Quaternion Involutions

Todd A. Ell*
5620 Oak View Court, Savage, MN 55378-4695, USA.
Email: T.Ell@IEEE.org.

Stephen J. Sangwine
Department of Electronic Systems Engineering,
University of Essex, Wivenhoe Park,
Colchester, CO7 9EU, United Kingdom.
Email: S.Sangwine@IEEE.org.

February 1, 2008

Abstract

An involution is usually defined as a mapping that is its own inverse. In this paper, we study quaternion involutions that have the additional properties of distribution over addition and multiplication. We review formal axioms for such involutions, and we show that the quaternions have an infinite number of involutions. We show that the conjugate of a quaternion may be expressed using three mutually perpendicular involutions. We also show that any set of three mutually perpendicular quaternion involutions is closed under composition. Finally, we show that projection of a vector or quaternion can be expressed concisely using involutions.

1 Introduction

Involutions are usually defined simply as self-inverse mappings. A trivial example is conjugation of a complex number, which is obviously self-inverse. In this paper we consider involutions of the quaternions, that is functions of a quaternion variable that are self-inverse. Quaternion conjugation is an obvious involution, but it is not the only quaternion involution. In fact, the quaternions have an infinite number of involutions, as we show. The paper begins by reviewing the classical basics of quaternions, and then presents axioms for involutions which go beyond the simple definition of a self-inverse mapping. Section 4 then presents the quaternion involutions, and section 5 presents their properties. Section 6 discusses the quaternion conjugate and shows that it may be expressed using three mutually perpendicular quaternion involutions. Finally, section 7 shows that the projection of a vector or quaternion may be expressed using involutions.

2 Basics of quaternions

A quaternion may be represented in Cartesian form $q = w + \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z$ where $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ are mutually perpendicular unit vectors obeying the multiplication rules below discovered by Hamilton in 1843 [1], and w, x, y, z, are real.

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = \boldsymbol{ijk} = -1 \tag{1}$$

The conjugate of a quaternion is given by $\overline{q} = w - \boldsymbol{i}x - \boldsymbol{j}y - \boldsymbol{k}z$.

The quaternion algebra $\mathbb{H}$ is a normed division algebra. The modulus of a quaternion is the square root of its norm: $|q| = \sqrt{w^2 + x^2 + y^2 + z^2}$, and every non-zero quaternion has a multiplicative inverse

*Dr T. A. Ell is a Visiting Fellow at the University of Essex, funded by grant number GR/S58621 from the United Kingdom Engineering and Physical Sciences Research Council.

given by its conjugate divided by its norm: $q^{-1} = \overline{q}/|q|^2 = (w - \boldsymbol{i}x - \boldsymbol{j}y - \boldsymbol{k}z)/(w^2 + x^2 + y^2 + z^2)$. For a more detailed exposition of the basics of quaternions, we refer the reader to Coxeter's 1946 paper [2].

An alternative and much more powerful representation for a quaternion is as a combination of a *scalar* and a *vector* part, analogous to a complex number, and this representation will be employed in the rest of the paper: $q = a + \boldsymbol{\mu}b$, where $\boldsymbol{\mu}$ is a unit vector, and a and b are real. b is the modulus of the vector part of the quaternion and $\boldsymbol{\mu}$ is its direction. In terms of the Cartesian representation:

$$a = w, \qquad b = \sqrt{x^2 + y^2 + z^2}, \qquad \boldsymbol{\mu} = \frac{\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z}{b} \tag{2}$$

Lemma 1. *The square of any unit vector is* -1.

Proof. Let $\boldsymbol{\mu}$ be an arbitrary unit vector as defined in equation 2. Its square is given by:

$$\begin{aligned}\boldsymbol{\mu}^2 &= \frac{(\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z)^2}{x^2 + y^2 + z^2} \\ &= \frac{\boldsymbol{i}^2x^2 + \boldsymbol{j}^2y^2 + \boldsymbol{k}^2z^2 + \boldsymbol{ij}xy + \boldsymbol{ji}xy + \boldsymbol{ik}xz + \boldsymbol{ki}xz + \boldsymbol{jk}yz + \boldsymbol{kj}yz}{x^2 + y^2 + z^2}\end{aligned}$$

Applying the rules in equation 1 we get: $\boldsymbol{\mu}^2 = \frac{-x^2-y^2-z^2}{x^2+y^2+z^2} = -1$ □

A corollary of Lemma 1 is that there are an infinite number of solutions to the equation $x^2 = -1$.

The conjugate of a quaternion in complex form is $\overline{q} = a - \boldsymbol{\mu}b$. Geometrically, this is obviously a reversal of the direction of the vector part. The quaternion conjugate has properties analogous to those of the complex conjugate, with one minor exception: the quaternion conjugate is an anti-involution whereas the complex conjugate is an involution. We define these terms in the next section, and we return to this point with Theorem 6 in section 6. The product of a quaternion with its conjugate gives the norm, or square of the modulus. This follows directly from Lemma 1: $(a+\boldsymbol{\mu}b)(a-\boldsymbol{\mu}b) = a^2 - \boldsymbol{\mu}^2b^2 = a^2 + b^2 = w^2 + x^2 + y^2 + z^2$.

3 Involutions

The formal definition of an involution is not easy to find (most mathematical reference works define it simply as a mapping which is its own inverse) but [3] gives a reasonably authoritative statement from which we reproduce the following axioms. It is clear from what follows in the paper that all three of these axioms are important, otherwise it is possible to define trivial self-inverse mappings, which have uninteresting properties. We denote an arbitrary involution by the mapping $x \rightarrow f(x)$.

Axiom 1. $f(f(x)) = x$. *An involution is its own inverse.*

Axiom 2. *An involution is linear:* $f(x_1 + x_2) = f(x_1) + f(x_2)$ *and* $\lambda f(x) = f(\lambda x)$ *where* λ *is a real constant.*

Axiom 3. $f(x_1x_2) = f(x_1)f(x_2)$. *If the terms on the right must be reversed (which can only be necessary if* x_1 *and* x_2 *do not commute), then* $f(\ )$ *is an* anti-involution.

4 Quaternion involutions

Involutions over the quaternion field have been published by Chernov [4], and used by Bülow [5, 6], but with an important difference from the involutions given in this paper: we show here that there are an infinite number of involutions over the quaternion field whereas Chernov and Bülow wrote of only three (plus conjugation). Chernov defined the following involutions, which we generalize in Theorem 1:

$$\begin{aligned}\alpha(q) &= -\boldsymbol{i}q\boldsymbol{i} = w + \boldsymbol{i}x - \boldsymbol{j}y - \boldsymbol{k}z \\ \beta(q) &= -\boldsymbol{j}q\boldsymbol{j} = w - \boldsymbol{i}x + \boldsymbol{j}y - \boldsymbol{k}z \\ \gamma(q) &= -\boldsymbol{k}q\boldsymbol{k} = w - \boldsymbol{i}x - \boldsymbol{j}y + \boldsymbol{k}z\end{aligned} \tag{3}$$

They also showed that the quaternion conjugate can be expressed in terms of these three involutions, a result that we generalize in Theorem 7.

Theorem 1. *The mapping $q \to -\boldsymbol{\nu} q \boldsymbol{\nu}$ where q is an arbitrary quaternion is an involution for any unit vector $\boldsymbol{\nu}$.*

Proof. Axiom 1 is easily shown to be satisfied using Lemma 1: $-\boldsymbol{\nu}(-\boldsymbol{\nu} q \boldsymbol{\nu})\boldsymbol{\nu} = (-1)q(-1) = q$. Axiom 2 is seen to be satisfied from: $-\boldsymbol{\nu}(q_1 + q_2)\boldsymbol{\nu} = -\boldsymbol{\nu} q_1 \boldsymbol{\nu} - \boldsymbol{\nu} q_2 \boldsymbol{\nu}$ (multiplication of quaternions is distributive over addition). Since reals commute with quaternions, the second part of the axiom is trivially seen.

Axiom 3 can be shown to be satisfied as follows:

$$\begin{aligned} f(q_1)f(q_2) &= (-\boldsymbol{\nu} q_1 \boldsymbol{\nu})(-\boldsymbol{\nu} q_2 \boldsymbol{\nu}) \\ &= \boldsymbol{\nu} q_1 \boldsymbol{\nu}\boldsymbol{\nu} q_2 \boldsymbol{\nu} \\ &= \boldsymbol{\nu} q_1 (-1) q_2 \boldsymbol{\nu} \\ &= -\boldsymbol{\nu} q_1 q_2 \boldsymbol{\nu} = f(q_1 q_2) \end{aligned}$$

□

Note that a mapping $q \to \boldsymbol{\nu}_1 q \boldsymbol{\nu}_2$ ($\boldsymbol{\nu}_1 \neq \boldsymbol{\nu}_2$) is its own inverse, but is not an involution as considered in this paper, because it does not satisfy axiom 3.

In what follows we introduce a new notation for involutions using an overbar with a subscript unit vector, thus $\overline{q}^{\boldsymbol{\nu}} = -\boldsymbol{\nu} q \boldsymbol{\nu}$. We refer to the direction defined by $\boldsymbol{\nu}$ as the *axis of involution*. The fact that we use an overbar to denote involutions as well as the conjugate is not coincidental and we shall see that there is a close relationship between involutions and conjugation.

5 Properties of quaternion involutions

Lemma 2. *The product of any two vectors with arbitrary non-zero norms is a quaternion. The scalar part is minus the inner or scalar product of the two vectors and the vector part is the vector product of the two vectors. Reversing the order of the product conjugates the resulting quaternion.*

Proof. Let $\boldsymbol{\mu}_1 = \boldsymbol{i}x_1 + \boldsymbol{j}y_1 + \boldsymbol{k}z_1$ and $\boldsymbol{\mu}_2 = \boldsymbol{i}x_2 + \boldsymbol{j}y_2 + \boldsymbol{k}z_2$. Their product is given by:

$$\begin{aligned} \boldsymbol{\mu}_1\boldsymbol{\mu}_2 &= (\boldsymbol{i}x_1 + \boldsymbol{j}y_1 + \boldsymbol{k}z_1)(\boldsymbol{i}x_2 + \boldsymbol{j}y_2 + \boldsymbol{k}z_2) \\ &= \boldsymbol{i}^2 x_1 x_2 + \boldsymbol{j}^2 y_1 y_2 + \boldsymbol{k}^2 z_1 z_2 \\ &\quad + \boldsymbol{ij} x_1 y_2 + \boldsymbol{ji} y_1 x_2 + \boldsymbol{ik} x_1 z_2 + \boldsymbol{ki} z_1 x_2 + \boldsymbol{jk} y_1 z_2 + \boldsymbol{kj} z_1 y_2 \\ &= -(x_1 x_2 + y_1 y_2 + z_1 z_2) \\ &\quad + \boldsymbol{i}(y_1 z_2 - z_1 y_2) + \boldsymbol{j}(z_1 x_2 - x_1 z_2) + \boldsymbol{k}(x_1 y_2 - y_1 x_2) \end{aligned}$$

Changing the order of the product changes the order of all the products of two unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$. Since $\boldsymbol{ij} = -\boldsymbol{ji}$ and so on, this negates all the components of the vector part of the result. The scalar part, $-(x_1x_2 + y_1y_2 + z_1z_2)$, is unchanged. Thus reversing the order of the product conjugates the result, as stated. The scalar and vector parts can be seen to be equal to minus the scalar product and the vector product respectively, according to standard definitions of these products. □

Lemma 3. *The product of two perpendicular vectors changes sign if the order of the product is reversed.*

Proof. Lemma 2 identified the scalar part of the result $-(x_1x_2 + y_1y_2 + z_1z_2)$ as minus the inner product of the two vectors. Since the inner product is zero in the case of perpendicular vectors, the product of perpendicular vectors is a vector and it is the vector product of the two vectors. It follows that this vector changes sign (reverses) if the order of the product is reversed. □

Theorem 2. *Composition: the composition of two perpendicular involutions is commutative. That is: $\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} = \overline{\overline{q}^{\boldsymbol{\nu}_2}}^{\boldsymbol{\nu}_1}$ where $\boldsymbol{\nu}_1 \perp \boldsymbol{\nu}_2$.*

Proof.

$$\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} = -\boldsymbol{\nu}_2(-\boldsymbol{\nu}_1 q \boldsymbol{\nu}_1)\boldsymbol{\nu}_2 = \boldsymbol{\nu}_2\boldsymbol{\nu}_1 q \boldsymbol{\nu}_1\boldsymbol{\nu}_2$$

and by Lemma 3 we can reverse the order of the pairs of unit vectors if we change their signs:

$$\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} = (-\boldsymbol{\nu}_1\boldsymbol{\nu}_2) q (-\boldsymbol{\nu}_2\boldsymbol{\nu}_1) = -\boldsymbol{\nu}_1(-\boldsymbol{\nu}_2 q \boldsymbol{\nu}_2)\boldsymbol{\nu}_1 = \overline{\overline{q}^{\boldsymbol{\nu}_2}}^{\boldsymbol{\nu}_1}$$

□

Theorem 3. *Double Composition: given a set of three mutually perpendicular unit vectors, $\boldsymbol{\nu}_1$, $\boldsymbol{\nu}_2$, $\boldsymbol{\nu}_3$, such that $\boldsymbol{\nu}_1\boldsymbol{\nu}_2 = \boldsymbol{\nu}_3$, then*

$$\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} = \overline{q}^{\boldsymbol{\nu}_3}$$

Proof.

$$\begin{aligned}\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} &= -\boldsymbol{\nu}_2(-\boldsymbol{\nu}_1 q\boldsymbol{\nu}_1)\boldsymbol{\nu}_2 = \boldsymbol{\nu}_2\boldsymbol{\nu}_1 q\boldsymbol{\nu}_1\boldsymbol{\nu}_2\\ &= -\boldsymbol{\nu}_1\boldsymbol{\nu}_2 q\boldsymbol{\nu}_1\boldsymbol{\nu}_2 \quad \text{by Lemma 3}\\ &= -\boldsymbol{\nu}_3 q\boldsymbol{\nu}_3 = \overline{q}^{\boldsymbol{\nu}_3}\end{aligned}$$

□

Corollary 1. *Triple Composition: the composition of three mutually perpendicular involutions is an identity.*

Proof. From Theorem 3: $\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} = \overline{q}^{\boldsymbol{\nu}_3}$. Apply an involution about $\boldsymbol{\nu}_3$ to both sides: $\overline{\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2}}^{\boldsymbol{\nu}_3} = \overline{\overline{q}^{\boldsymbol{\nu}_3}}^{\boldsymbol{\nu}_3} = q$ □

Thus we see that a set of three mutually perpendicular involutions (that is involutions about a set of three mutually perpendicular axes) is closed under composition of the involutions *and by Theorem 2 the order of the composition is unimportant.*

We now present a geometric interpretation of quaternion involution.

Theorem 4. *Given an arbitrary quaternion $q = a + \boldsymbol{\mu}b$, an involution $\overline{q}^{\boldsymbol{\nu}}$ leaves the scalar part of q (that is, a) invariant, and reflects the vector part of q (that is, $\boldsymbol{\mu}b$) across the line defined by the axis of involution $\boldsymbol{\nu}$. (Equivalently, the vector part of q is rotated by π radians about the axis of involution $\boldsymbol{\nu}$.)*

Proof.

$$\overline{q}^{\boldsymbol{\nu}} = -\boldsymbol{\nu}(a + \boldsymbol{\mu}b)\boldsymbol{\nu} = -\boldsymbol{\nu}a\boldsymbol{\nu} - \boldsymbol{\nu}\boldsymbol{\mu}b\boldsymbol{\nu} = -\boldsymbol{\nu}^2 a - \boldsymbol{\nu}\boldsymbol{\mu}\boldsymbol{\nu}b$$

and, since $\boldsymbol{\nu}$ is a unit vector, by Lemma 1:

$$\overline{q}^{\boldsymbol{\nu}} = a - \boldsymbol{\nu}\boldsymbol{\mu}\boldsymbol{\nu}b$$

We recognise $\boldsymbol{\nu}\boldsymbol{\mu}\boldsymbol{\nu}$ to be a reflection of $\boldsymbol{\mu}$ in the plane p normal to $\boldsymbol{\nu}$ as shown by Coxeter [2, Theorem 3.1]. Therefore $-\boldsymbol{\nu}\boldsymbol{\mu}\boldsymbol{\nu}$ is a reflection of $\boldsymbol{\mu}$ in the line defined by $\boldsymbol{\nu}$ as shown in Figure 1. The result of the reflection of the vector part remains a vector, and therefore the scalar part remains unchanged, as shown, and as stated. □

Corollary 2. *An involution applied to a quaternion with vector part parallel to the involution axis is an identity, i.e., $\overline{a + \boldsymbol{\mu}b}^{\boldsymbol{\nu}} = a + \boldsymbol{\mu}b$, where $\boldsymbol{\nu} \parallel \boldsymbol{\mu}$.*

Proof. From Theorem 4: $\overline{q}^{\boldsymbol{\nu}} = a - \boldsymbol{\nu}\boldsymbol{\mu}\boldsymbol{\nu}b$. Since both $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ are unit vectors, and are parallel, $\boldsymbol{\nu} = \pm\boldsymbol{\mu}$ and the result on the right reduces to $a + \boldsymbol{\mu}b = q$ in both cases. □

Corollary 3. *An involution applied to a quaternion with vector part perpendicular to the involution axis conjugates the quaternion, i.e., $\overline{a + \boldsymbol{\mu}b}^{\boldsymbol{\nu}} = a - \boldsymbol{\mu}b$, where $\boldsymbol{\nu} \perp \boldsymbol{\mu}$.*

Proof. From Theorem 4: $\overline{q}^{\boldsymbol{\nu}} = a - \boldsymbol{\nu}\boldsymbol{\mu}\boldsymbol{\nu}b$. If $\boldsymbol{\nu} \perp \boldsymbol{\mu}$ we can use Lemma 3 to reverse the order of the two unit vectors, thus obtaining $a + \boldsymbol{\nu}^2\boldsymbol{\mu}b = a - \boldsymbol{\mu}b = \overline{q}$ as stated. □

Lemma 4. *The product of two unit vectors is a quaternion with argument equal to the angle between the two vectors.*

The sign of the argument is significant, because it depends on the ordering of the two vectors (the angle is measured from the first vector to the second).

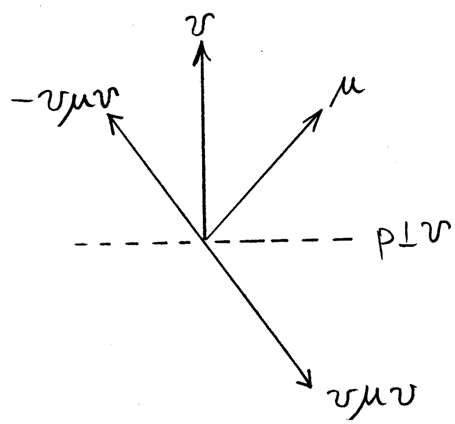

Figure 1: Reflection of a vector $\boldsymbol{\mu}$ in a line defined by a unit vector $\boldsymbol{\nu}$. p is a plane perpendicular to $\boldsymbol{\nu}$ seen edge on.

Proof. Lemma 2 identified the scalar and vector parts of the product of two vectors with minus the inner product, and the vector product respectively of the two vectors. Since these products are given for unit vectors by $\cos\theta$ and $\boldsymbol{\mu}\sin\theta$, where $\boldsymbol{\mu}$ is perpendicular to the plane containing the two vectors, we may write the product of two unit vectors as: $-\cos\theta + \boldsymbol{\mu}\sin\theta = -\exp(-\boldsymbol{\mu}\theta)$. □

Theorem 5. *The composition of two involutions is a rotation of the vector part of the quaternion operated upon about an axis normal to the plane containing the two axes of involution. The angle of rotation is twice the angle between the two involution axes. In the case where the two involutions are perpendicular the composite result is of course an involution, which is a rotation of the vector part of the quaternion by* π.

Proof. Let $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ be two unit vectors and q be an arbitrary quaternion. Then the composition of two involutions about $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ is given by:

$$\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} = -\boldsymbol{\nu}_2(-\boldsymbol{\nu}_1 q \boldsymbol{\nu}_1)\boldsymbol{\nu}_2 = \boldsymbol{\nu}_2\boldsymbol{\nu}_1 q \boldsymbol{\nu}_1\boldsymbol{\nu}_2$$

From Lemma 2 we can write the result on the right as $pq\overline{p}$, where $p = \boldsymbol{\nu}_2\boldsymbol{\nu}_1$ is a unit quaternion[1]. Separating $q = a + \boldsymbol{\mu}b$ into its scalar and vector parts we obtain:

$$\overline{\overline{q}^{\boldsymbol{\nu}_1}}^{\boldsymbol{\nu}_2} = p\overline{p}a + p\boldsymbol{\mu}b\overline{p} = a + (p\boldsymbol{\mu}\overline{p})b$$

We recognise the term $p\boldsymbol{\mu}\overline{p}$ as a rotation as given by Coxeter [2, Theorem 3.2]. The axis of rotation is given by the vector part of p, and the angle of rotation is twice the argument of p. Therefore, from Lemma 2 we know that the axis of rotation is perpendicular to the plane containing the two vectors, and from Lemma 4 we know that the angle of rotation is twice the angle between the two vectors. □

6 The quaternion conjugate

Theorem 6. *The quaternion conjugate is an anti-involution.*

Proof. The definition of the conjugate of a quaternion $q = a + \boldsymbol{\mu}b$ is $\overline{q} = a - \boldsymbol{\mu}b$. We have to show that this satisfies the three axioms given in section 3, and that in Axiom 3 we have to reverse the terms on the right-hand side.

It is easily seen that the quaternion conjugate satisfies Axiom 1. To demonstrate that the quaternion conjugate satisfies Axiom 2, let $q_1 = a + \boldsymbol{\mu}_1 b$ and $q_2 = c + \boldsymbol{\mu}_2 d$. Then $q_1 + q_2 = (a + c) + (\boldsymbol{\mu}_1 b + \boldsymbol{\mu}_2 d)$. Since reversing two vectors also reverses their sum, we see that the quaternion conjugate is distributive over addition, as required. The second part of the axiom is easily seen.

[1] p is a unit quaternion, because it is the product of two unit vectors. This follows from the fact that the quaternion algebra is a normed algebra.

To show that the quaternion conjugate satisfies Axiom 3, we state the required equality and then demonstrate that it is satisfied by expanding the left and right hand sides until identical:

$$\overline{q_1 q_2} = \overline{q_2}\,\overline{q_1}$$
$$\overline{(a+\boldsymbol{\mu}_1 b)(c+\boldsymbol{\mu}_2 d)} = (c-\boldsymbol{\mu}_2 d)(a-\boldsymbol{\mu}_1 b)$$

Expanding the left hand side, we employ Axiom 2:

$$ac - \boldsymbol{\mu}_1\boldsymbol{\mu}_2 bd - \boldsymbol{\mu}_1 bc - \boldsymbol{\mu}_2 ad = ac + \boldsymbol{\mu}_2\boldsymbol{\mu}_1 bd - \boldsymbol{\mu}_1 bc - \boldsymbol{\mu}_2 ad$$

and using Lemma 3 we change the order of the two vectors in the second term on the right-hand side to obtain the required result:

$$ac - \boldsymbol{\mu}_1\boldsymbol{\mu}_2 bd - \boldsymbol{\mu}_1 bc - \boldsymbol{\mu}_2 ad = ac - \boldsymbol{\mu}_1\boldsymbol{\mu}_2 bd - \boldsymbol{\mu}_1 bc - \boldsymbol{\mu}_2 ad$$

□

We now show that the quaternion conjugate can be expressed using the sum of three mutually perpendicular involutions.

Lemma 5. *The sum of three mutually perpendicular involutions applied to a vector negates the vector (reverses its direction). That is, given a set of three mutually perpendicular unit vectors as in Theorem 3 and an arbitary vector $\boldsymbol{\mu}$:*

$$\overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_1} + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_2} + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_3} = -\boldsymbol{\mu} \tag{4}$$

Proof. Let $\boldsymbol{\mu} = \boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 + \boldsymbol{\eta}_3$ where $\boldsymbol{\eta}_i \parallel \boldsymbol{\nu}_i, i \in \{1,2,3\}$. In other words, resolve $\boldsymbol{\mu}$ into three vectors[2] parallel to the three mutually perpendicular unit vectors $\boldsymbol{\nu}_1$, $\boldsymbol{\nu}_2$ and $\boldsymbol{\nu}_3$. Substitute this representation of $\boldsymbol{\mu}$ into the left-hand side of Equation 4:

$$\overline{\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 + \boldsymbol{\eta}_3}^{\boldsymbol{\nu}_1} + \overline{\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 + \boldsymbol{\eta}_3}^{\boldsymbol{\nu}_2} + \overline{\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 + \boldsymbol{\eta}_3}^{\boldsymbol{\nu}_3} = -\boldsymbol{\mu}$$

Axiom 2 allows us to apply the involutions separately to the three components:

$$\overline{\boldsymbol{\eta}_1}^{\boldsymbol{\nu}_1} + \overline{\boldsymbol{\eta}_2}^{\boldsymbol{\nu}_1} + \overline{\boldsymbol{\eta}_3}^{\boldsymbol{\nu}_1} + \overline{\boldsymbol{\eta}_1}^{\boldsymbol{\nu}_2} + \overline{\boldsymbol{\eta}_2}^{\boldsymbol{\nu}_2} + \overline{\boldsymbol{\eta}_3}^{\boldsymbol{\nu}_2} + \overline{\boldsymbol{\eta}_1}^{\boldsymbol{\nu}_3} + \overline{\boldsymbol{\eta}_2}^{\boldsymbol{\nu}_3} + \overline{\boldsymbol{\eta}_3}^{\boldsymbol{\nu}_3} = -\boldsymbol{\mu}$$

We now make use of Corollaries 2 and 3. In this case we are applying them to a vector, so the first states that an involution with axis parallel to a vector is an identity, and the second states that an involution with axis perpendicular to the vector reverses, or negates, the vector:

$$\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2 - \boldsymbol{\eta}_3 - \boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 - \boldsymbol{\eta}_3 - \boldsymbol{\eta}_1 - \boldsymbol{\eta}_2 + \boldsymbol{\eta}_3 = -\boldsymbol{\mu}$$

and cancelling out, we obtain: $-\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2 - \boldsymbol{\eta}_3 = -\boldsymbol{\mu}$, which is the assumption we made at the start of the proof. □

The following theorem is a generalization of a similar result given in [5, Definition 2.2, p.12].

Theorem 7. *Given a set of three mutually perpendicular unit vectors as in Theorem 3, the conjugate of q may be expressed as:*

$$\overline{q} = \frac{1}{2}\left(\overline{q}^{\boldsymbol{\nu}_1} + \overline{q}^{\boldsymbol{\nu}_2} + \overline{q}^{\boldsymbol{\nu}_3} - q\right) \tag{5}$$

Proof. Let $q = a + \boldsymbol{\mu} b$. Substituting this expression for q into the right-hand side of Equation 5 we obtain:

$$\overline{q} = \frac{1}{2}\left(\overline{a+\boldsymbol{\mu} b}^{\boldsymbol{\nu}_1} + \overline{a+\boldsymbol{\mu} b}^{\boldsymbol{\nu}_2} + \overline{a+\boldsymbol{\mu} b}^{\boldsymbol{\nu}_3} - (a+\boldsymbol{\mu} b)\right)$$

We now apply the three involutions separately to the components of q using Axiom 2, and noting from Theorem 4 that the scalar part a is invariant under involutions:

$$\overline{q} = \frac{1}{2}\left(a + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_1} b + a + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_2} b + a + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_3} b - a - \boldsymbol{\mu} b\right)$$

[2] The three vectors $\boldsymbol{\eta}_i$ are not, in general, of unit modulus.

Gathering terms together and factoring out b:

$$\overline{q} = a + \frac{1}{2}\left(\overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_1} + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_2} + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}_3} - \boldsymbol{\mu}\right) b$$

and the right-hand side is equal to $a - \boldsymbol{\mu} b$ by Lemma 5. □

7 Projection using involutions

Finally, we now demonstrate the utility of quaternion involutions by presenting formulae for projection of a vector into or perpendicular to a given direction. These results have been published in [7], but without explicit use of involutions.

Theorem 8. *An arbitrary vector $\boldsymbol{\mu}$ may be resolved into two components parallel to, and perpendicular to, a direction in 3-space defined by a unit vector $\boldsymbol{\nu}$:*

$$\boldsymbol{\mu}_{\parallel\boldsymbol{\nu}} = \frac{1}{2}(\boldsymbol{\mu} + \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}}) \qquad \boldsymbol{\mu}_{\perp\boldsymbol{\nu}} = \frac{1}{2}(\boldsymbol{\mu} - \overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}})$$

where $\boldsymbol{\mu}_{\parallel\boldsymbol{\nu}}$ is parallel to $\boldsymbol{\nu}$ and $\boldsymbol{\mu}_{\perp\boldsymbol{\nu}}$ is perpendicular to $\boldsymbol{\nu}$, and $\boldsymbol{\mu} = \boldsymbol{\mu}_{\parallel\boldsymbol{\nu}} + \boldsymbol{\mu}_{\perp\boldsymbol{\nu}}$.

Proof. From Theorem 4, $\overline{\boldsymbol{\mu}}^{\boldsymbol{\nu}}$ is the reflection of $\boldsymbol{\mu}$ in the line defined by $\boldsymbol{\nu}$ as shown in Figure 1. When $\boldsymbol{\mu}$ is added to its reflection the components of each perpendicular to $\boldsymbol{\nu}$ cancel, and the components parallel to $\boldsymbol{\nu}$ add to give twice the stated result. The factor of $\frac{1}{2}$ gives the result as stated. Similarly, half the difference between $\boldsymbol{\mu}$ and its reflection gives the component of $\boldsymbol{\mu}$ perpendicular to $\boldsymbol{\nu}$. □

Theorem 8 may be generalised to quaternions as well as vectors. Since the scalar part of a quaternion is invariant under an involution, the component of the quaternion 'parallel' to $\boldsymbol{\nu}$ includes the scalar part as well as the component of the vector part parallel to $\boldsymbol{\nu}$. In other words the 'parallel' component of the quaternion is that component which is in the same Argand plane as the axis of involution $\boldsymbol{\nu}$. The component of the quaternion perpendicular to $\boldsymbol{\nu}$ is a vector (the component of the vector part perpendicular to $\boldsymbol{\nu}$, and therefore perpendicular to the Argand plane of the 'parallel' component) since the subtraction cancels out the scalar part. As stated earlier, the representation $a + \boldsymbol{\mu} b$ is independent of the coordinate system in that it expresses the quaternion in terms of the direction in 3-space of the vector part. However, the quaternion can be rewritten in terms of a set of orthogonal basis vectors, $\boldsymbol{\nu}_1$, $\boldsymbol{\nu}_2$, and $\boldsymbol{\nu}_3$, without recourse to a numerical representation. The three projections across $\boldsymbol{\nu}_1$, $\boldsymbol{\nu}_2$, and $\boldsymbol{\nu}_3$ and the conjugate anti-involution provide the mechanism. That is, we may write a quaternion $q = a + \boldsymbol{\mu} b = a + \boldsymbol{b}$ as

$$q = a + \boldsymbol{\nu}_1\alpha + \boldsymbol{\nu}_2\beta + \boldsymbol{\nu}_3\gamma = a + \boldsymbol{b}_1 + \boldsymbol{b}_2 + \boldsymbol{b}_3$$

where $\boldsymbol{b}_i \parallel \boldsymbol{\nu}_i$ and α, β and γ are real:

$$a = \frac{1}{2}(q + \overline{q}); \qquad \boldsymbol{b} = \frac{1}{2}(q - \overline{q}); \qquad \boldsymbol{b}_i = \frac{1}{2}(\boldsymbol{b} + \overline{\boldsymbol{b}}^{\boldsymbol{\nu}_i}), \quad i \in \{1, 2, 3\}$$

References

[1] W. R. Hamilton, Elements of Quaternions, Longmans, Green and Co., London, 1866.

[2] H. S. M. Coxeter, Quaternions and reflections, American Mathematical Monthly 53 (1946) 136–146.

[3] M. Hazewinkel (Ed.), Encyclopaedia of mathematics: an updated and annotated translation of the Soviet 'Mathematical Encyclopaedia', Kluwer, Dordrecht, c1988–94.

[4] V. M. Chernov, Discrete orthogonal transforms with data representation in composition algebras, in: Proceedings Scandinavian Conference on Image Analysis, Uppsala, Sweden, 1995, pp. 357–364.

[5] T. Bülow, Hypercomplex spectral signal representations for the processing and analysis of images, Ph.D. thesis, University of Kiel, Germany (1999).

[6] T. Bülow, G. Sommer, Hypercomplex signals – a novel extension of the analytic signal to the multi-dimensional case, IEEE Trans. Signal Processing 49 (11) (2001) 2844–2852.

[7] T. A. Ell, S. J. Sangwine, Hypercomplex Wiener-Khintchine theorem with application to color image correlation, in: IEEE International Conference on Image Processing (ICIP 2000), Vol. II, Institute of Electrical and Electronics Engineers, Vancouver, Canada, 2000, pp. 792–795.