Hamilton and Quaternions: A Case Study in Mathematical Research

Jeff Suzuki

Department of Mathematics
Brooklyn College

The Value of History, Part One

The history of any subject gives useful insights into where it comes from and where it might go.

The history of mathematics is unique:

If it ever worked, it still works.

It's a useful and usable history.

The Value of History, Part Two

Even more importantly:

By examining how mathematicians of the past created new mathematics, we can gain insight into the creative process:

The history of mathematics provides case studies of how to do mathematical research.

Seeing i to i

We can interpret the complex number $a + bi$ algebraically by treating it as a polynomial in i with the peculiar property that $i^2 = -1$.

Seeing i to i

We can interpret the complex number $a + bi$ algebraically by treating it as a polynomial in i with the peculiar property that $i^2 = -1$.

We can also interpret $a + bi$ geometrically by identifying it with the point (a, b).

Seeing i to i

We can interpret the complex number $a + bi$ algebraically by treating it as a polynomial in i with the peculiar property that $i^2 = -1$.

We can also interpret $a + bi$ geometrically by identifying it with the point (a, b).

Or can we?

Seeing i to i

We can interpret the complex number $a + bi$ algebraically by treating it as a polynomial in i with the peculiar property that $i^2 = -1$.

We can also interpret $a + bi$ geometrically by identifying it with the point (a, b).

Or can we?

We can add, subtract, multiply, and divide algebraic quantiites.

Seeing i to i

We can interpret the complex number $a + bi$ algebraically by treating it as a polynomial in i with the peculiar property that $i^2 = -1$.

We can also interpret $a + bi$ geometrically by identifying it with the point (a, b).

Or can we?

We can add, subtract, multiply, and divide algebraic quantiites.

But how do you add or multiply two points?

Turnabout

We can express

$$a + bi = r(\cos\theta + i\sin\theta)$$

This gives us a useful interpretion of z as a geometric transformation, namely

- Rotation around the origin by angle θ,
- Scaling of the distance to the origin by factor r

A Product of The Times

Now consider $z_2 z_1$.

A Product of The Times

Now consider $z_2 z_1$.

z_1 corresponds to some rotation θ_1 and some scaling r_1.

A Product of The Times

Now consider $z_2 z_1$.

z_1 corresponds to some rotation θ_1 and some scaling r_1.

z_2 corresponds to some rotation θ_2 and some scaling r_2.

A Product of The Times

Now consider $z_2 z_1$.

z_1 corresponds to some rotation θ_1 and some scaling r_1.

z_2 corresponds to some rotation θ_2 and some scaling r_2.

So, viewed as a geometric transformation, $z_2 z_1$ should correspond to

A Product of The Times

Now consider $z_2 z_1$.

z_1 corresponds to some rotation θ_1 and some scaling r_1.

z_2 corresponds to some rotation θ_2 and some scaling r_2.

So, viewed as a geometric transformation, $z_2 z_1$ should correspond to

- A rotation of $\theta_1 + \theta_2$ (the argument rule),

A Product of The Times

Now consider $z_2 z_1$.

z_1 corresponds to some rotation θ_1 and some scaling r_1.

z_2 corresponds to some rotation θ_2 and some scaling r_2.

So, viewed as a geometric transformation, $z_2 z_1$ should correspond to

- A rotation of $\theta_1 + \theta_2$ (the argument rule),
- A scaling by a factor of $r_1 r_2$ (the modulus rule).

A Product of The Times

Now consider $z_2 z_1$.

z_1 corresponds to some rotation θ_1 and some scaling r_1.

z_2 corresponds to some rotation θ_2 and some scaling r_2.

So, viewed as a geometric transformation, $z_2 z_1$ should correspond to

- A rotation of $\theta_1 + \theta_2$ (the argument rule),
- A scaling by a factor of $r_1 r_2$ (the modulus rule).

But does it?

The Greatest Tragedy in Life

We know what the modulus and argument of the product $z_1 z_2$ should be, if we view them as geometric transformations.

The Greatest Tragedy in Life

We know what the modulus and argument of the product $z_1 z_2$ should be, if we view them as geometric transformations.

We can also compute $z_1 z_2$ algebraically and find the modulus and the argument of the product.

The Greatest Tragedy in Life

We know what the modulus and argument of the product $z_1 z_2$ should be, if we view them as geometric transformations.

We can also compute $z_1 z_2$ algebraically and find the modulus and the argument of the product.

But NOTHING guarantees the two will be the same!

The Greatest Tragedy in Life

We know what the modulus and argument of the product $z_1 z_2$ should be, if we view them as geometric transformations.

We can also compute $z_1 z_2$ algebraically and find the modulus and the argument of the product.

But NOTHING guarantees the two will be the same!

Unfortunately . . .

The Greatest Tragedy in Life

We know what the modulus and argument of the product $z_1 z_2$ should be, if we view them as geometric transformations.

We can also compute $z_1 z_2$ algebraically and find the modulus and the argument of the product.

But NOTHING guarantees the two will be the same!

Unfortunately ...

... they are.

Three's Company

In the early 19th century, several mathematicians asked the question:

Three's Company

In the early 19th century, several mathematicians asked the question:

Could a self-consistent arithmetic of hypercomplex numbers $a + bi + cj$ be developed, where j is another unit?

Three's Company

In the early 19th century, several mathematicians asked the question:

Could a self-consistent arithmetic of hypercomplex numbers $a + bi + cj$ be developed, where j is another unit?

In particular, we want to

Three's Company

In the early 19th century, several mathematicians asked the question:

Could a self-consistent arithmetic of hypercomplex numbers $a + bi + cj$ be developed, where j is another unit?

In particular, we want to

- Identify a hypercomplex number with a geometric transformation,

Three's Company

In the early 19th century, several mathematicians asked the question:

Could a self-consistent arithmetic of hypercomplex numbers $a + bi + cj$ be developed, where j is another unit?

In particular, we want to

- Identify a hypercomplex number with a geometric transformation,
- Define multiplication that preserves the modulus and argument rules,

Three's Company

In the early 19th century, several mathematicians asked the question:

Could a self-consistent arithmetic of hypercomplex numbers $a + bi + cj$ be developed, where j is another unit?

In particular, we want to

- Identify a hypercomplex number with a geometric transformation,
- Define multiplication that preserves the modulus and argument rules,
- Retain other nice properties like associativity, distributivity, commutativity.

The js Have It

The obvious step is to identify $x + iy + jz$ with the point (x, y, z).

The js Have It

The obvious step is to identify $x + iy + jz$ with the point (x, y, z).

i corresponds to the transformation taking $(1, 0, 0)$ to $(0, 1, 0)$:

The *j*s Have It

The obvious step is to identify $x + iy + jz$ with the point (x, y, z).

i corresponds to the transformation taking $(1, 0, 0)$ to $(0, 1, 0)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xy-plane.

The js Have It

The obvious step is to identify $x + iy + jz$ with the point (x, y, z).

i corresponds to the transformation taking $(1, 0, 0)$ to $(0, 1, 0)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xy-plane.

j corresponds to the transformation taking $(1, 0, 0)$ to $(0, 0, 1)$:

The js Have It

The obvious step is to identify $x + iy + jz$ with the point (x, y, z).

i corresponds to the transformation taking $(1, 0, 0)$ to $(0, 1, 0)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xy-plane.

j corresponds to the transformation taking $(1, 0, 0)$ to $(0, 0, 1)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xz-plane.

The *j*s Have It

The obvious step is to identify $x + iy + jz$ with the point (x, y, z).

i corresponds to the transformation taking $(1, 0, 0)$ to $(0, 1, 0)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xy-plane.

j corresponds to the transformation taking $(1, 0, 0)$ to $(0, 0, 1)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xz-plane.

In both cases, applying the transformation twice takes us to $(-1, 0, 0)$.

The js Have It

The obvious step is to identify $x + iy + jz$ with the point (x, y, z).

i corresponds to the transformation taking $(1, 0, 0)$ to $(0, 1, 0)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xy-plane.

j corresponds to the transformation taking $(1, 0, 0)$ to $(0, 0, 1)$:

A rotation of $\frac{\pi}{2}$ about the origin in the xz-plane.

In both cases, applying the transformation twice takes us to $(-1, 0, 0)$.

So

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

Hijinks

Algebraically,

$$(x + iy + jz)(x' + iy' + jz') = x'' + iy'' + jz''$$

requires defining

$$ij = \alpha + i\beta + j\gamma$$

Hijinks

Algebraically,

$$(x + iy + jz)(x' + iy' + jz') = x'' + iy'' + jz''$$

requires defining

$$ij = \alpha + i\beta + j\gamma$$

This should correspond to the image of $(1, 0, 0)$ if we

Hijinks

Algebraically,

$$(x + iy + jz)(x' + iy' + jz') = x'' + iy'' + jz''$$

requires defining

$$ij = \alpha + i\beta + j\gamma$$

This should correspond to the image of $(1, 0, 0)$ if we

- Rotate by $\frac{\pi}{2}$ in the xz-plane,

Hijinks

Algebraically,

$$(x + iy + jz)(x' + iy' + jz') = x'' + iy'' + jz''$$

requires defining

$$ij = \alpha + i\beta + j\gamma$$

This should correspond to the image of $(1, 0, 0)$ if we

- Rotate by $\frac{\pi}{2}$ in the xz-plane,
- Rotate by $\frac{\pi}{2}$ in the xy-plane

Hijinks

Algebraically,

$$(x + iy + jz)(x' + iy' + jz') = x'' + iy'' + jz''$$

requires defining

$$ij = \alpha + i\beta + j\gamma$$

This should correspond to the image of $(1, 0, 0)$ if we

- Rotate by $\frac{\pi}{2}$ in the xz-plane,
- Rotate by $\frac{\pi}{2}$ in the xy-plane

These transformations take us to $(0, 0, 1)$, suggesting

$$ij = j$$

Hijinks

Algebraically,

$$(x + iy + jz)(x' + iy' + jz') = x'' + iy'' + jz''$$

requires defining

$$ij = \alpha + i\beta + j\gamma$$

This should correspond to the image of $(1, 0, 0)$ if we

- Rotate by $\frac{\pi}{2}$ in the xz-plane,
- Rotate by $\frac{\pi}{2}$ in the xy-plane

These transformations take us to $(0, 0, 1)$, suggesting

$$ij = j$$

Huh?

Concrete Doesn't Hurt

Let's try a different product.

Concrete Doesn't Hurt

Let's try a different product.

Consider

$$\left(0 + i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2$$

(a point on the unit sphere that's "halfway" up from the y-axis)

The Algebra

By analogy with $(a+b+c)^2$, we have

$$\begin{aligned}\left(0 + i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2 &= \left(i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2 \\ &= \frac{1}{2}i^2 + \frac{1}{2}j^2 + ij \\ &= -1 + ij\end{aligned}$$

Consequently, if our transformation takes us to (x', y', z'), we want

$$x' + iy' + jz' = -1 + ij$$

A Breakdown

We might view the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a

- Rotation of $\frac{\pi}{2}$ in the *xy*-plane,
- Followed by a rotation of $\frac{\pi}{4}$ towards the *z*-axis

A Breakdown

We might view the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a

- Rotation of $\frac{\pi}{2}$ in the xy-plane,
- Followed by a rotation of $\frac{\pi}{4}$ towards the z-axis

If do that twice, we'll end at $(0, 0, 1)$, requiring

$$\left(0 + i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2 = j$$

A Breakdown

We might view the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a

- Rotation of $\frac{\pi}{2}$ in the *xy*-plane,
- Followed by a rotation of $\frac{\pi}{4}$ towards the *z*-axis

If do that twice, we'll end at $(0, 0, 1)$, requiring

$$\left(0 + i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2 = j$$

So

$$-1 + ij = j$$

which means $ij = 1 + j$.

Everything, Everywhere, All At Once

What if we considered the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a single rotation?

Everything, Everywhere, All At Once

What if we considered the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a single rotation?

This would be a rotation of $\frac{\pi}{2}$ in the plane passing through the two points and the origin.

Everything, Everywhere, All At Once

What if we considered the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a single rotation?

This would be a rotation of $\frac{\pi}{2}$ in the plane passing through the two points and the origin.

If we do that twice, we'll end at $(-1, 0, 0)$, requiring

$$\left(0 + i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2 = -1$$

Everything, Everywhere, All At Once

What if we considered the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a single rotation?

This would be a rotation of $\frac{\pi}{2}$ in the plane passing through the two points and the origin.

If we do that twice, we'll end at $(-1, 0, 0)$, requiring

$$\left(0 + i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2 = -1$$

But this means

$$-1 + ij = -1$$

and so $ij = 0$.

Everything, Everywhere, All At Once

What if we considered the rotation from $(1, 0, 0)$ to $\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ as a single rotation?

This would be a rotation of $\frac{\pi}{2}$ in the plane passing through the two points and the origin.

If we do that twice, we'll end at $(-1, 0, 0)$, requiring

$$\left(0 + i\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right)^2 = -1$$

But this means

$$-1 + ij = -1$$

and so $ij = 0$.

But this is a diffeerent answer from before!

Associativity

This means that, as defined, our hypercomplex numbers FAIL associativity:

Associativity

This means that, as defined, our hypercomplex numbers FAIL associativity:

If we view z as a single rotation in a plane,

$$z^2 = (-1, 0, 0)$$

Associativity

This means that, as defined, our hypercomplex numbers FAIL associativity:

If we view z as a single rotation in a plane,

$$z^2 = (-1, 0, 0)$$

But using $z = z_1 z_2$, where z_2 was a rotation by $\frac{\pi}{2}$ in the xy plane and z_1 was a rotation towards the z-axis by $\frac{\pi}{4}$, then

$$z_1(z_2 z) = (0, 0, 1)$$

Associativity

This means that, as defined, our hypercomplex numbers FAIL associativity:

If we view z as a single rotation in a plane,

$$z^2 = (-1, 0, 0)$$

But using $z = z_1 z_2$, where z_2 was a rotation by $\frac{\pi}{2}$ in the xy plane and z_1 was a rotation towards the z-axis by $\frac{\pi}{4}$, then

$$z_1(z_2 z) = (0, 0, 1)$$

In other words

$$\underbrace{(z_1 z_2)}_{z} z \neq z_1(z_2 z)$$

Points of Failure

With somewhat more effort, we can show that the distributive property also fails for these numbers.

Even worse:

Points of Failure

With somewhat more effort, we can show that the distributive property also fails for these numbers.

Even worse:

NEITHER of the possible products retain the modulus rule:

$$ij = 1 + j$$
$$ij = 0$$

and neither $(1, 0, 1)$ nor $(0, 0, 0)$ is on the unit sphere.

“I Can Only Add and Subtract Them . . . ”

So an algebra of hypercomplex numbers seems to require sacrificing one of

“I Can Only Add and Subtract Them . . . ”

So an algebra of hypercomplex numbers seems to require sacrificing one of

- Associativity,

“I Can Only Add and Subtract Them . . .”

So an algebra of hypercomplex numbers seems to require sacrificing one of

- Associativity,
- Distributivity,

"I Can Only Add and Subtract Them ..."

So an algebra of hypercomplex numbers seems to require sacrificing one of

- Associativity,
- Distributivity,
- The modulus rule,

"I Can Only Add and Subtract Them . . . "

So an algebra of hypercomplex numbers seems to require sacrificing one of

- Associativity,
- Distributivity,
- The modulus rule,

. . . as long as we take the obvious interpretation of i, j as rotations.

“I Can Only Add and Subtract Them . . . ”

So an algebra of hypercomplex numbers seems to require sacrificing one of

- Associativity,
- Distributivity,
- The modulus rule,

. . . as long as we take the obvious interpretation of i, j as rotations.

So could we salvage these rules if we interpret them as more complicated transformations?

“I Can Only Add and Subtract Them . . .”

So an algebra of hypercomplex numbers seems to require sacrificing one of

- Associativity,
- Distributivity,
- The modulus rule,

. . . as long as we take the obvious interpretation of i, j as rotations.

So could we salvage these rules if we interpret them as more complicated transformations?

Several nineteenth century mathematicians wrestled with this problem for a long time until . . .

A Walk to Remember

On Monday, October 16, 1843 William Rowan Hamilton (1805-1865) and his wife were walking along the Royal Canal in Dublin to a meeting of the Royal Irish Academy.

A Walk to Remember

On Monday, October 16, 1843 William Rowan Hamilton (1805-1865) and his wife were walking along the Royal Canal in Dublin to a meeting of the Royal Irish Academy.

Near the Brougham Bridge, Hamilton had a crucial insight:

A Walk to Remember

On Monday, October 16, 1843 William Rowan Hamilton (1805-1865) and his wife were walking along the Royal Canal in Dublin to a meeting of the Royal Irish Academy.

Near the Brougham Bridge, Hamilton had a crucial insight:

The modulus and argument rules, as well as distributivity, could be saved by introducing

A Walk to Remember

On Monday, October 16, 1843 William Rowan Hamilton (1805-1865) and his wife were walking along the Royal Canal in Dublin to a meeting of the Royal Irish Academy.

Near the Brougham Bridge, Hamilton had a crucial insight:

The modulus and argument rules, as well as distributivity, could be saved by introducing

- Anticommutativity, $ij = -ji$,

A Walk to Remember

On Monday, October 16, 1843 William Rowan Hamilton (1805-1865) and his wife were walking along the Royal Canal in Dublin to a meeting of the Royal Irish Academy.

Near the Brougham Bridge, Hamilton had a crucial insight:

The modulus and argument rules, as well as distributivity, could be saved by introducing

- Anticommutativity, $ij = -ji$,
- A third imaginary k, with $ij = k$

A Bit of Vandalism

Hamilton was so taken by his epiphany that he carved the fundamental equations into one of the bridge's stones.

The original carving is no longer there, but there's a plaque marking the location.

Enter the Quaternions

Hamilton used the term quaternion for numbers of the form $x + iy + jz + kw$.

(The term was generally used for any set of four objects)

PROCEEDINGS

OF

THE ROYAL IRISH ACADEMY.

1843. No. 42.

November 13.

SIR WM. R. HAMILTON, LL.D., President, in the Chair.

The Chair having been taken *pro tem.* by the Rev. H. Lloyd, D. D., Vice-President,

The President read a paper on a new Species of Imaginary Quantities, connected with a theory of Quaternions.

It is known to all students of algebra that an imaginary equation of the form $i^2 = -1$ has been employed so as to conduct to very varied and important results. Sir Wm. Hamilton proposes to consider some of the consequences which result from the following system of imaginary equations, or equations between *a system of three different imaginary quantities:*

$$i^2 = j^2 = k^2 = -1; \quad \text{(A)}$$
$$ij = k, \quad jk = i, \quad ki = j; \quad \text{(B)}$$
$$ji = -k, \quad kj = -i, \quad ik = -j; \quad \text{(C)}$$

no linear relation between i, j, k being supposed to exist, so that the equation

$$Q = Q',$$

in which

$$Q = w + ix + jy + kz,$$
$$Q' = w' + ix' + jy' + kz',$$

Fade to Black

In time, it was recognized that a quaternion $x + iy + jz + kw$ could be split into a real component x and an imaginary component $iy + jz + kw$.

And the imaginary component could be interpreted as a vector.

Over time, the vector component came to be considered more important and quaternions faded into the background.

Fade to Black

In time, it was recognized that a quaternion $x + iy + jz + kw$ could be split into a real component x and an imaginary component $iy + jz + kw$.

And the imaginary component could be interpreted as a vector.

Over time, the vector component came to be considered more important and quaternions faded into the background.

Until . . .

Turning Things Around

Recall we ran into a problem when describing rotations in three dimensions:

If we decomposed a rotation, associativity might fail.

In particular, defining a rotation using spherical or cylindrical coordinates leads to a problem called gimbal lock.

Using quaternions avoids gimbal lock!

So quaternions have become important again as an essential part of computer graphics.

The Algebra of Time

In November 1940, G. H. Hardy (1877-1947) wrote:

> *The great bulk of higher mathematics is useless.*

The Algebra of Time

In November 1940, G. H. Hardy (1877-1947) wrote:

> *The great bulk of higher mathematics is useless.*

He viewed this as a good thing:

The Algebra of Time

In November 1940, G. H. Hardy (1877-1947) wrote:

> *The great bulk of higher mathematics is useless.*

He viewed this as a good thing:

> *When the world is mad, a mathematician may find in mathematics an incomparable anodyne.*

The Algebra of Time

In November 1940, G. H. Hardy (1877-1947) wrote:

> *The great bulk of higher mathematics is useless.*

He viewed this as a good thing:

> *When the world is mad, a mathematician may find in mathematics an incomparable anodyne.*

Hardy's view corresponds to the divide between "pure" and "applied" mathematics.

The Algebra of Time

In November 1940, G. H. Hardy (1877-1947) wrote:

> *The great bulk of higher mathematics is useless.*

He viewed this as a good thing:

> *When the world is mad, a mathematician may find in mathematics an incomparable anodyne.*

Hardy's view corresponds to the divide between "pure" and "applied" mathematics.

But as quaternions show, even "useless" mathematical objects eventually find a use.

The Algebra of Time

In November 1940, G. H. Hardy (1877-1947) wrote:

The great bulk of higher mathematics is useless.

He viewed this as a good thing:

When the world is mad, a mathematician may find in mathematics an incomparable anodyne.

Hardy's view corresponds to the divide between "pure" and "applied" mathematics.

But as quaternions show, even "useless" mathematical objects eventually find a use.

(Hardy's own examples of "useless" mathematics were number theory and general relativity:

The Algebra of Time

In November 1940, G. H. Hardy (1877-1947) wrote:

> *The great bulk of higher mathematics is useless.*

He viewed this as a good thing:

> *When the world is mad, a mathematician may find in mathematics an incomparable anodyne.*

Hardy's view corresponds to the divide between "pure" and "applied" mathematics.

But as quaternions show, even "useless" mathematical objects eventually find a use.

(Hardy's own examples of "useless" mathematics were number theory and general relativity: the first is the basis of modern cryptography, and the second is critical for GPS navigation)