1 Multiplying Quaternions the Easy Way

Multiplying two complex numbers a + b I and c + d I is straightforward.

```
(a,b)(c,d) = (ac - bd, ad + bc)
```

For two quaternions, b I and d I become the 3-vectors B and D, where B = x I + y J + z K and similarly for D. Multiplication of quaternions is like complex numbers, but with the addition of the cross product.

$$\left(a, \vec{B}\right)\left(c, \vec{D}\right) = \left(ac - \vec{B}.\vec{D}, \ a\vec{D} + \vec{B}\,c + \vec{B}\,x\,\vec{D}\right)$$

Note that the last term, the cross product, would change its sign if the order of multiplication were reversed (unlike all the other terms). That is why quaternions in general do not commute.

If a is the operator d/dt, and B is the del operator, or d/dx I + d/dy J + d/dz K (all partial derivatives), then these operators act on the scalar function c and the 3-vector function D in the following manner:

$$\left(\frac{d}{dt}, \vec{\nabla}\right)\left(c, \vec{D}\right) = \left(\frac{dc}{dt} - \vec{\nabla}.\vec{D}, \ \frac{d\vec{D}}{dt} + \vec{\nabla}\,c + \vec{\nabla}\,x\,\vec{D}\right)$$

This one quaternion contains the time derivatives of the scalar and 3-vector functions, along with the divergence, the gradient and the curl. Dense notation :-)

2 Scalars, Vectors, Tensors and All That

According to my math dictionary, a tensor is ...

"An abstract object having a definitely specified system of components in every coordinate system under consideration and such that, under transformation of coordinates, the components of the object undergoes a transformation of a certain nature."

To make this introduction less abstract, I will confine the discussion to the simplest tensors under rotational transformations. A rank-0 tensor is known as a scalar. It does not change at all under a rotation. It contains exactly one number, never more or less. There is a zero index for a scalar. A rank-1 tensor is a vector. A vector does change under rotation. Vectors have one index which can run from 1 to the number of dimensions of the field, so there is no way to know a priori how many numbers (or operators, or ...) are in a vector. n-rank tensors have n indices. The number of numbers needed is the number of dimensions in the vector space raised by the rank. Symmetry can often simplify the number of numbers actually needed to describe a tensor.

There are a variety of important spin-offs of a standard vector. Dual vectors, when multiplied by its corresponding vector, generate a real number, by systematically multiplying each component from the dual vector and the vector together and summing the total. If the space a vector lives in is shrunk, a contravariant vector shrinks, but a covariant vector gets larger. A tangent vector is, well, tangent to a vector function.

Physics equations involve tensors of the same rank. There are scalar equations, polar vector equations, axial vector equations, and equations for higher rank tensors. Since the same rank tensors are on both sides, the identity is preserved under a rotational transformation. One could decide to arbitrarily combine tensor equations of different rank, and they would still be valid under the transformation.

There are ways to switch ranks. If there are two vectors and one wants a result that is a scalar, that requires the intervention of a metric to broker the transaction. This process in known as an inner tensor product or a contraction. The vectors in question must have the same number of dimensions. The metric defines how to form a scalar as the indices are examined one-by-one. Metrics in math can be anything, but nature imposes constraints on which ones are important in physics. An aside: mathematicians require the distance is non-negative, but physicists do not. I will be using the physics notion of a metric. In looking at events in spacetime (a 4-dimensional vector), the axioms of special relativity require the Minkowski metric, which is a 4x4 real matrix which has down the diagonal 1, -1, -1, -1 and zeros elsewhere. Some people prefer the signs to be flipped, but to be consistent with everything else on this site, I choose this convention. Another popular choice is the Euclidean metric, which is the same as an identity matrix. The result of general relativity for a spherically symmetric, non-rotating mass is the Schwarzschild metric, which has "non-one" terms down the diagonal, zeros elsewhere, and becomes the Minkowski metric in the limit of the mass going to zero or the radius going to infinity.

An outer tensor product is a way to increase the rank of tensors. The tensor product of two vectors will be a 2-rank tensor. A vector can be viewed as the tensor product of a set of basis vectors.

What Are Quaternions?

Quaternions could be viewed as the outer tensor product of a scalar and a 3-vector. Under rotation for an event in spacetime represented by a quaternion, time is unchanged, but the 3-vector for space would be rotated. The treatment of scalars is the same as above, but the notion of vectors is far more restrictive, as restrictive as the notion of scalars. Quaternions can only handle 3-vectors. To those familiar to playing with higher dimensions, this may appear too restrictive to be of interest. Yet physics on both the quantum and cosmological scales is confined to 3-spatial dimensions. Note that the infinite Hilbert spaces in quantum mechanics a function of the principle quantum number n, not the spatial dimensions. An infinite collection of quaternions of the form (En, Pn) could represent a quantum state. The Hilbert space is formed using the Euclidean product (q* q').

A dual quaternion is formed by taking the conjugate, because $q^* q = (t^2 + X.X, 0)$. A tangent quaternion is created by having an operator act on a quaternion-valued function

$$\left(\frac{\partial}{\partial t}, \vec{\nabla}\right)(f(q), \vec{F}(q)) = \left(\frac{\partial f}{\partial t} - \vec{\nabla}.\vec{F}, \frac{\partial \vec{F}}{\partial t} + \vec{\nabla}f + \vec{\nabla}x\vec{F}\right)$$

What would happen to these five terms if space were shrunk? The 3-vector F would get shrunk, as would the divisors in the Del operator, making functions acted on by Del get larger. The scalar terms are completely unaffected by shrinking space, because df/dt has nothing to shrink, and the Del and F cancel each other. The time derivative of the 3-vector is a contravariant vector, because F would get smaller. The gradient of the scalar field is a covariant vector, because of the work of the Del operator in the divisor makes it larger. The curl at first glance might appear as a draw, but it is a covariant vector capacity because of the right-angle nature of the cross product. Note that if time where to shrink exactly as much as space, nothing in the tangent quaternion would change.

A quaternion equation must generate the same collection of tensors on both sides. Consider the product of two events, q and q':

$$\left(t, \vec{X}\right)\left(t', \vec{X'}\right) = \left(t\,t' - \vec{X}.\vec{X'}, \ t\vec{X'} + \vec{X}t' + \vec{X}x\vec{X'}\right)$$

$$\text{scalars}: t, \ t', \ tt' - \vec{X}.\vec{X'}$$

$$\text{polar vectors}: \vec{X}, \ \vec{X'}, \ t\vec{X'} + \vec{X}t'$$

$$\text{axial vectors}: \vec{X}x\vec{X'}$$

Where is the axial vector for the left hand side? It is imbedded in the multiplication operation, honest :-)

$$\left(t', \vec{X'}\right)\left(t, \vec{X}\right) = \left(t't - \vec{X'}.\vec{X}, \ t'\vec{X} + \vec{X'}t + \vec{X'}x\vec{X}\right)$$
$$= \left(t\,t' - \vec{X}.\vec{X'}, \ t\vec{X'} + \vec{X}t' - \vec{X}x\vec{X'}\right)$$

The axial vector is the one that flips signs if the order is reversed.

Terms can continue to get more complicated. In a quaternion triple product, there will be terms of the form (XxX').X". This is called a pseudo-scalar, because it does not change under a rotation, but it will change signs under a reflection, due to the cross product. You can convince yourself of this by noting that the cross product involves the sine of an angle and the dot product involves the cosine of an angle. Neither of these will change under a rotation, and an even function times an odd function is odd. If the order of quaternion triple product is changed, this scalar will change signs for at each step in the permutation.

It has been my experience that any tensor in physics can be expressed using quaternions. Sometimes it takes a bit of effort, but it can be done.

Individual parts can be isolated if one chooses. Combinations of conjugation operators which flip the sign of a vector, and symmetric and antisymmetric products can isolate any particular term. Here are all the terms of the example from above

$$\left(t, \vec{X}\right)\left(t', \vec{X'}\right) = \left(t\,t' - \vec{X}.\vec{X'}, \ t\vec{X'} + \vec{X}t' + \vec{X}x\vec{X'}\right)$$

$$\text{scalars}: t = \frac{q + q^*}{2}, \ t' = \frac{q' + q'^*}{2}, tt' - \vec{X}.\vec{X'} = \frac{qq' + (qq')^*}{2}$$

$$\text{polar vectors}: \vec{X} = \frac{q - q^*}{2}, \ \vec{X'} = \frac{q' - q'^*}{2},$$

$$t\vec{X'} + \vec{X}t' = \frac{(qq' + (q'q)) - (qq' + (q'q))^*}{4}$$

$$\text{axial vectors}: \vec{X}x\vec{X'} = \frac{qq' - (q'q)}{2}$$

The metric for quaternions is imbedded in Hamilton's rule for the field.

$$\left(\vec{i}\right)^2 = \left(\vec{j}\right)^2 = \left(\vec{k}\right)^2 = \vec{i}\,\vec{j}\,\vec{k} = -1$$

This looks like a way to generate scalars from vectors, but it is more than that. It also says implicitly that i j = k, j k = i, and i, j, k must have inverses. This is an important observation, because it means that inner and outer tensor products can occur in the same operation. When two quaternions are multiplied together, a new scalar (inner tensor product) and vector (outer tensor product) are formed.

How can the metric be generalized for arbitrary transformations? The traditional approach would involve playing with Hamilton's rules for the field. I think that would be a mistake, since that rule involves the fundamental definition of a quaternion. Change the rule of what a quaternion is in one context and it will not be possible to compare it to a quaternion in another context. Instead, consider an arbitrary transformation T which takes q into q'

$$\mathtt{q} \rightarrow \mathtt{q'} = \mathtt{T\,q}$$

T is also a quaternion, in fact it is equal to q' q^-1. This is guaranteed to work locally, within neighborhoods of q and q'. There is no promise that it will work globally, that one T will work for any q. Under certain circumstances, T will work for any q. The important thing to know is that a transformation T necessarily exists because quaternions are a field. The two most important theories in physics, general relativity and the standard model, involve local transformations (but the technical definition of local transformation is different than the idea presented here because it involves groups).

This quaternion definition of a transformation creates an interesting relationship between the Minkowski and Euclidean metrics.

Let T = I, the identity matrix
$$\frac{\mathtt{I\,q\,I\,q} + (\mathtt{I\,q\,I\,q})^*}{2} = \left(\mathtt{t}^2 - \vec{\mathtt{x}}.\vec{\mathtt{x}},\ 0\right)$$
$$(\mathtt{I\,q})^*\mathtt{I\,q} = \left(\mathtt{t}^2 + \vec{\mathtt{x}}.\vec{\mathtt{x}},\ 0\right)$$

In order to change from wrist watch time (the interval in spacetime) to the norm of a Hilbert space does not require any change in the transformation quaternion, only a change in the multiplication step. Therefore a transformation which generates the Schwarzschild interval of general relativity should be easily portable to a Hilbert space, and that might be the start of a quantum theory of gravity.

So What Is the Difference?

I think it is subtle but significant. It goes back to something I learned in a graduate level class on the foundations of calculus. To make calculus rigorous requires that it is defined over a mathematical field. Physicists do this be saying that the scalars, vectors and tensors they work with are defined over the field of real or complex numbers.

What are the numbers used by nature? There are events, which consist of the scalar time and the 3-vector of space. There is mass, which is defined by the scalar energy and the 3-vector of momentum. There is the electromagnetic potential, which has a scalar field phi and a 3-vector potential A.

To do calculus with only information contained in events requires that a scalar and a 3-vector form a field. According to a theorem by Frobenius on finite dimensional fields, the only fields that fit are isomorphic to the quaternions (isomorphic is a sophisticated notion of equality, whose subtleties are appreciated only by people with a deep understanding of mathematics). To do calculus with a mass or an electromagnetic potential has an identical requirement and an identical solution. This is the logical foundation for doing physics with quaternions.

Can physics be done without quaternions? Of course it can! Events can be defined over the field of real numbers, and then the Minkowski metric and the Lorentz group can be deployed to get every result ever confirmed by experiment. Quantum mechanics can be defined using a Hilbert space defined over the field of complex numbers and return with every result measured to date.

Doing physics with quaternions is unnecessary, unless physics runs into a compatibility issue. Constraining general relativity and quantum mechanics to work within the same topological algebraic field may be the way to unite these two separately successful areas.

3 Inner and Outer Products of Quaternions

A good friend of mine has wondered what is means to multiply two quaternions together (this question was a hot topic in the nineteenth century). I care more about what multiplying two quaternions together can do. There are two basic ways to do this: just multiply one quaternion by another, or first take the transpose of one then multiply it with the other. Each of these products can be separated into two parts: a symmetric (inner product) and an antisymmetric (outer product) components. The symmetric component will remain unchanged by exchanging the places of the quaternions, while the antisymmetric component will change its sign. Together they add up to the product. In this section, both types of inner and outer products will be formed and then related to physics.

The Grassman Inner and Outer Products

There are two basic ways to multiply quaternions together. There is the direct approach.

$$\left(t, \vec{x}\right)\left(t', \vec{x'}\right) = \left(t\,t' - \vec{x}.\vec{x'}, t\vec{x'} + \vec{x}t' + \vec{x}x\vec{x'}\right)$$

I call this the Grassman product (I don't know if anyone else does, but I need a label). The inner product can also be called the symmetric product, because it does not change signs if the terms are reversed.

$$\mathbf{even}\left(\left(t, \vec{x}\right), \left(t', \vec{x'}\right)\right) \equiv$$

$$\equiv \frac{\left(t, \vec{x}\right)\left(t', \vec{x'}\right) + \left(t', \vec{x'}\right)\left(t, \vec{x}\right)}{2} = \left(t\,t' - \vec{x}.\vec{x'}, t\vec{x'} + \vec{x}t'\right)$$

I have defined the anticommutator (the bold curly braces) in a non-standard way, including a factor of two so I do not have to keep remembering to write it. The first term would be the Lorentz invariant interval if the two quaternions represented the same difference between two events in spacetime (i.e. t1=t2=delta t,...). The invariant interval plays a central role in special relativity. The vector terms are a frame-dependent, symmetric product of space with time and does not appear on the stage of physics, but is still a valid measurement.

The Grassman outer product is antisymmetric and is formed with a commutator.

$$\mathbf{odd}\left(\left(t, \vec{x}\right), \left(t', \vec{x'}\right)\right) \equiv$$

$$\equiv \frac{\left(t, \vec{x}\right)\left(t', \vec{x'}\right) - \left(t', \vec{x'}\right)\left(t, \vec{x}\right)}{2} = \left(0, \vec{x}x\vec{x'}\right)$$

This is the cross product defined for two 3-vectors. It is unchanged for quaternions.

The Euclidean Inner and Outer Products

Another important way to multiply a pair of quaternions involves first taking the transpose of one of the quaternions. For a real-valued matrix representation, this is equivalent to multiplication by the conjugate which involves flipping the sign of the 3-vector.

$$\left(t, \vec{x}\right)^{*}\left(t', \vec{x'}\right) = \left(t, -\vec{x}\right)\left(t', \vec{x'}\right)$$

$$= \left(t\,t' + \vec{x}.\vec{x'}, t\vec{x'} - \vec{x}t' - \vec{x}x\vec{x'}\right)$$

Form the Euclidean inner product.

$$\frac{\left(t, \vec{x}\right)^{*}\left(t', \vec{x'}\right) + \left(t', \vec{x'}\right)^{*}\left(t, \vec{x}\right)}{2} = \left(t\,t' + \vec{x}.\vec{x'}, \vec{0}\right)$$

The first term is the Euclidean norm if the two quaternions are the same (this was the reason for using the adjective "Euclidean"). The Euclidean inner product is also the standard definition of a dot product.

Form the Euclidean outer product.

$$\frac{\left(t,\vec{x}\right)^{*}\left(t',\vec{x'}\right) - \left(t',\vec{x'}\right)^{*}\left(t,\vec{x}\right)}{2} = \left(0, t\vec{x'} - \vec{x}t' - \vec{x}x\vec{x'}\right)$$

The first term is zero. The vector terms are an antisymmetric product of space with time and the negative of the cross product.

Implications

When multiplying vectors in physics, one normally only considers the Euclidean inner product, or dot product, and the Grassman outer product, or cross product. Yet, the Grassman inner product, because it naturally generates the invariant interval, appears to play a role in special relativity. What is interesting to speculate about is the role of the Euclidean outer product. It is possible that the antisymmetric, vector nature of the space/time product could be related to spin. Whatever the interpretation, the Grassman and Euclidean inner and outer products seem destine to do useful work in physics.

4 Quaternion Analysis

Complex numbers are a subfield of quaternions. My hypothesis is that complex analysis should be self-evident within the structure of quaternion analysis.

The challenge is to define the derivative in a non-singular way, so that a left derivative always equals a right derivative. If quaternions would only commute... Well, the scalar part of a quaternion does commute. If, in the limit, the differential element converged to a scalar, then it would commute. This idea can be defined precisely. All that is required is that the magnitude of the vector goes to zero faster than the scalar. This might initially appears as an unreasonable constraint. However, there is an important application in physics. Consider a set of quaternions that represent events in spacetime. If the magnitude of the 3-space vector is less than the time scalar, events are separated by a timelike interval. It requires a speed less than the speed of light to connect the events. This is true no matter what coordinate system is chosen.

Defining a Quaternion

A quaternion has 4 degrees of freedom, so it needs 4 real-valued variables to be defined:

$$q = (a_0, a_1, a_2, a_3)$$

Imagine we want to do a simple binary operation such as subtraction, without having to specify the coordinate system chosen. Subtraction will only work if the coordinate systems are the same, whether it is Cartesian, spherical or otherwise. Let e0, e1, e2, and e3 be the shared, but unspecified, basis. Now we can define the difference between two quaternion q and q' that is independent of the coordinate system used for the measurement.

$$dq =$$
$$q' - q = ((a_0' - a_0)e_0, (a_1' - a_1)e_1/3, (a_2' - a_2)e_2/3, (a_3' - a_3)e_3/3)$$

What is unusual about this definition are the factors of a third. They will be necessary later in order to define a holonomic equation later in this section. Hamilton gave each element parity with the others, a very reasonable approach. I have found that it is important to give the scalar and the sum of the 3-vector parity. Without this "scale" factor on the 3-vector, change in the scalar is not given its proper weight.

If dq is squared, the scalar part of the resulting quaternion forms a metric.

$$dq\char`\^2 = \left(da_0{}^2 e_0{}^2 + da_1{}^2 \frac{e_1{}^2}{9} + da_2{}^2 \frac{e_2{}^2}{9} + da_3{}^2 \frac{e_3{}^2}{9},\right.$$

$$\left. 2\, da_0 da_1 e_0 \frac{e_1}{3}, 2\, da_0 da_2 e_0 \frac{e_2}{3}, 2\, da_0 da_3 e_0 \frac{e_3}{3}\right)$$

In order to make a simple and direct connection to the Schwarzschild metric of general relativity, let e1, e2, and e3 form an independent, dimensionless, orthogonal basis for the 3-vector such that:

$$-\frac{1}{e_1{}^2} = -\frac{1}{e_2{}^2} = -\frac{1}{e_3{}^2} = e_0{}^2$$

This unusual relationship between the basis vectors is consistent with Hamilton's choice of 1, i, j, k if e0^2 = 1. For that case, calculate the square of dq:

$$dq^2 = \left(da_0{}^2 e_0{}^2 - \frac{da_1{}^2}{9e_0{}^2} - \frac{da_2{}^2}{9e_0{}^2} - \frac{da_3{}^2}{9e_0{}^2}, 2\, da_0 \frac{da_1}{3}, 2\, da_0 \frac{da_2}{3}, 2\, da_0 \frac{da_3}{3}\right)$$

The scalar part is known in physics as the Minkowski interval between two events in flat spacetime. If e0^2 does not equal one, then the metric would apply to a non-flat peacetime. A metric that has been measured experimentally is the Schwarzchild metric of general relativity. Set e0^2 = (1 - 2 GM/c^2 R), and calculate the square of dq:

$$dq^2 = \left(da_0{}^2\left(1 - \frac{2\,G\,M}{c^2\,R}\right) - \frac{dA.dA}{9\left(1 - \frac{2\,G\,M}{c^2\,R}\right)}, \, 2\,da_0\frac{da_1}{3}, \, 2\,da_0\frac{da_2}{3}, \, 2\,da_0\frac{da_3}{3}\right)$$

This is the Schwarzchild metric of general relativity. Notice that the 3-vector is unchanged (this may be a defining characteristic). There are very few opportunities for freedom in basic mathematical definitions. I have chosen this unusual relationships between the squares of the basis vectors to make a result from physics easy to express. Physics guides my choices in mathematical definitions :-)

An Automorphic Basis for Quaternion Analysis

A quaternion has 4 degrees of freedom. To completely specify a quaternion function, it must also have four degrees of freedom. Three other linearly-independent variables involving q can be defined using conjugates combined with rotations:

$$q^* = a_0 e_0 - a_1 e_1/3 - a_2 e_2/3 - a_3 e_3/3$$

$$q^* = -a_0 e_0 + a_1 e_1/3 - a_2 e_2/3 - a_3 e_3/3 = (e_1\, q\, e_1)^*$$

$$q^{*2} \equiv -a_0 e_0 - a_1 e_1/3 + a_2 e_2/3 - a_3 e_3/3 = (e_2\, q\, e_2)^*$$

The conjugate as it is usually defined (q^*) flips the sign of all but the scalar. The q*1 flips the signs of all but the e1 term, and q*2 all but the e2 term. The set q, q*, q*1, q*2 form the basis for quaternion analysis. The conjugate of a conjugate should give back the original quaternion.

$$(q^*)^* = q, \ (q^{*1})^{*1} = q, \ (q^{*2})^{*2} = q$$

Something subtle but perhaps directly related to spin happens looking at how the conjugates effect products:

$$(q\, q')^* = q'^*\, q^*$$

$$(q\, q')^{*1} = -q'^{*1}\, q^{*1}, \ (q\, q')^{*2} = -q'^{*2}\, q^{*2}$$

$$(q\, q'\, q\, q')^{*1} = q'^{*1}\, q^{*1} q'^{*1}\, q^{*1}$$

The conjugate applied to a product brings the result directly back to the reverse order of the elements. The first and second conjugates point things in exactly the opposite way. The property of going "half way around" is reminiscent of spin. A tighter link will need to be examined.

Future Timelike Derivative

Instead of the standard approach to quaternion analysis which focuses on left versus right derivatives, I concentrate on the ratio of scalars to 3-vectors. This is natural when thinking about the structure of Minkowski spacetime, where the ratio of the change in time to the change in 3-space defines five separate regions: timelike past, timelike future, lightlike past, lightlike future, and spacelike. There are no continuous Lorentz transformations to link these regions. Each region will require a separate definition of the derivative, and they will each have distinct properties. I will start with the simplest case, and look at a series of examples in detail.

Definition: The future timelike derivative:

Consider a covariant quaternion function f with a domain of H and a range of H. A future timelike derivative to be defined, the 3-vector must approach zero faster than the positive scalar. If this is not the case, then this definition cannot be used. Implementing these requirements involves two limit processes applied sequentially to a differential quaternion D. First the limit of the three vector is taken as it goes to zero, (D - D*)/2 -> 0. Second, the limit of the scalar is taken, (D + D*)/2 -> +0 (the plus zero indicates that it must be approached with a time greater than zero, in other words, from the future). The net effect of these two limit processes is that D->0.

$$\frac{\partial\, f(q, q^*, q^{*1}, q^{*2})}{\partial\, q} =$$

$$= \text{limit as } \left(d, \vec{0}\right) - > +0$$

$$\left(\text{limit as}\left(d, \vec{D}\right) - > \right.$$
$$\left. \left(d, \vec{0}\right)\left(f\left(q + \left(d, \vec{D}\right), q^*, q^{*1}, q^{*2}\right) - f(q, q^*, q^{*1}, q^{*2})\right)\left(d, \vec{D}\right)^{-1}\right)\right)$$

The definition is invariant under a passive transformation of the basis.

The 4 real variables a0, a1, a2, a3 can be represented by functions using the conjugates as a basis.

$$f(q, q^*, q^{*1}, q^{*2}) = a_0 = \frac{e_0\,(q + q^*)}{2}$$

$$f = a_1 = \frac{e_1\,(q + q^{*1})}{(-2/3)} = \frac{(q + q^{*1})\,e_1}{(-2/3)}$$

$$f = a_2 = \frac{e_2\,(q + q^{*2})}{(-2/3)} = \frac{(q + q^{*2})\,e_2}{(-2/3)}$$

$$f = a_3 = \frac{e_3\,(q + q^* + q^{*1} + q^{*2})}{(2/3)} = \frac{(q + q^* + q^{*1} + q^{*2})\,e_3}{(2/3)}$$

Begin with a simple example:

$$f(q, q^*, q^{*1}, q^{*2}) = a_0 = \frac{e_0\,(q + q^*)}{2}$$

$$\frac{\partial\, a_0}{\partial\, q} =$$

$$\frac{\partial\, a_0}{\partial\, q^*} = \lim\left(\lim\left(\left(e_0\left(\left(q + \left(d, \vec{D}\right) + q^*\right) - (q + q^*)\right)\right)\left(2\left(d, \vec{D}\right)\right)^{-1}\right)\right) = \frac{e_0}{2}$$

$$\frac{\partial\, a_0}{\partial\, q^{*1}} = \frac{\partial\, a_0}{\partial\, q^{*2}} = 0$$

The definition gives the expected result.

A simple approach to a trickier example:

$$f = a_1 = \frac{e_1\,(q + q^{*1})}{(-2/3)}$$

$$\frac{\partial\, a_1}{\partial\, q} = \frac{\partial\, a_1}{\partial\, q^{*1}} =$$

$$\lim\left(\lim\left(\left(e_1\left(\left(q + \left(d, \vec{D}\right) + q^{*1}\right) - (q + q^{*1})\right)\right)\left((-2/3)\left(d, \vec{D}\right)\right)^{-1}\right)\right) = -\frac{3\,e_1}{2}$$

$$\frac{\partial\, a_1}{\partial\, q^*} = \frac{\partial\, a_1}{\partial\, q^{*2}} = 0$$

So far, the fancy double limit process has been irrelevant for these identity functions, because the differential element has been eliminated. That changes with the following example, a tricky approach to the same result.

$$f(q, q^*, q^{*1}, q^{*2}) = a_1 = \frac{(q + q^{*1})\,e_1}{(-2/3)}$$

$$\frac{\partial\, a_1}{\partial\, q} = \frac{\partial\, a_1}{\partial\, q^{*1}} =$$

$$= \lim\left(\lim\left(\left(\left(q + \left(d, \vec{D}\right) + q^{*1}\right) - (q + q^{*1})\right)e_1\left((-2/3)\left(d, \vec{D}\right)\right)^{-1}\right)\right) =$$

$$= \lim\left(\lim\left(\left(d, \vec{D}\right)e_1\left((-2/3)\left(d, \vec{D}\right)\right)^{-1}\right)\right) =$$

$$= \lim\left(\left(d, \vec{0}\right)e_1\left((-2/3)\left(d, \vec{0}\right)\right)^{-1}\right) = -\frac{3\,e_1}{2}$$

Because the 3-vector goes to zero faster than the scalar for the differential element, after the first limit process, the remaining differential is a scalar so it commutes with any quaternion. This is what is required to dance around the e1 and lead to the cancellation.

The initial hypothesis was that complex analysis should be a self-evident subset of quaternion analysis. So this quaternion derivative should match up with the complex case, which is:

$$z = a + b\,i,\ b = (Z - Z^*)/2i$$

$$\frac{\partial b}{\partial z} = -\frac{i}{2} = -\frac{\partial b}{\partial z^*}$$

These are the same result up to two subedits. Quaternions have three imaginary axes, which creates the factor of three. The conjugate of a complex number is really doing the work of the first quaternion conjugate q*1 (which equals -z*), because z* flips the sign of the first 3-vector component, but no others.

The derivative of a quaternion applies equally well to polynomials.

$$\text{let } f = q^2$$

$$\frac{\partial f}{\partial q} = \lim\left(\lim\left(\left(\left(q + \left(d, \vec{D}\right)\right)^2 - q^2\right)\left(d, \vec{D}\right)^{-1}\right)\right) =$$

$$= \lim\left(\lim\left(\left(q^2 + q\left(d, \vec{D}\right) + \left(d, \vec{D}\right)q + \left(d, \vec{D}\right)^2 - q^2\right)\left(d, \vec{D}\right)^{-1}\right)\right) =$$

$$= \lim\left(\lim\left(q + \left(d, \vec{D}\right)q\left(d, \vec{D}\right)^{-1} + \left(d, \vec{D}\right)\right)\right) =$$

$$= \lim\left(2q + \left(d, \vec{0}\right)\right) = 2q$$

This is the expected result for this polynomial. It would be straightforward to show that all polynomials gave the expected results.

Mathematicians might be concerned by this result, because if the 3-vector D goes to -D nothing will change about the quaternion derivative. This is actually consistent with principles of special relativity. For timelike separated events, right and left depend on the inertial reference frame, so a timelike derivative should not depend on the direction of the 3-vector.

Analytic Functions

There are 4 types of quaternion derivatives and 4 component functions. The following table describes the 16 derivatives for this set

$\backslash$ ∂	a_0 e_0	a_1 e_1	a_2 e_2	a_3 e_3
$\frac{\partial}{\partial q}$	$\frac{e_0}{2}$	$-2/3$	$-2/3$	$\frac{e_3}{2/3}$
$\frac{\partial}{\partial q^*}$	$\frac{e_0}{2}$	0	0	$\frac{e_3}{2/3}$
$\frac{\partial}{\partial q^{*1}}$	0	$\frac{e_1}{-2/3}$	0	$\frac{e_3}{2/3}$
$\frac{\partial}{\partial q^{*2}}$	0	0	$\frac{e_2}{-2/3}$	$\frac{e_3}{2/3}$

This table will be used extensively to evaluate if a function is analytic using the chain rule. Let's see if the identity function $w = q$ is analytic.

$$\text{Let } w = q = a_0 e_0 + a_1 \frac{e_1}{3} + a_2 \frac{e_2}{3} + a_3 \frac{e_3}{3}$$

Use the chain rule to calculate the derivative will respect to each term:

$$\frac{\partial w}{\partial a_0}\frac{\partial a_0}{\partial q} = e_0 \frac{e_0}{2} = \frac{1}{2}$$

$$\frac{\partial w}{\partial a_1}\frac{\partial a_1}{\partial q} = \frac{e_1}{3}\frac{e_1}{(-2/3)} = \frac{1}{2}$$

$$\frac{\partial w}{\partial a_2}\frac{\partial a_2}{\partial q} = \frac{e_2}{3}\frac{e_2}{(-2/3)} = \frac{1}{2}$$

$$\frac{\partial w}{\partial a_3}\frac{\partial a_3}{\partial q} = \frac{e_3}{3}\frac{e_3}{(2/3)} = -\frac{1}{2}$$

Use combinations of these terms to calculate the four quaternion derivatives using the chain rule.

$$\frac{\partial w}{\partial q} =$$

$$\frac{\partial w}{\partial a_0}\frac{\partial a_0}{\partial q} + \frac{\partial w}{\partial a_1}\frac{\partial a_1}{\partial q} + \frac{\partial w}{\partial a_2}\frac{\partial a_2}{\partial q} + \frac{\partial w}{\partial a_3}\frac{\partial a_3}{\partial q} = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} - \frac{1}{2} = 1$$

$$\frac{\partial w}{\partial q^*} = \frac{\partial w}{\partial a_0}\frac{\partial a_0}{\partial q^*} + \frac{\partial w}{\partial a_3}\frac{\partial a_3}{\partial q^*} = \frac{1}{2} - \frac{1}{2} = 0$$

$$\frac{\partial w}{\partial q^{*1}} = \frac{\partial w}{\partial a_1}\frac{\partial a_1}{\partial q^{*1}} + \frac{\partial w}{\partial a_3}\frac{\partial a_3}{\partial q^{*1}} = \frac{1}{2} - \frac{1}{2} = 0$$

$$\frac{\partial w}{\partial q^{*2}} = \frac{\partial w}{\partial a_2}\frac{\partial a_2}{\partial q^{*2}} + \frac{\partial w}{\partial a_3}\frac{\partial a_3}{\partial q^{*2}} = \frac{1}{2} - \frac{1}{2} = 0$$

This has the derivatives expected if w=q is analytic in q.

Another test involves the Cauchy-Riemann equations. The presence of the three basis vectors changes things slightly.

$$\text{Let } u = a_0 e_0, \ \vec{v} = a_1\frac{e_1}{3} + a_2\frac{e_2}{3} + a_3\frac{e_3}{3}$$

$$\frac{\partial u}{\partial a_0}\frac{e_1}{3} = \frac{\partial \vec{v}}{\partial a_1}e_0, \ \frac{\partial u}{\partial a_0}\frac{e_2}{3} = \frac{\partial \vec{v}}{\partial a_2}e_0, \ \frac{\partial u}{\partial a_0}\frac{e_3}{3} = \frac{\partial \vec{v}}{\partial a_3}e_0$$

This also solves a holonomic equation.

$$\left(\frac{\partial u}{\partial a_0} + \frac{\partial \vec{v}}{\partial a_1} + \frac{\partial \vec{v}}{\partial a_2} + \frac{\partial \vec{v}}{\partial a_3}\right) \cdot (e_0 + e_1 + e_2 + e_3) =$$

$$e_0 e_0 + \frac{e_1}{3}e_1 + \frac{e_2}{3}e_2 + \frac{e_3}{3}e_3 = 0$$

There are no off diagonal terms to compare.

This exercise can be repeated for the other identity functions. One noticeable change is that the role that the conjugate must play. Consider the identity function w = q*1. To show that this is analytic in q*1 requires that one always works with basis vectors of the q*1 variety.

$$\text{Let } u = -a_0 e_0, \ \vec{v} = +a_1\frac{e_1}{3} - a_2\frac{e_2}{3} - a_3\frac{e_3}{3}$$

$$\frac{\partial u}{\partial a_0}\left(-\frac{e_1}{3}\right) = \frac{\partial \vec{v}}{\partial a_1}e_0, \ \frac{\partial u}{\partial a_0}\frac{e_2}{3} = \frac{\partial \vec{v}}{\partial a_2}e_0, \ \frac{\partial u}{\partial a_0}\frac{e_3}{3} = \frac{\partial \vec{v}}{\partial a_3}e_0$$

This also solves a first conjugate holonomic equation.

$$\left(\frac{\partial u}{\partial a_0} + \frac{\partial \vec{v}}{\partial a_1} + \frac{\partial \vec{v}}{\partial a_2} + \frac{\partial \vec{v}}{\partial a_3}\right) \cdot (e_0 + e_1 + e_2 + e_3)^{*1} =$$

$$-e_0 (-e_0) + \frac{e_1}{3}e_1 - \frac{-e_2}{3}e_2 - \frac{-e_3}{3}e_3 = 0$$

Power functions can be analyzed in exactly the same way:

$$\text{Let } w = q^2 = a_0^2 e_0^2 + a_1^2\frac{e_1^2}{9} + a_2^2\frac{e_2^2}{9} +$$

$$a_3^2\frac{e_3^2}{9} + 2a_0 a_1 e_0\frac{e_1}{3} + 2a_0 a_2 e_0\frac{e_2}{3} + 2a_0 a_3 e_0\frac{e_3}{3}$$

$$u = a_0^2 e_0^2 + a_1^2\frac{e_1^2}{9} + a_2^2\frac{e_2^2}{9} + a_3^2\frac{e_3^2}{9}$$

$$\vec{v} = 2a_0 a_1 e_0\frac{e_1}{3} + 2a_0 a_2 e_0\frac{e_2}{3} + 2a_0 a_3 e_0\frac{e_3}{3}$$

12

$$\frac{\partial u}{\partial a_0}\frac{e_1}{3} = \frac{2\,a_0{e_0}^2 e_1}{3} = \frac{\partial \vec{v}}{\partial a_1}e_0$$

$$\frac{\partial u}{\partial a_0}\frac{e_2}{3} = \frac{2\,a_0{e_0}^2 e_2}{3} = \frac{\partial \vec{v}}{\partial a_2}e_0$$

$$\frac{\partial u}{\partial a_0}\frac{e_3}{3} = \frac{2\,a_0{e_3}^2}{3} = \frac{\partial \vec{v}}{\partial a_3}$$

This time there are cross terms involved.

$$\frac{\partial u}{\partial a_1}e_0 = \frac{2\,a_1 e_0 {e_1}^2}{9} = \frac{\partial \left(\vec{v}\right)_1}{\partial a_0}\frac{e_1}{3}$$

$$\frac{\partial u}{\partial a_2}e_0 = \frac{2\,a_2 e_0 {e_2}^2}{9} = \frac{\partial \left(\vec{v}\right)_2}{\partial a_0}\frac{e_2}{3}$$

$$\frac{\partial u}{\partial a_3}e_0 = \frac{2\,a_3 e_0 {e_3}^2}{9} = \frac{\partial \left(\vec{v}\right)_3}{\partial a_0}\frac{e_3}{3}$$

At first glance, one might think these are incorrect, since the signs of the derivatives are suppose to be opposite. Actually they are, but it is hidden in an accounting trick :-) For example, the derivative of u with respect to a1 has a factor of e1^2, which makes it negative. The derivative of the first component of V with respect to a0 is positive. Keeping all the information about signs in the e's makes things look non-standard, but they are not.

Note that these are three scalar equalities. The other Cauchy-Riemann equations evaluate to a single 3-vector equation. This represents four constraints on the four degrees of freedom found in quaternions to find out if a function happens to be analytic.

This also solves a holonomic equation.

$$\left(\frac{\partial u}{\partial a_0} + \frac{\partial \vec{v}}{\partial a_1} + \frac{\partial \vec{v}}{\partial a_2} + \frac{\partial \vec{v}}{\partial a_3}\right)\cdot(e_0 + e_1 + e_2 + e_3) =$$

$$= 2\,a_0{e_0}^3 + \frac{2\,a_0 e_0 e_1}{3}e_1 + \frac{2\,a_0 e_0 e_2}{3}e_2 + \frac{2\,a_0 e_0 e_3}{3}e_3 = 0$$

Since power series can be analytic, this should open the door to all forms of analysis. (I have done the case for the cube of q, and it too is analytic in q).

4 Other Derivatives

So far, this work has only involved future timelike derivatives. There are five other regions of spacetime to cover. The simplest next case is for past timelike derivatives. The only change is in the limit, where the scalar approaches zero from below. This will make many derivatives look time symmetric, which is the case for most laws of physics.

A more complicated case involves spacelike derivatives. In the spacelike region, changes in time go to zero faster than the absolute value of the 3-vector. Therefore the order of the limit processes is reversed. This time the scalar approaches zero, then the 3-vector. This creates a problem, because after the first limit process, the differential element is (0, D), which will not commute with most quaternions. That will lead to the differential element not cancelling. The way around this is to take its norm, which is a scalar.

A spacelike differential element is defined by taking the ratio of a differential quaternion element D to its 3-vector, D - D*. Let the norm of D approach zero. To be defined, the three vector must approach zero faster than its corresponding scalar. To make the definition non-singular everywhere, multiply by the conjugate. In the limit D D*/((D - D*)(D - D*))* approaches (1, 0), a scalar.

$$\frac{\partial f(q,q^*,q^{*1},q^{*2})}{\partial q}\frac{\partial f(q,q^*,q^{*1},q^{*2})^*}{\partial q} =$$

$$= \texttt{limit as } \left(0, \vec{D}\right) -> 0 \left(\texttt{limit as } \left(d, \vec{D}\right) -> \left(0, \vec{D}\right)\right)$$

$$\left(\left(f\left(q + \left(d, \vec{D}\right), q^*, q^{*1}, q^{*2}\right) - f(q, q^*, q^{*1}, q^{*2})\right) \left(d, \vec{D}\right)^{-1}\right.$$
$$\left.\left(f\left(q + \left(d, \vec{D}\right), q^*, q^{*1}, q^{*2}\right) - f(q, q^*, q^{*1}, q^{*2})\right)^* \left(d, \vec{D}\right)^{-1*}\right)\right)$$

To make this concrete, consider a simple example, f = q^2. Apply the definition:

$$\texttt{Norm}\left(\frac{\partial q^2}{\partial q}\right) = \texttt{limit}\left(\left(0, \vec{D}\right) -> 0 \left(\texttt{limit as } \left(d, \vec{D}\right) -> \left(0, \vec{D}\right)\right)\right.$$

$$\left(\left(\left(\left(a, \vec{B}\right) + \left(d, \vec{D}\right)\right)^2 - \left(a, \vec{B}\right)^2\right)\right.$$
$$\left.\left(d, \vec{D}\right)^{-1}\left(\left(\left(a, \vec{B}\right) + \left(d, \vec{D}\right)\right)^2 - \left(a, \vec{B}\right)^2\right)^* \left(d, \vec{D}\right)^{-1*}\right)\right) =$$

$$= \texttt{lim}\left(\left(\left(a, \vec{B}\right) + \left(0, \vec{D}\right)\left(a, \vec{B}\right)\left(0, -\vec{D}\right)/\texttt{norm}\left(\left(0, \vec{D}\right)\right) + \left(0, \vec{D}\right)\right)\right.$$
$$\left.\left(\left(a, \vec{B}\right) + \left(0, \vec{D}\right)\left(a, \vec{B}\right)\left(0, -\vec{D}\right)/\texttt{norm}\left(\left(0, \vec{D}\right)\right) + \left(0, \vec{D}\right)\right)^*\right) =$$

The second and fifth terms are unitary rotations of the 3-vector B. Since the differential element D could be pointed anywhere, this is an arbitrary rotation. Define:

$$\left(a, \vec{B}'\right) = \left(0, \vec{D}\right)\left(a, \vec{B}\right)\left(0, -\vec{D}\right)/\texttt{norm}\left(\left(0, \vec{D}\right)\right)$$

Substitute, and continue:

$$= \texttt{lim}\left(\left(\left(a, \vec{B}\right) + \left(a, \vec{B}'\right) + \left(0, \vec{D}\right)\right)\left(\left(a, \vec{B}\right) + \left(a, \vec{B}'\right) + \left(0, \vec{D}\right)\right)^*\right) =$$

$$= \texttt{lim}\left(4a^2 + 2\vec{B}.\vec{B} + 2\vec{B}.\vec{B}' + 2\vec{D}.\vec{B} + 2\vec{D}.\vec{B}', \vec{0}\right)$$

$$= \left(4a^2 + 2\vec{B}.\vec{B} + 2\vec{B}.\vec{B}', \vec{0}\right) <= \left|2q\right|^2$$

Look at how wonderfully strange this is! The arbitrary rotation of the 3-vector B means that this derivative is bound by an inequality. If D is in direction of B, then it will be an equality, but D could also be in the opposite direction, leading to a destruction of a contribution from the 3-vector. The spacelike derivative can therefore interfere with itself. This is quite a natural thing to do in quantum mechanics. The spacelike derivative is positive definite, and could be used to define a Banach space.

Defining the lightlike derivative, where the change in time is equal to the change in space, will require more study. It may turn out that this derivative is singular everywhere, but it will require some skill to find a technically viable compromise between the spacelike and timelike derivative to synthesis the lightlike derivative.

5 Topological Properties of Quaternions

(section under development)

Topological Space

If we choose to work systematically through Wald's "General Relativity", the starting point is "Appendix A, Topological Spaces". Roughly, topology is the structure of relationships that do not change if a space is distorted. Some of the results of topology are required to make calculus rigorous.

In this section, I will work consistently with the set of quaternions, H^1, or just H for short. The difference between the real numbers R and H is that H is not a totally ordered set and multiplication is not commutative. These differences are not important for basic topological properties, so statements and proofs involving H are often identical to those for R.

First an open ball of quaternions needs to be defined to set the stage for an open set. Define an open ball in H of radius (r, 0) centered around a point (y, Y) [note: small letters are scalars, capital letters are 3-vectors] consisting of points (x, X) such that

$$\sqrt{((x - y, X - Y)^* (x - y, X - Y))} < (r, 0)$$

An open set in H is any set which can be expressed as a union of open balls.

[p. 423 translated] A quaternion topological space (H,T) consists of the set H together with a collection T of subsets of H with these properties:

1.The union of an arbitrary collection of subsets, each in T, is in T

2.The intersection of a finite number of subsets of T is in T

3.The entire set H and the empty set are in T

T is the topology on H. The subsets of H in T are open sets. Quaternions form a topology because they are what mathematicians call a metric space, since q* q evaluates to a real positive number or equals zero only if q is zero. Note: this is not the meaning of metric used by physicists. For example, the Minkowski metric can be negative or zero even if a point is not zero. To keep the same word with two meanings distinct, I will refer to one as the topological metric, the other as an interval metric. These descriptive labels are not used in general since context usually determines which one is in play.

An important component to standard approaches to general relativity is product spaces. This is how a topology for R^n is created. Events in spacetime require R^4, one place for time, three for space. Mathematicians get to make choices: what would change if work was done in R^2, R^3, or R^5? The precision of this notion, together with the freedom to make choices, makes exploring these decisions fun (for those few who can understand what is going on :-)

By working with H, product spaces are unnecessary. Events in spacetime can be members of an open set in H. Time is the scalar, space the 3-vector. There is no choice to be made.

Open Sets

The edges of sets will be examined by defining boundaries, open and closed sets, and the interior and closure of a set.

I am a practical guy who likes pragmatic definitions. Let the real numbers L and U represent arbitrary lower and upper bounds respectively such that L < U. For the quaternion topological space (H, T), consider an arbitrary induced topology (A, t) where x and a are elements of A. Use inequalities to define:

```
an open set :  (L, 0) < (x - a)*(x - a) < (U, 0)

a closed set :  (L, 0) <= (x - a)*(x - a) <= (U, 0)

a half open set :  (L, 0) <= (x - a)*(x - a) < (U, 0)

or (L, 0) < (x - a)*(x - a) <= (U, 0)

a boundary :  (L, 0) = (x - a)*(x - a)
```

The union of an arbitrary collection of open sets is open.

The intersection of a finite number of open sets is open.

The union of a finite number of closed sets is closed.

The intersection of an arbitrary number of closed sets is closed.

Clearly there are connections between the above definitions

```
open set union boundary - > closed set
```

This creates complementary ideas. [Wald, p.424]

The interior of A is the union of all open sets contained within A.

The interior equals A if and only if A is open.

The closure of A is the intersection of all closed sets containing A.

The closure of A equals A if and only if A is closed.

Define a point set as the set where the lower bound equals the upper bound. The only open set that is a point set is the null set. The closed point set is H. A point set for the real numbers has only one element which is identical to the boundary. A point set for quaternions has an infinite number of elements, one of them identical to the boundary.

What are the implications for physics?

With quaternions, the existence an open set of events has nothing to do with the causality of that collection of events.

```
an open set :  (L, 0) < (x - a)*(x - a) < (U, 0)

timelike events : scalar((x - a)²) > (0, 0)

lightlike events : scalar((x - a)²) = (0, 0)

spacelike events : scalar((x - a)²) < (0, 0)
```

A proper time can have exactly the same absolute value as a pure spacelike separation, so these two will be included in the same sets, whether open, closed or on a boundary.

There is no correlation the reverse way either. Take for example a collection of lightlike events. Even though they all share exactly the same interval - namely zero - their absolute value can vary all over the map, not staying within limits.

Although independent, these two ideas can be combined synergistically. Consider an open set S of timelike intervals.

```
S = {x, a e H, a fixed; U,
    L e R | (L, 0) < (x - a)*(x - a) < (U, 0), and scalar ((x - a)²) > 0}
```

The set S could depict a classical world history since they are causally linked and have good topological properties. A closed set of lightlike events could be a focus of quantum electrodynamics. Topology plus causality could be the key for subdividing different regions of physics.

Hausdorff Topology

This property is used to analyze compactness, something vital for rigorously establishing differentiation and integration.

[Wald p424] The quaternion topological space (H, T) is Hausdorff because for each pair of distinct points a, b E H, a not equal to b, one can find open sets Oa, Ob E T such that a E Oa, b I Ob and the intersection of Oa and Ob is the null set.

For example, find the half-way point between a and b. Let that be the radius of an open ball around the points a and b:

```
let (r,0) = (a - b)*(a - b)/4

Oa = {a, x e H, a is fixed, r e R | (a - x)*(a - x) < r}

Ob = {b, x e H, b is fixed, r e R | (b - x)*(b - x) < r}
```

Neither set quite reaches the other, so their intersection is null.

Compact Sets

In this section, I will begin an investigation of compact sets of quaternions. I hope to share some of my insights into this subtle but significant topic.

First we need the definition of a compact set of quaternions.

[Translation of Wald p. 424] Let A be a subset of the quaternions H. Set A could be opened, closed or neither. An open cover of A is the union of open sets {Oa} that contains A. A union of open sets is open and could have an infinite number of members. A subset of {Oa} that still covers A is called a subcover. If the subcover has a finite number of elements it is called a finite subcover. The set A subset of H is compact if every open cover of A has a finite subcover.

Let's find an example of a compact set of quaternions. Consider a set S composed of points with a finite number of absolute values:

```
S = {x1, x2, ..., xn e H; a1, a2, ..., an e R,
    n is finite | (x1 * x1)^0.5 = (a1, 0), (x2 * x2)^0.5 = (a2, 0), ...}
```

The set S has an infinite number of members, since for any of the equalities, specifying the absolute value still leaves three degrees of freedom (if the domain had been x E R, then S would have had a finite number of elements). The set S can be covered by an open set {O} which could have an infinite number of members. There exists a subset {C} of {O} that is finite and still covers S. The subset {C} would have one member for each absolute value.

$$C = \left\{ y \, e \, \{O\}, e \, e \, R, e > 0 \, \middle| \, (a1 - e) < \sqrt{y^*y} < (a1 + e, 0), \right.$$
$$\left. (a2 - e) < \sqrt{y^*y} < (a2 + e, 0), ..., \text{one } y \text{ exists for each inequality} \right\}$$

Every set of quaternions composed of a finite number of absolute values like the set S is compact.

Notice that the set S is closed because it consists of a boundary without an interior. The link between compact, closed and bound set is important, and will be examined next

A compact set is a statement about the ability to find a finite number of open sets that cover a set, given any open cover. A closed set is the interior of a set plus the boundary of that set. A set is bound if there exists a real number M such that the distance between a point and any member of the set is less than M.

For quaternions with the standard topology, in order to have a finite number of open sets that cover the set, the set must necessarily include its boundary and be bound. In other words, to be compact is to be closed and bound, to be closed and bound is to be compact.

[Wald p. 425] Theorem 1 (Heine-Borel). A closed interval of quaternions S:

$$S = \left\{ x \in H, a, b \in R, a < b \,\middle|\, (a, 0) <= \sqrt{x^* x} <= (b, 0) \right\}$$

with the standard topology on H is compact.

Wald does not provide a proof since it appears in many books on analysis. Invariably the Heine-Borel Theorem employs the domain of the real numbers, x E R. However, nothing in that proof changes by using quaternions as the domain.

[Wald p. 425] Theorem 2. Let the topology (H, T) be Hausdorff and let the set A subset of H be compact. Then A is closed.

Theorem 3. Let the topology (H, T) be compact and let the set A subset of H be closed. Then A is compact.

Combine these theorems to create a stronger statement on the compactness of subsets of quaternions H.

Theorem 4. A subset A of quaternions is compact if and only if it is closed and bounded.

The property of compactness is easily proved to be preserved under continuous maps.

Theorem 5. Let (H, T) and (H', T') be topological spaces. Suppose (H, T) is compact and the function f: H -> H' is continuous. The f[H] = {h' E H' | h' = f(h)} is compact. This creates a corollary by theorem 4.

Theorem 6. A continuous function from a compact topological space into H is bound and its absolute value attains a maximum and minimum values.

[end translation of Wald]

Rˆ1 versus Rˆn

It is important to note that these theorems for quaternions are build directly on top of theorems for real numbers, Rˆ1. Only the domain needs to be changed to Hˆ1. Wald continues with theorems on product spaces, specifically Tychonoff's Theorem, so that the above theorems can be extended to Rˆn. In particular, the product space Rˆ4 should have the same topology as the quaternions.

Hopefully, subtlety matters in the discussion of the logical foundations of general relativity. Both Rˆ1 and Hˆ1 have a rule for multiplication, but Hˆ1 has an antisymmetric component. This is a description of a difference. Rˆ4 does not come equipped with a rule for multiplication, so it is qualitatively different, even if topologically similar to the quaternions.