EXTENDED ESSAY

MATHEMATICS

Title: Hypercomplex representations of the seemingly ordinary

Research Question: To what extent is a four-dimensional number system best suited to define rotation in a three-dimensional space?

Word Count: 3612

Table of Contents:

Introduction

With the advent of computer programs aiding every part of our life and the seemingly unavoidable revolution of the Metaverse describing rotations in three-dimensional space has never been more important than today. Over the past hundred of years mathematicians developed several ways of rigorously defining a rotation, the most interesting one in my opinion is through numbers called Quaternions. Just as the complex numbers can describe rotations on a plane, even higher dimensional numbers can be used to describe rotations on multiple planes. This paper will attempt to find the most suitable way to describe the rotations by comparing two most common techniques.

The Complex Prerequisite

In order to introduce the basics of quaternion algebra, a short summary of the complex algebra must be outlined, as the quaternion field is a sort of extension over the complex field.

Definition 1:

I define a complex number z as an expression of the form:

$$z \in \mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$$

The real part of the complex number is represented by a and the imaginary part of the complex number is represented by b**.** The symbol i stands for the imaginary unit, the square root of -1,

$$i = \sqrt{-1},$$

$$i^2 = -1.$$

Definition 2:

Two complex numbers are equal only when their real and imaginary parts are equal, therefore for $z_1 = a_1 + b_1 i \in \mathbb{C}$ and $z_2 = a_2 + b_2 i \in \mathbb{C}$,

$$z_1 = z_2 \Leftrightarrow a_1 = a_2 \wedge \, b_1 = b_2$$

Operations with complex numbers:

Addition and subtraction of complex numbers is defined as the piecewise sum or difference:

$$z_1 \pm z_2 = (a_1 + b_1 i) \pm (a_2 + b_2 i) = (a_1 \pm a_2) + (b_1 \pm b_2)i$$

Multiplication and division of the complex numbers is handled by standard algebraic multiplication while keeping in mind the special property of i being the square root of -1.

$$z_1 z_2 = a_1 a_2 + a_1 b_2 i + b_1 a_2 i + b_1 b_2 i^2 = a_1 a_2 - b_1 b_2 + (a_1 b_2 + b_1 a_2)i$$

An important thing to define is the concept of a mathematical **norm**.

Definition 3:

A norm is a function over the domain of real or complex numbers with a range of nonnegative real numbers. These are the properties that the norm is required to obey:

1. It commutes with scaling – It means that $\forall \varphi \in \mathbb{R}$

$$Norm(\varphi z) = |\varphi| Norm(z)$$

2. It obeys a form of the triangle inequality – so for any triangle of lengths a, b and c where $a \leq c$ and $b \leq c$ the sum of the lengths of any two sides must be greater or equal than the length of the remaining side.

$$\forall z_1, z_2 \in \mathbb{C} \;\; Norm(z_1 + z_2) \leq Norm(z_1) + Norm(z_2)$$

3. $Norm(z) \; = \; 0$ only for $z = 0$.

Definition 4:

The concept of a norm may be thought of as a distance from the origin in higher dimensions. For the first dimension, the real number set, for all real values of x

$$Norm(x) = |x| = \sqrt{x^2}$$

Definition 5:

In the complex dimension, for all complex values of z

$$Norm(z) = |z| = \sqrt{a^2 + b^2}$$

In general, on an n-dimensional space the Norm has been defined by Euclid as the square root of the sum of the squares of each component of direction. It is a direct consequence of the Pythagorean theorem applied to higher dimensions.

Definition 6:

Let χ be an element of an $n-$ dimensional vector space.

$$Norm(\chi) = \sqrt{x_1^2 + \cdots + x_n^2}$$

A normed algebra is a set of operations over a set of numbers in which an idea of a norm of an element may be worked out.

The three-dimensional endeavors

Inspired by the geometrical representation of the complex numbers on the Argand plane as points on a two-dimensional plane, an Irish mathematician William Rowan Hamilton began research to extend the complex numbers to the three-dimensional space, through his "Theory of Triplets." It seemed only natural that since a number can be represented in 2 dimensions, it should also be possible to be represent it in a 3-dimensional space. In modern language, what Hamilton was looking for is a **3-dimensional normed division algebra**, but in today's times this algebra has been proven impossible to exist.

I shall now recreate Hamilton's endeavors to investigate why a system of 3-dimensional normed division algebra is impossible.

Definition 7:

Let T be an expression of the form

$$T \in \mathbb{T} = \{\alpha + \beta i + \gamma j :\ \alpha, \beta, \gamma \in \mathbb{R}\}$$

Let the unit of 1, i and j be vectors of length 1 in directions that are mutually perpendicular to each other, akin to the common notation of a (i, j, k) vector.

Definition 8:

The norm of T can easily be defined as:

$$Norm(T) = \sqrt{\alpha^2 + \beta^2 + \gamma^2}$$

I observe that defining the norm of a triplet didn't cause any problems. I shall now investigate the divisibility requirement of our new number system. Suppose that there exists a 3-dimensional number T such that for all $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2 \in \mathbb{R}$ there exist real numbers A, B, C such that

$$(\alpha_1 + \beta_1 i + \gamma_1 j)(\alpha_2 + \beta_2 i + \gamma_2 j) = A + Bi + Cj = T$$

Hamilton required that it would be possible to multiply the two triplets' term by term, and that the product of the triplets will be equal to the product of the lengths. He called the second requirement the "law of the moduli". The law of moduli can be written algebraically as follows, for $T_1, T_2 \in \mathbb{T}$:

$$|T_1T_2| = |T_1||T_2|$$

$$Norm(T_1T_2) = Norm(T_1)Norm(T_2)$$

I shall now multiply the two numbers according to the standard rules of algebra, while keeping in mind that $i^2 = j^2 = -1$.

$$T_1T_2 = \alpha_1\alpha_2 + \alpha_1\beta_2 i + \alpha_1\gamma_2 j + \alpha_2\beta_1 i + \beta_1\beta_2 i^2 + \beta_1\gamma_2 ij + \alpha_2\gamma_1 j + \beta_2\gamma_1 ij + \gamma_1\gamma_2 j^2$$

Upon collecting like terms

$$T_1T_2 = \alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2 + (\alpha_1\beta_2 + \beta_1\alpha_2)i + (\alpha_1\gamma_2 + \beta_2\gamma_1)j + (\beta_1\gamma_2 + \beta_2\gamma_1)ij$$

I substitute A, B and C for the coefficients of the real and imaginary units

$$T_1T_2 = A + Bi + Cj + (\beta_1\gamma_2 + \beta_2\gamma_1)ij$$

A this point the *ij* term does not fit into what I am looking to obtain. I need to find a way to express

$$(\beta_1\gamma_2 + \beta_2\gamma_1)ij$$

In the form of

$$a' + ib' + jc',$$

Where $a', b', c' \in \mathbb{R}$.

The first attempt at transforming the unruly term came from observing that since

$$i^2 = j^2 = -1$$

Then the square of the product of the squares of *i* and *j* has to equal 1.

$$(ij)^2 = i^2 j^2 = (-1)(-1) = 1$$

Therefore, taking the square root

$$\sqrt{(ij)^2} = 1 \ \lor \ \sqrt{(ij)^2} = -1$$

Unfortunately, with neither assumption will the resulting triplet obey the law of moduli, as shown here

$$|T_1||T_2| = \sqrt{\alpha_1^2 + \beta_1^2 + \gamma_1^2 + \alpha_2^2 + \beta_2^2 + \gamma_2^2}$$

Proof:

Assuming $ij = 1$

$$T_1T_2 = \alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2 + \beta_1\gamma_2 + \beta_2\gamma_1 + (\alpha_1\beta_2 + \alpha_2,\beta_1)i + (\alpha_1\gamma_2 + \beta_2\gamma_1)j$$

$$|T_1T_2| = \sqrt{(\alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2 + \beta_1\gamma_2 + \beta_2\gamma_1)^2 + (\alpha_1\beta_2 + \alpha_2,\beta_1)^2 + (\alpha_1\gamma_2 + \beta_2\gamma_1)^2}$$

$$Norm(T_1T_2) \neq Norm(T_1)Norm(T_2)$$

Assuming ij = -1

$$T_1T_2 = \alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2 - \beta_1\gamma_2 - \beta_2\gamma_1 + (\alpha_1\beta_2 + \alpha_2,\beta_1)i + (\alpha_1\gamma_2 + \beta_2\gamma_1)j$$

$$|T_1T_2| = \sqrt{(\alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2 - \beta_1\gamma_2 - \beta_2\gamma_1)^2 + (\alpha_1\beta_2 + \alpha_2,\beta_1)^2 + (\alpha_1\gamma_2 + \beta_2\gamma_1)^2}$$

$$Norm(T_1T_2) \neq Norm(T_1)Norm(T_2)$$

□

For the second attempt of making sense of the ij term I shall consider the simplest case of a square.

$$(\alpha + \beta i + \gamma j)^2 = \alpha^2 - \beta^2 - \gamma^2 + (2\alpha\beta)i + (2\alpha\gamma)j + (2\beta\gamma)ij$$

I calculate the sum of the squares of the coefficients of the unit vectors of the right side, ignoring the ij term

$$(\alpha^2 - \beta^2 - \gamma^2)^2 + (2\alpha\beta)^2 + (2\alpha\gamma)^2 = \alpha^4 + \beta^4 + \gamma^4 + 2\alpha^2\beta^2 + 2\alpha^2\gamma^2 + 2\beta^2\gamma^2$$
$$= (\alpha^2 + \beta^2 + \gamma^2)^2$$

Therefore, the law of moduli is fulfilled if I let

$$ij = 0$$

Assuming $ij = 0$

$$T^2 = \alpha^2 - \beta^2 - \gamma^2 + (2\alpha\beta)i + (2\alpha\gamma)j + (2\beta\gamma)ij$$

$$T^2 = \alpha^2 - \beta^2 - \gamma^2 + (2\alpha\beta)i + (2\alpha\gamma)j$$

$$|T^2| = \sqrt{(\alpha^2 - \beta^2 - \gamma^2)^2 + (2\alpha\beta)^2 + (2\alpha\gamma)^2}$$

Although this makes mathematical sense, the assumption did not sit right with Hamilton. *"But this seemed odd and uncomfortable"* he writes in a letter. The assumption also brings with itself a contradiction; if $ij = 0$, then the absolute value of the product ij would be zero, which would contradict the law of the moduli. Hamilton had to experiment with something new: He gave up the commutative property of $AB = BA$ and let $ij = -ji$. He introduced a new constant k, and let

$$ij = k$$

$$ji = -k$$

Let's multiply the triplets again, while restricting $\beta_1 = \beta_2 = \beta$ and $\gamma_1 = \gamma_2 = \gamma$

$$(\alpha_1 + \beta i + \gamma j)(\alpha_2 + \beta i + \gamma j) = \alpha_1\alpha_2 + \alpha_1\beta i + \alpha_1\gamma j + \alpha_2\beta i - \beta^2 + \beta\gamma k + \alpha_2\gamma j - \beta\gamma k - \gamma^2$$

Upon collecting like terms, I obtain:

$$(\alpha_1\alpha_2 - \beta^2 - \gamma^2) + (\alpha_1 + \alpha_2)\beta i + (\alpha_1 + \alpha_2)\gamma j + (\beta\gamma - \beta\gamma)k = (\alpha_1\alpha_2 - \beta^2 - \gamma^2) + (\alpha_1 + \alpha_2)\beta i + (\alpha_1 + \alpha_2)\gamma j$$

The “artificial” imaginary unit k vanishes, as its coefficient turns out to be zero. This assumption satisfies the law of the moduli and confirms that $ij = -ji$. Hamilton now examines the true value of k. I shall now investigate the general product of two triplets again utilizing our new discovery.

$$(\alpha_1 + \beta_1 i + \gamma_1 j)(\alpha_2 + \beta_2 i + \gamma_2 j)$$

$$(\alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2) + (\alpha_1\beta_2 + \beta_1\alpha_2)i + (\alpha_1\gamma_2 + \gamma_1\alpha_2)j + (\beta_1\gamma_2 - \gamma_1\beta_2)k$$

Assume $k = 0$

$$(\alpha_1 + \beta_1 i + \gamma_1 j)(\alpha_2 + \beta_2 i + \gamma_2 j) = (\alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2) + (\alpha_1\beta_2 + \alpha_2\beta_1)i + (\alpha_1\gamma_2 + \alpha_2\gamma_1)j$$

Unfortunately, the law of the moduli under this assumption is unsatisfied, as the identity

$$(\alpha_1^2 + \beta_1^2 + \gamma_1^2)(\alpha_2^2 + \beta_2^2 + \gamma_2^2) = (\alpha_1\alpha_2 - \beta_1\beta_2 - \gamma_1\gamma_2)^2 + (\alpha_1\beta_2 + \alpha_2\beta_1)^2 + (\alpha_1\gamma_2 + \alpha_2\gamma_1)^2$$

does not hold, the left-hand side exceeds the right-hand side by $(\beta_1\gamma_2 - \beta_2\gamma_1)$. At this point in his research, it dawned on Hamilton that a fourth dimension is needed to calculate products and factors of triplets. He “promoted” the constant k to the rank of a third distinct imaginary symbol, where:

$$k = ij = -ji.$$

Now that I have derived a way to multiply the imaginary units i and j, I shall derive a way to add the imaginary unit k to the mix as well. Using the associative law $a(bc) = (ab)c$ I have:

$$ik = i(ij) = (ii)j = i^2 j = -j$$

I also conclude that

$$ki = -(ji)i = -j(ii) = -ji^2 = j$$

It is important that I substitute k as $-ji$, and not as ij in this case. Since multiplication is now defined as non-commutative, I cannot change the order of the symbols – I only freely control where the brackets are placed.

Similarly,

$$jk = -j(ji) = -(jj)i = -j^2 i = i.$$

And

$$kj = (ij)j = i(jj) = ij^2 = -i.$$

What remains to be derived is the square of k,

$$k^2 = kk = -(ji)(ij) = -ji^2 j = jj = j^2 = -1.$$

Therefore,

$$i^2 = j^2 = k^2 = ijk = -1.$$

What I derived is an extension of the complex numbers into not three, but a four-dimensional space – This set of rules is known today as Hamilton's "Theory of Quaternions.", the **4-dimensional normed division algebra**. Although Hamilton gave up on his journey to discover a system of 3-dimensional algebra because of numerous attempts of trial and error, there exists a mathematical proof that there can't exist such a system discovered by a French mathematician Adrien-Marie Legendre. What's more, a German mathematician Adolf Hurwitz proved in 1898 that such a system of rules can exist only in $1, 2, 4$ and 8 dimensions. These proofs however are beyond the scope of this essay.

The Hypercomplex Sine Qua Non

Definition 9:

Let the set of all quaternions be denoted by $\mathbb{H}$ (H is used in honor of Hamilton).

$$\mathbb{H} = \{w + ix + jy + kz : \ w, x, y, z \in \mathbb{R}\}$$

Let Q be a quaternion of the form

$$Q \in \mathbb{H} = \{w + \ ix + \ jy + \ kz \ : \ w, x, y, z \in \mathbb{R}\}$$

Let the coefficients w, x, y, z be called the four *constituents* of the quaternion Q, and the quantities i, j, k be called *imaginary units*. The imaginary units of a quaternion cannot be obtained by multiplying another imaginary unit by a real scalar, just as you cannot obtain i in the complex plane by multiplying a real number by a real scalar. The four quaternion constituents are all set perpendicular to each other in four dimensions, therefore for two quaternions to be equal all of their individual constituents have to be equal analogously to the complex case examined in Definition 2

Operations on Quaternions:

Addition and subtraction of quaternions is defined as the **piecewise** sum or difference of the constituents, so each constituent is added or subtracted from the corresponding constituent of the other quaternion.

$$Q_1 \pm Q_2 = w_1 \pm w_2 + i(x_1 \pm x_2) + j(y_1 \pm y_2) + k(z_1 \pm z_2)$$

A notion of a quaternion conjugate analogous to the complex conjugate has also been developed, defined by changing the sign of each of the imaginary part of the quaternions.

$$Q^* = w - ix - jy - kz$$

I observe that multiplying a quaternion by its conjugate results in a real number of the form

$$QQ^* = Q^*Q = w^2 + x^2 + y^2 + z^2$$

I are now going to prove that quaternion algebra is normed.

Proof:

I will take the square root of the product of a quaternion by its conjugate

$$\sqrt{QQ^*} = \sqrt{w^2 + x^2 + y^2 + z^2}$$

I observe that the result follows the formula for the 4-dimensional Euclidian norm, as defined in Definition 5:

$$\sqrt{w^2 + x^2 + y^2 + z^2} = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$$

This is a norm because it holds the three required properties. Let $QQ^* = \boldsymbol{x}$

1. $\sqrt{\boldsymbol{x}} = 0 \iff Q = 0$
2. $\sqrt{\lambda \boldsymbol{x}} = |\lambda|\sqrt{\boldsymbol{x}}$
3. $\sqrt{(\boldsymbol{x_1} + \boldsymbol{x_2})} \leq \sqrt{\boldsymbol{x_1}} + \sqrt{\boldsymbol{x_2}}$, as demonstrated below:

$$\left(\sqrt{\boldsymbol{x_1} + \boldsymbol{x_2}}\right)^2 \leq \left(\sqrt{\boldsymbol{x_1}} + \sqrt{\boldsymbol{x_2}}\right)^2$$

$$\boldsymbol{x_1} + \boldsymbol{x_2} \leq \boldsymbol{x_1} + 2\sqrt{\boldsymbol{x_1}\boldsymbol{x_2}} + \boldsymbol{x_2}$$

$$0 \leq 2\sqrt{\boldsymbol{x_1}\boldsymbol{x_2}}$$

□

Multiplication of a quaternion by a scalar is defined as multiplying each constituent by said scalar in this fashion:

$$\varphi Q = \varphi w + i(\varphi x) + j(\varphi y) + k(\varphi z)$$

We've spent a bulk of the essay on deriving the multiplication of quaternions and I shall now prove that these assumptions describe a 4-dimensional normed division algebra.

Proof:

We've already proven that a notion of a norm exists in the quaternion space, now I will prove that there exists a quaternion Q such that

$$Q_1Q_2 = A + Bi + Cj + Dk = Q$$

$$(w_1 + x_1i + y_1j + z_1k)(w_2 + x_2i + y_2j + z_2k) =$$

$$(w_1w_2 + w_1x_2i + w_1y_2j + w_1z_2k + w_2x_1i + x_1x_2i^2 + x_1y_2ij + x_1z_2ik + w_2y_1j + x_2y_1ji + y_1y_2j^2 + y_1z_2jk + w_2z_1k + x_2z_1ki + z_1y_2kj + z_1z_2k^2)$$

I shall now apply the special identities I derived and am operating under, namely

$$i^2 = j^2 = k^2 = -1$$

And

$$ij = k = -ji$$

To obtain

$$(w_1w_2 + w_1x_2i + w_1y_2j + w_1z_2k + x_1w_2i - x_1x_2 + x_1y_2k - x_1z_2j + y_1w_2j - y_1x_2k - y_1y_2 + y_1z_2i + z_1w_2k + z_1x_2j - z_1y_2i - z_1z_2)$$

After collecting like terms

$$(w_1w_2 - x_1x_2 - y_1y_2 - z_1z_2) + (w_1x_2 + x_1w_2 + y_1z_2 - z_1y_2)i + (w_1y_2 - x_1z_2 + y_1w_2 + z_1x_2)j + (w_1z_2 + x_1y_2 - y_1x_2 + z_1w_2)k$$

I obtain an expression in the form

$$A + Bi + Cj + Dk$$

$$\begin{cases} A = w_1w_2 - x_1x_2 - y_1y_2 - z_1z_2 \\ B = w_1x_2 + x_1w_2 + y_1z_2 - z_1y_2 \\ C = w_1y_2 - x_1z_2 + y_1w_2 + z_1x_2 \\ D = w_1z_2 + x_1y_2 - y_1x_2 + z_1w_2 \end{cases}$$

I observe

$$\frac{A + Bi + Cj + Dk}{w_1 + x_1i + y_1j + z_1k} = w_2 + x_2i + y_2j + z_2k$$

Thus proving, that the 4-dimensional normed algebra is a division algebra.

□

The task of multiplying two generic quaternions is very straightforward yet tedious to complete by hand and would take up the whole page, thus I will not write the whole process out, yet Hamilton carried it out and observed, that the result does obey the law of the modulus. I shall investigate how quaternions may be represented geometrically.

The Hamiltonian quaternions exist as sort of a combination of a scalar and vector, and adequate notation has been introduced. I will call the scalar part of the number $Q \in \mathbb{H}$
Q_S and the vector part Q_V. Thus,

$$Q = Q_S + Q_V$$

$$w + ix + jy + kz = (w) + (ix + jy + kz)$$

When the quaternion consists of only the vector part, it can also be called a *pure quaternion*. (This nomenclature is similar to how I denote some complex numbers as *purely imaginary*) I define an analog of the dot product and cross product of vectors onto pure quaternions.

For $A = ix_1 + jy_1 + kz_1,$ and $B = ix_2 + jy_2 + kz_2$

The scalar part of the product is analogous to the dot product,

$$(AB)_S = x_1x_2 + y_1y_2 + z_1z_2$$

And the vector part of the product is analogous to the cross product,

$$(AB)_V = i(y_1z_2 - z_1y_2) + j(z_1x_2 - x_1z_2) + k(x_1y_2 - y_1x_2)$$

Therefore:

$$AB = (AB)_S + (AB)_V$$

$$= x_1x_2 + y_1y_2 + z_1z_2 + i(y_1z_2 - z_1y_2) + j(z_1x_2 - x_1z_2) + k(x_1y_2 - y_1x_2)$$

I can associate each quaternion with its norm – the modulus, and an angle – analogous to the complex argument. This is analogous to the complex polar form of $z = r(cos\theta + isin\theta)$, this notion will be developed later in the essay.

Let Q be a quaternion of the form

$$Q = w + xi + yj + zk$$

Let the modulus $|Q|$ be defined as the norm of Q

Definition 10:

$$|Q| = \sqrt{QQ^*} = \sqrt{w^2 + x^2 + y^2 + z^2}$$

I define a unit quaternion U as a quaternion of $Norm(U) = 1$.

To find a unit quaternion of a quaternion Q, I divide the quaternion by its norm.

$$U_Q = \frac{Q}{|Q|}$$

Proof:

To make sure we're comfortable with using that result, let $U_Q = \lambda_0 + \lambda_1 i + \lambda_2 j + \lambda_3 k$ be a quaternion with unknown real constituents

$$U_Q = (\lambda_0 + \lambda_1 i + \lambda_2 j + \lambda_3 k) = \frac{w + xi + yj + zk}{\sqrt{w^2 + x^2 + y^2 + z^2}}$$

$$(\lambda_0 + \lambda_1 i + \lambda_2 j + \lambda_3 k)(\sqrt{w^2 + x^2 + y^2 + z^2}) = w + xi + yj + zk$$

Using the piece-wise multiplication by a scalar rule

$$\left(\lambda_0(\sqrt{w^2+x^2+y^2+z^2}) + \lambda_1(\sqrt{w^2+x^2+y^2+z^2})i + \lambda_2(\sqrt{w^2+x^2+y^2+z^2})j + \lambda_3(\sqrt{w^2+x^2+y^2+z^2})k\right) = w + xi + yj + zk$$

For the two quaternions to be equal, their corresponding constituents must be equal

$$\begin{cases} \lambda_0\left(\sqrt{w^2+x^2+y^2+z^2}\right) = w \\ \lambda_1\left(\sqrt{w^2+x^2+y^2+z^2}\right) = x \\ \lambda_2\left(\sqrt{w^2+x^2+y^2+z^2}\right) = y \\ \lambda_3\left(\sqrt{w^2+x^2+y^2+z^2}\right) = z \end{cases}$$

After dividing all sides by $|Q|$ I obtain the exact values of the constituents of U_Q

$$\begin{cases} \lambda_0 = \dfrac{w}{\sqrt{w^2+x^2+y^2+z^2}} \\ \lambda_1 = \dfrac{x}{\sqrt{w^2+x^2+y^2+z^2}} \\ \lambda_2 = \dfrac{y}{\sqrt{w^2+x^2+y^2+z^2}} \\ \lambda_3 = \dfrac{z}{\sqrt{w^2+x^2+y^2+z^2}} \end{cases} = \begin{cases} \lambda_0 = \dfrac{w}{|Q|} \\ \lambda_1 = \dfrac{x}{|Q|} \\ \lambda_2 = \dfrac{y}{|Q|} \\ \lambda_3 = \dfrac{z}{|Q|} \end{cases}$$

Therefore U_Q is of the form

$$U_Q = \frac{w}{|Q|} + \frac{x}{|Q|}i + \frac{y}{|Q|}j + \frac{z}{|Q|}k$$

$$U_Q = \frac{Q}{|Q|}$$

Let's see if $Norm(U_Q)$ really does equal 1

$$|U_Q| = \frac{w^2}{w^2+x^2+y^2+z^2} + \frac{x^2}{w^2+x^2+y^2+z^2} + \frac{y^2}{w^2+x^2+y^2+z^2} + \frac{z^2}{w^2+x^2+y^2+z^2} = \frac{w^2+x^2+y^2+z^2}{w^2+x^2+y^2+z^2} = 1$$

$\square$

Just as a set of all unit complex numbers forms a unit circle of radius 1, the set of all

$$U_Q \in \mathbb{H} : Norm(U_Q) = 1$$

forms a four-dimensional unit sphere with radius 1. I may also define a reciprocal Q^{-1} of the non-zero quaternion Q. To find the reciprocal of a quaternion I divide its conjugate by the square of its norm. I know that multiplying a quaternion by its conjugate results in a real number of the form $w^2 + x^2 + y^2 + z^2$, which is also the form of the square of a norm of a quaternion. A reciprocal multiplied by the original quaternion must result 1, therefore I define a quaternion reciprocal as such

$$Q^{-1} = \frac{Q^*}{|Q|^2}$$

Multiplying both sides by Q yields 1.

$$Q^{-1}Q = \frac{Q^*Q}{|Q|^2} = \frac{|Q|^2}{|Q|^2} = 1$$

The Applications to Rotations.

Soon after the invention of quaternions mathematicians recognized that they could be used to represent rotations in the three-dimensional space, similarly to how complex numbers describe rotation on the two-dimensional plane. Before I get into the specifics, I need to establish how a three-dimensional rotation matrix operates. Let's start by examining rotations on the two-dimensional plane. A rotation in the $x - y$ plane by an angle θ measured counterclockwise from the positive x-axis is represented by this $2x2$ rotation matrix:

$$M = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$$

So a point (x, y) rotated by an angle θ in the same way becomes:

$$\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cos\theta \cdot x - sin\theta \cdot y \\ sin\theta \cdot x + cos\theta \cdot y \end{pmatrix}$$

For example, the point $(1,0)$ rotated by $\frac{\pi}{2}$ becomes:

$$\begin{pmatrix} \cos\frac{\pi}{2} & -sin\frac{\pi}{2} \\ sin\frac{\pi}{2} & cos\frac{\pi}{2} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} cos\frac{\pi}{2} \cdot 1 - sin\frac{\pi}{2} \cdot 0 \\ sin\frac{\pi}{2} \cdot 1 + cos\frac{\pi}{2} \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Each point p in space has three coordinates, (p_x, p_y, p_z). Each of the coordinates can be moved while rotating the point. This is the three-dimensional equivalent of the two-dimensional rotation matrix, which describes a counterclockwise rotation by an angle θ about the z-axis, perpendicular to the $x - y$ plane:

$$R_z(\theta) = \begin{pmatrix} cos\theta & -sin\theta & 0 \\ sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The axis of rotation and the angle of rotation are the arguments of M. A general form of a three-dimensional rotation matrix is denoted by $R_n\theta$, where n is one of the three axes of rotation, and a rotation angle θ. Although matrices are generally multiplicatively non-commutative, if both rotations are taken with respect to the same axis, this formula may be applied:

$$(R_n\theta_1)(R_n\theta_2) = R_n(\theta_1 + \theta_2)$$

The traditional rotational system

The Euler Angles is the most used system of describing rotations in three dimensions. Let us investigate their working. Euler's rotation theorem states that *any rotation may be described using three angles*. The Euler Angles are a sequence of three angles, which are usually denoted by $\phi,\ \theta$ and ψ. Here is how they are defined:

Consider two sets of the x, y and z axes. One of them is fixed in place, while the other set of axes is free to move and rotate in 3D space. Initially, the two sets of axes coincide. The stationary set will be referred to as the fixed frame, and the second set will be referred to as the mobile frame. To define a rotation, I add another frame to the system, with the origin at the same place as the other two frames. That frame is the final rotation I wish to describe. In order to rotate the mobile frame to the final frame, I use rotation matrices to rotate the mobile frame in this order:

1. About the x axis of the fixed frame by ϕ.
2. About the y axis of the fixed frame by θ.

3. About the z axis of the fixed frame by ψ.

Although you can use any order you want granted that you use it consistently, this order is most often used by convention. Using simple trigonometry of arctangents, I can get the exact angles I need to rotate by, and using the rotation matrices I get the three steps in that order in the language of matrices:

1. $R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\phi & -sin\phi \\ 0 & sin\phi & cos\phi \end{bmatrix}$

2. $R_y(\theta) = \begin{bmatrix} cos\theta & 0 & sin\theta \\ 0 & 1 & 0 \\ -sin\theta & 0 & cos\theta \end{bmatrix}$

3. $R_z(\psi) = \begin{bmatrix} cos\psi & -sin\psi & 0 \\ sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$

These three steps in this order can be represented by one matrix after matrix multiplication:

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\phi & -sin\phi \\ 0 & sin\phi & cos\phi \end{bmatrix} \begin{bmatrix} cos\theta & 0 & sin\theta \\ 0 & 1 & 0 \\ -sin\theta & 0 & cos\theta \end{bmatrix} \begin{bmatrix} cos\psi & -sin\psi & 0 \\ sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} cos\theta & 0 & sin\phi \\ sin\phi sin\theta & cos\phi & -sin\phi cos\theta \\ -cos\phi sin\theta & sin\phi & cos\phi cos\theta \end{bmatrix} \begin{bmatrix} cos\psi & -sin\psi & 0 \\ sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} cos\theta cos\psi & -cos\theta sin\psi & sin\theta \\ sin\phi sin\theta cos\psi + cos\phi sin\psi & cos\phi cos\psi - sin\phi sin\theta cos\theta sin\psi & -sin\phi cos\theta \\ sin\phi sin\psi - cos\phi sin\theta cos\psi & cos\phi sin\theta sin\psi + sin\phi cos\psi & cos\phi cos\theta \end{bmatrix}$$

This way of describing rotations is simpler than using 4-dimensional algebra, yet it falls victim to a problem known as the gimbal lock. If when performing the rotations two of the axes are rotated into a parallel configuration, the system loses one axis of freedom. Let us demonstrate this problem by examining what happens as $\theta = \frac{\pi}{2}$.

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\phi & -sin\phi \\ 0 & sin\phi & cos\phi \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} cos\psi & -sin\psi & 0 \\ sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & 0 & 1 \\ sin\phi cos\psi + cos\phi sin\psi & -sin\phi sin\psi + cos\phi cos\psi & 0 \\ -cos\phi cos\psi + sin\phi sin\psi & cos\phi sin\psi + sin\phi cos\psi & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & 0 & 1 \\ \sin(\phi + \psi) & \cos\,(\phi + \psi) & 0 \\ -\cos(\phi + \psi) & \sin\,(\phi + \psi) & 0 \end{bmatrix}$$

I observe that changing the values of both ψ and ϕ has the same effect – rotating by the combined angle $(\phi + \psi)$. The rotation remains around the z axis, as the last column and the first row in the matrix won't change. As long as $\theta = \frac{\pi}{2}$, and will not have different roles. When rotating by θ around the y-axis, both other angles decide about rotation around the z-axis. I lost the ability to rotate around the x-axis. It is possible to setup the rotations in other order than the $x -> y -> z$ convention, but even then, there will always be a different such a critical angle that causes the gimbal lock. This problem causes real issues when applying Euler angles-based rotations to solve engineering problems, one of the most famous examples happened during the Apollo 11 Moon mission. The spacecraft rotated in a way that made the axes align and lost a degree of freedom in its movement. The spacecraft had to be urgently realigned manually. The problem of a gimbal lock proves to be a danger for some use cases.

The quaternion rotational system

Let's examine the alternative way to define rotations, free of the problems plaguing the Euler angles system. Any vector in the three-dimensional space can be expressed as a pure quaternion.

$$Q_{pure} = 0 + xi + yj + zk$$

A rotation is expressed in terms of unit quaternions ($Q \in \mathbb{H}: |Q| = 1$). A rotation from one coordinate set A to another coordinate set B is given by multiplying the initial coordinate frame by a unit quaternion from the left, and by the conjugate of the unit quaternion from the right. This order matters, as quaternion multiplication is non-commutative.

$$Q_B = U_Q Q_A U_Q^*$$

This operation results in a pure quaternion Q_B, which can be checked by applying standard quaternion multiplication rules, as demonstrated here:

$$\begin{aligned} U_Q Q_A U_Q^* &= (\lambda_0 + \lambda_1 i + \lambda_2 j + \lambda_3 k)(xi + yj + zk)(\lambda_0 - \lambda_1 i - \lambda_2 j - \lambda_3 k) \\ &= \left(x(\lambda_0^2 + \lambda_1^2 - \lambda_2^2 - \lambda_3^2) + 2y(\lambda_1\lambda_2 - \lambda_0\lambda_3) + 2z(\lambda_0\lambda_2 + \lambda_1\lambda_3)\right) i \\ &\quad + \left(2x(\lambda_0\lambda_3 + \lambda_1\lambda_2) + y(\lambda_0^2 - \lambda_1^2 + \lambda_2^2 - \lambda_3^2) + 2z(\lambda_2\lambda_3 - \lambda_0\lambda_1)\right) j \\ &\quad + \left(2x(\lambda_1\lambda_3 - \lambda_0\lambda_2) + 2y(\lambda_0\lambda_1 + \lambda_2\lambda_3) + z(\lambda_0^2 - \lambda_1^2 - \lambda_2^2 + \lambda_3^2)\right) k \end{aligned}$$

This product may be represented as a 3 by 3 rotation matrix M

$$M = \begin{bmatrix} (\lambda_0^2 + \lambda_1^2 - \lambda_2^2 - \lambda_3^2) & 2(\lambda_1\lambda_2 - \lambda_0\lambda_3) & 2(\lambda_0\lambda_2 + \lambda_1\lambda_3) \\ 2(\lambda_0\lambda_3 + \lambda_1\lambda_2) & (\lambda_0^2 - \lambda_1^2 + \lambda_2^2 - \lambda_3^2) & 2(\lambda_2\lambda_3 - \lambda_0\lambda_1) \\ 2(\lambda_1\lambda_3 - \lambda_0\lambda_2) & 2(\lambda_0\lambda_1 + \lambda_2\lambda_3) & (\lambda_0^2 - \lambda_1^2 - \lambda_2^2 + \lambda_3^2) \end{bmatrix}$$

I know that the norm of U_Q is equal to one, from which I can derive the four following equations:

$$\lambda_0^2 + \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1$$

$$-\lambda_2^2 - \lambda_3^2 = \lambda_0^2 + \lambda_1^2 - 1$$

$$-\lambda_1^2 - \lambda_3^2 = \lambda_0^2 + \lambda_2^2 - 1$$

$$-\lambda_1^2 - \lambda_2^2 = \lambda_0^2 + \lambda_3^2 - 1$$

Substituting these equations into the matrix, it can be simplified into this form:

$$M = 2\begin{bmatrix} \lambda_0^2 + \lambda_1^2 - 0.5 & \lambda_1\lambda_2 - \lambda_0\lambda_3 & \lambda_0\lambda_2 + \lambda_1\lambda_3 \\ \lambda_0\lambda_3 + \lambda_1\lambda_2 & \lambda_0^2 + \lambda_2^2 - 0.5 & \lambda_2\lambda_3 - \lambda_0\lambda_1 \\ \lambda_1\lambda_3 - \lambda_0\lambda_2 & \lambda_0\lambda_1 + \lambda_2\lambda_3 & \lambda_0^2 + \lambda_3^2 - 0.5 \end{bmatrix}$$

Given such a rotation matrix M, in order to compute the quaternion which represents the same rotation I need to compute its *Trace*, that is the sum of the elements on the diagonal line of M.

$$\begin{aligned} Trace(M) &= M_{11} + M_{22} + M_{33} \\ &= 2(\lambda_0^2 + \lambda_1^2 - 0.5 + \lambda_0^2 + \lambda_2^2 - 0.5 + \lambda_0^2 + \lambda_3^2 - 0.5) = 2(3\lambda_0^2 + \lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 1.5) \\ &= 2(3\lambda_0^2 + (1 - \lambda_0^2) - 1.5) \\ &= 2(2\lambda_0^2 - 0.5) \\ &= 4\lambda_0^2 - 1 \end{aligned}$$

Solving this equation gives us λ_0:

$$|\lambda_0| = \sqrt{\frac{Trace(M) + 1}{4}}$$

In order to obtain λ_1, λ_2 and λ_3, I substitute λ_0 into M_{11}, M_{22}, and M_{33} respectively:

$$|\lambda_1| = \sqrt{\left(\frac{M_{11}}{2} + \frac{1 - Trace(M)}{4}\right)}$$

$$|\lambda_2| = \sqrt{(\frac{M_{22}}{2} + \frac{1 - Trace(M)}{4})}$$

$$|\lambda_3| = \sqrt{(\frac{M_{33}}{2} + \frac{1 - Trace(M)}{4})}$$

I now have all the tools necessary to translate a rotation matrix into a quaternion: Let's look at the general case of representing the counterclockwise rotation by an angle about the z-axis in the form of a quaternion

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

I apply the previously derived formulas

$$Trace(R) = 2cos\theta + 1$$

$$|\lambda_0| = \sqrt{\frac{2cos\theta + 1 + 1}{4}} = \sqrt{\frac{1 + cos\theta}{2}} = \cos\frac{\theta}{2},$$

$$|\lambda_1| = |\lambda_2| = \sqrt{\frac{cos\theta}{2} + \frac{1 - (2cos\theta + 1)}{4}} = 0,$$

$$|\lambda_3| = \sqrt{\frac{1}{2} + \frac{1 - (2cos\theta + 1)}{4}} = \sqrt{\frac{1 - cos\theta}{2}} = \sin\frac{\theta}{2}.$$

Therefore, the unit quaternion which rotates a point by an angle ψ around the z axis is:

$$U_{Q\mathrm{z}} = \cos\frac{\psi}{2} + ksin\frac{\psi}{2}.$$

In a similar way I compute the quaternions corresponding to rotations around the y and x axes:

$$U_{Qy} = \cos\frac{\theta}{2} + jsin\frac{\theta}{2},$$

$$U_{Q\mathrm{x}} = \cos\frac{\phi}{2} + isin\frac{\phi}{2}.$$

Let us rotate the point (1,0,0) by $\frac{\pi}{2}$ around the z-axis as an example. If our assumptions are correct, our point p should end up at (0,1,0)

The rotation matrix is:

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Using the previously derived formulas:

$$Trace(R) = 0 + 0 + 1 = 1$$

$$|\lambda_0| = \sqrt{\frac{1+1}{4}} = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2},$$

$$|\lambda_1| = \sqrt{\frac{R_{11}}{2} + \frac{1 - Trace(R)}{4}} = \sqrt{0+0} = 0,$$

$$|\lambda_2| = \sqrt{\frac{R_{22}}{2} + \frac{1 - Trace(R)}{4}} = \sqrt{0+0} = 0,$$

$$|\lambda_3| = \sqrt{\frac{R_{33}}{2} + \frac{1 - Trace(R)}{4}} = \sqrt{\frac{1}{2} + 0} = \frac{\sqrt{2}}{2}.$$

Therefore:

$$U_Q = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}k$$

I can obtain the same result using the other formula we've derived:

$$U_{Q\hat{z}} = \cos\frac{\pi}{4} + k sin\frac{\pi}{4} = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}k$$

This is the pure unit quaternion which defines this particular rotation! With the help of the fourth dimension, I can compute rotations. The point p in quaternion notation is $0 + i + 0j + 0k$, therefore I compute the rotation as:

$$U_Q \cdot i \cdot U_Q^* =$$

$$\left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}k\right) i \left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}k\right) = \left(\frac{\sqrt{2}}{2}i - \frac{\sqrt{2}}{2}ki\right)\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}k\right) = \left(\frac{\sqrt{2}}{2}i - \frac{\sqrt{2}}{2}j\right)\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}k\right)$$

$$= \frac{1}{2}(i - ik + j - jk) = \frac{1}{2}(i + j + j - i) = j$$

The pure quaternion j may be represented as the point $(0,1,0)$, which is what I wanted to obtain.

Conclusion

The research question investigated in this essay was: **To what extent is a four-dimensional number system best suited to define rotation in a three-dimensional space?** By demonstrating in detail why there cannot exist a three-dimensional triplet extension of the complex numbers which behaves in a mathematically interesting and applicable manner we've arrived at a conclusion that the four-dimensional quaternion number system is the first natural extension of the complex numbers. Investigating their mathematical properties yielded very interesting results regarding their non-commutativity during multiplication, akin to the later introduced concept of matrix multiplication. Nevertheless, the numbers have been proven to be operational under division and to obey a form of a norm function projecting onto the real number line. Defining a unit quaternion as a quaternion with a norm equal to 1 allowed us to form a four-dimensional unit sphere which was essential upon expanding the topic to the applications to three-dimensional rotation. Every rotation in three-dimensions has been mapped to a certain point on the four-dimensional unit sphere, which could be accessed by multiplying an arbitrary point by the rotation quaternion from the left and its reciprocal from the right. This method of describing rotations does not fall short under very specific circumstances which can not be said for the widely used Euler angles system.

In conclusion, although Euler angles have their merit in their ease of use and low-entry level, the technique falls short under very specific circumstances known as the gimbal lock. The Quaternion rotations are immune to these errors and can be used to describe rotations just as precisely. Quaternions are being used in modelling 3D space in virtual reality and it is very successful. Quaternions are a stellar example of mathematics discovered many decades ago finding use cases in modern times. The problem investigated could be expanded by further research on representing rotations in four-dimensional space and by investigating the applications of higher-dimensional number systems such as the aforementioned 8-dimensional number system.

Bibliography

1. Bannon, Thomas. "The Origin of Quaternions". *The College Mathematics Journal*, vol 46, no. 1, 2015, pp. 43-50. *Informa UK Limited*, https://doi.org/10.4169/college.math.j.46.1.43.

2. Ben-Ari, Mordechai. "*A Tutorial on Euler Angles and Quaternions*". 2014.

3. Hamilton, William Rowan. "LXXVIII. On Quaternions; Or on A New System of Imaginaries in Algebra". *The London, Edinburgh, And Dublin Philosophical Magazine and Journal of Science*, vol 25, no. 169, 1844, pp. 489-495. *Informa UK Limited*, https://doi.org/10.1080/14786444408645047.

4. Hanson, Andrew J. *Visualizing Quaternions*. Morgan Kaufmann Publishers Is an Imprint of Elsevier, 2006.

5. Savin, Diana. "About Division Quaternion Algebras and Division Symbol Algebras". *Carpathian Journal of Mathematics*, vol 32, no. 2, 2016, pp. 233-240. *Technical University of Cluj Napoca, North University Center of Baia Mare*, https://doi.org/10.37193/cjm.2016.02.11.

6. Van Der Waerden, B. L. "Hamilton's Discovery of Quaternions". *Mathematics Magazine*, vol 49, no. 5, 1976, pp. 227-234. *Informa UK Limited*,

7. Weiner, Joel L., and George R. Wilkens. "Quaternions And Rotations In E^4". *The American Mathematical Monthly*, vol 112, no. 1, 2005, p. 69. JSTOR, https://doi.org/10.2307/30037387.

8. Williams, Hugh. "On Quaternions and Octonions: Their Geometry, Arithmetic, And Symmetry, By John H. Conway and Derek A. Smith. Pp. 159. $29. 2003. ISBN 1 5688 134 9 (A. K. Peters).". *The Mathematical Gazette*, vol 88, no. 512, 2004, pp. 357-357. *Cambridge University Press (CUP)*, https://doi.org/10.1017/s002555720017545x.

9. Wisnesky, Ryan. "The Forgotten Quaternions". *Philosophy 300*, 2004.

10. Bushman, Nathan. “Hypercomplex Numbers and Early Vector Systems: A History”. 2020. *The Ohio State University,* Master’s.