of that particular theorem requires one to utilize the notion of a *reciprocal function* which, unlike the case of functions $f : \mathbb{C} \to \mathbb{C}$, require considerable development in order to ensure well-definedness and the presence of desirable, analogous properties. To that end:

Definition 4.7. Let $B = B(\mathbf{0}, R)$, let $f : B \to \mathbb{H}$ be a regular function, and let f^c, f^s be the regular conjugate and regular symmetrization, respectively, of f. Then the **regular reciprocal** $f^{-*} : B \setminus \mathbb{Z}_{f^s} \to \mathbb{H}$ is defined to be

$$\begin{aligned} f^{-*q} &= f^{-s}(q) f^c(q) \\ &= \frac{1}{f^s(q)} f^c(q). \end{aligned}$$

For completeness, consider the following proposition concerning the properties of the regular reciprocal of a function f.

Proposition 4.4. Let $B = B(\mathbf{0}, R)$, let $f, g : B \to \mathbb{H}$ be regular functions, and let f^c, f^s be the regular conjugate and regular symmetrization, respectively, of f. Then if f has no zeros on B,:

1. f^{-*} is regular on its domain.
2. $f^{-*} * f = 1 = f f^{-*}$.

3. $(f*g)(q) = f(q)\, g\left(f(q)^{-1} q f(q)\right)$ for all $q \in B \setminus Z_f$. In particular, $f^{-*}(q) = f\left(f^c(q)^{-1} q f^c(q)\right)^{-1}$ for all $q \in B \setminus Z_f$.

Proof. The statement of the first proposition is a specific instance of a more general fact, namely that the function $h : B \setminus Z_g \to \mathbb{H}$ defined as $h(q) = (1/g(q))f(q)$ is regular on its domain. This can be proven directly from the condition for Cullen regularity given in definition 4.1.

For the second claim, note that f having no zeros on B directly corresponds to f^s having no zeros on B, whereby it follows that both f^{-s} and f^{-*} are well-defined on B. Moreover, because $f^{-s}(q) * g(q) = g(q) * f^{-s}(q)$ for all regular functions $g : B \to \mathbb{H}$, it follows that

$$\begin{aligned} &f^{-*} * f = f^{-s} * f^c * f = f^{-s} * f^s = 1, \text{ and} \\ &f * f^{-*} = f * f^{-s} * f^c = f^{-s} * f * f^c = f^{-s} * f^s = 1. \end{aligned}$$

For (3), note that the second assertion follows by way of trivial calculation from the first. For the first assertion, let $f(q) = \sum_{n=0}^{\infty} q^n a_n$ and $g(q) = \sum_{n=0}^{\infty} q^n b_n$, from which it follows that

$$\begin{aligned} (f * g)(q) &= \sum_{n=0}^{\infty} q^n \sum_{k=0}^{n} a_k b_{n-k} \\ &= \sum_{k,\ell=0}^{\infty} q^{\ell+k} a_k b_\ell \text{ by letting } \ell = n-k \\ &= \sum_{\ell=0}^{\infty} q^\ell f(q) b_\ell, \text{ as } q^{\ell+k} a_k = q^\ell q^k a_k \\ &= \sum_{\ell=0}^{\infty} f(q) f(q)^{-1} q^\ell f(q) b_\ell, \text{ by multiplying by } 1 = f(q)f(q)^{-1} \\ &= f(q) \sum_{\ell=0}^{\infty} \left[f(q)^{-1} q f(q)\right]^\ell b_\ell, \text{ since } f(q)^{-1} q^\ell f(q) = [f(q)^{-1} q f(q)]^\ell \\ &= f(q)\, g(f(q)^{-1} q f(q)) \text{ by definition of } g. \end{aligned}$$

Hence, the result is shown. □

Upon deriving a well-defined notion of quaternionic reciprocity, the tools are in place to begin working towards proving the open mapping theorem. Doing so requires proof of the minimum modulus theorem, which itself is dependent upon both the identity and maximum modulus theorems. Note, however, that the statements given in theorems 4.6 and 4.7, respectively, can be extended to include a larger class of functions on domains which subsume the domains stated therein. Rather than stating and proving these reformulations rigorously, consider the following.

Proposition 4.5. The identity and maximum modulus theorems can be extended as follows.

1. The conclusion of the identity theorem holds for functions f and g which coincide on a domain $\Omega \cap \mathbb{R}$ where $\Omega \subset \mathbb{H}$ is a domain which intersects the real axis and for which $\Omega \cap L_I$ is connected for any complex line $L_I = \mathbb{R} + I\mathbb{R}$

2. The conclusion of the maximum modulus theorem holds for regular functions $f : \Omega \to \mathbb{H}$ where Ω is the ball $B(\mathbf{0}, R)$ minus a closed set E consisting of isolated points and isolated 2-spheres of the form $x + yS$.

Proof. Both of these theorems are proved succinctly in [24] and in more detail in sources cited therein. □

A common notation adopted for the expression $f^c(q)^{-1}qf^c(q)$ in statement 3 of proposition 4.4 is $T_f(q)$. Using this language, that proposition says that $f^{-*}(q) = f(T_f(q))^{-1}$, and hence the modulus of the reciprocal f^{-*} of f is equal to the reciprocal of the modulus of the expression $f(T_f(q))^{-1}$. This observation plays an important role in the first of the major lemmas used to derive the Open Mapping Theorem.

Lemma 4.7 (Weak Minimum Modulus Theorem)**.** Let $B = B(\mathbf{0}, R)$ and let $f : B \to \mathbb{H}$ be a regular function. If $|f|$ has a local minimum $p = x + yI \in B$, then either f is constant on B or f has a zero in the 2-sphere $x + yS$.

Proof. Assume that $|f|$ has a local minimum at a point $p \in B$. Suppose that f is nonzero for all points p_0 in $\Sigma = x + yS$ and consider the regular reciprocal f^{-*} of f. By previous arguments, f^{-*} is regular on the domain $\Omega = B \setminus Z_{f^s}$ where Z_{f^s} is the zero set of the symmetrization f^s of f. By the previous observation that $|f^{-*}| = |f(T_f(q))|^{-1}$, letting $p = T_f(p')$ implies that $|f|$ has a minimum at p precisely when $|f \circ T_f|$ has a minimum at p'[12], which in turn happens if and only if $|f^{-*}|$ has a *maximum* at p'. Thus, by the Maximum Modulus Theorem, f^{-*} is constant on its domain, thereby proving that $|f|$ has a local minimum at p if and only if f^{-*} is constant in its domain, which by definition occurs if and only if f is also constant on its domain. □

For real numbers x and $y > 0$, the 2-sphere $\Sigma = x + yS$ mentioned in Lemma 4.7 is of particular importance when characterizing the behavior of regular functions f and their zero sets Z_f. One way to illustrate the role of Σ in these theories is by way of proving the remaining fundamental results; another is by way of the following definition concerning them.

Definition 4.8. Let $B = B(\mathbf{0}, R)$, let x and $y > 0$ be real numbers for which $\Sigma \subset B$ where $\Sigma = x + yS$, and let $f : B \to \mathbb{H}$ be a regular function. The 2-sphere Σ is said to be **degenerate** for f if the restriction $f\big|_\Sigma$ of f to Σ is constant. The union D_f of all degenerate spheres for f is called the **degenerate set** of f.

[12] This stems from the fact that T_f is a diffeomorphism of $B \setminus Z_{f^s}$ onto itself, a fact verified in depth in [24].

It's easily shown that for a regular function $f : B \to \mathbb{H}$ and for a nondegenerate sphere $\Sigma = x + yS \subset B$, the restriction $f\big|_\Sigma$ of f to Σ is an affine map of Σ onto a sphere of the form $b + Sc$ for $b, c \in \mathbb{H}$. In particular, this proves that in such a situation, the modulus $\big|f\big|_\Sigma\big|$ of the restriction $f\big|_\Sigma$ has precisely one global maximum, one global minimum, and *no* other extreme point. These two observations, along with the weak version of the Minimum Modulus Theorem, allow for a stronger version to be stated. Upon stating and proving the Strong Minimum Modulus Theorem, the last of the major tools needed to establish the Open Mapping Theorem will be in place.

Theorem 4.9 (Strong Minimum Modulus Theorem)**.** Let $B = B(\mathbf{0}, R)$ and let $f : B \to \mathbb{H}$ be a regular function. If $|f|$ has a local minimum point $p \in B$, then either $f(p) = 0$ or f is constant on its domain.

Proof. Suppose that f is non-constant and that $|f|$ has a local minimum at a point $p = x + yI$ of B. By the weak version of the theorem, this implies that $f(p') = 0$ for some $p' \in x + yS$. Hence, $|f|$ has a minimum at p'. By the above observation, $|f|$ cannot have two local minimum points on the sphere $x + yS$ unless $x + yS$ is degenerate for f. In particular, then, either $|f|$ has one local minimum and $p = p'$ *or* if $|f|$ *does* have a second local minimum, then $x + yS$ is degenerate for f and so f is constant on $x + yS$. In the case where $p = p'$, $f(p) = f(p') = 0$, thus proving the result. □

Having stated a full version of the minimum modulus theorem, the stage is now set for the culminating result. Unsurprisingly, the proof the Open Mapping Theorem for quaternion functions of a quaternion variable draws heavily from the algebraic and geometric properties of both the zero sets and the degenerate sets of these functions. Before proceeding, it's worth noting that the degenerate set D_f of a regular quaternion-valued function f defined on the ball $B = B(\mathbf{0}, R)$ is actually a closed subset of $B \setminus \mathbb{R}$. One way to see this is to note that $D_f = C_f \setminus R$ where C_f is the union of all "spheres" (here, "spheres" refers to both 2-spheres in $\mathbb{H}$ and real singletons) on which f is constant. Even more is true, however, because in addition to being a closed subset of $B \setminus \mathbb{R}$, the degenerate set D_f for non-constant regular functions f has the property that its interior D_f° is actually empty, whereby it follows that the entirety of D_f is actually its boundary ∂D_f for non-constant f. Both of these observations are used below.

This section concludes with the establishment of the Open Mapping Theorem for Cullen regular functions. For various other results concerning Cullen regularity, the reader is directed to the literature cited throughout this section, as well as to [25] for a more expansive presentation.

Theorem 4.10 (Open Mapping Theorem)**.** Let $f : B(\mathbf{0}, R) \to \mathbb{H}$ be a non-constant regular function with degenerate set D_f. Then $f : B(\mathbf{0}, R) \setminus \overline{D_f} \to \mathbb{H}$ maps open sets to open sets.

Proof. Let Ω denote the domain $B(\mathbf{0}, R) \setminus \overline{D_f}$ and let U be an open set of Ω with $p_0 \in f(U)$. Let $\varepsilon > 0$. In order to prove that f maps open sets to open sets, it suffices to prove that there exists a ball $B(p_0, \varepsilon)$ in the image $f(U)$ of U. To that end, let $q_0 \in U$ so that $f(q_0) = p_0$.

Because $U \subset \Omega$ where $\Omega = B(\mathbf{0}, R) \setminus \overline{D_f}$, it follows that $U \cap D_f = \varnothing$, whereby it also follows that q_0 must be an isolated zero of the function $f(q) - p_0$.[13] As an isolated point, it follows that there exists an entire neighborhood around q_0 for which the function $f(q) - p_0$ is bigger than some prescribed positive value. To be more precise, there exists a value $r > 0$ for which $\overline{B(q_0, r)} \subset U$ and for which $f(p) - q_0 \neq 0$ for all $q \in \overline{B(q_0, r)} \setminus \{q_0\}$. Moreover, given the arbitrary value ε above, r can be chosen so that $|f(q) - p_0| > 3\varepsilon$ for all q on the circle $|q - q_0| = r$. Thus, for all p in $B(p_0, \varepsilon)$,

$$\begin{aligned} |f(q) - p| &\geq ||f(q) - p_0| - |p - p_0|| \text{ by the triangle inequality} \\ &\geq |3\varepsilon - \varepsilon| \text{ whenever } |q - q_0| = r, \text{ by the above arguments} \\ &= 2\varepsilon. \end{aligned}$$

Moreover, $|f(q_0) - p| = |p_0 - p| \leq \varepsilon$ for all such q_0, from which it follows that

$$|f(q_0) - p| < \min_{|q - q_0| = r} |f(q) - p|$$

and hence that $|f(q) - p|$ has a local minimum in $B(q_0, r)$. Hence, f non-constant implies that $f(q) - p$ is non-constant, which by the minimum modulus theorem implies that $f(q) - p$ must equal zero at some value $q \in B(q_0, r)$, i.e. that $f(q) = p$ for some $q \in B(q_0, r) \subset U$. Therefore, $f(B(q_0, r)) \subset B(f(q_0), \varepsilon)$, thereby proving that f maps open sets to open sets. □

As was the case for complex-valued functions of a complex variable, the Open Mapping Theorem simplifies the proofs of several important results and also yields an abundance of other, independently worthwhile knowledge concerning the topology of the host space. For example, Theorem 4.10 immediately implies that any regular function $f : B(\mathbf{0}, R) \to \mathbb{H}$ which is non-constant on every 2-sphere $x + yS$, $y > 0$, is automatically open. In particular, given any so-called "circular" subset[14] $U \subset B(\mathbf{0}, R)$, $f(U)$ is necessarily open in $\mathbb{H}$, thereby allowing easy recognition of a large number of open subsets of the quaternion space $\mathbb{H}$.

5 Applications: General Relativity

As mentioned in Chapter 1, one of the motivating ideas behind Hamilton's defining of the quaternion group was to represent algebraically the quotient of vectors in Euclidean 3-space. As such, the quaternions serve as a useful basis for describing rotations of vectors in three- and four-dimensional Euclidean space. It may come as no surprise, then, that the behavior of the quaternion algebra $\mathbb{H}$ and of quaternion-valued functions of a quaternion variable is inexorably linked to many modern scientific applications, notable among which is the study of general relativity. Heuristically, general relativity is the geometrically-represented theory of universal gravitation first published by Einstein in the early 20th century. Because

[13] Otherwise, if q_0 weren't isolated, then f would equal zero on an entire neighborhood U_0 of q_0, thereby implying that $q_0 \in U_0 \subset D_f$.

[14] As defined in [24], a domain U is *circular* if for every quaternion $x + yI$, $y > 0$, in U, the entire 2-sphere $x + yS$ is also contained in U.

of its geometric roots, general relativity is largely mathematical, and so applications of quaternionic theory to the theory of general relativity will be the focus of this particular segment of the exposition—one which is based in principle on the presentation found in the online textbook [26]. To begin, some general framework is necessary.

5.1 Introduction

The emergence of the theory of general relativity dates back to the year 1915, to the work of Einstein. In that year, Einstein published for the first time versions of his field equations (from a tensor perspective), which were shown to describe the fundamental interaction between gravity and the resulting curvature of spacetime due to the presence of matter and energy. More precisely, the Einstein field equations (EFE for short) are a system of ten partial differential equations which "equate local spacetime curvature with the local energy and momentum within that spacetime" [27]. In order to understand the physical universe through the perspective of general relativity, it's necessary to first understand the structure of the universe (that is, of spacetime) as a mathematical object.

Simply put, spacetimes are any descriptions of the arenas in which physical events take place which are, necessarily, independent of any observer. Because space is generally regarded to exist in a three-dimensional context, definitions of events in spacetimes are associated to vectors consisting of four real numbers, namely the three associated to the aforementioned viewpoint of space plus an additional one to describe the one-dimensional time component of the event. From a mathematical perspective, spacetime in the sense of Einstein's general relativity model is therefore most logically associated to Minkowski's model of spacetime—a model which, in contrast to four-dimensional Euclidean space consisting only of "space-like" dimensions, consists of three "space-like" dimensions and one "time-like" dimension. Minkowski space is a robust example with properties of mathematical interest from lots of different perspectives: In particular, Minkowski space M is a pseudo-Euclidean space which is also a pseudo-Riemannian (specifically, a Lorentzian) manifold equipped with a nondegenerate, symmetric bilinear form with metric signature $(-,+,+,+) = (1,3)$. These properties ensure that M is equipped with an "inner product" η and a resulting "metric"[15], etc. These notions will be defined rigorously below.

5.2 Preliminaries: Background, Notation, and Definitions

5.2.1 General Preliminaries

As explained in [26], distance in the world of spacetime models is measured "with a clock" by moving a clock along a reference path between two events in spacetime and describing the "length" of the resulting path (that is, the distance) as the measured time that's lapsed. More generally, this process describes a way by which the separation between two events

[15]These fail the nonnegativity conditions associated with inner products and metrics, respectively, and so they are, at best, pseudo-inner products and pseudo-metrics, respectively.

can be measured in an interval-invariant fashion which is unlike the purely spatial distance utilized in traditional Euclidean space models. As mentioned above, the use of Minkowski space as a model for spacetime complicates the naive use of the term "interval" significantly, so before proceeding, some basic definitions are needed.

Let s^2 denote the interval between two events in space time. Then for space and time components $\mathbf{r}$ and t, respectively, a the exact value for s^2 is given by

$$s^2 = (\Delta \mathbf{r})^2 - c^2 (\Delta t)^2, \tag{5.2.1}$$

where c is the speed of light and $\Delta \mathbf{r}$, Δt denote the differences in the space- and time-components, respectively, of the two events. Note that, a priori, there's no way to know whether s^2 is positive or negative using the definition in (5.2.1): This fact provides the basis necessary to make the following definitions.

Definition 5.1. Let s^2 denote the spacetime interval defined in (5.2.1). This interval is said to be **time-like** provided $s^2 < 0$, **light-like** if $s^2 = 0$, and **space-like** if $s^2 > 0$.

Remark 5.1. Note for completeness that an alternative formulation of definition 5.1 consists of examining explicitly the behavior of the Minkowski inner-product $\eta : M \times M \to \mathbb{R}$ and to define time-like, light-like, and space-like behaviors of individual vectors $\mathbf{v} \in M$ based on whether $\eta(\mathbf{v}, \mathbf{v}) < 0$, $\eta(\mathbf{v}, \mathbf{v}) = 0$, and $\eta(\mathbf{v}, \mathbf{v}) > 0$, respectively. In this case, defining an orientation of the time dimension allows time-like and light-like vectors $\mathbf{v}$ to be further quantified as *future directed* and *past directed* based on whether the time component of $\mathbf{v}$ is positive or negative. These notions will be largely unnecessary moving forward.

Given a time-like interval, the *proper time interval* $\Delta\tau$ is the quantity given when the path of an observer traveling between the two events with a clock intersects each event as that particular event occurs. Similarly, given a space-like interval, the space-like separation $\Delta\sigma$ called the *proper distance* is used. Mathematically, expressions for $\Delta\tau$ and $\Delta\sigma$ are

$$\Delta\tau = \sqrt{(\Delta t)^2 - \frac{(\Delta \mathbf{r})^2}{c^2}} \text{ and } \Delta\sigma = \sqrt{(\Delta \mathbf{r})^2 - c^2 (\Delta t)^2}, \tag{5.2.2}$$

both of which are positive given their respective interval-types.

Additional behavior of Minkowski space (and therefore, of spacetime models) can be derived by examining M in the somewhat more general context of pseudo-Riemannian geometry. As mentioned above, M is an example of a *Lorentzian Manifold*, that is, a manifold of dimension n with metric signature of the form $(1, n-1)$. For a given pseudo-Riemannian manifold X with metric signature $(1, n-1)$, a "metric" $m : X \times X \to \mathbb{R}$ can be defined by way of differentials as

$$m = dx_1^2 - dx_2 - \ldots - dx_n^2.$$

Adopting this definition to the particular instance of Minkowski space yields a "metric" m of the form $m = dx_1 - dx_2 - dx_3 - dx_4$ or, equivalently, $m = dt - dx - dy - dz$ for vectors written in the form $\mathbf{v} = (t, x, y, z)$. This particular metric allows the above qualification of vector behavior to be reconciled by examining whether $m(\mathbf{u}, \mathbf{v}) > 0$, $m(\mathbf{u}, \mathbf{v}) = 0$, or $m(\mathbf{u}, \mathbf{v}) < 0$ for vectors $\mathbf{u}, \mathbf{v}$ in the tangent space T_pM at a given point $\mathbf{p} \in M$.

One final consideration in the association of spacetime to pseudo-Riemannian geometry that's worth mentioning before moving on to a quaternion-specific investigation is its curvature. Indeed, the debate as to how best to quantify the shape of the universe is an active one in the field of cosmology, with new theories being proposed regularly. Physically, curvature models of the universe are derived by examining locally the magnitude of the density parameter Ω, defined to be the ratio of the observed density ρ to the "critical density" ρ_c within the so-called Friedmann equations[16,17]. On the other hand, curvature analysis from a mathematical perspective happens by way of considering the Riemann curvature tensor, the definition of which will conclude this section and serve as a segue into the discussion of general relativity from a quaternionic perspective.

Given a differentiable manifold M, a *vector field* on M is an assignment of a tangent vector $\mathbf{v}_p \in T_pM$ to each point $\mathbf{p} \in M$. Given such a manifold, let $C^\infty(M, TM)$ denote the collection of vector fields on M. An *affine connection* on M is a bilinear map

$$\begin{aligned} C^\infty(M, TM) \times C^\infty(M, TM) &\to C^\infty(M, TM) \\ (X, Y) &\mapsto \nabla_X Y \end{aligned}$$

which satisfies the following criteria for all smooth functions $f \in C^\infty(M, \mathbb{R})$ and for all vector fields X, Y on M:

1. $\nabla_{fX} Y = f \nabla_X Y$, i.e., $C^\infty(M, \mathbb{R})$-linearity in the first coordinate, and
2. $\nabla_X (fY) = \mathrm{d}f(X)Y + f\nabla_X Y$, i.e. satisfaction of the Leibniz rule in the second coordinate.

Moreover, given two vector fields X and Y, the *Lie Bracket* $[X, Y]$ of the two is a third vector field defined by

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

For a (pseudo-)Riemannian manifold M with metric g, the affine connection ∇ on M is called a *Levi-Civita connection* provided that

1. $\nabla g = 0$, i.e. ∇ preserves the metric properties of g, and
2. $\nabla_X Y - \nabla_Y X = [X, Y]$, that is to say ∇ is torsion-free.

[16] The Friedmann equations are a set of equations derived from the EFE by Alexander Friedmann in the 1920s which are theorized to govern universal expansion

[17] A mathematical expression for Ω is

$$\Omega = \frac{8\pi G\rho}{3H^2},$$

where G is Newton's gravitational constant and H is the so-called Hubble parameter.

Note that the Fundamental Theorem of Riemannian Geometry guarantees the existence and uniqueness of a Levi-Civita connection. Finally, for vectors $\mathbf{u}, \mathbf{v} \in T_pM$, the *Riemannian curvature tensor* is a linear transformation $R(\mathbf{u}, \mathbf{v}) : T_pM \to T_pM$ defined on each $\mathbf{w} \in T_pM$ by

$$R(\mathbf{u}, \mathbf{v})\mathbf{w} = \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\mathbf{w} - \nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\mathbf{w} - \nabla_{[\mathbf{u},\mathbf{v}]}\mathbf{w}. \tag{5.2.3}$$

Though seemingly abstract, the Riemannian curvature tensor can be thought of as a measure of magnitude of the *holomony* of the manifold, that is the magnitude by which vectors $\mathbf{v}$ in (pseudo-)Riemannian manifolds fail to maintain their spatial orientation when transported in parallel fashion around a loop. More precisely, because all vectors in Euclidean space are known to maintain their original orientation given such a transport, the Riemannian curvature tensor can be thought of as a measure of how "non-Euclidean" a (pseudo-)Riemannian manifold M is in terms of its curvature. This tool is, therefore, fundamentally centered at the heart of general relativity as a tool to measure the amount by which spacetime bends when subjected to the gravitational effects of an event therein. Later, quaternionic vector expressions analogous to the one in (5.2.3) will be given.

5.2.2 Use of Quaternions

One commonly-used framework for general relativity involves a minor modification of the standard Hamiltonian definition of quaternions. These modified quaternions are known as *Pauli quaternions* after their inventor, Austrian physicist Wolfgang Pauli, who utilized the modified system in his study of particle spin in quantum mechanics and electromagnetic field theory. Before jumping in fully, however, a little background is required.

A square $n \times n$ matrix $A = (a_{ij})$, $i, j = 1, \ldots, n$, is called *Hermitian* or *conjugate symmetric* if $a_{ij} = \overline{a_{ji}}$ for all i, j. One of the fundamental quantifiers describing the geometric properties of universal gravitation are the so-called *Pauli Matrices* σ_i, $i = 1, \ldots, 4$, which are defined recursively in terms of the Kronecker delta δ_{ij} as

$$\sigma_j = \begin{pmatrix} \delta_{j3} & \delta_{j1} - i\delta_{j2} \\ \delta_{j1} + i\delta_{j2} & -\delta_{j3} \end{pmatrix}, j = 1, 2, 3. \tag{5.2.4}$$

Expanding (5.2.4) gives an explicit representation for the Pauli spin matrices, namely

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

with a fourth matrix σ_4 defined by the relation $\sigma_4\sigma_j = \sigma_j\sigma_4 = i\sigma_j$ for $j = 1, 2, 3$. It's easily verified that under this identification, $\sigma_4 = \text{diag}_{2\times 2}(i)$, that is the 2×2 diagonal matrix with diagonal entries $a_{11} = a_{22} = i$.

The collection of the matrices σ_j corresponds directly a redefined quaternion number system known as the *Pauli quaternions* whose basis elements are of the form s_i, $i = 1, \ldots, 4$, subject to the multiplicative identities induced by the matrix products $\sigma_i\sigma_j$ for $i, j = 1, \ldots, 4$. In particular, then,

$$\begin{aligned} s_1 s_2 &= -s_2 s_1 = i s_3 \\ s_3 s_1 &= -s_1 s_3 = i s_2 \\ s_2 s_3 &= -s_3 s_2 = i s_1 \\ s_4 s_k &= s_k s_4 = i s_k,\ k = 1, \ldots, 4 \\ s_k^2 &= -1 = -s_4^2,\ k = 1, 2, 3. \end{aligned}$$

Evidently, basis elements s_i of these new Pauli quaternions differ from the basis elements e_i of the Hamilton quaternions by a multiplicative factor of i. From this point forward, use of the unqualified term "quaternion" will be used to denote "Pauli quaternion," whereas the the Hamiltonian quaternions will be specified by way of the modifier "Hamilton" or "Hamiltonian". Also note that, due to the minor modification involved in converting standard Hamiltonian quaternions to the Pauli quaternions, many of the identities listed in chapter 2 follow through with only minor adjustments. For this reason, such identities are omitted here with the exception of the ones needed to proceed.

Given a general quaternion $q = q_1 s_1 + q_2 s_2 + q_3 s_3 + q_4 s_4$, the *norm* of q is defined to be $\|q\| = \text{Norm}(q) = \sum_{i=1}^{4} q_i^2$, while its *magnitude* $|q| = \text{Mag}(q) = \sqrt{\text{Norm}(q)}$. A quaternion q satisfying $\text{Norm}(q) = 1$ will be called *unimodular*. One should be aware that there exists an operator U that constructs a unimodular quaternion $U(q)$ in the same direction as q which has the form

$$U(q) = \frac{q}{\sqrt{\text{Norm}(q)}}.$$

Note that any general quaternion $q = q_1 s_1 + q_2 s_2 + q_3 s_3 + q_4 s_4$ can be written in vector notation as $q = \begin{pmatrix} q_1, & q_2, & q_3, & q_4 \end{pmatrix}^T$. In general, Pauli quaternions are assumed to have coefficients $q_i \in \mathbb{C}$ rather than the Hamiltonian convention that $q_i \in \mathbb{R}$[18]. For such an element, the *Hermitian conjugate* $q^\dagger$ of q is the quaternion $q^\dagger = \begin{pmatrix} q_1^*, & q_2^*, & q_3^*, & -q_4^* \end{pmatrix}^T$, where q_i^* denotes the complex conjugate of q_i. From this definition, a quaternion q can be called *Hermitian* if $q^\dagger = q$.

For an arbitrary quaternion q, an *orthogonal transformation* T of q is an operator, the result of which is an element of the form $q' = aqb$, where a, b are unimodular. The collection of all such transformations forms a group with "multiplication" $T_2 T_1$ of any such transformations defined to be a quaternion $q'' = a_2 q' b_2 = a_2 a_1 q b_1 b_2$, details of which (including representations of this group) can be found in [28]. A specific example of these arises when q and q' are Hermitian and $b = a^\dagger$, in which case the transformation has the form

$$q' = aqa^\dagger,\ a \text{ unimodular},\ aa^\dagger = 1,$$

[18] Generally, quaternions of this type are called *biquaternions*, though no such distinction shall be made herein.

and is called the *Lorentz Transformation*. From a physical perspective, the importance of orthogonal transformations lies in the fact that said transformations represent rotations, reflections, or combinations thereof [26]. Moreover, the collection of all Lorentz Transformations (called the *Lorentz Group*) is the group of homogeneous transformations which conserve the quantity $c^2t^2 - x^2 - y^2 - z^2$ [28]. Hence, it's unsurprising that unimodular quaternions and the transformations thereof should be expected to play a fundamental role in the study of motion in spacetime.

Along with the observations above, one should be aware that, given any two hermitian quaternions q, q', there is a unimodular quaternion q_L connecting them in the sense that $q' = q_L q q_L^\dagger$. In particular, [26] shows that q_L has the form

$$\begin{aligned} q_L &= \sqrt[Q]{U(q)\overline{U(q')}} \\ &= -i\sqrt{i\,\mathrm{Mag}(q)}\left\{s_4 \cos\left(\frac{\mathrm{Arg}_Q(q)}{2}\right) + U(\mathrm{Im}(q))\sin\left(\frac{\mathrm{Arg}_Q(q)}{2}\right)\right\}, \end{aligned} \tag{5.2.5}$$

where

$$\mathrm{Im}(q) = \frac{1}{2}(q - \overline{q}) \text{ and } \mathrm{Arg}_Q(q) = \arccos\left(\frac{q_4}{\mathrm{Mag}(q)}\right).$$

The formula (5.2.5) will be used throughout the study of general relativity, as will the forms and operators defined above. To prove that the above framework consists of more than generalized abstraction and nonsense, consider the following example, proposed in the framework of [26] though first deduced theoretically by Silberstein in [29]. First, note that the location of an event in four-dimensional spacetime can be represented by the Hermitian quaternion $q_e = \begin{pmatrix} x, & y, & z, & ict \end{pmatrix}^T$ and that the omnipresent c denotes the speed of light.

Example 5.1. Consider a physical model with two observers, Observer A (O_A) and Observer B (O_B). Say that O_A spots an event at the origin of a frame at rest and makes its observation at time t_A. The Hermitian quaternion $q_A = \begin{pmatrix} 0, & 0, & 0, & ict_A \end{pmatrix}^T$ describes the event from the perspective of O_A. If, at the same time (call it t_B for time relative to O_B), O_B is traveling at a velocity v_B relative to frame A, then the events relative to B exist at a point in spacetime denoted quaternionically by $q_B = \begin{pmatrix} vt_B, & 0, & 0, & ict_B \end{pmatrix}^T$. By way of equation (5.2.5), the unimodular quaternion q_L that transforms q_A into q_B can be computed. The result of this computation is the now-famous *Lorentz Contraction*, the quaternion-form of which is given by

$$q_L(v) = \begin{pmatrix} \sqrt{1-\gamma}, & 0 & 0 & \sqrt{1+\gamma} \end{pmatrix}^T, \text{ where } \gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \text{ is the Lorentz Factor.}$$

It's important to note that q_L represents a transform whose behavior is to make things traveling at velocity v appear at rest. Heuristically, for velocities v not near the speed of light, $\gamma \approx 1$ implies that the only perceived "distinction" between q_A and q_B exists in the time dimension, as no tangible curvature of the spacetime would result. □

5.2.3 Functions and Differential Forms

Note that both (5.2.5) and Example 5.1 illustrate something suggested in Section 5.2.1, namely that the field of general relativity can be intimately linked to and naturally extended by the utilization of quaternion functions of a quaternionic variable. As [26] indicates, the ability to map a quaternion q in spacetime to another quaternion $Q_f(q)$ is fundamentally essential to the field of general relativity, though as pointed out in [30], the algebra of spacetime itself is particularly conducive to the restriction to those functions whose domains are Hermitian quaternions. Therefore, from here on, the unmodified term "quaternion(s)" will therefore refer to *Hermitian quaternion(s)* unless otherwise noted.

For this section, the notation adopted will be as follows: A function Q_f has as its input a quaternion q. Assuming that $Q_f(q_0) = Q_{f_0}$ for some quaternion q_0, the value of Q_f at a neighboring point $q_0 + dq$ is given by $Q_{f_0} + dQ$ where dQ is the quaternion 1-form given by the gradient formula

$$dQ = \sum_{i=1}^{4} \left(\tfrac{\partial Q_1}{\partial q_i}, \quad \tfrac{\partial Q_2}{\partial q_i}, \quad \tfrac{\partial Q_3}{\partial q_i}, \quad \tfrac{\partial Q_4}{\partial q_i} \right)^T dq_i.$$

The use of quaternion-valued 1-forms and their exterior derivatives follow the standard usage outlined in Appendix 1.1, upon careful observation of the fact that quaternion-valued terms therein fail to commute. In addition, higher-level quaternion differential forms will be used below to succinctly describe motion in spacetime, though throughout, careful attention will paid to the order with which operations are performed.

Recall that inn section 5.2.1, it was shown how to derive a quaternion q_L defining a Lorentz transformation between two Hermitian quaternions q, q'. [28] takes special care to note that in the form previously given, the Lorentz transformation resulting is "pure," i.e. that it contains no spatial rotations. More generally, a Lorentz transformation between two quaternions q and q' consists of a pure transform composed with a three-dimensional rotation $q \mapsto rq\overline{r}$, the result of which is either of the two transformations[19]

$$q' = b(rq\overline{r})b^\dagger \text{ or } q' = r(bqb^\dagger)\overline{r}.$$

Worth noting is that these spatial rotations can be defined in terms of quaternion differential operators. Indeed, let H represent an arbitrary Hermitian quaternion field. As above, values of H in a neighborhood of some point q_0 is given by the 1-form $dH = \sum_{i=1}^4 H_i dq_i$. Given a second unimodular quaternion field R defined differentially by $I + dR = \sum_{i=1}^4 R_i dq_i$, the effect of rotating H by R is the resulting quaternion 2-form $R \wedge H + H \wedge R^\dagger$, whose component-wise form can be derived from the identity

$$\begin{aligned} R \wedge H = {} & (R_1H_2 - R_2H_1)\, dq_1dq_2 + (R_1H_3 - R_3H_1)\, dq_1dq_3 + (R_1H_4 - R_4H_1)\, dq_1dq_4 \\ & + (R_2H_3 - R_3H_2)\, dq_2dq_3 + (R_2H_4 - R_4H_2)\, dq_2dq_4 + (R_3H_4 - R_4H_3)\, dq_3dq_4. \end{aligned}$$

[19] In general, these will be different because of noncommutativity.

Now that the basic tools needed to understand the process of "shifting" one quaternion-defined event in spacetime to another, it suffices to jump more fully into the adaptation of general relativity to a perspective that's fully quaternionic. The objective will be to state some fundamental results from the field of general relativity and to see how a quaternion-based perspective can be used to succinctly reduce the complexity of these results. That goal will be the driving force behind the remainder of this section.

5.3 General Relativity Redux, Quaternion Style

The remainder of this chapter will be dedicated to reformulating significant general relativity results and translating them into the language of quaternions. The results themselves will be treated as self-contained so that any necessary background to understand them is included where necessary.

5.3.1 Einstein's Field Equations

As mentioned in section 5.1, the publications of the Einstein Field Equations (EFE) served as the establishment of general relativity as a field of modern science. The EFE may be written concisely in the tensor form

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{5.3.1}$$

where the constants in (5.3.1) are as follows:

- $R_{\mu\nu}$ is the Ricci curvature tensor, which describes the amount by which the volume of a geodesic ball in spacetime deviates from that of a standard ball in Euclidean space.
- R is the scalar curvature, which assigns to each point $\mathbf{p}$ in spacetime a real number describing the relativistic geometry of the manifold near $\mathbf{p}$.
- $g_{\mu\nu}$ is the metric tensor, which captures all geometric and causal structure of spacetime.
- Λ is the energy density.
- G and c are the gravitational constant and the speed of light, respectively.
- $T_{\mu\nu}$ is the stress-energy tensor, which describes the density and flux of energy and momentum in spacetime.

In the original derivation of the EFE, $\Lambda = 0$ was assumed, whereby (5.3.1) is reduced to the simpler tensor equation

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = kT_{\mu\nu}, \tag{5.3.2}$$

where k is used to denote the constant $8\pi Gc^{-4}$. The form (5.3.2)of the EFE is the one that will be focused on in this section, where it will be proven that this particular equation can be re-expressed as a system of two equations in which quaternion variables $q_\mu(\mathbf{x})$ and $\overline{q_\nu}(\mathbf{x})$ are utilized in the place of $g_{\mu\nu}(\mathbf{x})$. This construction will largely follow the exposition in [31].

Let q_μ be a quaternion whose basis elements are (σ_4, σ_k), while $\overline{q_\mu}$ is its "time-reversed field" with basis quaternions $(-\sigma_4, \sigma_k)$, where σ_i denotes the ith Pauli matrix, $i = 1, \ldots, 4$ and $k = 1, 2, 3$. One can verify that the metric tensor $g_{\mu\nu}$ can be written in terms of quaternions q_μ and q_ν as

$$\sigma_4 g_{\mu\nu}(\mathbf{x}) = -\frac{1}{2}(q_\mu \overline{q_\nu} + q_\nu \overline{q_\mu})(\mathbf{x}), \tag{5.3.3}$$

where σ_4 denotes the fourth Pauli matrix. Solving (5.3.3) for $g_{\mu\nu}$ yields

$$g_{\mu\nu}(\mathbf{x}) = -\frac{1}{4}\operatorname{Tr}\left(q_\mu \overline{q_\nu} + q_\nu \overline{q_\mu}\right)(\mathbf{x}), \tag{5.3.4}$$

where Tr denotes the trace operator. Evidently, equation (5.3.4) is bilinear in both q_μ and $\overline{q_\nu}$, which suggests that factoring (5.3.2) into a pair of quaternion-valued expressions of q_μ and $\overline{q_\nu}$ may be possible. Worth noting is that the expression of q_μ, $\overline{q_\mu}$ as matrices with unit quaternion entries yields an interesting invariant, $q_\mu \overline{q_\mu} = \overline{q_\mu} q_\mu = -4$. For a more rigorous treatment of this derivation, see [31].

Using the index notation for tensors, let $R_{\mu\kappa\rho\lambda}$ denote the aforementioned Riemann curvature tensor. By exploiting the constant behavior of the product $q_\mu \overline{q_\mu}$ with respect to derivatives, it follows that the behavior of derivatives near a point are uniquely determined by two quantities, namely changes in spinor space and changes in coordinate space. Let η denote a covariant spinor field and write $[q_\mu]$ coordinate space transformation. In [31], Sachs defines the spin-affine connection Ω_ρ, and by investigating the behavior of Ω_ρ relative to the form (5.3.4) and the quantities $[q_\mu]$ and η, it can be shown that

$$R_{\mu\kappa\rho\lambda} R_\kappa = [q_\mu]_{\rho\lambda} - [q_\mu]_{\lambda\rho}. \tag{5.3.5}$$

By then denoting the spin curvature by $K_{\rho\lambda}$ and performing a series of calculations, the author derives component-wise quantifications for the Riemann curvature tensor, the Ricci tensor, and the scalar curvature tensor of the forms:

$$R_{\mu\kappa\rho\lambda} = \frac{1}{4}\operatorname{Tr}\left[K_{\rho\lambda}(q_\mu \overline{q_\kappa} - q_\kappa \overline{q_\mu}) + h.c\right], \tag{5.3.6}$$

$$R_{\kappa\rho} = \frac{1}{4}\operatorname{Tr}\left[K_{\rho\lambda}(q_\lambda \overline{q_\kappa} - q_\kappa \overline{q_\lambda}) + h.c.\right], \text{ and} \tag{5.3.7}$$

$$R = \frac{1}{2}\operatorname{Tr}[\overline{q_\rho} K_{\rho\lambda} q_\lambda + h.c.], \tag{5.3.8}$$

where "h.c." denotes the hermitian conjugate matrix fields. Note, then, that in the framework of quaternion-valued spinors, equations (5.3.4), (5.3.6), (5.3.7), and (5.3.8) give derivations of many of the useful tensorial quantities used in the derivation of the original EFE. After some notational framework is established, the so-called Palitini technique is applied with respect to the Lagrangian $\mathcal{L}_G = R\sqrt{-g}$, from which the following quaternionic field equations result:

$$\frac{1}{4}\left(K_{\rho\lambda}q_\lambda + q_\lambda K^\dagger_{\rho\lambda}\right) + \frac{1}{8}Rq_\rho = K\mathcal{T}_\rho \tag{5.3.9}$$

$$-\frac{1}{4}\left(K^\dagger_{\rho\lambda}\overline{q_\lambda} + \overline{q_\lambda}K_{\rho\lambda}\right) + \frac{1}{8}R\overline{q_\rho} = K\left(\tau\mathcal{T}\right)_\rho, \tag{5.3.10}$$

where $\mathcal{T}_\rho$ is the vector field obtained in the Palitini method by minimizing the spatial variables of the Lagrangian, $\tau\mathcal{T}_\rho = \varepsilon\mathcal{T}^*_\rho\varepsilon$, $*$ denotes complex conjugation, and ε is the 2×2 matrix

$$\varepsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

One way to verify the equivalence of the field equations (5.3.9) and (5.3.10) to the original tensor formulation of the EFE is to recover the latter from the prior. This is done in moderate detail by Sachs, and as such, the reader is directed to [31] to verify. More generally, because of the hurried approach taken here for brevity, the reader is directed to [31] and to the sources cited therein for a treatment of the derivation which is thoroughly more rigorous and detailed.

5.3.2 The Schwarzschild Metric

As mentioned in section 5.3.1, the independent variable for the original tensorial EFE is a metric tensor $g_{\mu\nu}$. In particular, then, solutions to the EFE are themselves components of specific metric tensors. Because of the astounding level of difficulty of the EFE, exact solutions can be found only by making use of miscellaneous simplifying assumptions. The Schwarzschild metric is one such solution.

Generally considered to be one of the simplest solutions of the EFE, the Schwarzschild metric describes the shape of spacetime within a local vicinity of a non-rotating, spherically-symmetric object. Beginning with equation (5.3.1), assume $\Lambda = 0$ to obtain (5.3.2). Further assume that the metric solution is vacuum. Under these particular hypotheses, Schwarzschild derived as an exact solution to the EFE the quantity $c^2(\Delta\tau)^2 = -\left(\Delta s\right)^2$, which, in spherical coordinates (t, r, θ, ϕ), has the form

$$c^2(\Delta\tau)^2 = \left(1 - \frac{r_s}{r}\right)c^2dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1}dr^2 - r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right), \tag{5.3.11}$$

where s, $\Delta\tau$ are as in (5.2.1) and (5.2.2), respectively. Additional constants in (5.3.11) are:

- r is the radial component, measured as the normalized circumference of the body.
- θ is the "colatitude," that is, the angle from North, in radians.
- ϕ is the longitude, in radians.

- r_s is the so-called *Schwarzschild radius* of the massive body, defined to be $r_s = 2GMc^{-2}$ where G is the gravitational constant and M is the mass of the object itself.

While the solution as written in (5.3.11) is given in spherical coordinates and representative of a tensorial viewpoint of the metric, it's shown in [26] that an alternative formulation involving quaternion-valued vectors and differential forms is easily attained. The goal of this section is to derive that formulation.

One way to express the metric whose coordinate-wise representation is given in (5.2.1) is to demonstrate the actual displacement corresponding to changes in quaternionic coordinates. To that end, start with the "scaling quaternion" σ, which has the form

$$\sigma = \begin{pmatrix} \dfrac{dr}{\sqrt{1 - r_s/r}} \\ r\,d\theta \\ r \sin\theta\, d\phi \\ i\sqrt{1 - r_s/r}\,dt \end{pmatrix}. \tag{5.3.12}$$

The form sought is obtained in part by considering the effects of moving in the spacetime environment given: In particular, the change in σ by such a movement can be thought of as a rotation being acted upon by a second quaternion denoted Ω, where Ω is the 1-form determined by the relation

$$d\sigma = \Omega \wedge \sigma + \sigma \wedge \Omega^\dagger.$$

Using the expression for σ in (5.3.12), it follows that $d\sigma$ has the form

$$d\sigma = \begin{pmatrix} 0 \\ dr\, d\theta \\ r\cos\theta\, d\theta d\phi + \sin\theta\, dr d\phi \\ \dfrac{ir_s\, dr dt}{2r^2\sqrt{1 - r_s/r}} \end{pmatrix}, \tag{5.3.13}$$

where $(d\sigma)_1 = 0$ because $d(dr) = d^2 r = 0$. Combining the above expression defining Ω with (5.3.13) yields that

$$\Omega = \begin{pmatrix} \dfrac{r_s\, dt + 2ir^2\cos\theta\, d\phi}{4r^2} \\ \dfrac{i\sqrt{1 - r_s/r}}{2}\sin\theta\, d\phi \\ \dfrac{i\sqrt{1 - r_s/r}}{2}\, d\theta \\ 0 \end{pmatrix}. \tag{5.3.14}$$

Next, define the curvature quaternion Θ to be the 2-form $\Theta = d\Omega - \Omega \wedge \Omega$, where $d\Omega$ denotes the exterior derivative of Ω (see section 1.1 for particulars). Note that this particular curvature quantity can, in some ways, be reconciled with the previously-mentioned Riemannian curvature tensor R_{ijmn}, as

$$\Theta = i \sum_{m,n=1}^{4} \begin{pmatrix} R_{23mn} \\ R_{31mn} \\ R_{12mn} \\ 0 \end{pmatrix} \sigma_m \wedge \sigma_n + \sum_{m,n=1}^{4} \begin{pmatrix} R_{14mn} \\ R_{24mn} \\ R_{34mn} \\ 0 \end{pmatrix} \sigma_m \wedge \sigma_n.$$

Given the expressions for $d\sigma$ and Ω in (5.3.13) and (5.3.14) above, Θ for the Schwarzschild metric can be easily derived. In particular,

$$\Theta = \begin{pmatrix} -\dfrac{r_s\, drdt + ir^2 r_s \sin\theta\, d\theta d\phi}{2r^3} \\ -\dfrac{(r_s^2 - rrs)\, d\theta dt + irr_s \sin\theta\, drd\phi}{4r^3\sqrt{1 - r_s/r}} \\ -\dfrac{(r - rr_s)^2\, \sin\theta\, d\phi dt - irr_s\, drd\theta}{4r^3\sqrt{1 - r_s/r}} \\ 0 \end{pmatrix}. \tag{5.3.15}$$

Writing Θ as the sum $\Theta = \sum_{m,n} q_{\Omega mn} dx_m dx_n$ of differential 2-forms with quaternion coefficients demonstrates the precise geometrical meaning of Θ. In particular, the $q_{\Omega mn}$ represent rotation quaternions signifying the amount of rotational displacement a vector $\mathbf{v}$ undergoes when translated in a parallel manner along the rectangle $dx_m dx_n$. To better illustrate how this quantity highlights properties of the underlying manifold, scale each of the 2-forms above so that they have unit distance and write the resulting form Θ' as

$$\Theta' = \sum_{m,n} q'_{\Omega mn} \sigma_m \wedge \sigma_n.$$

Doing so ensures that displacement of vectors $\mathbf{v}$ now takes place along rectangles of uniform size, whereby the underlying geometrical gems of the host manifold—properties like symmetry, for example—become more evident. Performing this rescaling to Θ from (5.3.15) above yields

$$\Theta' = \frac{ir_s}{4r^3} \begin{pmatrix} 2(drdt - d\theta d\Phi) \\ -(d\theta dt - drd\phi) \\ -(d\phi dt - drd\theta) \\ 0 \end{pmatrix}.$$

5.4 Miscellaneous Results

The results detailed in sections 5.3.1 and 5.3.2 serve only as the tip of the expansive iceberg that is a quaternionic perspective of the physical universe. Many fundamental results in both classical and modern physics can be adapted in such a way as to fit the framework of quaternions and quaternion-valued functions. Because of the incalculable breadth of this merger in perspective, detailing any significant portion thereof is an exercise in futility. For that reason, this section will be centered on mentioning key results and directing the reader to literature that's more focused and specific.

In section 5.3.1, it was shown that the Einstein Field Equations can be adapted to the framework of quaternions and quaternionic functions. As mentioned in the introduction, the EFE can, for all intents and purposes, be considered the emergence of the entire field of General Relativity. It stands to reason, then, that any number of other keynote, fundamental results in physics can be adapted in a similar manner: For that reason, the remainder of this treatise will be focused on outlining such similarities.

Consider, first, the case of *Maxwell's Equations*. In the classical sense, Maxwell's equations are a set of partial differential equations largely considered to be the foundation of electrodynamics, optics, and electromagnetism. As it stands, the term "Maxwell's Equations" is used to describe a vast number of vector field results pertaining to electricity and magnetism, including but not limited to Gauss's Law, the Maxwell-Faraday equation, and Ampére's circuital law. As seen above, integration and differentiation along vector fields suggests strongly that a quaternionic perspective may well be adopted, and as the authors of [32] show, that's precisely the case. In addition, Christianto and Smarandache are able to use the freshly-derived quaternionic perspective to derive new results, more precisely, the potential for expressing momentum and energy (and, therefore, mass) as quaternion numbers. The authors conclude by hinting that further advances in that line of thought are to be pursued.

Related to Maxwell's Electromagnetism Theory is a system of generalized Maxwell equations applicable to the framework of relativistic spacetime. The so-called *Maxwell Equations in curved spacetime* are a system of highly-generalized expressions derived to expand the perspective of Maxwell's equations into the realm of electromagnetic fields in curved spacetime. Upon investigation, the derivations of this new set of results is done from a generalized tensorial perspective, the adaptation of which to a Quaternion-space (Q-space) viewpoint can be done using an algorithmic redefinition similar to the one used in the case of the EFE. A commonly-adopted standpoint regarding spacetime seems to be that the natural utilization of quaternions to describe spatial rotations makes them an excellent piece of machinery for the study of General Relativity.

Besides the overwhelming generalizations to spacetime and its related frameworks, many significant, specific results in that field conform naturally to a Q-space viewpoint. One family of results come intrinsically from the EFE as a cluster of solutions (i.e., metrics) in spacetime. In addition to the Schwarzschild metric described in section 5.3.2, several other (generally more complex) solutions to the EFE exist which describe interactions in spacetime in the presence of various other parameters. The Schwarzschild metric describes the behavior of spacetime in the presence of uncharged, non-rotating bodies: Other metrics describe the behavior of spacetime in the presence of bodies that are non-rotating and charged (the Reissner-Nordström metric), rotating and uncharged (the Kerr metric), and rotating and charged (the Kerr-Newman metric). Though seemingly different from the case shown in section 5.3.2, each of the above mentioned metrics are presented as solutions to the EFE in standard spherical coordinates. For that reason, the procedure utilized in section 5.3.2 can be copied nearly verbatim so as to present a wide array of spacetime metrics in terms of quaternion functions and Q-space numbers.

Besides the aforementioned results, there are many broad-reaching attempts to completely characterize the behavior of the physical universe from the standpoint of quaternion calculus. The previously-referenced magnum opus [28] by Girard is one such work—a work that attacks a manifold array of physical applications (both physics and otherwise) from a quaternion perspective. Perhaps even more impressive is the colossal collection [33] by Doug Sweester. In that work, Sweester presents a self-contained work that highlights both the mathematical beauty of the quaternion system, as well as an extensive array—in both breadth and depth—of physical applications. Beginning with Newton's laws of classical mechanics and blazing a trail that winds through special relativity, electromagnetism, quantum mechanics, and gravitation, Sweester's goal of making a quaternion viewpoint accessible and appreciated to the masses does nothing if not illustrate the true power of the framework of quaternionic analysis.

6 References

[1] Md. Shah Alam, *Comparative Study of Quaternions and Mixed Numbers*, "Journal of Theoretics Vol. 3-6", Journal of Theoretics Inc., 2001.

[2] Tony Sudbury, *Quaternionic Analysis*,"Mathematical Proceedings of the Cambridge Philosophical Society Vol. 85", Cambridge University Press, 1979.

[3] Henri Cartan, *Elementary theory of analytic functions of one or several complex variables*, Addison-Wesley Publishing Company, 1963.

[4] C.G. Cullen, *An integral theorem for analytic intrinsic functions on quaternions*, "Duke Mathematical Journal Vol. 32 (1965)," pp 139-148.

[5] C.A. Deavours, *The Quaternion Calculus*, "American Mathematics Monthly Vol. 80, No. 9 (1973)," pp 995-1008. Accessed online via: http://www.jstor.org/stable/2318774.

[6] Thi Ngoc Ha Vu, *Helmholtz Operator in Quaternionic Analysis*, 2005. Accessed online via: http://www.diss.fu-berlin.de/2005/61/.

[7] Louis Brand, *The Roots of a Quaternion*, "The American Mathematical Monthly Vol. 49, No. 8 (Oct. 1942)," pp 519-520. Accessed online via: http://www.jstor.org/stable/2302858.

[8] Simo Särkkä, *Notes on Quaternions*, Technical Report for Aalto University, Espoo, Finland, Posted June 28, 2007. Accessed online via: http://www.lce.hut.fi/ ssarkka/pub/quat.pdf.

[9] David Eberly, *Quaternion Algebra and Calculus*, Last Modified August 18, 2010. Accessed online via: http://www.lce.hut.fi/ ssarkka/pub/quat.pdf.

[10] Giorgio Balzarotti and Gianluca Marsiglia, Forum post dated January 28, 2005. Accessed online via: http://greenspun.com/bboard/q-and-a-fetch-msg.tcl?msg_id=00ARhE.

[11] R.S. Palais, *The Classification of Real Division Algebras*, "The American Mathematical Monthly, Vol. 75, No. 4 (Apr., 1968)," pp 366-368, Mathematical Association of America.

[12] V.C.A. Ferraro, *On Functions of Quaternions*, "Proceedings of the Royal Irish Academy Section A: Mathematical and Physical Sciences Vol. 44 (1937-1938)," pp 101-108, The Royal Irish Academy. Accessed online via: http://www.jstor.org/stable/20490730.

[13] Stephen Bedding and Keith Briggs, *Iteration of Quaternion Functions*, "The American Mathematical Monthly, Vol. 103, No. 8 (Oct., 1996)," pp 654-664, Mathematical Association of America. Accessed online via: http://www.jstor.org/stable/2974877.

[14] Graziano Gentili and Daniele C. Struppa, *A new theory of regular functions of a quaternionic variable*, "Advances in Mathematics, Vol. 216 (2007)," pp 279-301, Elseiver Inc. Accessed online via: https://neo.math.unifi.it/users/gentili/lavoripdf/AdvancesCP.pdf.

[15] Alessandro Perotti, *A differential criterium for regularity of quaternionic functions*, preprint submitted to Elseiver Science (2003). Accessed online via: http://www.science.unitn.it/ perotti/regcr.pdf.

[16] S. DeLeo and P.P. Rotelli, *Quaternionic Analyticity*, "Applied Mathematics Letters, Vol. 16 (2003)," pp 1077-1081, Elseiver Inc. Accessed online via: www.ime.unicamp.br/ deleo/Pub/p35.pdf.

[17] Daniel Alayón-Solarz, *On Calculus with a Quaternionic Variable and its Characteristic Cauchy-Riemann Type Equations*, preprint submitted to the arXiv (2009). Accessed online via: arxiv.org/pdf/0903.2899.

[18] Omar Dzagnidze, *On the Differentiability of Quaternion Functions*, preprint submitted to the arXiv (2012). Accessed online via: arxiv.org/pdf/1203.5619.

[19] Guy Laville and Ivan Ramadanoff, *Holomorphic Cliffordian Functions*, preprint submitted to the arXiv (2005). Accessed online via: http://arxiv.org/pdf/math/0502066.pdf.

[20] E.C. Titchmarsh, *The Theory of Functions, Second Edition*, Oxford University Press, London, England, 1939.

[21] R. Fueter, *Über vierfachperiodische Funktionen*, "Montsh. Math. Phys., Vol.48 (1939)," pp 161-169.

[22] Lars V. Ahlfors, *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable, Third Edition*, McGraw-Hill Inc., New York, 1979.

[23] Gareth A Jones and David Singerman, *Complex Functions: An Algebraic and Geometric Viewpoint*, Cambridge University Press, Cambridge, England, 1997.

[24] Graziano Gentili and Caterina Stoppato, *The Open Mapping Theorem for Regular Quaternionic Functions*, preprint submitted to the ArXiV (2008). Accessed online via: http://arxiv.org/pdf/0802.3861.pdf.

[25] Graziano Gentilli, Caterina Stoppato, and Daniele C. Struppa, *Regular Functions of a Quaternionic Variable*, Springer, 2013.

[26] George Raetz, *Quaternions and General Relativity*, Revised edition (May 20,2000), National Cheng Kung University. Available online at: http://ckw.phys.ncku.edu.tw/public/pub/WebSources/Quarternions/www.pcisys.net/_bestwork.1/index-2.html.

[27] Charles W. Misner, Kip S. Thorne, and John Archibald. *Gravitation*, W.H. Freeman Publishers, San Francisco, CA, 1973.

[28] P.R. Girard, *The Quaternion Group and Modern Physics*, "European Journal of Physics, Vol. 5, No. 25 (1984)," pp 25-32. Accessed online at: http://davidsenouf.com/wp-content/uploads/2012/01/The-quaternion-group-and-modern-physics-Girard.pdf.

[29] L. Silberstein, *Quaternionic Forms of Relativity*, "Philosophical Magazine Series 6, Vol. 23, No. 137 (May 1912)," pp 790-809. Accessed online via: http://www.theworld.com/ sweetser/quaternions/ps/Silberstein-Relativity.pdf.

[30] Mark Ashdown, *Imaginary Numbers are not Real - the Geometric Algebra of Spacetime*, Published February 9, 1993. Accessed online via: http://www.mrao.cam.ac.uk/ clifford/introduction/intro/intro.html.

[31] M. Sachs, *On Factorization of Einstein's Formalism into a Pair of Quaternion Field Equations*, International Atomic Energy Agency, International Center for Theoretical Physics, July 1966.

[32] Vic Chrisitianto and Florentin Smarandachey, *A Derivation of Maxwell Equations in Quaternion Space*, "Progress in Physics, Vol. 2 (April 2010)," pp 23-27. Accessed online via: http://www.ptep-online.com/index_files/2010/PP-21-06.PDF.

[33] Douglas B. Sweester, *Doing Physics with Quaternions*, 2005. Accessed online via: http://www.theworld.com/ sweetser/quaternions/ps/book.pdf.

[34] Reyer Sjamaar, *Manifolds and Differential Forms*, Department of Mathematics, Cornell University, Ithaca, New York, 2006. Accessed online via: http://www.math.cornell.edu/ sjamaar/papers/manifold.pdf.

[35] Michael Spivak, *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*, Addison-Wesley Publishing Company, 1965.

[36] Elias M. Stein and Rami Shakarchi, *Complex Analysis*, "Princeton Lectures in Analysis II", Princeton University Press, 2003.

Appendix 1: Preliminaries

Contained in this section is a list of the main topics necessary to deduce a thorough understanding of the material presented in section ??. While the goal is for this paper to be thorough and self-contained, it's worth noting that a completely-thorough exposition of all the topics regularly employed in the study of quaternionic analysis is simply infeasible. Therefore, in addition to the topics listed in this section, it's assumed that the reader has a background that includes at least fundamental understanding of areas such as real analysis and basic group, ring, and field theory.

1.1 Differential Geometry and Exterior Calculus

Note that f is said to be a "quaternion-valued function of a quaternion variable" whenever $f : \mathbb{H} \to \mathbb{H}$. In particular, then, because $\mathbb{H}$ is algebraically isomorphic to four-dimensional Euclidean space $\mathbb{R}^4$, any analysis done on such functions must be done in the language of "calculus on manifolds". Thus, it's important to have at least a basic understanding of the language of Differential Geometry and of the Exterior Calculus utilized therein in order to understand analytic properties of such functions. It should be noted that there are many exceptional guides to manifold calculus. This overview will largely follow presentations given in [34] and [35].

1.1.1 Multi-indices and k-forms

To begin, note that a *multi-index of degree k on $\mathbb{R}^4$* is defined to be a vector $I = (i_1, \ldots, i_k)$ consisting of k integer entries ranging between 1 and 4. As such, given a collection of smooth functions f_I, multi-indices can be used to denote sums of the form

$$\alpha = \sum_I f_I \, dx_I, \tag{A1}$$

where the f_I are called *coefficients* of α and $dx_I = dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}$ is called the *differential* of α. An expression with the form in (A1) is called a *k-form* in $\mathbb{R}^4$. Without loss of generality, the form in (A1) still makes sense if the multi-index I is assumed to be *increasing* in the sense that $i_j < i_k$ for all $j < k$.

Traditionally, the term dx_i is thought of as "the infinitesimal increment in the variable x_i," from which it follows that the expression dx_I represents the oriented volume of an infinitesimal k-dimensional rectangular block with sides $dx_{i_1}, \ldots, dx_{i_k}$ [34]. In particular, then, the wedge product $\wedge$ denotes the orientation-preserving product which satisfies the anticommutativity property, written here with respect to $\mathbb{R}^n$:

$$dx_1 \wedge \cdots dx_j \wedge dx_i \wedge \cdots dx_n = -dx_1 \wedge \cdots dx_i \wedge dx_j \wedge \cdots dx_n. \tag{A2}$$

Note that equation (A2) immediately implies that $dx_i \wedge dx_i = 0$ for all i. Thus, for all $k > 4$, every k-form on $\mathbb{R}^4$ is identically zero. This fact yields that all multi-indices on $\mathbb{R}^4$ can be assumed to have the form $I = (i_1, i_2, i_3, i_4)$ without loss of generality, and that each i_k can be assumed to satisfy the increasing condition $1 \le i_1 < i_2 < i_3 < i_4 \le 4$. Some additional properties of the wedge product as noted in [35] are as follows:

$$\begin{aligned}(dx_i + dx_j) \wedge dx_k &= dx_i dx_k \wedge dx_j dx_k \\ dx_i \wedge (dx_j + dx_k) &= dx_i \wedge dx_j + dx_i \wedge dx_k \\ \alpha dx_i \wedge dx_j &= dx_i \wedge \alpha dx_j = \alpha(dx_i \wedge dx_j) \\ (dx_i \wedge dx_j) \wedge dx_k &= dx_i \wedge (dx_j \wedge dx_k).\end{aligned}$$

Before defining the exterior derivative, it's worth noting k-forms on $\mathbb{R}^n$ can be combined via multiplication for general n. Indeed, if $\alpha = \sum_I f_I dx_I$ and $\beta = \sum_J f_J dx_J$ are k- and ℓ-forms, respectively, on $\mathbb{R}^n$, then their product $\alpha\beta$ is a $(k+\ell)$-form on $\mathbb{R}^n$ of the form

$$\alpha\beta = \sum_{I,J} f_I g_J dx_I dx_J.$$

It's easily shown that the reverse-product $\beta\alpha$ has the form $\beta\alpha = (-1)^{k\ell}\alpha\beta$, a fact that partially stems from the alternating property of the wedge products dx_I and dx_J.

1.1.2 Exterior Derivatives

In a k-form $\alpha = \sum_I f_I dx_I$, the coefficient functions f_I are assumed to be smooth, i.e., differentiable. In particular, then, it follows that a sort of derivative operator called "the exterior derivative" can be applied to α. The concept of exterior derivative was developed by French mathematician Élie Cartan around 1900 (see [34]) and is described briefly below.

For a smooth function (i.e., a 0-form) $f : \mathbb{R}^n \to \mathbb{R}$, define df to be the 1-form

$$\begin{aligned} df &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \qquad \text{(A3)} \\ &= \sum_{i=1}^{n} D_i f dx_i, \text{ where } D_i f = \frac{\partial f}{\partial x_i}.\end{aligned}$$

In a natural way, equation (A3) can be extended so as to compute $d\alpha$ for an arbitrary k-form $\alpha = \sum_I f_I dx_I$ simply by defining

$$d\alpha = \sum_I df_I dx_I. \qquad \text{(A4)}$$

The operator d in (A4) is a first-order partial differential operator known as *the exterior derivative* and it assigns to every k-form α an associated $(k+1)$-form $d\alpha$ by way of the sum described above. One useful property of the exterior derivative d is that it's the unique $\mathbb{R}$-linear map assigning $(k+1)$-forms to k-forms subject to the following properties:

1. df satisfies equation (A3) for smooth 0-forms f
2. $d(df) = 0$ for any smooth function f
3. For any k-form α, $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k(\alpha \wedge d\beta)$ for all forms β.

Proofs of properties 2 and 3 above can be found in [34] and [35] in varying amounts of detail.

Note that while the concept of the exterior derivative d can be abstracted and generalized to much larger degrees (see [35] for such details), the exposition above will be largely sufficient for understanding the topics written about in this paper.

1.1.3 Integration

The topic of integration on general manifolds is at the center of a colossal amount of literature, so much so that any meaningful exposition easily falls far beyond the scope of this particular manuscript. For that reason, the topics chosen here are, at best, sparse: They are chosen equally from [34] and [35], though these and several more advanced topics are also investigated in [6].

Given an arbitrary k-form $\alpha = \sum_I f_I dx_I$ defined on some open set V of $\mathbb{R}^m$ with coordinates identified by $y_1, \ldots, y_m$, it's possible to derive a "new" k-form on $\mathbb{R}^m$ by replacing the "old" variables y_j with "new" variables $x_1, \ldots, x_n$. In particular, suppose that for each $j = 1, \ldots, m$, the variable y_j is given in terms of $x_i, \ldots, x_n$ by way of functions

$$y_j = \phi_j(x_1, \ldots, x_n).$$

It follows, then, that the *pullback* of α along ϕ is a k-form $\phi^*\alpha$ obtained by substituting the functions ϕ_j into the formula for α for all j. More explicitly, the k-form $\phi^*\alpha$ has the form

$$\phi^*\alpha = \sum_I (\phi^* f_I)(\phi^* dy_I), \tag{A5}$$

where, for $I = (i_1, \ldots, i_k)$, the quantities in (A5) have the forms

$$\begin{aligned} \phi^* f_I &= f_I \circ \phi, \text{ and} \\ \phi^* dy_I &= \phi^*(dy_{i_1} \wedge \cdots \wedge dy_{i_k}) = d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k}. \end{aligned}$$

Mathematically, then, the idea of the pullback is to "convert" a map $\phi : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^m$ into a map $\phi^* : \Lambda^k(V) \to \Lambda^k(U)$, where $\Lambda^k(X)$ denotes the set of all k-forms on the space X. This idea is fundamental in defining integration in the general sense.

By way of defining the pullback operator of a k-form, one can define the integration of arbitrary k-forms over k-dimensional parameterized regions. More precisely, let U be an open subset of $\mathbb{R}^n$ for general n, let α be a k-form defined on U, and define a region M to be a rectangular block in $\mathbb{R}^k$ of the form $M = [a_1, b_1] \times \cdots \times [a_k, b_k]$. A smooth map $c : M \to U$ is called a *k-dimensional parameterized path*, and the pullback $c^*\alpha$ is a k-form on M. By arbitrarily denoting $c^*\alpha$ so that

$$c^*\alpha = g(t)\, dt_1 \wedge \cdots dt_k \text{ for some function } g : M \to \mathbb{R},$$

one can define the *integral of α over c* to be the value given by

$$\begin{aligned} \int_c \alpha &= \int_M c^*\alpha \\ &= \int_{a_k}^{b_k} \cdots \int_{a_2}^{b_2} \int_{a_1}^{b_1} g(t) dt_1 \wedge \cdots \wedge dt_k. \end{aligned}$$

This idea can be generalized one step farther by considering integration over *chains*. To do so, one first considers *k-cubes*, which are smooth maps $c : [0,1]^k \to U$ that a priori may or may not have self-intersections. From here, a *k-chain* in U is defined to be a formal linear combination

$$c = a_1 c_1 + \cdots + a_n c_n$$

of k-cubes in U (where $a_i \in \mathbb{R}$ are real coefficients and c_i are k-cubes) and integration of an arbitrary k-form α over such a chain is defined by the equation

$$\int_c \alpha = \sum_{i=1}^{n} a_i \int_{c_i} \alpha.$$

Perhaps the biggest accomplishment of generalized integration is the theorem of Stokes, from which many other fundamental results and so-called "classical theorems" can be deduced. Here are some additional definitions related to chains which are fundamental to the statement of Stokes' theorem:

Definition 1.1. The **boundary** ∂c of a k-chain $c = \sum_i a_i c_i$ is defined to be the corresponding linear combination of the boundaries of its k-cubes, namely $\partial c = \sum_i a_i \partial c_i$. In particular, ∂ is a linear map sending k-chains to $(k-1)$-chains.

Definition 1.2. A k-cube c_i is **degenerate** if $c_i(t_1, \ldots, t_k)$ is independent of t_j for some j. A k-chain c is said to be **degenerate** if it is the sum of degenerate k-cubes.

Definition 1.3. A k-chain c is said to be **a cycle** if ∂c is degenerate as a $(k-1)$-chain. Moreover, a k-chain c is defined to be a **boundary** if $c = \partial b + c'$ for some $(k+1)$-chain b and some degenerate k-chain c'.

This section is concluded by stating Stokes' Theorem, a theorem from which many of the keynote results and "classical theorems" in vector calculus can be deduced. For brevity, the proof is withheld, though the reader is redirected to meaningful proofs in other literature.

Theorem 1.1 (Stoke's Theorem)**.** Let α be a $(k-1)$-form on an open subset U of $\mathbb{R}^n$ and let c be a k-chain in U. Then

$$\int_c d\alpha = \int_{\partial c} \alpha.$$

Proof. This is proven in both [34] and [35] and in both instances, the proof is enlightening. [34] provides a more computational proof, while [35] proves the statement in a somewhat more general setting. □

1.2 Complex Analysis Primer

The purpose of this section is to recall some of the fundamental ideas whose quaternion-valued analogues are presented in section **??**. The outline followed here largely comes from [36]. In this section, the group $\mathbb{C}$ denotes the usual complex numbers, $\mathbb{C} = \{a + bi \,|\, a, b \in \mathbb{R}, i^2 = -1\}$, and all functions f are assumed to be complex-valued functions of a complex variable, i.e., $f : \mathbb{C} \to \mathbb{C}$.

The function f is said to be *complex-differentiable*[20] at $z_0 \in \mathbb{C}$ provided that the limit

$$\lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

taken along all possible paths $h \to 0$ exists as a number $\lambda \in \mathbb{C}$. In this case, $f'(z_0) = \lambda$ is called the *derivative* of f at z_0 and all the "standard derivative rules" can be easily derived. In the case where f is complex-differentiable at every point z_0 in an open neighborhood U of $\mathbb{C}$, f is said to be *holomorphic on* U. Also related to differentiability and holomorphicity, a function f is said to be *complex-analytic* on U if f can be given as a convergent power series of the form $f(z) = \sum_{n=1}^{\infty} a_n (z - z_0)^n$ about every point $z_0 \in U$. Note that every holomorphic function is complex-analytic.

One important test for analyticity of a complex-valued function f is the satisfaction by f of the so-called *Cauchy-Riemann Equations*. More precisely, if $f = u + iv$ is a complex-valued function defined in an open set U and if u and v are continuously differentiable and satisfy the equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}, \tag{A1}$$

then f is holomorphic on U. The equations in (A1) will be important in the discussion of quaternion-valued functions of a quaternion variable, so much so that the existence of an alternate formulation is also worthwhile. Recall that the *Wirtinger Derivatives* give partial derivatives of $z, \overline{z}$ with respect to those of x, y under the identities

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \text{ and } \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

[20]When at no risk of confusion, the adjective "complex-" will be dropped for brevity.

Under these operators, the equations in (A1) can be condensed to a single equation of Wirtinger form, namely

$$\frac{\partial f}{\partial \bar{z}} = 0.$$

Finally, a related notion that's also relevant in the study of functions $f : \mathbb{H} \to \mathbb{H}$ is the notion of harmonicity: A function f is said to be *harmonic* if $\Delta f = 0$, where $\Delta f = \partial^2 f/\partial x^2 + \partial^2 f/\partial y^2$.

One technique that's fundamental in the study of complex-valued functions of a complex variable is that of forming *path integrals*. Indeed, if $\gamma : [a, b] \to \mathbb{C}$ is a smooth curve in $\mathbb{C}$ and if f is a continuous function on γ, then the integral of f along γ is given by

$$\int_\gamma f(z)\, dz = \int_a^b f(\gamma(t))\gamma'(t)\, dt. \tag{A2}$$

The result in (A2) can be extended to the case where γ is instead piecewise smooth, at which point the integral on the right hand side of (A2) is simply replaced by the sums of the integrals of f along the smooth components of γ. A path $\gamma : [a, b] \to \mathbb{C}$ is said to be *closed* if $\gamma(a) = \gamma(b)$, and a function F holomorphic in a region U is said to be a *primitive* of f if $F'(z) = f(z)$ for all $z \in U$. Some fundamental results can now be stated.

Theorem 1.2 (Goursat's Theorem)**.** If U is an open set in $\mathbb{C}$ and if $\Delta \subset U$ is a triangle whose interior is also contained in U, then

$$\int_\Delta f(z)\, dz = 0$$

whenever f is holomorphic in U.

Remark 1.1. From a complex analytic perspective, the foremost significance of Goursat's Theorem is its use in the proof of Cauchy's Theorem. Hidden subtlety within the proof of the theorem, however, is a method of dissecting a region (Δ, in this case) into smaller regions Δ_i^k in an iterative fashion so that the diameter of Δ_i^n tends to 0 for sufficiently large i. This fundamental idea will be utilized in the framework of quaternion-valued theory when proving the analogous Cauchy-Feuter Integral Formula.

Theorem 1.3 (Cauchy's Theorem)**.** If U is an open, simply-connected region of $\mathbb{C}$ and if $f : U \to \mathbb{C}$ is holomorphic, then

$$\int_\gamma f(z)\, dz = 0$$

for any closed curve γ in that disc.

Theorem 1.4 (Cauchy's Integral Formula)**.** If f is holomorphic in an open set U, then f has infinitely many (complex) derivatives in U and, if $C \subset U$ is a circle whose interior is contained in U, then

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_C \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, d\zeta$$

for all z in the interior of C.

Theorem 1.5 (Liouville's Theorem)**.** If f is bounded and is holomorphic on all of $\mathbb{C}$, then f is constant.

Theorem 1.6 (Open Mapping Theorem)**.** If f is holomorphic and non-constant in a region $U \subset \mathbb{C}$, then f is an open map, i.e. $f(\mathcal{O}) \subset \mathbb{C}$ is open for all open sets $\mathcal{O} \subset \mathbb{C}$.

Theorem 1.7 (Maximum Modulus Principle)**.** If f is a non-constant holomorphic function in a region U, then f cannot attain a maximum in U. More precisely, $|f(z)|$ attains its maximum on the boundary ∂U of U and not at any interior point of U.

Appendix 2: Notation

Here's an encyclopedic list of notation used throughout this paper.

$\mathbb{H}$ - Group of Hamilton Quaternions

$\mathbb{R}$ - The real numbers

$\mathbb{C}$ - The complex numbers

$f : \mathbb{H} \to \mathbb{H}$ - A quaternion-valued function of a quaternion variable.

$\mathbb{F}$ - An arbitrary field.

D - Unless otherwise stated, a division algebra over a field $\mathbb{F}$.

$\| \cdot \|_D$ - A norm over D, provided D is a normed vector space.

i, j, k - Imaginary basis elements for $\mathbb{H}$

$\|\cdot\|_{\mathbb{H}}$ - The quaternionic norm function defined as $\|(a, b, c, d)\|_{\mathbb{H}} = \sqrt{a^2 + b^2 + c^2 + d^2}$

$\overline{q}$ - The quaternionic conjugate of $q \in \mathbb{H}$

$\operatorname{Re} q$ - The real-part $a \in \mathbb{R}$ of $q = (a, b, c, d) \in \mathbb{H}$

$\operatorname{Im} q$ - The "imaginary part" $bi + cj + dk$ of $q = (a, b, c, d) \in \mathbb{H}$

q^{-1} - The inverse of a nonzero quaternion q.

$\mathbb{O}$ - The Octonians.

df - Differential 1-form for $f : \mathbb{H} \to \mathbb{H}$ satisfying $df = dq(df/dq)$.

$\partial f/\partial x_j$ - The jth partial derivative of f.

$\overline{\partial}_\ell$,$\overline{\partial}_r$ - The Left/Right Cauchy-Riemann-Fueter operators.

∂_C - The Cullen differential operator.

∂_ℓ, ∂_r - The Cauchy-Riemann-Fueter "complement" operators.

Δf - The gradient of f.

$R_F(\mathbb{H})$ - The space of Fueter regular functions $f : \mathbb{H} \to \mathbb{H}$.

$R_C(\mathbb{H})$ - The space of Cullen regular functions.

$\mathrm{Ho}(\mathbb{H})$ - The space of holomorphic functions.

$\wedge$ - The wedge product of differential forms.

dq - For $q = t + ix + jy + kz$, $dq = dt + idx + jdy + kdz$.

Dq - The volume 3-form.

ε_{ijk} - The Levi-Civita symbol.

δ_{ij} - The Kronecker Delta.

$\partial\Omega$ - The boundary of a region Ω.

K° - The interior of a region K, defined as $\overline{K} \setminus \partial K$.

$v(C)$ - The volume of a region C in $\mathbb{R}^n$.

S - The unit sphere of purely imaginary Hamilton quaternions.

L_I - $\mathbb{R} + \mathbb{R}_I$ for some $I \in S$.

$\partial_I f$ - The I-derivative of f.

$\sum_{n=0}^{\infty} q^n a_n$ - A quaternionic power series.

Z_f - The zero set of f.

$B(\mathbf{0}, R)$ - The ball in $\mathbb{R}^n$ centered at $\mathbf{0} \in \mathbb{R}^n$ with radius $R > 0$.

$f * g$ - The regular product of f and g.

f^C - The regular conjugate of f.

f^S - The symmetrization of f.

f^{-*} - The reciprocal of f.

D_f - The degenerate set of f.

c - The speed of light, unless otherwise noted.

$(\Delta \mathbf{r})^2$, $(\Delta t)^2$ - The differences in space- and time-components of spacetime events.

s^2 - A spacetime interval

$\Delta\tau$, $\Delta\sigma$ - The proper time interval and proper distance, respectively.

η - The Minkowski (pseudo-)inner product.

M - A manifold, unless otherwise noted.

T_pM - The tangent space at a given point $\mathbf{p} \in M$.

$C^\infty(M, TM)$ - The collection of all vector fields on M.

$\nabla_X Y$ - Usually, an affine connection.

$[X, Y]$ - The Lie Bracket of X and Y.

$R(\mathbf{u}, \mathbf{v})w$ - The Riemannian curvature tensor of $\mathbf{w}$.

σ_i - The Pauli spin matrices.

s_i - The bases for the Pauli quaternions.

Norm(q), Mag(q) - The norm, magnitude of a (Pauli) quaternion q.

$U(q)$ - The unimodular operator acting on q.

q_i^* - The complex conjugate of a (bi-)quaternion.

$q^\dagger$ - The hermitian conjugate of a (bi-)quaternion.