Quaternions And Their Importance

0525921

23.04.07

Contents

1 Introduction to Quaternions

Definition 1. *A quaternion is defined to be an expression* $x_0 + x_1 i + x_2 j + x_3 k$, *with* $i^2 = j^2 = k^2 = -1$

$$\begin{array}{ccc} ij = k & jk = i & ki = j \\ ji = -k & kj = -i & ik = -j \end{array}$$

with $(x_0, x_1, x_2, x_3) \in \mathbb{R}^4$

This description was given by Hamilton, as he famously inscribed on the Brougham Bridge in a flash of ingenuity.

Definition 2 (Alternate notation)**.** *For ease, the quaternion* $q = a + bi + cj + dk \in \mathbb{H}$ *will now be written as* $q = [a, b, c, d] = [a, \mathbf{v}] = a + \mathbf{v}$, *where* $\mathbf{v}$ *is the vector* $Im(q) = (b, c, d) \in \mathbb{R}^3$.

1.1 The Initial Work

For years, Hamilton sought to achieve a "theory of triplets"[1] analogous to the theory of the standard complex numbers. To this purpose, he considered the existence of another number, akin to the i of complex theory in the sense that it was also perpendicular to the line 1, but a distinct square root of unity in its own right.

Calling this number j, he investigated the square of a linear combinations of these two square roots of -1 to achieve

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

Extending from the complex numbers, with i and j as the orthonormal basis vectors, the Euclidean norm of the square of a vector should be the square of the norm of the original linear combination...

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

...therefore, if he wanted conform to the Euclidean norm, the *2ijbc* term wouldn't be involved in the norming process. Also, we notice that $(a^2 - b^2 - c^2, 2ab, 2ac)$ is the so-called "square-point"[1], the point achieved from doubling the rotation direction from the positive x-axis, and giving it the squared magnitude of α. Hamilton perceived that this exclusion of the ij term may be achieved by letting $ij = k$, $ji = -k$ (he considered having $ij = 0$ i.e. zero divisors, but he decided it may be "de trop"). An investigation into another basic product gives...

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

...and this correctly gives the co-ordinates of the "product-point", so justifying the anti-commutative assumption. By trying the product of two general triplets, he noticed that the modulus of the product would only equal the product of the moduli if you included the k component when taking the modulus. This implies that k symbolises a separate spatial dimension altogether - leading to four dimensions being needed to represent triplets. This is the discovery that led to the the appellation of **quaternions** - from the Latin *quaternio*, meaning 'set of four', and the remark, "we must admit, in some sense, a fourth dimension of space for the purpose of calculating with triplets"[1]

1.2 Binary operations on quaternions

As has been implicitly stated, a quaternion can be written as a linear combination of the 3 distinct square roots of unity in the way $q = a + bi + cj + dk$, $a, b, c, d \in \mathbb{R}$. Define $A = x_0 + x_1 i + x_2 j + x_3 j = A_t + \mathbf{a}$, $B = y_0 + y_1 i + y_2 j + y_3 j = B_t + \mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$

Definition 3 (Addition)**.** *Addition of quaternions is defined component-wise.* $A + B = (x_0 + y_0) + (x_1 + y_1 i) + (x_2 + y_2)j + (x_3 + y_3)j$

Definition 4 (Conjugation)**.** *The conjugate of a quaternion* $A = x_0 + x_1 i + x_2 j + x_3 j$ *is defined to be* $\bar{A} = x_0 - x_1 i - x_2 j - x_3 j = A_t - \mathbf{a}$

Definition 5 (Inverse)**.** *The inverse of a quaternion q is defined to be*

$$q^{-1} = \frac{\bar{q}}{q\bar{q}}$$

Definition 6 (Modulus)**.** *The modulus of a quaternion is just the square root of the quaternion dot product,* $\|A\| = \sqrt{q \cdot q} = \sqrt{A_t^2 + \mathbf{a} \cdot \mathbf{a}}$

Exercise 1.1. *Show that* $\| \cdot \| : \mathbb{H} \to \mathbb{R}$ *defined by* $\|q\| = \sqrt{q \cdot q}$ *is a norm on* $\mathbb{H}$, *if we consider* $\sqrt{q \cdot q}$ *to mean "take the principal square root of* $q \cdot q$*".*

Working 1.2. $q := a + bi + cj + dk$

$\sqrt{q \cdot q} \geq 0$ *by definition of the square root.*

$\sqrt{q \cdot q} = 0 \Leftrightarrow q \cdot q = 0 \Leftrightarrow a^2 + b^2 + c^2 + d^2 = 0 \Leftrightarrow a = b = c = d = 0$, *as* $a, b, c, d \in \mathbb{R}$

For $e \in \mathbb{R}$, $\|eq\| = \sqrt{(ea)^2 + (eb)^2 + (ec)^2 + (ed)^2} = \sqrt{e^2(a^2 + b^2 + c^2 + d^2)} = \sqrt{e^2}\sqrt{a^2 + b^2 + c^2 + d^2} = |e|\|q\|$

Define two quaternions $q = [\alpha, a, b, c] = [\alpha, \mathbf{v}]$, $q\prime = [\beta, d, e, f] = [\beta, \mathbf{w}]$. *Now* $\|qq\prime\| = \|[\alpha, \mathbf{v}][\beta, \mathbf{w}]\| = \sqrt{[\alpha\beta - \mathbf{v} \cdot \mathbf{w}, \alpha\mathbf{w} + \beta\mathbf{v} + \mathbf{v} \times \mathbf{w}]}$

$= \sqrt{\alpha^2\beta^2 + (\mathbf{v} \cdot \mathbf{w})^2 - 2\alpha\beta(\mathbf{v} \cdot \mathbf{w}) + |(\alpha\mathbf{w} + \beta\mathbf{v} + \mathbf{v} \times \mathbf{w})|^2}$

This can be rather inelegantly expanded, and then reduced into the form

$= \sqrt{\alpha^2\beta^2 + \beta^2|\mathbf{v}|^2 + \alpha^2|\mathbf{w}|^2 + |\mathbf{v}|^2|\mathbf{w}|^2} = \sqrt{(\alpha^2 + |\mathbf{v}|^2)(\beta^2 + |\mathbf{w}|^2)}$

$= \sqrt{\alpha^2 + |\mathbf{v}|^2}\sqrt{\beta^2 + |\mathbf{w}|^2} = \|q\|\|q\prime\|$

1.2.1 Quaternion Products

Definition 7 (Grassman product)**.** *Let* A, B *be quaternions. The Grassman product is defined to be the quaternion Q s.t.*

$$Q = AB = A_t B_t - \mathbf{a} \cdot \mathbf{b} + A_t \mathbf{b} + B_t \mathbf{a} + \mathbf{a} \times \mathbf{b}$$

Exercise 1.3. *Find* $Q = AB$ *when* $A = i + 2k$, $B = 4 - 2i + j + k$

Working 1.4 (Answer)**.** $A_t B_t = 0$, $\mathbf{a} \cdot \mathbf{b} = -2 + 2 = 0$, $A_t \mathbf{b} = 0$, $B_t \mathbf{a} = 4i + 8k$, $\mathbf{a} \times \mathbf{b} = -2i - 5j + k \Rightarrow Q = 0 - 0 + 0 + 4i + 8k - 2i - 5j + k = 0 + 2i - 5j + 9k$

Also, $(i + 2k)(4 - 2i + j + k) = 4i - 2i^2 + ij + ik + 8k - 4ki + 2kj + 2k^2 = 4i + 2 + k - j + 8k - 4j - 2i - 2 = 0 + 2i - 5j + 9k$

Theorem 1.1. $(\mathbb{H}, +, \times)$ *is a non-commutative division ring, with* $\times$ *being the Grassman product, usually implied by juxtaposition.*

(a) Show ($\mathbb{H}$,+) is an abelian group. .
Define $q_i = [a_i, \mathbf{v}_i]$ for $i \in \{1,2,3\}$
(ai) Associativity... inherited from real numbers. (aii) Identity... $0_q = [0, \mathbf{0}]$
(aiii) Inverse... $-q_1 = [-a_1, -\mathbf{v}_1 \Rightarrow q_1 + (-q_1) = -q_1 + q_1 = [0, \mathbf{0}] = 0_q$
(aiv) Commutativity... $q_1+q_2 = [a_1+a_2, \mathbf{v}_1+\mathbf{v}_2] = [a_2+a_1, \mathbf{v}_2+\mathbf{v}_1] = q_2+q_1$ □

(b) Associativy of × over $\mathbb{H}$. .
(bi) $q_1(q_2q_3) = [a_1, \mathbf{v}_1][a_2a_3 - \mathbf{v}_2 \cdot \mathbf{v}_3, a_2\mathbf{v}_3 + a_3\mathbf{v}_2 + \mathbf{v}_2 \times \mathbf{v}_3] = [a_1a_2a_3 - a_1\mathbf{v}_2 \cdot \mathbf{v}_3 - a_2\mathbf{v}_1 \cdot \mathbf{v}_3 - a_3\mathbf{v}_1 \cdot \mathbf{v}_2 + \mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3), a_1a_2\mathbf{v}_3 + a_1a_3\mathbf{v}_2 + a_1\mathbf{v}_2 \times \mathbf{v}_3 + a_2a_3\mathbf{v}_1 - (\mathbf{v}_2 \cdot \mathbf{v}_3)\mathbf{v}_1 + a_2\mathbf{v}_1 \times \mathbf{v}_3 + a_3\mathbf{v}_1 \times \mathbf{v}_2 + \mathbf{v}_1 \times (\mathbf{v}_2 \times \mathbf{v}_3)]$

(bii) $(q_1q_2)q_3 = [a_1a_2 - \mathbf{v}_1 \cdot \mathbf{v}_2, a_1\mathbf{v}_2 + a_2\mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2][a_3, \mathbf{v}_3] = [a_1a_2a_3 - a_3\mathbf{v}_1 \cdot \mathbf{v}_2 - a_1\mathbf{v}_2 \cdot \mathbf{v}_3 - a_2\mathbf{v}_1 \cdot \mathbf{v}_3 + (\mathbf{v}_1 \times \mathbf{v}_2) \cdot \mathbf{v}_3, a_1a_2\mathbf{v}_3 - (\mathbf{v}_1 \cdot \mathbf{v}_2)\mathbf{v}_3 + a_1a_3\mathbf{v}_2 + a_2a_3\mathbf{v}_1 + a_3\mathbf{v}_1 \times \mathbf{v}_2 + a_1\mathbf{v}_2 \times \mathbf{v}_3 + a_2\mathbf{v}_1 \times \mathbf{v}_3 + (\mathbf{v}_1 \times \mathbf{v}_2) \times \mathbf{v}_3$

If we expand $\mathbf{v}_1 \times (\mathbf{v}_2 \times \mathbf{v}_3)$ and $(\mathbf{v}_1 \times \mathbf{v}_2) \times \mathbf{v}_3$ using Lagrange's formula, then we get the term $-\mathbf{v}_1(\mathbf{v}_2 \cdot \mathbf{v}_3) + \mathbf{v}_2(\mathbf{v}_1 \cdot \mathbf{v}_3) - \mathbf{v}_3(\mathbf{v}_1 \cdot \mathbf{v}_2)$ in both equations.

This leaves us with quaternion multiplication being associative iff $^{(a)}\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3) = (\mathbf{v}_1 \times \mathbf{v}_2) \cdot \mathbf{v}_3$, which holds by the (scalar) triple product formula. □

(c) Distributivity. .
$(q_1 + q_2)q_3 = [a_1 + a_2, \mathbf{v}_1 + \mathbf{v}_2][a_3, \mathbf{v}_3] = [a_1a_3 + a_2a_3 - \mathbf{v}_1 \cdot \mathbf{v}_3 - \mathbf{v}_2 \cdot \mathbf{v}_3, a_1\mathbf{v}_3 + a_2\mathbf{v}_3 + a_3\mathbf{v}_1 + a_3\mathbf{v}_2 + (\mathbf{v}_1 + \mathbf{v}_2) \times \mathbf{v}_3] = [(a_1a_3 - \mathbf{v}_1 \cdot \mathbf{v}_3) + (a_2a_3 - \mathbf{v}_2 \cdot \mathbf{v}_3), (a_1\mathbf{v}_3 + a_3\mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_3) + (a_2\mathbf{v}_3 + a_3\mathbf{v}_2 + \mathbf{v}_2 \times \mathbf{v}_3)] = q_1q_3 + q_2q_3$

This is clearly true for $q_1(q_2+q_3)$, as all the operations involved in quaternion multiplication are themselves distributive. □

(d) Identity for × over $\mathbb{H}$. .
The identity for quaternion multiplication is $1_q =$, as can be seen by:

$$[1, \mathbf{0}][a_1, \mathbf{v}_1] = 1 \cdot a_1 - \mathbf{0} \cdot \mathbf{v}_1 + 1\mathbf{v}_1 + a_1\mathbf{0} + \mathbf{0} \times \mathbf{v}_1 = [a_1, \mathbf{v}_1]$$

$$[a_1, \mathbf{v}_1][1, \mathbf{0}] = a_1 \cdot 1 - \mathbf{v}_1 \cdot \mathbf{0} + a_1\mathbf{0} + 1\mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{0} = [a_1, \mathbf{v}_1]$$

□

(e) Inverses for × over $\mathbb{H} \setminus \{0\}$. .

$$q^{-1} := \frac{\bar{q}}{q\bar{q}} = [\frac{a}{a^2 + \mathbf{v} \cdot \mathbf{v}}, \frac{-\mathbf{v}}{a^2 + \mathbf{v} \cdot \mathbf{v}}]$$

$qq^{-1} = [a, \mathbf{v}][\frac{a}{a^2+\mathbf{v}\cdot\mathbf{v}}, \frac{-\mathbf{v}}{a^2+\mathbf{v}\cdot\mathbf{v}}] = \frac{a^2+\mathbf{v}\cdot\mathbf{v}-a\mathbf{v}+a\mathbf{v}-\mathbf{v}\times\mathbf{v}}{a^2+\mathbf{v}\cdot\mathbf{v}} = \frac{a^2+\mathbf{v}\cdot\mathbf{v}}{a^2+\mathbf{v}\cdot\mathbf{v}} = [1, \mathbf{0}]$ □

Remark 1.2. *The quaternions are one of only four normed division algebras over the reals, the other three being $\mathbb{R}, \mathbb{C}$ and $\mathbb{O}$, the octonions.*

1.3 A Different Approach

1.3.1 Geometric interpretation (classical)

Definition 8. *Denote a vector from A to B, A,B$\in \mathbb{R}^n$, as α. Then $\alpha = T(\alpha)U(\alpha)$ where $T(\alpha)$ is the* tensor *{i.e. magnitude, a 4-D equivalent of the modulus} of the vector, and $U(\alpha)$ is the* versor *{i.e. unit vector in direction of α}.*

Given 2 vectors α and β, there exists an operator q s.t. $q\beta = \alpha$ (Clifford). By analogy, we can define $q = \alpha\beta^{-1}$. β^{-1} is defined with $\beta\beta^{-1} = 1$, giving β^{-1} the reciprocal tensor to β, and a versor in the opposite direction. An example of such an operator would be one to firstly rotate β to make it parallel to the versor of α, and then alter the tensor to having the same tensor as α. This operator will involve four numbers: one to change the tensor, one for the angle of rotation, and two for the direction of the axis about which the rotation is occurring. We call this operator q a *quaternion* [2].

A versor is defined to be, more specifically, a unit vector being employed to rotate a vector in a plane perpendicular to said unit vector, the unit of rotation being $\frac{\pi}{2}$.

As should be clear, when i operates as a versor on j, using the right-handed rule to define the positive direction, the result is the vector k. This can be written as $ij = k$, and similar operations will achieve the two analogues. Also, using the fact that these square roots of -1 are left-associative with their inverses, we can...

Exercise 1.5. - Prove that $j^2 = -1$

Working 1.6 (Answer)**.** $ji = -k \Rightarrow j = -ki^{-1}$, $jk = i \Rightarrow j = ik^{-1}$
$\Rightarrow j^2 = jj = (-ki^{-1})(ik^{-1}) = -(k)(ii^{-1})(k) = -(kk^{-1}) = -1$

...and a similar result can be achieved for i, k, and their negatives, thus affirming associativity, and also geometrically obtaining Hamilton's intital equations.

1.3.2 "The Difficulty of Quaternions"

Hamilton remarked himself [5] that the "difficulty in the geometrical interpretation of Quaternions" lies with their ability to operate as both vectors and versors. This can be made evident with the following false logic: Given that $i^2 = j^2 = k^2 = -1$, we know that $i = j = k = \sqrt{-1}$. Therefore, we can use $i = jk$ to "prove" $\sqrt{-1} = \sqrt{-1}.\sqrt{-1} = -1$. The falsehood arises from the use of $\sqrt{-1}$ as a substitution for the *vectors* i and k, where it represents only their ability to rotate a vector through $\frac{\pi}{2}$ as described before. It would also be incorrect to write $i = \sqrt{-1}.k$ as $\sqrt{-1}$ represents an "indeterminate" versor as opposed to the specific versor j.[1]

1.3.3 Geometric interpretation (modern)

i could be interpreted, in the context of $\mathbb{C}$ to be the operator "rotate by $\frac{\pi}{2}$ in a positive direction". If, however, we try to extend this analogy to "rotate by $\frac{\pi}{2}$ in a right-handed direction around i-axis" over the quaternionic numbers, $\mathbb{H}$, we run into problems: for example, rotating an object around i and then j by $\frac{\pi}{2}$ doesn't give any multiple of rotations around the k-axis[6].

This difficulty can be overcome by making the rotations have angle π. This creates a new problem, that a double rotation (a rotation of 2π) around the i-axis seemingly gives $i^2 = 1$. This, however, can be resolved with the concept

[1]drawn from *"The Outlines of Quaternions"*

of *spinors*, mathematical objects that change sign [2] under a rotation of 2π - to retrieve $ij = k$, we need to adopt the convention that ij means "j followed by i". Though this seems a counter-intuitive notion, there are exercises to exhibit the unclear logic of spinors [3].

It is a well-known property of rotations in three dimensions that the composition of two rotations will be a new rotation about some axis. It is also known that the composition of two translations is also a translation, defined by the *triangle law.* Combining these two facts, it is possible to combine these two facts to form a geometrical backing for the spinorial suggestion.

Consider our 'vectors' to be "oriented arcs of great circles drawn on a sphere" [4]. We can use them to represent rotations in their direction, about an axis perpendicular to the plane of the great circle and through its centre. Obtaining a similar triangle law for rotations, however, needs the rotation represented by an arc to have twice the angle of the arc (w.l.o.g. we can assume the sphere to have a radius of 1, therefore $s = r\theta = \theta \Rightarrow$ the rotation symbolised is twice the length of the arc)

Theorem 1.3. *The product of simple rotations about the vertices of a spherical triangle, through twice the angles of that triangle, is the identity*

Proof. If P,Q,R are reflections in the sides of the triangle. then the three rotations are PQ,QR,RP (PQ=P followed by Q) $PQ.QR.RP = PQ^2R^2P = PP = 1$ □

2 Quaternions and their use for rotations

To find a formula for rotations in three dimensions, we can try to draw an analogy from two-dimensional rotations in $\mathbb{C}$. In the complex plane, you could perform an anti-clockwise rotation by α with the function $f(w) = zw$ for $z = e^{\alpha i}$.

Lemma 2.1. *The exponential of a quaternion* $q = [a, \mathbf{v}]$ *is* $exp(q) = exp(a)[\cos(m), \mathbf{L}\sin(m)]$*, where* m *is a scalar and* $\mathbf{L}$ *a vector such that* $m\mathbf{L}$=$\mathbf{v}$

Proof. The exponential of a vector is defined as the exponential of a scalar, with $exp(\mathbf{v}) = \Sigma_{n=0}^{\infty}(\frac{\mathbf{v}^n}{n!})$, with $\mathbf{v}^{n+2} := (\mathbf{v}\cdot\mathbf{v})\mathbf{v}^n$.

Now, if I consider the exponential of a pure quaternion $\mathbf{u}$ with $\mathbf{u}^2 = -1$, the same derivation as for Euler's formula can be used to achieve $exp(\mathbf{u}x) = \cos x + \mathbf{u}\sin x$ for $x \in \mathbb{R}$. Also, any vector $\mathbf{v}$= (x, y, z) can be written as $\frac{\mathbf{v}}{|\mathbf{v}|}$ to give it the property

$$(\frac{\mathbf{v}}{|\mathbf{v}|})^2 = (\frac{xi + yj + zk}{\sqrt{x^2 + y^2 + z^2}})^2 = \frac{-x^2 + abk - acj - abk - b^2 + bci + acj - bci - c^2}{x^2 + y^2 + z^2}$$

$$= \frac{-x^2 - y^2 - z^2}{x^2 + y^2 + z^2} = -1$$

. Combining these two facts, we get that any quaternion $\mathbf{q}$ can be written as

$$exp(q) = exp(a + \mathbf{v}) = exp(a)exp(\mathbf{v}) = exp(a)exp(\frac{\mathbf{v}}{|\mathbf{v}|}|\mathbf{v}|)$$

[2] definition obtained from [6]

[3] http://en.wikipedia.org/wiki/Quaternions/_and_spatial_rotation and [6] pp. 205-206

[4] definition from [6]

$$= exp(a)(\cos \mathbf{v} + \frac{\mathbf{v}}{|\mathbf{v}|}\sin \mathbf{v}) = exp(a)[\cos \mathbf{v}, \frac{\mathbf{v}}{|\mathbf{v}|}\sin \mathbf{v}]$$

□

Corollary 2.2. *Any quaternion q can be written $q = R.exp([0, \mathbf{u}]\theta)$ for $R, \theta \in \mathbb{R}$ and a 3-vector* $\mathbf{u}$

Corollary 2.3. $\log q = \log R.exp([0, \mathbf{u}]\theta) = [\log R, u\theta]$

Corollary 2.4. $p^q = exp(p \log q)$

As we will see later, with the exponential of a quaternion, we can achieve a rotation around $\mathbf{v}$ by conjugating with the exponential of a pure quaternion $[0,\mathbf{v}]$.

Another approach to rotate a vector $\mathbf{v}$ by an angle θ about a unitary axis $\mathbf{w}$ is using the formula $\mathbf{v\prime} = \mathbf{w}(\mathbf{v}\cdot\mathbf{w}) + (\mathbf{v} - \mathbf{w}(\mathbf{v}\cdot\mathbf{w}))\cos\theta + (\mathbf{w}\times\mathbf{v})\sin\theta$[5]. The rotation of a vector $\mathbf{v} \in \mathbb{R}^3$ should be performed by a function $f : \mathbb{R}^3 \to \mathbb{R}^3$ i.e. $Re(f) = 0$. The operation of "conjugation by q" gives $f(\mathbf{v}) = q[0, \mathbf{v}]q^{-1}$ with $Re(f(\mathbf{v})) = Re(q[0, \mathbf{v}]q^{-1}) = Re([0, \mathbf{v}]qq^{-1}) = Re([0, \mathbf{v}]) = 0$ using both associativity of quaternion multiplication and the fact that $q \in \mathbb{H}_1$ as defined later.

Theorem 2.5. $[0, \mathbf{v\prime}] = q[0, \mathbf{v}]\bar{q}$ *defines the vector achieved by rotating the vector $\mathbf{v}$ by an angle θ about the unitary axis $\mathbf{w}$, using the quaternion $q = [\cos\frac{\theta}{2}, \mathbf{w}\sin\frac{\theta}{2}]$ and the Grassman product.*

Proof.

$$[0, \mathbf{v\prime}] = q[0, \mathbf{v}]\bar{q} = [\cos\frac{\theta}{2}, \hat{\mathbf{w}}\sin\frac{\theta}{2}](0 + (\sin\frac{\theta}{2})\mathbf{v}\cdot\hat{\mathbf{w}} + \cos\frac{\theta}{2}\mathbf{v} - (\sin\frac{\theta}{2}\mathbf{v}\times\hat{\mathbf{w}})$$

$$= [\cos\frac{\theta}{2}, \hat{\mathbf{w}}\sin\frac{\theta}{2}][(\sin\frac{\theta}{2})\mathbf{v}\cdot\hat{\mathbf{w}}, (\cos\frac{\theta}{2})\mathbf{v} - (\sin\frac{\theta}{2})\mathbf{v}\times\hat{\mathbf{w}}$$

$$= (\sin\frac{\theta}{2}\cos\frac{\theta}{2})\mathbf{v}\cdot\hat{\mathbf{w}} - (\sin\frac{\theta}{2}\cos\frac{\theta}{2})\mathbf{v}\cdot\hat{\mathbf{w}} + (\sin\frac{\theta}{2})^2\hat{\mathbf{w}}\cdot(\mathbf{v}\times\hat{\mathbf{w}}) + (\cos\frac{\theta}{2})^2\mathbf{v}$$

$$-(\sin\frac{\theta}{2}\cos\frac{\theta}{2})\mathbf{v}\times\hat{\mathbf{w}} + (\sin\frac{\theta}{2})^2(\mathbf{v}\cdot\hat{\mathbf{w}})\hat{\hat{\mathbf{w}}} + (\sin\frac{\theta}{2}\cos\frac{\theta}{2})\hat{\mathbf{w}}\times\mathbf{v} - (\sin\frac{\theta}{2})^2\hat{\mathbf{w}}\times(\mathbf{v}\times\hat{\mathbf{w}})$$

Now, $\hat{\mathbf{w}}\cdot(\mathbf{v}\times\hat{\mathbf{w}}) = \mathbf{0}$ as $\mathbf{v}\times\hat{\mathbf{w}} \perp \hat{\mathbf{w}}$, $\mathbf{v}\times\hat{\mathbf{w}} = -\hat{\mathbf{w}}\times\mathbf{v}$ and evaluating $\hat{\mathbf{w}}\times(\mathbf{v}\times\hat{\mathbf{w}})$ with Lagrange's formula, we know it will equal:

$$= (\cos\frac{\theta}{2})^2\mathbf{v} + (2\sin\frac{\theta}{2}\cos\frac{\theta}{2})\hat{\mathbf{w}}\times\mathbf{v} + (\sin\frac{\theta}{2})^2(\mathbf{v}\cdot\hat{\mathbf{w}})\hat{\mathbf{w}} - (\sin\frac{\theta}{2})^2((\hat{\mathbf{w}}\cdot\hat{\mathbf{w}})\mathbf{v} - (\hat{\mathbf{w}}\cdot\mathbf{v})\hat{\mathbf{w}})$$

$$= (\cos\frac{\theta}{2})^2\mathbf{v} + (\sin\theta)\hat{\mathbf{w}}\times\mathbf{v} + 2(\sin\frac{\theta}{2})^2(\mathbf{v}\cdot\hat{\mathbf{w}})\hat{\mathbf{w}} - (\sin\frac{\theta}{2})^2\mathbf{v}$$

Also, using the double angle formula for $\cos(2\frac{\theta}{2})$ and $\sin(2\frac{\theta}{2})$

$$= (\cos\theta)\mathbf{v} + (\sin\theta)\hat{\mathbf{w}}\times\mathbf{v} + (1 - \cos\theta)(\mathbf{v}\cdot\hat{\mathbf{w}})\hat{\mathbf{w}}$$
$$= \hat{\mathbf{w}}(\mathbf{v}\cdot\hat{\mathbf{w}}) + (\mathbf{v} - \hat{\mathbf{w}}(\mathbf{v}\cdot\hat{\mathbf{w}}))\cos\theta + (\hat{\mathbf{w}}\times\mathbf{v})\sin\theta$$

□

[5] by Rodrigues' rotation formula

...which, as you can see, gives the formula $[0, v\prime] = q[0, v]\bar{q}$, using the Grassman product as the binary operation implied by element juxtaposition.

Definition 9. *The set of unit quaternions is denoted by* $\mathbb{H}_1$

Lemma 2.6. *For any quaternions* $q_1 = [a_1, \mathbf{v_1}], q_2 = [a_2, \mathbf{v_2}]$ *with* $q_1, q_2 \in \mathbb{H}$ *we have* $\overline{q_1 q_2} = \overline{q_2 q_1}$

Proof. $\overline{q_2 q_1} = [a_2 a_1 - \mathbf{v_2} \cdot \mathbf{v_1}, -a_2\mathbf{v_1} - a_1\mathbf{v_2} - \mathbf{v_2} \times \mathbf{v_1}]$ $\overline{q_1 q_2} = [a_1 a_2 - (-\mathbf{v_1} \cdot -\mathbf{v_2}), -a_1\mathbf{v_2} - a_2\mathbf{v_1} + (-\mathbf{v_1}) \times (-\mathbf{v_2})]$ so $\overline{q_1 q_2} = \overline{q_2 q_1}$ by the anti-symmetric property of the cross product and the symmetric property of the dot product. □

Theorem 2.7. *The composition of two rotations is a rotation.*

Proof. Rotate a vector $\mathbf{v}$ with a quaternion $q_1 \in \mathbb{H}_1$ to give $\mathbf{v}\prime = q_1 \mathbf{v} \overline{q_1}$

Rotate this vector $\mathbf{v}\prime$ with a quaternion $q_2 \in \mathbb{H}_1$ to give $\mathbf{v}\prime\prime = q_2(q_1\mathbf{v}\overline{q_1})\overline{q_2} = (q_2 q_1)\mathbf{v}(\overline{q_1 q_2}) = (q_2 q_1)\mathbf{v}(\overline{q_2 q_1})$ i.e. rotating by q_1 then q_2 gives a rotation by the quternion $q_2 q_1$ □

Exercise 2.1. *Show that the quaternion* q^{-1} *gives the opposite rotation to* q

Working 2.2. *A rotation quaternion will be of the form* $q = [\cos\frac{\theta}{2}, w \sin\frac{\theta}{2}]$.

The inverse of a quaternion is defined, as before, by $q^{-1} = \frac{\bar{q}}{q\bar{q}}$. *Therefore, we get, using* $\bar{q} = [\cos\frac{\theta}{2}, -w\sin\frac{\theta}{2}]$

$$q\bar{q} = (\cos\frac{\theta}{2})^2 - (\sin\frac{\theta}{2})^2(-w \cdot w) + \cos\frac{\theta}{2}\sin\frac{\theta}{2}(-w) + \cos\frac{\theta}{2}\sin\frac{\theta}{2}(w) + (\sin\frac{\theta}{2})^2(-w \times w)$$

$$= (\cos\frac{\theta}{2})^2 + (\sin\frac{\theta}{2})^2(1) + (\sin\frac{\theta}{2})^2(\mathbf{0}) = (\cos\frac{\theta}{2})^2 + (\sin\frac{\theta}{2})^2 = 1$$

$$\Rightarrow q^{-1} = \frac{\bar{q}}{q\bar{q}} = \bar{q} = [\cos\frac{\theta}{2}, -w\sin\frac{\theta}{2}] = [\cos\frac{-\theta}{2}, w\sin\frac{-\theta}{2}]$$

using the odd and even properties for sin *and* cos *respectively.*

Exercise 2.3. *Show that the quaternion* $-q$ *gives the same rotation to* q

Working 2.4. $q = [\cos\frac{\theta}{2}, w\sin\frac{\theta}{2}]$

Using the rotation formula from before (reference somehow?) we know that $[0, \mathbf{v}\prime] = q[0, \mathbf{v}]\bar{q}$

We can see if we use the quaternion $-q$, *we get*

$$(-q)[0, \mathbf{v}]\overline{(-q)} = (-1)(q)[0, \mathbf{v}]\overline{(-1)(q)} = (-1)^2(q)[0, \mathbf{v}]\bar{q} = q[0, \mathbf{v}]\bar{q} = [0, \mathbf{v}\prime]$$

2.1 Quaternions vs. Orthogonal Matrices

Quaternions are used even today in computer graphics, for three main reasons. Firstly, the quaternion representation uses less than half the amount of digits as if using an orthogonal matrix. Secondly, this compactness of notation is compounded by the fact that you can easily read off the angle and axis of rotation. Thirdly, it's easier to compensate for rounding errors using quaternions than for orthogonal matrices, as you only have to re-divide by the modulus of the quaternion to keep it unitary.

2.2 Slerp

Definition 10 (lifted from [8])**.** *In computer graphics, Slerp is shorthand for spherical linear interpolation, in the context of quaternion interpolation for the purpose of animating 3D rotation. It refers to constant speed motion along a unit radius great circle arc, given the ends and an interpolation parameter between 0 and 1. It is called spherical-linear because the two quaternion rotations are interpolated uniformly along a geodesic in the surface of the 3-sphere.*

Definition 11. *The quaternion used for a starting rotation given by a and ending with rotation b, for* $a, b \in \mathbb{H}$, *is* $q = (ba^{-1})^t a$. *This can be written* $Slerp(t, a, b) = (ba^{-1})^t a$

Remark 2.8. *The quaternion product* ba^{-1} *can be greatly simplified by use of the fact that, for a unit quaternion* $u = [\cos\theta, w\sin\theta]$ *and* $u^t = [\cos t\theta, w\sin t\theta]$

From the definition you can see that $t = 0$ gives rotation a, $t = 1$ the rotation b, and $t \in (0, 1)$ gives all intermediate rotations.

2.3 Squad

Definition 12. *Let the quaternions* p, a, b, q *be the ordered vertices of a quadrilateral. The quaternion used for the starting rotation a, the ending rotation b and the path defined by p and q. The spherical cubic interpolation is defined to be* $Squad(t, a, b, p, q) = Slerp(2t(1-t), Slerp(t, a, b), Slerp(t, p, q)$

Squads can be used to not only define the initial and final rotations, but also the initial and final rotation speeds.

2.4 Euler Angles

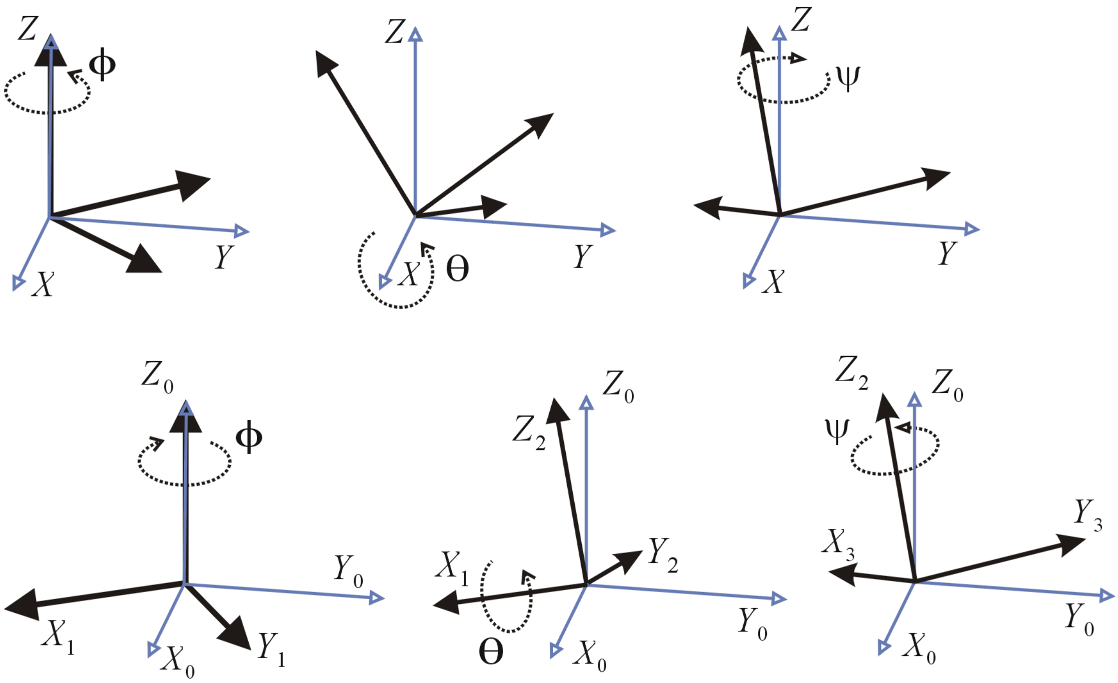
If you consider the rotation of an object in 3-space, you can break it into a rotation about the x-axis, the y-axis, and the z-axis.[6] It is also known that one can write a rotation in the form of a matrix. Therefore, if you consider the matrices A, B, C as rotation matrices, then you can define $D := ABC$ as the combination of those rotations. The three rotations giving these three matrices are called the *E*uler angles. One common notation is ZXZ - other names are the 3-1-3 or x-convention notation. The rotation given by the Euler angles (ϕ, θ, ψ) being the rotation by ϕ about the z-axis, $\theta \in [0, \pi]$ about the x-axis, and ψ around the z-axis again. Another common notation is the so-called "general" convention, where the rotations are taken about ϕ, θ and ψ about "successive body-fixed axes"[7] as shown below.

2.5 Quaternions vs. Euler Angles

Quaternions are a better alternative for a few reasons. Firstly, it is simple to combine two rotations using the Grassman product. Secondly, they use less computer power to renormalise (so they still compare to valid rotations throughout the computing process) than rotation matrices. Thirdly, they *g*imbal lock

[6] by Euler's rotation formula

[7] Wikipedia entry on *Rotation representations*



(when the original and new planes get aligned, so a degree of freedom is lost).

Above: Figures for the x-convention and the general notation

3 Other uses of quaternions

3.1 Hyperspherical co-ordinates

Generalising the spherical co-ordinates of a 2-sphere, we can achieve the hyper-spherical co-ordinates of a unit 3-sphere as follows:

$$\mathbf{r}(\psi,\theta,\varphi) = \begin{pmatrix} \cos\psi \\ \cos\varphi\sin\theta\sin\psi \\ \sin\varphi\sin\theta\sin\psi \\ \cos\theta\sin\varphi \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

By generalising Euler's fo rmula for quaternions, we can achieve

$$q = e^{\tau\varphi} = \cos\varphi + \tau\sin\varphi$$

where τ is a unit imaginary quaternion. All the unit imaginary quaternions lie on the unit 2-sphere in $\mathrm{Im}(\mathbb{H})$ so that any τ can be written

$$\tau = \cos\varphi\sin\theta i + \sin\varphi\sin\theta j + \cos\theta k \quad \Rightarrow \quad q = e^{\tau\varphi} = x_0 + x_1 i + x_2 j + x_3 k$$

with the x_i defined as before, thus giving a neater way to define the 3-sphere [8].

[8] http://en.wikipedia.org/wiki/3-sphere

3.2 Pauli matrices

Definition 13. *The Pauli matrices are the matrices* $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{C}^{2\times 2}$ *that are Hermitian* ($\sigma_i = \sigma_i^*$) *and unitary* ($\sigma_i^* = \sigma_i^{-1}$) *where*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

If you consider the real linear span S of $\{I_2, \sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1\}$, you will realise that S is isomorphic to $\mathbb{H}$, with the mappings $1 \mapsto 1, i \mapsto \sigma_1\sigma_2, j \mapsto \sigma_3\sigma_1, k \mapsto \sigma_2\sigma_3$. As the unit quaternions are isomorphic to SU(2), then it gives a way of describing SU(2) via tha Pauli matrices.

3.3 Dense notation

Quaternions are useful as a way to condense notation for vector properties. This has already been illustrated by the Grassman product containing the dot and cross products. An extension of this is when you consider an operator quaternion $q = [\frac{d}{dt}, \nabla]$ and the Grassman product with a general quaternionic function (the addition of a scalar function and a 3-vector function) [9]:

$$[\frac{d}{dt}, \nabla][b, \mathbf{c}] = [\frac{db}{dt} - \nabla \cdot \mathbf{c}, \frac{d\mathbf{c}}{dt} + \nabla b + \nabla \times \mathbf{c}] = [\frac{db}{dt} - div(\mathbf{c}), \frac{d\mathbf{c}}{dt} + grad(b) + curl(\mathbf{c})]$$

[9] Idea taken from http://world.std.com/~sweetser/quaternions/qindex/qindex.html

References

[1] Letter from Sir William R. Hamilton to John T. Graves, Esq. *On Quaternions*, October 17, 1843, published in *London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, vol. xxv (1844), pp 489-95.

[2] Kelland and Tait, *Introduction to Quaternions*, 3rd edition, 1904

[3] Hume, HWL *The Outlines of Quaternions*

[4] Joly, Charles Jasper *A Manual of Quaternions*

[5] Graves, R. P. *Life of Sir William Rowan Hamilton*, 1882

[6] Penrose, Roger *The Road To Reality*

[7] Brown, Simon *Representing Rotations in Quaternion Arithmetic*, http://www.sjbrown.co.uk/?article=quaternions

[8] Wikipedia entry on spherical linear interpolation, http://en.wikipedia.org/wiki/Slerp