The Origin of Quaternions
Author(s): Thomas Bannon
Source: *The College Mathematics Journal*, Vol. 46, No. 1 (January 2015), pp. 43-50
Published by: Mathematical Association of America
Stable URL: http://www.jstor.org/stable/10.4169/college.math.j.46.1.43
Accessed: 07-01-2018 22:17 UTC

The Origin of Quaternions

Thomas Bannon

Thomas Bannon (tbannon@adelphi.edu) is a retired New York City high school mathematics teacher. He currently works as an adjunct at Adelphi University in Garden City, New York. He has a B.S. in mathematics education from St. Francis College, a M.A. in mathematics from Queens College, a M.S. in computer science from the New York Institute of Technology, and a D. A. in mathematics from Adelphi University under the direction of Robert Bradley. All his free time is taken up enjoying his new granddaughter, Abby.

On October 16, 1843 while walking with his wife along the Royal Canal in Dublin, Sir William Rowan Hamilton had an insight into a problem he had been working on for over a decade. Sir William was trying unsuccessfully to develop a theory of triplets. His goal was to define operations on ordered triplets that would obey the laws of real number arithmetic but he was unable to find a satisfactory definition of multiplication. He believed his insight would lead to a solution of this problem. He was so elated that he immediately carved onto the side of a bridge the basic equations that governed the behavior of the mathematical entity that he called quaternions. This dramatic scene ensured Hamilton's contribution would be remembered long after quaternions faded in importance.

What makes Hamilton's discovery doubly interesting though is that the next day he wrote a letter to a fellow Irish mathematician and friend John T. Graves. In this letter he spelled out the thought processes including the dead ends that led to his eureka moment. This letter gives a valuable insight into the deliberations of a great mathematician as he struggles with a difficult problem. Fortunately this letter was published [**3**] and is now available electronically.

Complex numbers

Before we look at this letter though we have to consider what led Hamilton to this problem in the first place. By mid-nineteenth century mathematicians were aware of the geometric properties of complex numbers. These were developed first by the Norwegian born, Danish mathematician Caspar Wessel and later independently by the French amateur mathematician Jean-Robert Argand. In the letter Hamilton states that he is familiar with the English mathematician John Warren's work on the subject [**7**]. Warren had identified line segments extending from the origin with complex numbers and showed how geometric constructions could be used to perform arithmetic operations on these segments. Hence complex numbers could be identified with a directed line segment or vector from the origin to a point in the plane. We can see that associated with each complex number is an angle and a complex number can be

http://dx.doi.org/10.4169/college.math.j.46.1.43
MSC: 01A55, 17A35

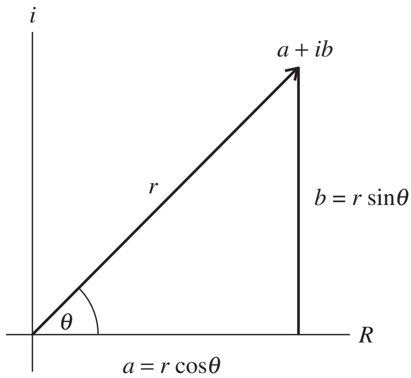

Figure 1. Polar form of the complex number $a + ib = r\cos\theta + ir\sin\theta$.

written in terms of that angle and the length (modulus) of the line segment. That is, $a + ib = r(\cos\theta + i\sin\theta)$ where $r = \sqrt{a^2 + b^2}$; see Figure 1.

Once this is done we can multiply complex numbers as follows.

$$\begin{aligned}
&[r_1(\cos\theta_1 + i\sin\theta_1)][r_2(\cos\theta_2 + i\sin\theta_2)]\\
&\quad = r_1r_2[\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)]\\
&\quad = r_1r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)]
\end{aligned}$$

This remarkable result means that a complex number of unit length $\cos\theta + i\sin\theta$ can be used as a rotation operator for two dimensional vectors. A vector represented by a complex number can be rotated by an angle θ simply by multiplying its complex number by $\cos\theta + i\sin\theta$.

$$(\cos\theta + i\sin\theta)(r\cos\phi + ir\sin\phi) = r\cos(\theta + \phi) + ir\sin(\theta + \phi)$$

In particular multiplication by the imaginary number $i = \cos(\pi/2) + i\sin(\pi/2)$ results in a rotation of $\pi/2$, as shown in Figure 2.

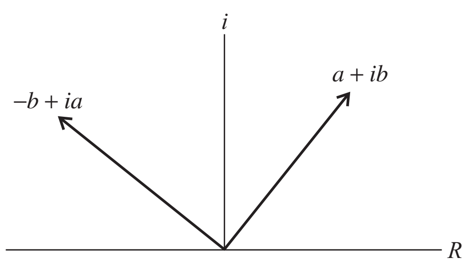

Figure 2. Multiplication of the complex number $a + bi$ by i.

One of Hamilton's early successes dealt with complex numbers. He replaced a complex number with a couplet or an ordered pair of real numbers. He then defined addition and multiplication of these couplets using the same rules as those for complex numbers without ever explicitly mentioning the imaginary number i. He showed that with these rules his couplets formed what we would now call a field. He noted that one of the reasons this could be done was that the "law of the modulus" was observed. In terms of complex numbers this meant that when two complex numbers were multiplied the product of their lengths equals the length of their product. In terms of Hamilton's couplets this law states that in order for multiplication of couplets to be validly defined as $(a, b)(x, y) = (ax - by, ay + bx)$ it must be true that $(a^2 + b^2)(x^2 + y^2) = (ax - by)^2 + (ay + bx)^2$, which is easily verified.

After his success with couplets Hamilton decided to see if he could define an algebra of triplets. He undoubtedly hoped that such an algebra would yield a rotation operator for three dimensions. Using hindsight we can see that his project was doomed to failure. A rotation in two dimensions is a rotation about a single axis normal to the plane. In three dimensions rotations can take place about three axes. In moving from two dimensions to three dimensions there is an increase of two degrees of freedom not one. While a two dimensional entity (a complex number) can be a rotation operator for a two dimensional vector, the analogous rotation operator for a three dimensional vector must have two additional dimensions and be a four dimensional entity.

He also assumed that multiplication of his triplets would be commutative. This was a reasonable but incorrect assumption for him to make. He had no reason to suspect that he was dealing with noncommutative multiplication here and his realization that he was might be classified as a mini-eureka moment.

The idea of the modulus of the complex number was extended to triplets and later to quaternions by Hamilton. Thus the modulus of a triplet (a, b, c) becomes $\sqrt{a^2 + b^2 + c^2}$ and that of a quaternion (a, b, c, d) becomes $\sqrt{a^2 + b^2 + c^2 + d^2}$. Hamilton insisted that a valid definition of multiplication must obey the law of the modulus which states that the product of the moduli of the factors must equal the modulus of the product.

Early attempts

We are now ready to follow the evolution of Hamilton's ideas as described in his letter to Graves. If you have not done so yet, now might be a good time to download the letter and use what follows as a guide to Hamilton's ideas. He begins by multiplying two triplets of the form $(a + ib + jc)$ where i and j are two different imaginary roots of -1. He also assumes multiplication is commutative so $ij = ji$.

$$\begin{aligned}&(a + ib + jc)(x + iy + jz)\\&\quad = ax - by - cz + i(ay + bx) + j(az + cx) + ij(bz + cy)\end{aligned}$$

He states that at this point he does not know how to handle the ij term. He then simplifies the problem by squaring a triplet and uses the law of the modulus to check if the multiplication is legitimate.

$$(a + ib + jc)^2 = a^2 - b^2 - c^2 + 2iab + 2jac + 2ijbc$$

$$(a^2 + b^2 + c^2)^2 = (a^2 - b^2 - c^2)^2 + (2ab)^2 + (2ac)^2$$

He notes that the law of the modulus will only hold here if the last term $2ijbc$ disappears. He toys with the idea of setting $ij = 0$ but feels uncomfortable doing so. He then states that he "perceives" that the last term would also disappear if $ij = -ji$. He decides to abandon his assumption that the multiplication is commutative and let $ij = k$ and $ji = -k$ leaving open the possibility that $k = 0$.

To further test his hypothesis he multiplies the special case

$$(a + ib + jc)(x + ib + jc) = ax - b^2 - c^2 + i(a + x)b + j(a + x)c + k(bc - bc)$$

and sees that the law of the modulus is satisfied. He also notes that analogous to complex numbers, an angle can be associated with each of the factors such that the sum of the angles of the factors equals the angle of the product. He does not supply all the details but this could be done as follows. Rewrite the product as

$$\left(a + \sqrt{b^2 + c^2}\,\vec{u}\right)\left(x + \sqrt{b^2 + c^2}\,\vec{u}\right) = ax - b^2 - c^2 + (a + x)\sqrt{b^2 + c^2}\,\vec{u}$$

where $\vec{u}$ is a unit vector in the direction $ib + jc$. Then let

$$\theta_1 = \tan^{-1}\frac{\sqrt{b^2 + c^2}}{a}, \quad \theta_2 = \tan^{-1}\frac{\sqrt{b^2 + c^2}}{x}, \quad \theta_3 = \tan^{-1}\frac{(a + x)\sqrt{b^2 + c^2}}{ax - b^2 - c^2}$$

which are related by

$$\tan(\theta_1 + \theta_2) = \frac{\frac{\sqrt{b^2+c^2}}{a} + \frac{\sqrt{b^2+c^2}}{x}}{1 - \frac{b^2+c^2}{ax}} = \frac{(a + x)\sqrt{b^2 + c^2}}{ax - b^2 - c^2} = \tan\theta_3.$$

We can only imagine that promising results like this are what kept Hamilton working on this problem for ten years.

He then finds that if he multiplies two general triplets using $ij = k$ and $ji = -k$,

$$\begin{aligned}(a + ib + jc)&(x + iy + jz)\\ &= ax - by - cz + i(ay + bx) + j(az + cx) + k(bz - cy),\end{aligned}$$

then the law of the modulus is satisfied,

$$\begin{aligned}(a^2 + b^2 + c^2)&(x^2 + y^2 + z^2)\\ &= (ax - by - cz)^2 + (ay + bx)^2 + (az + cx)^2 + (bz - cy)^2.\end{aligned}$$

The eureka moment

At this point he realizes that he may be able solve his problem by introducing a fourth dimension. He can treat k as another distinct imaginary unit with $k^2 = -1$, $ij = k$ and $ji = -k$. He dubs these four dimensional objects quaternions. Making use of the insight that and assuming that his three imaginary units would behave analogously when multiplied, he posits that $ij = -ji = k$, $jk = -kj = i$, and $ki = -ik = j$.

Hamilton carves the equation $i^2 = j^2 = k^2 = ijk = -1$ onto the bridge.

Up until now these are assumptions. He then undertakes the straightforward but tedious task of multiplying two general quaternions using these assumptions and verifying that the result obeys the law of the modulus. He is delighted to find that the terms not involving squares cancel and the remaining terms are just those needed to verify the law of the modulus, thus justifying his assumptions.

Hamilton then defines division. At this point he knows that quaternions under the operations he has defined obey all the laws of ordinary arithmetic except that multiplication is not commutative. In modern terminology they form a noncommutative division ring. The algebra works. He then turns his attention to the geometry.

He associates each quaternion with a modulus and an angle. For a quaternion $q = a + ib + jc + kd$ we define its modulus as $\mu = \sqrt{a^2 + b^2 + c^2 + d^2}$. Then $q = \mu(q_0 + iq_1 + jq_2 + kq_3)$ with $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$.

If we write $\vec{v} = iq_1 + jq_2 + kq_3$, $|\vec{v}| = \sqrt{q_1^2 + q_2^2 + q_3^2}$, and $\vec{u} = \vec{v}/|\vec{v}|$, then $q = \mu(q_0 + |\vec{v}|\,\vec{u})$ and, because $q_0^2 + |\vec{v}|^2 = 1$, there exists a ρ which Hamilton calls the amplitude of the quaternion such that $q = \mu(\cos\rho + (\sin\rho)\,\vec{u})$.

Hamilton then defines two other angles, the colatitude ϕ and the longitude φ. This allows him to write the unit vector $\vec{u} = i\cos\phi + j\sin\phi\cos\varphi + k\sin\phi\sin\varphi$; see Figure 3. Putting this all together we get the quaternion in the form

$$q = \mu\cos\rho + i\mu\sin\rho\cos\phi + j\mu\sin\rho\sin\phi\cos\varphi + k\mu\sin\rho\sin\phi\sin\varphi.$$

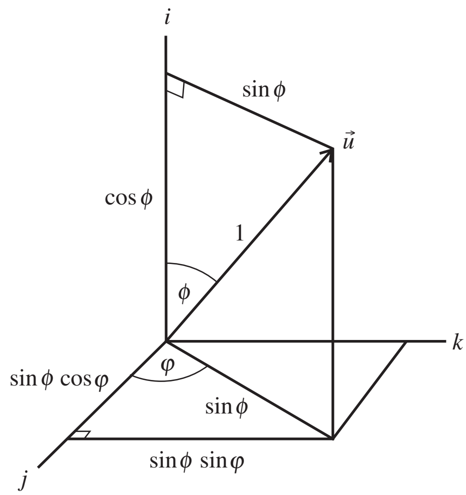

Figure 3. Coordinates of the unit vector $\vec{u}$.

Hamilton then looks at some special cases which arise in the multiplication of quaternions. If the amplitude of the quaternion is $\rho = \pi/2$, then the scalar part vanishes, leaving what he here calls a pure imaginary and will later be called a pure quaternion. He shows that the square of a pure imaginary is real and negative.

Quaternions versus vectors

Hamilton will later identify a pure quaternion with a vector and his followers will declare vector analysis to be superfluous because quaternions do all that is necessary with vectors. However, vectors would later be preferred to quaternions because they generalize easily to higher dimensions and they are easier to understand. Indeed most modern books dealing with quaternions, for example [**5**], explain them in terms of vectors. But we cannot underestimate the importance of quaternion theory in the development of vector analysis.

The cross and dot products arise naturally in the multiplication of quaternions. In the modern notation of vector analysis Hamilton's product of two pure imaginaries $p = 0 + \vec{p}$ and $q = 0 + \vec{q}$ can be written as $pq = -\vec{p}\cdot\vec{q} + \vec{p}\times\vec{q}$. At the end of the letter when Hamilton multiplies two pure imaginaries, he describes the result as follows.

> ...the product-line is perpendicular to the plane of the factors; its length is the product of their lengths multiplied by the sine of the angle between them: and the real part of the product, with its sign changed, is the same product of the lengths of the factors multiplied by the cosine of their inclination.

As we see he describes what would later be called the cross and dot products in detail.

We can be sure that Hamilton was confident that he would be able to find something analogous to the rotation operator of the complex numbers to rotate a vector. But he will not succeed completely. The problem is that multiplying a unit quaternion and a pure quaternion does not always result in a pure quaternion. Using modern notation the product of a quaternion $p = p_0 + \vec{p}$ and a pure quaternion $q = 0 + \vec{q}$ can be written as $pq = -\vec{p} \cdot \vec{q} + p_0 \vec{q} + \vec{p} \times \vec{q}$. We see that in order for the product to be a pure quaternion the dot product $\vec{p} \cdot \vec{q}$ must equal zero which means that the operator would have to be perpendicular to the vector.

Rodrigues's rotation operator

There is however a way of defining a quaternion rotation operator. Olinde Rodrigues, a French mathematician, solved the problem of rotating a three dimensional vector in 1840, three years before Hamilton discovered quaternions. His solution translated directly into defining a rotation operator for quaternions.

In terms of quaternions the solution is this. We pre-multiply the pure quaternion by a unit quaternion $q = a + ib + jc + kd$ and post-multiply it by the conjugate $q' = a - ib - jc - kd$. We will use the symbol $q(\cdot)q'$ to denote this operator. The product is a pure quaternion. The result is that the vector is rotated about the axis determined by $ib + jc + kd$. However, the amount of rotation is twice the amplitude ρ of the operator.

This led to a dilemma more philosophical than mathematical which plagued the early development of quaternions. There is a question as to which angle should naturally be associated with a quaternion. Should we keep ρ with its original definition as the amplitude of the quaternion or should we let ρ become the amount of rotation performed by the general operator? If we take the second option a unit quaternion would be written as

$$q = \cos\frac{\rho}{2} + i\sin\frac{\rho}{2}\cos\phi + j\sin\frac{\rho}{2}\sin\phi\cos\varphi + k\sin\frac{\rho}{2}\sin\phi\sin\varphi.$$

This is essentially what Rodrigues did although of course he did not formulate it in terms of quaternions. Hamilton came up with the same solution independently of Rodrigues (there is no evidence that he was aware of Rodrigues's work) but never published it. Even though the second option simplifies matters when working with the general rotation operator, Hamilton chose to stay with his original definition of ρ. For more information on Rodrigues's contributions see [**1**, introduction].

It is easy to surmise why Hamilton chose the first option. He considered quaternions first and foremost as algebraic entities. Their geometric properties would always come second in his mind. With the first option the value of ρ is $\pi/2$ for i, with the second $\rho = \pi$. Associating i with a rotation of $\pi/2$ works so well in the complex plane, it was difficult to abandon this idea. It was commonly believed that a rotation of 2π must restore an object to its original state. Hence, a rotation of 2π should be equivalent to one of 4π. Yet if the imaginary unit i is associated with a rotation of π, then $i^2 = -1$ is associated with a rotation of 2π and $i^4 = 1$ with one of 4π. Algebraically, i^2 could not be considered equivalent to i^4. The algebra and geometry do not agree.

Associating the imaginary units with rotations of $\pi/2$ could seem justified by the fact that multiplying i times the pure quaternion j results in a rotation of j by the amount $\pi/2$ to the pure quaternion k, mimicking what occurs in the complex plane See Figure 4. This however only works because i is perpendicular to j.

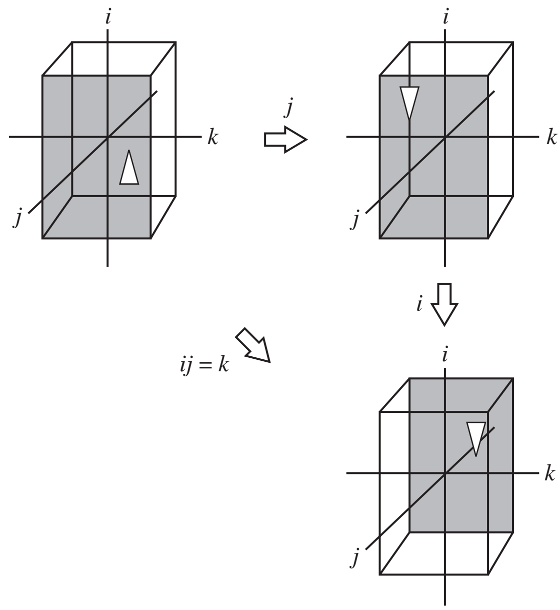

Figure 4. Rotations by π.

There is a compelling reason to associate all quaternions with twice the value of Hamilton's ρ and hence to associate the imaginary units with rotations of π. Suppose we wish to find an operator for a rotation that composes two rotations. The operator $r(\cdot)r'$ which composes the rotation performed by $p(\cdot)p'$ followed by the one performed by $q(\cdot)q'$ can be easily found by multiplying $r = qp$. (Note that the order of the multiplication is important since multiplication of quaternions is not commutative.) All of these operators lead to rotations with twice the value of Hamilton's ρ. So if we wish to compose a rotation about j followed by a rotation about i using Hamilton's equation $ij = k$, the amount of rotation associated with an imaginary unit must be π not $\pi/2$.

Later physicists discovered objects called spinors which have the property that a rotation by 2π places them in a negative state and a rotation of 4π restores them to their original state. The algebra and geometry of quaternion rotations agree in a reality, just not the familiar one. For more information on quaternion rotations and their relation to spinors see [**6**].

Quaternions today

Quaternions had a good run. Vectors gradually replaced quaternions in most applications but applications for quaternions still crop up. They are still used as an example of a non-commutative division ring and in the study of rotation groups [**1**]. Graves built on Hamilton's work to develop octonions [**2**]. Quaternions have recently been used in computer graphics because they solve the problem known as gimbal lock and make continuous rotations easier to program [**4**]. They will always ultimately be remembered though for the image of an exuberant man carving their equations onto a bridge.

Summary. We discuss a letter that Hamilton wrote a letter the day after he discovered quaternions. Describing what led to his discovery, it gives a valuable insight into the deliberations of a great mathematician as he struggles with a difficult problem.

References

1. A. Altmann, *Rotations, Quaternions, and Double Groups*. Dover, Mineola, NY, 2005.
2. J. C. Baez, The octonions, *Bull. Amer. Math. Soc. (N. S.)* **39** (2002) 145–205, http://dx.doi.org/10.1090/S0273-0979-01-00934-X.
3. W. R. Hamilton, (Copy of a) Letter from Sir William R. Hamilton to John Graves, Esq. on quaternions, *Philosophical Magazine, 3rd series* **25** (1884) 489–95, http://www.maths.tcd.ie/pub/HistMath/People/Hamilton/QLetter/QLetter.pdf.
4. A. Hanson, *Visualizing Quaternions*. Morgan Kaufmann, San Francisco, 2006.
5. J. Kuipers, *Quaternions and Rotation Sequences*. Princeton Univ. Press, Princeton, 1999.
6. R. Penrose, *The Road to Reality*. Knopf, New York, 2005.
7. J. Warren, *Treatise on the Geometric Representation of the Square Roots of Negative Quantities*. J. Smith, Cambridge, 1828.

Celebrants of Pi Day will be able to celebrate π correct to four decimal places on 3/14/15. Indeed, those with digital watches should revel at 9:26:53a.m., when they can claim nine places. Those with watches that indicate hundredths of a second can extend this, however fleetingly, two more places, and those of us who are continuously inclined can enjoy infinite agreement, which will not occur again for a century.

—John Lander Leonard, 1935–2014,
Emeritus, Department of Mathematics, University of Arizona

