QUATERNIONS AND BEYOND

NEIL MAKUR

Abstract. We briefly discuss the history of quaternions, and the story of their discovery. We go on to look at some of the definitions of quaternions, such as the algebra stemming from the original equation, and the treatment of the imaginary component as a vector. We then study the structure of quaternions, like how they fail to form a field, but form a division ring. In particular, we look at one very interesting subgroup of the quaternions. We end with covering some applications of quaternions, such as their application in quantum mechanics, and their role in one proof of Lagrange's Four Square Theorem.

1. History

Quaternions were first discovered by Sir William Rowan Hamilton while searching how to rotate in 3 dimensions. Similar to how $\mathbb{C}$ can be used to rotate objects in 2 dimensions by multiplying them by $e^{i\theta}$, Hamilton searched for a way to rotate objects in 3 dimensions using numbers of the form $a + bi + cj$, $a, b, c \in \mathbb{R}$. As the story goes, every day Hamilton's son would ask him "Well Papa, can you multiply triplets?", to which he would reply with "No, I can only add and subtract them.". Then, as he was walking along the Broome Bridge, Hamilton had a flash of insight, and carved the defining equation of quaternions into a stone

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

2. Definitions

2.1. **Using the Defining Equation.** We have already seen the defining equation of quaternions, that is $i^2 = j^2 = k^2 = ijk = -1$. Using these three values (i, j, k), and real values, we define a quaternion as follows.

Definition 2.1. A quaternion is a number of the form $a + bi + cj + dk$ where $a, b, c, d \in \mathbb{R}$, and i, j, and k satisfy the equation $i^2 = j^2 = k^2 = ijk = -1$. Addition and multiplication of quaternions are defined as outlined below, with multiplication distrubuting over addition.

We denote the quaternions as $\mathbb{H}$. Addition is done term-by-term, similar to complex numbers. Numerically,

$$(a + bi + cj + dk) + (e + fi + gk + hk) = (a + e) + (b + f)i + (c + g)j + (d + h)k.$$

Defining multiplication is a bit harder. To do this, firstly notice that $\frac{1}{i}, \frac{1}{j}, \frac{1}{k}$ are equal to $-i, -j, -k$ respectively. It is also worth noting that real numbers commute with quaternions in multiplication. However, when trying to figure out quaternion multiplication, we must not make the assumption that they commute, as they do not. If we look at the equation $i^2 = ijk$, we can see that $jk = i$. We can also see the following

$$jk = i \Rightarrow jki = i^2 = j^2 \Rightarrow ki = j$$

Date: April 2020.

Finally, if we consider the equation $k^2 = ijk$, we can see that $ij = k$. We can also show that that quaternions do not commute. Multiplying the equations $ij = k$, $jk = i$ and $ki = j$ by i, j, and k on the left respectively, we get that $ik = -j$, $ji = -k$ and $kj = -i$. Putting this all together, we get the following table.

×	i	j	k
i	-1	k	-j
j	-k	-1	i
k	j	-i	-1

This allows us to multiply quaternions using the regular properties of multiplication distributing over addition, and using the fact that the general form of a quaternion is $a + bi + cj + dk$ where $a, b, c, d \in \mathbb{R}$. After computation, the result is that

$$(a + bi + cj + dk)(e + fi + gj + hk) = (ae - fb - cg - dh)$$
$$+(af + be + ch - dg)i + (ag - bh + ce + df)j + (ah + bg - cf + de)k$$

2.2. **Using the Defining Equation-Other Operations.** Along with addition and multiplication, there are other operations that can be done on quaternions. One of these is conjugation. The conjugate of a quaternion, q is denoted $\overline{q}$, and is defined as $\overline{a + bi + cj + dk} = a - bi - cj - dk$. You can also take the norm of a quaternion using the formula

$$|a + bi + cj + dk| = \sqrt{a^2 + b^2 + c^2 + d^2},$$

or that $|q| = \sqrt{q\overline{q}}$. The function $N(q)$ denotes the squared norm of q. There are also inverses of quaternions,

$$q^{-1} = \frac{\overline{q}}{q\overline{q}} = \frac{\overline{q}}{|q|^2} = \frac{\overline{q}}{N(q)}.$$

Let us look at some examples of these. Let us consider the quaternion $q = 2 + 7i + j + 8k$. The conjugate of this is $\overline{q} = 2 - 7i - j - 8k$. Meanwhile, $N(q) = 2^2 + 7^2 + 1^2 + 8^2 = 118$, and $|q| = \sqrt{N(q)} = \sqrt{118}$. The inverse of q is $\frac{2}{118} - \frac{7}{118}i - \frac{1}{118}j - \frac{8}{118}k$.

2.3. **Using Vectors.** Let us consider operations with the unit vectors, $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ and the cross product. We know that $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}$, $\hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}}$, $\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}$, and that $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$. These properties are remarkably similar to quaternions, except that $\mathbf{v} \times \mathbf{v} = 0$, while $i^2 = j^2 = k^2 = -1$. However, there is an alternate definition of quaternions defined as $q = [s, \mathbf{v}]$ where the $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ components of $\mathbf{v}$ correspond to the i, j, k components of q respectively, and s corresponds to the real component. This new definition of quaternions defines addition as $[s, \mathbf{v}] + [s' + \mathbf{v'}] = [s + s', \mathbf{v} + \mathbf{v'}]$. It defines multiplication as $[s, \mathbf{v}] \cdot [s', \mathbf{v'}] = [ss' - \mathbf{v} \cdot \mathbf{v'}, \mathbf{v} \times \mathbf{v'} + s\mathbf{v'} + s'\mathbf{v}]$. Showing that the two definitions of quaternion multiplication that we have outlined are the same is left as an exercise to the reader. This notation of quaternions is useful when describing rotations in 3 dimensions, where we ignore s, and imagine quaternions as vectors.

2.4. **Using Vectors-Other Operations.** The non-addition and non-multiplication operations that we defined have analogs in this new definition. If $q = [s, \mathbf{v}]$, then $\overline{q} = [s, -\mathbf{v}]$. $N(q) = q\overline{q}$ is equal to $s^2 + \mathbf{v} \cdot \mathbf{v}$, and $|q| = \sqrt{N(q)} = \sqrt{s^2 + \mathbf{v} \cdot \mathbf{v}}$. Furthermore, $q^{-1} = [s/(s^2 + \mathbf{v} \cdot \mathbf{v}), -\mathbf{v}/(s^2 + \mathbf{v} \cdot \mathbf{v})]$. Although these are not immediately obvious, the fact that these definitions are equivalent is clear after a little vector algebra.

3. The Structure of Quaternions

3.1. **Quaternions as Groups.** Similar to the complex numbers, the quaternions, form an abelian group under addition. In fact, $\mathbb{H} \cong \mathbb{C} \times \mathbb{C}$. This is not very interesting, and we will not discuss it here. However, quaternions do form a group under multiplication, which we will show here.

Claim 3.1. *The quaternions form a group under multiplication.*

Proof. To show that the quaternions form a group, we will check the group axioms. First, it is easy to see that the identity of the quaternions is 1. To see closure, we will look at the formula that we discussed earlier

$$(a+bi+cj+dk)(e+fi+gj+hk) = (ae-fb-cg-dh)$$

$$+i(af+be+ch-dg)+j(ag-bh+ce+df)+k(ah+bg-cf+de)$$

This shows that the product of two quaternions is a quaternion. To see inverses, notice that $q^{-1} = \frac{1}{q}$, which we know is $\frac{\bar{q}}{N(q)}$. Finally, showing associativity requires a lot of algebra, but it can be shown that $(q_1q_2)q_3 = q_1(q_2q_3)$ where $q_1, q_2, q_3 \in \mathbb{H}$. The proof of this fact is left as an exercise to the reader. ■

3.2. **An Interesting Subgroup.** We can also see that the subset $\{\pm 1, \pm i, \pm j, \pm k\}$ forms a group under multiplication. This can be seen by noticing that the negative of each element is included, and by noticing that the product of two of i, j, k with one another is plus or minus the last element. We also already know that the associative property holds, and the identity, 1 is included in the group. We can also see that the proper subgroups of this are $\{1\}$, $\{\pm 1\}$, $\{\pm 1, \pm i\}$, $\{\pm 1, \pm j\}$, and $\{\pm 1, \pm k\}$. However, these subgroups have an interesting property.

Claim 3.2. *All subgroups of $\{\pm 1, \pm i, \pm j, \pm k\}$ are normal. That is, if $H \leq \{\pm 1, \pm i, \pm j, \pm k\}$, then $H \lhd \{\pm 1, \pm i, \pm j, \pm k\}$ (although this property is normally unremarkable, note that the group is nonabelian).*

Proof. It is easy to see that the subgroups $\{1\}$ and $\{\pm 1\}$ are normal subgroups. To show that the rest form a subgroup, we must show that $iji = j, iki = k, jij = i, jkj = k, kik = i, kjk = j$. Notice that we have excluded the negatives that come with taking the inverse of an element. However, this is okay because whenever i, j, or k is in a subgroup, its negative is also, and so the action of conjugation still results in a element of the same subgroup. Showing only these six facts allows us to show that every subgroup is normal, because it is trivial to show that all other elements remain in the subgroup under conjugation, as reals commute with quaternions. These six equations do not require too much thought and can be seen easily as seen below.

$$iji = (ij)i = (k)i = j \qquad iki = i(ki) = i(j) = k$$
$$jkj = (jk)j = (i)j = k \qquad jij = j(ij) = j(k) = i$$
$$kik = (ki)k = (j)k = i \qquad kjk = k(jk) = k(i) = j$$

Because reals commute with quaternions, this means that all subgroups of the group $\{\pm 1, \pm i, \pm j, \pm k\}$ are normal. ■

3.3. **Planes of Quaternions.** Let us consider a purely imaginary quaternion $u = u_1 i + u_2 j + u_3 k$ such that $N(u) = 1$. We will define the u-plane as below.

Definition 3.3. The u-plane is the subgroup of the quaternions spanned by the real numbers and a purely imaginary quaternion u with $N(u) = 1$. In other words, the u-plane is all quaternions that take the form $a + bu$ where $a, b \in \mathbb{R}$.

Note that the term "u-plane" is not a real term, and is only being defined so that our explanations are clear. However let us look at u^2. Some algebra reveals the following:

$$u^2 = (u_1 i + u_2 j + u_3 k)(u_1 i + u_2 j + u_3 k) =$$
$$-u_1^2 + u_1 u_2 k - u_1 u_3 j - u_1 u_2 k - u_2^2 + u_1 u_2 i +$$
$$u_1 u_3 j - u_1 u_2 i - u_3^2 = -(u_1^2 + u_2^2 + u_3^2) = -1.$$

This means that we can define an isomorphism ϕ from the u-plane to $\mathbb{C}$ with $\phi(a+bu) = a+bi$. This is because addition in the u plane is done termwise (which is not that hard to see), and because we have just shown that $(a + bu)(c + du) = (ac - bd) + (ad + bc)u$ (not exactly, but from $u^2 = -1$, this fact is trivial). To see that ϕ is a homomorphism, we can see the following:

$$\phi((a + bu)(c + du)) = \phi((ac - bd) + (ad + bc)u) =$$
$$(ac - bd) + (ad + bc)i = (a + bi)(c + di) = \phi(a + bu)\phi(c + du).$$

We can also see that

$$\phi((a + bu) + (c + du)) = \phi((a + c) + (b + d)u) = (a + c) + (b + d)i =$$
$$(a + bi) + (c + di) = \phi(a + bu) + \phi(c + du).$$

It is not hard to see that ϕ is injective and surjective, and therefore an isomorphism. Therefore, all u-planes (with u varying) are isomorphic to one another. An interesting fact that comes from this is that, even though $\mathbb{H}$ does not form a field, as we will see below, there are subsets of $\mathbb{H}$ that do.

3.4. **Quaternions as Rings.** We have seen that quaternions form a group under multiplication and addition. However, they do not form a field because they do not commute with each other using multiplication. Despite this, another algebraic structure that they form is a ring.

Definition 3.4. A ring is a set of elements together with two operations, addition ($+$) and multiplication ($\times$) such that the following axioms hold.

The set forms an abelian group under addition (we call the identity 0)
Multiplication is associative
The set is closed under multiplication
There exists an identity for multiplication (called 1)
Multiplication distributes over addition

Claim 3.5. *The quaternions form a ring.*

Proof. Similar to groups, we will check the ring axioms. We know that the quaternions form an abelian group under addition. We also know that the quaternions form a group under multiplication, so multiplication is associative, has an identity, and is closed under multiplication. Finally, we know that multiplication distributes over addition, so all the ring axioms have been met. ■

There is also a stronger algebraic structure, called a division ring, defined as follows.

Definition 3.6. A division ring is a ring such that there are multiplicative inverses for every element but 0. In other words, a division ring forms an abelian group under addition, a group under multiplication (excluding 0), and multiplication distributes over addition.

It is easy to see that the quaternions are a division ring, as they are a ring, and the inverse of a non-zero quaternion is given by $q^{-1} = \frac{\overline{q}}{N(q)}$.

4. Applications

4.1. **Rotations.** Hamilton originally came up with quaternions to describe rotations in three dimensions. To do this, we consider space as being spanned by the imaginary quaternions. In other words, for $r \in \mathbb{R}^3$, r can be represented as $xi + yj + zk$ where $x, y, z \in \mathbb{R}$. We can also look at analogs of Euler's formula in this new definition of $\mathbb{R}^3$. Similar to how

$$\cos(\phi) + i\sin(\phi) = e^{i\phi}$$

we say

$$\cos(\phi) + j\sin(\phi) = e^{j\phi}$$

and

$$\cos(\phi) + k\sin(\phi) = e^{k\phi}.$$

Although this seems a bit hand-wavy, note that the original proof of Euler's Formula uses the Taylor Series of e^x, $\sin(x)$ and $\cos(x)$, and substitutes $x = i\theta$. However, because $i^2 = j^2 = k^2$, we can just as easily substitute $x = j\theta$ or $x = k\theta$, and get the same result. We can do more than this, however. If we let $u = u_1 i + u_2 j + u_3 k$ such that $u_1^2 + u_2^2 + u_3^2 = N(u) = 1$, then we say that $\cos(\phi) + u\sin(\phi) = e^{u\phi}$. Proving this fact rigorously is actually not that difficult. If we look at u^2, we see the following:

$$u^2 = (u_1 i + u_2 j + u_3 k)(u_1 i + u_2 j + u_3 k) =$$
$$-u_1^2 + u_1 u_2 k - u_1 u_3 j - u_1 u_2 k - u_2^2 + u_1 u_2 i +$$
$$u_1 u_3 j - u_1 u_2 i - u_3^2 = -(u_1^2 + u_2^2 + u_3^2) = -1.$$

Therefore, letting $x = u\theta$ in the Taylor Series of e^x will give the same result as letting $x = i\theta$, as $i^2 = u^2$. This analog of Euler's Formula allows us to describe rotations as follows.

Theorem 4.1. *If u is any unit quaternion, and v is any quaternion, then the formula $e^{u\phi} v e^{-u\phi}$ rotates v about the axis in the direction of u by 2ϕ.*

Proof. We shall prove the theorem for $u = i$ because there is nothing special about the i direction (of course, it is part of the basis normally used for quaternions, but any basis works, so there really is nothing special). We shall start with the equation

$$e^{i\phi}(v_1 i + v_2 j + v_3 k)e^{-i\phi}.$$

We can pull out a j and get that this is equal to

$$e^{i\phi}(v_1 i + (v_2 + v_3 i)j)e^{-i\phi}.$$

From here, we can distribute to get

$$e^{i\phi}(v_1 i)e^{-i\phi} + e^{i\phi}(v_2 + v_3 i)je^{-i\phi}.$$

Next, notice that

$$j(\cos(\phi) - i\sin(\phi)) = \cos(\phi) + i\sin(\phi))j$$

(this is because $ji = -ij$). This means that we can rewrite our equation as

$$e^{i\phi}(v_1 i)e^{-i\phi} + e^{i\phi}(v_2 + v_3 i)e^{i\phi} j.$$

Looking at two numbers with no j or k terms, we can see that they commute (they are part of the i-plane, which we know is isomorphic to $\mathbb{C}$). This means that our product is equal to

$$v_1 i + e^{2i\phi}(v_2 + v_3 i)j.$$

After some thought, we can see that this does indeed rotate about the i axis by 2ϕ. ■

4.2. **Rotations in Quantum Mechanics.** This usage of quaternions has applications in quantum mechanics. In order to understand this, we will discuss some basics. In quantum mechanics, all things are described by a wavefunction, ψ. Therefore, if we consider what happens to the wavefunction of a particle, we will see what happens to the particle itself. First, we define

$$R_{u,\phi}(\mathbf{v}) = e^{u\phi/2}\mathbf{v}e^{-u\phi/2}.$$

One of the features of an electron is that is has a spin of half. Using quaternions, we can see that this is true. If we consider the electron's wavefunction, ψ, Pauli found that $\psi : \mathbb{R}^3 \to \mathbb{H}$. Rotating particles is normally done by $R_{u,\phi}(\psi(\mathbf{v})) = \psi(R_{u,\phi}(\mathbf{v}))$. However, Pauli found that for an electron,

$$R_{u,\phi}(\psi(\mathbf{v})) = e^{u\phi/2}\psi(R_{u,\phi}(\mathbf{v})).$$

If we consider $\phi = 2\pi$, we can see that

$$R_{u,2\pi}(\psi(\mathbf{v})) = e^{u2\pi/2}\psi(R_{u,2\pi}(\mathbf{v})) = e^{u\pi}\psi(\mathbf{v}) = -\psi(\mathbf{v})$$

This means that a rotation by 2π actually changes an electron. Note, however, that a rotation by 4π does not, leading to the notion of half-integer spin.

4.3. **Lagrange's Four Square Theorem.** Here, we will prove Lagrange's Four Square Theorem, which states that any integer can be written as the sum of 4 integer squares. However, to do this, we must prove a lemma first.

Lemma 4.2. *If x and y can both be written as the sum of four squares, then their product can also be written as the sum of four squares.*

Proof. We shall first show that $N(q_1q_2) = N(q_1)N(q_2)$. The rest will follow. If we let $q_1 = r_1e^{u\theta}$ and $q_2 = r_2e^{v\phi}$. Then $q_1q_2 = r_1r_2e^{u\theta+v\phi}$, and so $N(q_1q_2) = r_1r_2 = N(q_1)N(q_2)$. Therefore, if x is the sum of four squares, we can write it as the squared norm of a quaternion, say q_1, with integer coefficients. We can do the same with y, say q_2. So, if we want to find xy, we can rewrite this as $N(q_1)N(q_2)$, which is equal to $N(q_1q_2)$. Because q_1 and q_2 both have integer coefficients, q_1q_2 does too. Therefore, $N(q_1q_2) = xy$ is the sum of four squares. ■

While proving Lagrange's Four Square Theorem, we will need to use the Hurwitz quaternions, so it is worth talking about them

Definition 4.3. The Hurwitz quaternions are quaternions with either all integer or all half-integer coefficients. They take the form

$$\frac{E_0}{2}(1 + i + j + k) + E_1 i + E_2 j + E_3 k$$

where $E_0, E_1, E_2, E_3 \in \mathbb{Z}$.

Let us look at some examples of Hurwitz quaternions. One example is the quaternion $\frac{1}{2}+\frac{1}{2}i+\frac{1}{2}j+\frac{1}{2}k$, as it can be expressed using $E_0 = 1, E_{1,2,3} = 0$. We can also see that $3+i+4j+k$ is a Hurwitz quaternion by letting $E_0 = 6, E_1 = -2, E_2 = 1, E_3 = -2$. Another example is i, with $E_{0,2,3} = 0$ and $E_1 = 1$.However, notice that $\frac{1}{2}+5i+\frac{7}{2}j+4k$ is not an Hurwitz quaternion, as it can not be expressed in the form $\frac{E_0}{2}(1+i+j+k)+E_1i+E_2j+E_3k$. Note that the square of the norm of a Hurwitz quaternions is an integer.

Finally, we need one last theorem to prove Lagrange's Four Square Theorem, which we will take for granted. The theorem states that any odd prime p can be broken down into the product of two Hurwitz quaternions, α and β such that $N(\alpha), N(\beta) \neq 1$. A proof of this can be found at [Wik20]. Now, we are finally ready to prove Lagrange's Four Square Theorem.

Theorem 4.4. *Any integer can be written as the sum of four integer squares.*

Proof. We know that if two numbers can be written as the sum of four squares, then their product can also. This means that proving this theorem comes down to proving that all primes can be written as the sum of four squares. Firstly, we can see that $2 = 1^2+1^2+0^2+0^2$. This means that we now have to show this fact for all odd primes. We know that all odd primes p can be written as the product of two Hurwitz quaternions, α and β, such that $N(\alpha), N(\beta) \neq 1$. Therefore, looking at the equation $p = \alpha\beta$, we can take N of both sides to get $N(p) = p^2 = N(\alpha\beta) = N(\alpha)N(\beta)$. From here, we can see that both $N(\alpha)$ and $N(\beta)$ are equal to p. If α has integer coefficients, then we have proved that p is the sum of four squares. However, it is also possible that α has half-integer coefficients. If this is the case, we will choose

$$\omega = \frac{\pm 1 + \pm i \pm j \pm k}{2}$$

such that $\gamma = \omega+\alpha$ has even integer coefficients. Then, we know that $p = N(\alpha) = N(\gamma-\omega) = (\gamma-\omega)(\overline{\gamma-\omega})$. Note that $|\omega| = 1$, and so we can safely multiply by $\overline{\omega}\omega$. This gives us that $p = (\gamma-\omega)(\overline{\gamma-\omega})\overline{\omega}\omega = \overline{\omega}(\gamma-\omega)(\overline{\gamma-\omega})\omega$, because real numbers commute with quaternions. This is equal to $\overline{\omega}(\gamma-\omega)(\overline{\gamma}-\overline{\omega})\omega$. Multiplying out, we get $p = (\overline{\omega}\gamma-\overline{\omega})(\overline{\gamma}\omega-\omega) = (\overline{\omega}\gamma-1)(\overline{\gamma}\omega-1)$. Note that γ was defined to have even integer coefficients, and ω was defined to have half-integer coefficients, so $\overline{\omega}\gamma-1$ must have integer coefficients. Therefore, we can let $\alpha_1 = \overline{\omega}\gamma-1$ and $\beta_1 = \overline{\gamma}\omega-1$, and repeat our process with the equation $p = \alpha_1\beta_1$. However, we are now guaranteed that α_1 has integer coefficients, and so p, and thus any number, can be written as the sum of four squares. ∎

References

[Wik20] Wikipedia contributors. Lagrange's four-square theorem, 2020. [Online; accessed 1-June-2020].