Chapter 4

Quaternions
4 Quaternions

4.1 Definition and Properties

Quaternions are an extension of the complex numbers. Rather than introducing just one value whose square is $-1$ we introduce three, so we define $i$, $j$ and $k$ such that

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

Moreover we insist that these are different from one another and we relate them as follows:

$$ij = k, \quad jk = i \quad \text{and} \quad ki = j$$

From these rules we get some other facts:
\[ ji = -k, \ kj = -i, \ ik = -j \] and \[ ijk = -1 \]

At this point notice that for example \( ij = -ji \) so multiplication as defined is not commutative when it includes \( i, j \) and \( k \).

We can then define a quaternion.

**Definition 4.1.0.1.** A quaternion is a number of the form:

\[ q = s + ai + bj + ck \quad \text{with} \quad s, a, b, c \in \mathbb{R} \]

We extend addition, subtraction and multiplication to the quaternions by obeying the above rules as well as distributivity and associativity. A *pure quaternion* is a quaternion with scalar part equal to 0.

The set of quaternions is denoted \( \mathbb{H} \).

Warning: It’s written above but can’t be emphasized enough: Quaternion multiplication is not commutative!

**Example 4.1.** if \( q_1 = 2 + i \) and \( q_2 = 3 + 4j \) then:

\[
q_1q_2 = (2 + i)(3 + 4j) \\
= 2(3 + 4j) + i(3 + 4j) \\
= 6 + 8j + 3i + 4ij \\
= 6 + 8j + 3i + 4k
\]

**Example 4.2.** If \( q_1 = 2 + 3i - 2j + 1k \) and \( q_2 = 1 - 1i + 4j + 5k \) then:

\[
q_1q_2 = (2 + 3i - 2j + 1k)(1 - 1i + 4j + 5k) \\
= 2(1 - 1i + 4j + 5k) \\
+ 3i(1 - 1i + 4j + 5k) \\
- 2j(1 - 1i + 4j + 5k) \\
+ 1k(1 - 1i + 4j + 5k) \\
= 2 - 2i + 8j + 10k \\
+ 3i - 3i^2 + 12ij + 15ik \\
- 2j + 2ji - 8j^2 - 10jk \\
+ 1k - 1ki + 4kj + 5k^2 \\
= 2 - 2i + 8j + 10k \\
+ 3i - 3(-1) + 12(k) + 15(-j) \\
- 2j + 2(-k) - 8(-1) - 10(i) \\
+ 1k - 1(j) + 4(-i) + 5(-1) \\
= 8 - 13i - 10j + 21k
\]
It’s worth doing one or two of these just to settle the rules in your head.

**Exercise 4.1.** If $q_1 = 2 + 3i + 5k$ and $q_2 = 1 - 2j + 3k$ find $q_1q_2$ and $q_2q_1$.

The use of $i, j$ and $k$ are not arbitrary, we can use vector operations like cross products and dot products on the non-scalar part and in fact pure quaternions will denoted by vector notation simply to emphasize this point.

This will in fact make calculations like the above much nicer.

**Example 4.3.** If $q_1 = 2i + 3j - 1k$ and $q_2 = 5i + 4j + 6k$ then $q_1 \cdot q_2 = 16$ and $q_1 \times q_2 = 22i - 17j - 7k$. Notice these are both quaternions, the first is just a scalar and the second is pure.

For this reason often quaternions are broken into the scalar term and the vector term and so a quaternion can be written:

$$q = s + v$$

where $s \in \mathbb{R}$ and $v = a_i + b_j + c_k$.

In fact the cross and dot products simplify quaternion multiplication quite a bit as demonstrated by the following:

**Theorem 4.1.0.1.** For quaternions $q_1 = s_1 + v_1$ and $q_2 = s_2 + v_2$ we have:

$$q_1q_2 = (s_1s_2 - v_1 \cdot v_2) + (s_1v_2 + s_2v_1 + v_1 \times v_2)$$

Note: The parentheses are there to distinguish the scalar and vector parts.

**Proof.** This is just brute force calculation. 

**Exercise 4.2.** Do the brute brute force calculation for the above theorem.

As a special case of this we have:

**Corollary 4.1.0.1.** For pure quaternions $v$ and $w$ we have:

$$vw = v \times w - v \cdot w$$

$$wv = w \times v - w \cdot v = -(v \times w) - v \cdot w$$

**Proof.** The first of these follows immediately from the previous theorem when $s_1 = s_2 = 0$.

The second follows from the commutativity of the dot product and the anti-commutativity of the cross product.
It’s worth noting that we can now see that quaternion multiplication is neither commutative nor anticommutative. In other words in general $vw \neq wv$ and in general $vw \neq -wv$.

It follows from this theorem that we can calculate the dot product and cross product from the quaternion product, as the following shows:

**Theorem 4.1.0.2.** For vectors $v$ and $w$ we have:

\[
\begin{align*}
    v \times w &= \frac{1}{2} (vw - wv) \\
    v \cdot w &= -\frac{1}{2} (vw + wv)
\end{align*}
\]

*Proof.* These follow by adding or subtracting the two equations in the previous theorem and then dividing by $\frac{1}{2}$. \hfill \square

**Definition 4.1.0.2.** The *conjugate* of a quaternion $q = s + ai + bj + ck$ is denoted $q^*$ and is defined by:

\[q^* = s - ai - bj - ck\]

We don’t write $\bar{q}$ since $q$ already involves a vector and this could cause confusion. A better way to write this might be:

\[(s + v)^* = s - v\]

**Theorem 4.1.0.3.** If $q = s + ai + bj + ck \in \mathbb{H}$ then

\[qq^* = s^2 + a^2 + b^2 + c^2 = |q|^2\]

*Proof.* Brute force. \hfill \square

**Exercise 4.3.** Work out the brute force.

**Theorem 4.1.0.4.** If $q_1, q_2 \in \mathbb{H}$ then $(q_1 q_2)^* = q_2^* q_1^*$

*Proof.* Brute force. \hfill \square

**Exercise 4.4.** Work out the brute force.

**Definition 4.1.0.3.** The *magnitude* of a quaternion $q = s + ai + bj + ck$ is:

\[|q| = \sqrt{s^2 + a^2 + b^2 + c^2}\]

Note that if $q = s + v$ then $|q|^2 = s^2 + |v|^2$. 
4.1. DEFINITION AND PROPERTIES

**Definition 4.1.0.4.** A unit quaternion is a quaternion with magnitude 1.

**Theorem 4.1.0.5.** It follows that $|q|^2 = qq^*$ and for a unit quaternion $qq^* = 1$.

*Proof.* Immediate from previous theorems. \(\square\)

**Exercise 4.5.** Elaborate the above proof.

As with complex conjugation this allows us to divide quaternions in a sensible manner and allows us to write the result as an obvious quaternion. We do this in general by:

$$
\frac{q_1}{q_2} = \frac{q_1 q_2^*}{|q_2|^2} = \ldots
$$

**Example 4.4.** To divide $1 + 2i + 0j - 3k$ by $2 + 0i + 4j - 1k$ we proceed as follows:

$$
\frac{1 + 2i + 0j - 3k}{2 + 0i + 4j - 1k} = \frac{(1 + 2i + 0j - 3k)(2 - 0i - 4j + 1k)}{2^2 + 0^2 + 4^2 + (-1)^2}
$$

$$
= \ldots
$$

$$
= 5 - 8i - 6j - 13k
$$

$$
= \frac{21}{21}i - \frac{8}{21}j - \frac{13}{21}k
$$

**Exercise 4.6.** Calculate the result of dividing $1 + 1i + 2j$ by $3i - 1j + 5k$.

Unit quaternions are interesting in the sense that they are all square roots of $-1$. So by constructing $\mathbb{H}$ by introducing three new square roots of $-1$ we actually have gained infinitely many.

**Theorem 4.1.0.6.** $q \in \mathbb{H}$ is a unit pure quaternion then $q^2 = -1$.

*Proof.* For a general quaternion $q = s + ai + bj + ck$ we have by brute force:

$$
q^2 = (s^2 - a^2 - b^2 - c^2) + 2asi + 2bsj + 2csk
$$

If $q$ is a unit pure quaternion then $s = 0$ and $a^2 + b^2 + c^2 = 1$ and the result follows immediately.
On the other hand suppose \( q^2 = -1 \). Then we have all of:

\[
\begin{align*}
  s^2 - a^2 - b^2 - c^2 &= -1 \\
  2as &= 0 \\
  2bs &= 0 \\
  2cs &= 0
\end{align*}
\]

We cannot have \( s \neq 0 \) since that would imply \( a = b = c = 0 \) from the last three which contradicts the first. Thus we must have \( s = 0 \) in which case the first yields \( a^2 + b^2 + c^2 = 1 \) and therefore \( |q| = 1 \).

Consider what this states. If we think of unit pure quaternions as vectors (which they are) then they form the sphere of radius 1 centered at the origin. So in \( \mathbb{H} \) there are a sphere’s worth of square roots of \(-1\).

**Definition 4.1.0.5.** A quaternion \( q \) is invertible if there is another quaternion, denoted \( q^{-1} \), such that \( qq^{-1} = q^{-1}q = 1 \).

**Theorem 4.1.0.7.** All nonzero quaternions are invertible.

**Proof.** The inverse of \( q = s + ai + bj + ck \) is:

\[
q^{-1} = \frac{q^*}{|q|^2} = \frac{s - ai - bj - ck}{|q|^2}
\]

Observe that:

\[
q \left( \frac{q^*}{|q|^2} \right) = \frac{qq^*}{|q|^2} = \frac{|q|^2}{|q|^2} = 1
\]

**Exercise 4.7.** Calculate the inverse of \( 2 + 4i - 2j + 3k \).

**Corollary 4.1.0.2.** If \( q = s + ai + bj + ck \) is a unit quaternion then

\[
q^{-1} = s - ai - bj - ck = q^*
\]

To close out the section here’s a useful formula which is true of vectors and hence of pure quaternions:

**Theorem 4.1.0.8.** (Lagrange’s Formula aka triple product expansion)

Given vectors (pure quaternions) \( a, b \) and \( c \) we have:

\[
a \times (b \times c) = (a \cdot c)b - (a \cdot b)c
\]

**Proof.** Omitted. This is brute force.

\[\square\]
4.2 Graphical Representation of Quaternions

The only quaternions we represent graphically are pure quaternions, meaning those of the form $ai + bj + ck$. We represent these either as vectors or as points, depending on how we’re using them.

It may seem odd that we even need to have scalars to begin with but as we’ve seen, we can’t have $i$, $j$, and $k$ without having scalars since dot and quaternion products often yield results with a scalar part.

What basically happens is that our standard transformations end up taking pure quaternions to pure quaternions and the scalar part can be assigned to be zero and left that way. This will be elaborated on as we go.

4.3 Translations

If a point $xi + yj + zk$ is to be translated in 3D space we simply add or subtract another quaternion.

Example 4.5. To shift $2i + 3j - 1k$ by 5 in the $x$-direction, 2 in the $y$-direction, and 7 in the $z$-direction we simply do:

$$2i + 3j - 1k \mapsto 2i + 3j - 1k + 5i + 2j + 7k = 7i + 5j + 6k$$

4.4 Rotations

4.4.1 Lines through the Origin

It turns out that extending complex numbers to quaternions allows rotations to extend to three dimensions in a very convenient way. It permits us to easily construct a formula for rotation about an arbitrary axis.

First a well-known formula. While this formula does the job it is complicated from an algebraic point of view, meaning it’s fine for doing a simple calculation but it’s not the type of calculation we want to carry about.

Theorem 4.4.1.1. (Rodrigues Rotation Formula)
Suppose $\hat{u}$ is a unit vector and $v$ is some vector. Then the result of rotating $v$ around $\hat{u}$ by an angle $\theta$ counterclockwise with regards to the right-hand rule equals:

$$\text{Rot}(v) = (1 - \cos \theta)(\hat{u} \cdot v)\hat{u} + (\cos \theta)v + (\sin \theta)(\hat{u} \times v)$$
Proof. We begin by breaking \( \mathbf{v} \) into components, one perpendicular to \( \hat{\mathbf{u}} \) and one parallel to \( \hat{\mathbf{u}} \):

\[
\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel
\]

In order to rotate \( \mathbf{v} \) we leave \( \mathbf{v}_\parallel \) alone, rotate \( \mathbf{v}_\perp \) and then add \( \mathbf{v}_\parallel + \text{Rot}(\mathbf{v}_\perp) \).
That is:

\[
\text{Rot}(\mathbf{v}) = \mathbf{v}_\parallel + \text{Rot}(\mathbf{v}_\perp)
\]

The reason for this is illustrated by this picture:

![Diagram showing the rotation of \( \mathbf{v} \) into \( \text{Rot}(\mathbf{v}) \) with \( \mathbf{v}_\parallel \) and \( \mathbf{v}_\perp \) components.

The calculation for \( \text{Rot}(\mathbf{v}_\perp) \) is a specific example of the 2D case from Chapter 2 which used with \( \mathbf{v}_\perp \) tells us that:

\[
\text{Rot}(\mathbf{v}_\perp) = (\cos \theta) \mathbf{v}_\perp + (\sin \theta)(\hat{\mathbf{u}} \times \mathbf{v}_\perp)
\]

If we use this along with the facts that:

\[
\mathbf{v}_\parallel = (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}}
\]
\[
\mathbf{v}_\perp = \mathbf{v} - (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}}
\]

So now we calculate:
4.4. ROTATIONS

\[ \text{Rot}(v) = v_\parallel + \text{Rot}(v_\perp) \]
\[ = v_\parallel + (\cos \theta)v_\perp + (\sin \theta)(\hat{u} \times v_\perp) \]
\[ = (\hat{u} \cdot v)\hat{u} + (\cos \theta)(v - (\hat{u} \cdot v)\hat{u}) + (\sin \theta)(\hat{u} \times (v - (\hat{u} \cdot v)\hat{u})) \]
\[ = (1 - \cos \theta)\hat{u} \cdot v + (\sin \theta)(\hat{u} \times v) \]

It's worth taking a minute to verify that all this makes sense. Each term is independently a scalar times a vector so the end result is a linear combination of \( \hat{u}, v \) and \( \hat{u} \times v \).

**Exercise 4.8.** Use RRF to calculate the result of rotating \( v = 2i + 3j - 1k \) about \( u = 3i + 4j - 5k \) by \( 7\pi/6 \) radians. Note that \( u \) has not been normalized so do this first.

And now our theorem:

**Theorem 4.4.1.2.** Suppose \( \hat{u} \) is a unit vector and \( v \) is some vector. Then the result of rotating \( v \) around \( \hat{u} \) by an angle \( \theta \) counterclockwise with regards to the right-hand rule can be obtained by letting:

\[ p = \cos \left( \frac{\theta}{2} \right) + \sin \left( \frac{\theta}{2} \right) \hat{u} \]

and then doing:

\[ v_{\text{Rot}} = pvp^{-1} = pvp^* \]

**Proof.** This is just calculation. First note that since \( |p| = 1 \) that \( p^{-1} = p^* \). Then consider:

\[ pvp^* = \left( \cos \left( \frac{\theta}{2} \right) + \sin \left( \frac{\theta}{2} \right) \hat{u} \right) v \left( \cos \left( \frac{\theta}{2} \right) - \sin \left( \frac{\theta}{2} \right) \hat{u} \right) \]
\[ = \cos^2 \left( \frac{\theta}{2} \right) v + \sin \left( \frac{\theta}{2} \right) \cos \left( \frac{\theta}{2} \right) (\hat{u}v - v\hat{u}) - \sin^2 \left( \frac{\theta}{2} \right) \hat{u}v\hat{u} \]
\[ = \cos^2 \left( \frac{\theta}{2} \right) + 2\sin \left( \frac{\theta}{2} \right) \cos \left( \frac{\theta}{2} \right) (\hat{u} \times v) - \sin^2 \left( \frac{\theta}{2} \right) (-2(\hat{u} \cdot v)\hat{u} + v) \]
\[ = \left( \cos^2 \left( \frac{\theta}{2} \right) - \sin^2 \left( \frac{\theta}{2} \right) \right) v + 2\sin \left( \frac{\theta}{2} \right) \cos \left( \frac{\theta}{2} \right) (\hat{u} \times v) + 2\sin^2 \left( \frac{\theta}{2} \right) (\hat{u} \cdot v) \hat{u} \]
\[ = (\cos \theta) v + (\sin \theta)(\hat{u} \times v) + (1 - \cos \theta)(\hat{u} \cdot v) \hat{u} \]
\[ = v_{\text{Rot}} \]
Note: There are two substitutions here that may not be clear:

(a) Earlier we showed that
\[ \mathbf{v} \times \mathbf{w} = \frac{1}{2}(\mathbf{vw} - \mathbf{wv}) \]
In this case we get
\[ \hat{\mathbf{u}} \times \mathbf{v} = \frac{1}{2}(\hat{\mathbf{u}}\mathbf{v} - \mathbf{v}\hat{\mathbf{u}}) \]
and so
\[ \hat{\mathbf{u}}\mathbf{v} - \mathbf{v}\hat{\mathbf{u}} = 2(\hat{\mathbf{u}} \times \mathbf{v}) \]

(b) Earlier we showed that
\[ \mathbf{v} \cdot \mathbf{w} = -\frac{1}{2}(\mathbf{vw} + \mathbf{wv}) \]
In this case we get
\[ \hat{\mathbf{u}} \cdot \mathbf{v} = -\frac{1}{2}(\hat{\mathbf{u}}\mathbf{v} + \mathbf{v}\hat{\mathbf{u}}) \]
and so
\[ \hat{\mathbf{u}}\mathbf{v} = -2(\hat{\mathbf{u}} \cdot \mathbf{v}) - \mathbf{v}\hat{\mathbf{u}} \]
and so
\[ \hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}} = (-2(\hat{\mathbf{u}} \cdot \mathbf{v}) - \mathbf{v}\hat{\mathbf{u}})\hat{\mathbf{u}} = -2(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} - \mathbf{v}\hat{\mathbf{u}}\hat{\mathbf{u}} = -2(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} + \mathbf{v} \]
keeping in mind that
\[ \hat{\mathbf{u}}\hat{\mathbf{u}} = \hat{\mathbf{u}} \times \hat{\mathbf{u}} - \hat{\mathbf{u}} \cdot \hat{\mathbf{u}} = 0 - \hat{\mathbf{u}} \cdot \hat{\mathbf{u}} \]

This should and should not surprise you. In \( \mathbb{C} \) it was multiplication by \( \cos \theta + i\sin \theta \) which did rotation and so this \( \mathbf{p} \) should remind you a little of that.

In this case it’s not a simple multiplication but rather a pair of multiplications. There’s some elegant beauty in the fact that each of those multiplication involves half the required overall angle.

It is also worth noting that although \( \mathbf{p} \) is not a pure quaternion the result of the calculation \( \mathbf{p}\mathbf{v}\mathbf{p}^* \) where \( \mathbf{v} \) is a pure quaternion results in a pure quaternion.

**Example 4.6.** To rotate \( \mathbf{v} = 2\mathbf{i} + 1\mathbf{j} \) by \( \theta = \pi / 3 \) radians about \( \hat{\mathbf{u}} = \frac{1}{\sqrt{2}}\mathbf{j} + \frac{1}{\sqrt{2}}\mathbf{k} \) we set:
\[ \mathbf{p} = \cos(\pi/6) + \sin(\pi/6) \left( \frac{1}{\sqrt{2}}\mathbf{j} + \frac{1}{\sqrt{2}}\mathbf{k} \right) = \frac{\sqrt{3}}{2} + \frac{1}{2\sqrt{2}}\mathbf{j} + \frac{1}{2\sqrt{2}}\mathbf{k} \]
4.5. REFLECTIONS

and then the result is:

\[ pvp^* = \left( \frac{\sqrt{3}}{2} + \frac{1}{2\sqrt{2}} + \frac{1}{2\sqrt{2}} \right) \left( 2i + 1j \right) \left( \frac{\sqrt{3}}{2} - \frac{1}{2\sqrt{2}} - \frac{1}{2\sqrt{2}} \right) \]
\[ \approx \ldots \text{Matlab} \ldots \]
\[ \approx 0 + 0.88763i + 1.7247j + 1.1124k \]

**Exercise 4.9.** Use the above formula to rotate \( v = i \) about \( \hat{u} = k \) by \( \pi/2 \) radians. Does this correspond to your expectations? Hint: What should it do?

**Exercise 4.10.** Use the above formula to rotate \( v = 1i + 1j + 2k \) about \( u = 2i + 0j + 3k \) by \( 2\pi/3 \) radians.

4.4.2 Lines Not Through the Origin

To rotate about a line not through the origin the process is simple. We take a point on the line and translate that point to the origin, then rotate, then translate back. Note that the direction vector for the axis does not change.

Thus if we have a line containing point \( v_0 \) with unit direction vector \( \hat{u} \) then rotation about this line can be done by:

\[ v \mapsto p(v - v_0)p^* + v_0 \]

where \( p \) is as before for rotations.

**Exercise 4.11.** Find the result of rotating \( v = 2i + 3j + 1k \) by \( \pi/4 \) about the line passing through \((0, 0, 1)\) with direction \( \bar{u} = 3i + 2j + 0k \).

Note that you need to make \( \bar{u} \) into \( \hat{u} \).

4.5 Reflections

4.5.1 In Planes Through the Origin

We'll focus first on planes through the origin since other planes may be dealt with through translation.

First we must clarify how we are to represent a plane, but this is easy. Since pure quaternions are equivalent to vectors we may take the standard Calculus 3 approach and simply choose a unit pure quaternion which will represent the normal vector to the plane.

It turns out we get a particularly nice formula:
Theorem 4.5.1.1. Given a plane $\mathcal{P}$ through the origin represented by the unit pure quaternion (unit normal vector) $\mathbf{n}$ the reflection of the vector $\mathbf{v}$ is given by:

$$\mathbf{v} \mapsto \hat{\mathbf{n}}\mathbf{v}\hat{\mathbf{n}}$$

Proof. Any vector may be decomposed into the sum of two vectors, one in $\mathcal{P}$ (perpendicular to $\mathbf{n}$) and one perpendicular to $\mathcal{P}$ (a multiple of $\mathbf{n}$), using standard vector projection. Reflecting in $\mathcal{P}$ involves leaving the part in $\mathcal{P}$ alone and negates the perpendicular part.

We'll look at these two parts independently.

Suppose first that $\mathbf{w} \in \mathcal{P}$. Then observe that:

$$\hat{\mathbf{n}}\mathbf{w}\hat{\mathbf{n}} = \hat{\mathbf{n}}(\mathbf{w} \times \hat{\mathbf{n}} - \mathbf{w} \cdot \hat{\mathbf{n}})$$

$$= \hat{\mathbf{n}}(\mathbf{w} \times \hat{\mathbf{n}} - 0)$$

$$= \hat{\mathbf{n}} \times (\mathbf{w} \times \hat{\mathbf{n}}) - \hat{\mathbf{n}} \cdot (\mathbf{w} \times \hat{\mathbf{n}})$$

$$= (\mathbf{w} \times \hat{\mathbf{n}}) - (\hat{\mathbf{n}} \cdot \mathbf{w})\hat{\mathbf{n}}$$

$$= (1)\mathbf{w} - (0)\hat{\mathbf{n}}$$

$$= \mathbf{w}$$

Suppose second that $\mathbf{w}$ is perpendicular to $\mathcal{P}$ so that $\mathbf{w} = \alpha \hat{\mathbf{n}}$ for $\alpha \in \mathbb{R}$. Then:

$$\hat{\mathbf{n}}\mathbf{w}\hat{\mathbf{n}} = \hat{\mathbf{n}}(\alpha \hat{\mathbf{n}})\hat{\mathbf{n}}$$

$$= \alpha \hat{\mathbf{n}}\hat{\mathbf{n}}\hat{\mathbf{n}}$$

$$= \alpha \hat{\mathbf{n}}(\hat{\mathbf{n}} \times \hat{\mathbf{n}} - \mathbf{n} \cdot \hat{\mathbf{n}})$$

$$= \alpha \hat{\mathbf{n}}(0 - 1)$$

$$= -\alpha \hat{\mathbf{n}}$$

$$= -\mathbf{v}$$
4.5. REFLECTIONS

So now for any \( \mathbf{v} \) we write \( \mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel \) and then:

\[
\hat{n}\mathbf{v}\hat{n} = \hat{n}(\mathbf{v}_\perp + \mathbf{v}_\parallel)\hat{n} \\
= \hat{n}\mathbf{v}_\perp\hat{n} + \hat{n}\mathbf{v}_\parallel\hat{n} \\
= -\mathbf{v}_\perp + \mathbf{v}_\parallel
\]

This result is the reflection. □

Note that if we have \( \mathbf{n} \) not normalized then to normalize we simply divide by \(|\mathbf{n}|\) and the formula can be rewritten as:

\[
\mathbf{v} \mapsto \left(\frac{1}{|\mathbf{n}|^2}\right)\mathbf{n}\mathbf{v}\mathbf{n}
\]

**Example 4.7.** To reflect \( \mathbf{v} = 3\mathbf{i} + 1\mathbf{j} + 1\mathbf{k} \) in the plane through the origin with normal vector \( \mathbf{n} = 2\mathbf{i} + 2\mathbf{j} - 1\mathbf{k} \) we calculate:

\[
3\mathbf{i} + 1\mathbf{j} + 1\mathbf{k} \mapsto \left(\frac{1}{3}\right)(2\mathbf{i} + 2\mathbf{j} - 1\mathbf{k})(3\mathbf{i} + 1\mathbf{j} + 1\mathbf{k})(2\mathbf{i} + 2\mathbf{j} - 1\mathbf{k}) \\
\mapsto \ldots \\
\mapsto \left(\frac{1}{3}\right)(-1\mathbf{i} - 19\mathbf{j} + 23\mathbf{k})
\]

**Exercise 4.12.** Calculate the result of reflecting \( \mathbf{v} = 3\mathbf{i} + 2\mathbf{j} + 1\mathbf{k} \) in the plane through the origin with normal vector \( \hat{n} = \mathbf{i} \). Is this what you expect?

**Exercise 4.13.** Calculate the result of reflecting \( \mathbf{v} = 15\mathbf{i} + 10\mathbf{j} - 20\mathbf{k} \) in the plane through the origin with normal vector \( \mathbf{n} = 1\mathbf{i} + 1\mathbf{j} + 2\mathbf{k} \).

4.5.2 Two Reflections (Still) Make a Rotation

It ought to seem reasonable at this point that if we reflect in two planes through the origin, one after the other, that the result is a rotation about the axis formed by the intersection of the two. Let’s check that this is the result. Suppose two planes \( \mathcal{P}_1 \) and \( \mathcal{P}_2 \) have unit normal vectors \( \hat{n}_1 \) and \( \hat{n}_2 \) respectively and meet at an angle of \( \theta \).

The axis formed by the intersection of the two has vector \( \hat{n}_1 \times \hat{n}_2 \) but this is probably not a unit vector.

The unit vector \( \hat{u} \) would be:

\[
\hat{u} = \frac{\hat{n}_1 \times \hat{n}_2}{|\hat{n}_1 \times \hat{n}_2|}
\]
and would satisfy:

\[
\hat{n}_1 \times \hat{n}_2 = \frac{\hat{n}_1 \times \hat{n}_2}{|\hat{n}_1 \times \hat{n}_2|} |\hat{n}_1 \times \hat{n}_2|
= \hat{u} |\hat{n}_1 \times \hat{n}_2|
= \hat{u} |\hat{n}_1||\hat{n}_2| \sin \theta
= \hat{u} (1)(1) \sin \theta
\]

Keeping in mind also that:

\[
\hat{n}_1 \cdot \hat{n}_2 = |\hat{n}_1||\hat{n}_2| \cos \theta = \cos \theta
\]

The double-reflection will then be:

\[
v \mapsto -\hat{u} v \hat{u}
\]

This is exactly equal to a rotation of 2\(\theta\) radians about the axis \(\hat{u}\).

The direction of rotation here is by the right-hand rule applied to the vector \(\hat{u}\). This vector arose from \(\hat{n}_1 \times \hat{n}_2\) and so the direction of \(\hat{u}\) is such that the right-hand rule rotates \(P_1\) toward \(P_2\) the short-way around, meaning closing the smaller angle between them.

### 4.5.3 In Lines Through the Origin

It’s also possible in three dimensions to reflect through a line as shown in the theorem below.

**Theorem 4.5.3.1.** Given a line \(L\) through the origin represented by the unit pure quaternion \(\hat{u}\) the reflection of the vector \(v\) is given by:

\[
v \mapsto -\hat{u} v \hat{u}
\]

**Proof.** Notice how similar this is to the previous theorem. This is not a coincidence and the proof is very similar, read that one first!

We decompose \(v\) into the sum of two vectors, one perpendicular to \(\hat{u}\) and one parallel to (a multiple of) \(\hat{u}\). Here’s where the proof differs. In this case reflecting in \(L\) involves leaving the parallel part intact and negating the perpendicular part.
The result follows.

As with rotation, if the vector $\mathbf{u}$ is not a unit vector then we can factor out the normalization:

$$
\mathbf{v} \mapsto \frac{1}{|\mathbf{u}|^2} \mathbf{uvu}
$$

**Exercise 4.14.** Find the result when the vector $\mathbf{v} = 10\mathbf{i} + 12\mathbf{j} + 8\mathbf{k}$ is reflected in the axis $\hat{\mathbf{u}} = \mathbf{k}$. Is the result what you expect?

**Exercise 4.15.** Find the result when the vector $\mathbf{v} = 10\mathbf{i} + 12\mathbf{j} + 8\mathbf{k}$ is reflected in the axis $\mathbf{u} = 5\mathbf{i} + 1\mathbf{j} + 2\mathbf{k}$.

### 4.5.4 Reflections in Other Planes and Lines

To reflect in a plane not through the origin the process is simple. We take a point on the plane and translate that point to the origin, then reflect, then translate back. Note that the normal vector for the plane does not change.

Thus if $\hat{\mathbf{n}}$ is the unit normal vector for the plane and $\mathbf{v}_0$ is a point on the plane then reflection in the plane will be given by:

$$
\mathbf{v} \mapsto \hat{\mathbf{n}}(\mathbf{v} - \mathbf{v}_0)\hat{\mathbf{n}} + \mathbf{v}_0
$$

Reflection in a line works similarly:

$$
\mathbf{v} \mapsto -\hat{\mathbf{u}}(\mathbf{v} - \mathbf{v}_0)\hat{\mathbf{u}} + \mathbf{v}_0
$$
Exercise 4.16. Find the result when \( \mathbf{v} = 3\mathbf{i} + 3\mathbf{j} + 10\mathbf{k} \) is reflected in the plane \( 2x + 4y + 4z = 12 \) with normal vector arising from the coefficients.

Exercise 4.17. Find the result when \( \mathbf{v} = 3\mathbf{i} + 3\mathbf{j} + 10\mathbf{k} \) is reflected in the line through \((1, 1, 2)\) with axis \( \mathbf{u} = \mathbf{i} + 2\mathbf{j} + 2\mathbf{k} \).

4.5.5 Summary

It’s worth summarizing to notice how similar all these formulas are. We have the following:

- Rotation about a line \( \hat{\mathbf{u}} \) \( \mathbf{v} \mapsto \mathbf{pvp}^* \) where \( \mathbf{p} = \cos(\theta/2) + \sin(\theta/2)\hat{\mathbf{u}} \)
- Reflection in a plane \( \hat{\mathbf{n}} \) \( \mathbf{v} \mapsto \hat{\mathbf{n}}\mathbf{v}\hat{\mathbf{n}} \)
- Reflection in a line \( \hat{\mathbf{u}} \) \( \mathbf{v} \mapsto -\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}} \)

This is the beauty in using quaternions and will be similar in geometric algebra. Geometric transformations are represented by calculations which are algebraically speaking quite simple. In this case multiplication of quaternions gives us rotation and two different reflections in extremely similar forms.

4.6 Transformations of Lines and Planes

4.6.1 Representations

The most direct way to store a line in \( \mathbb{R}^3 \) which does not pass through the origin is with an anchor point and a direction \((\mathbf{v}_0, \mathbf{L})\). Then the line consists of all points of the form:

\[
\mathbf{v}(t) = \mathbf{v}_0 + t\mathbf{L}
\]

Likewise the most direct way to store a plane in \( \mathbb{R}^3 \) which does not pass through the origin is with an anchor point and a normal vector \((\mathbf{v}_0, \hat{\mathbf{n}})\). Then the line consists of all points \( \mathbf{v}(t) \) satisfying:

\[
\hat{\mathbf{n}} \cdot (\mathbf{v} - \mathbf{v}_0) = 0
\]

4.6.2 Transformations

To transform a line we simply apply the transformation in the appropriate manner.

(a) To translate the line parametrized by \((\mathbf{v}_0, \mathbf{L})\) we translate the anchor point \( \mathbf{v}_0 \). The direction doesn’t change so \( \mathbf{L} \) is not touched. That is:

\[
(\mathbf{v}_0, \mathbf{L}) \mapsto (\mathbf{v}_0 + q, \mathbf{L})
\]
(b) To rotate the line parametrized by \((v_0, L)\) we rotate both the anchor point and the vector. That is:

\[
(v_0, L) \mapsto (pv_0, pLp^*)
\]

If it’s not clear why this is, simply observe that the original line passes through the points \(v_0\) and \(v_0 + L\) and so the rotated line must pass through the points \(pv_0p^*\) and \(p(v_0 + L)p^* = pv_0p^* + pLp^*\) and hence has direction vector \(pLp^*\). This is precisely our statement.

(c) Reflections in both a plane and in a line follow the same approach as rotations for the same reason.

(d) Rotations of lines about axes not through the origin and reflections of lines in planes and lines not through the origin must be done using translations as we did with points. Likewise with rotations and reflections of planes.

**Exercise 4.18.** Find the result when the line \((2i + 0j + 1k, 3i + 1j + 1k)\) is rotated about the axis through the origin with \(u = 2i + 2j + 1k\).

**Exercise 4.19.** Find the result when the line \((2i + 0j + 1k, 3i + 1j + 1k)\) is rotated about the axis through \((10, 10, 0)\) with \(u = 2i + 2j + 1k\).

Note: First translate so the axis passes through the origin, then rotate, then translate back.

**Exercise 4.20.** Find the result when the line \((2i + 0j + 1k, 3i + 1j + 1k)\) is reflected in the plane through the origin with normal vector \(n = 1i + 1j + 2k\).

**Exercise 4.21.** Find the result when the line \((2i + 0j + 1k, 3i + 1j + 1k)\) is reflected in the plane through \((4, 3, 0)\) with normal vector \(n = 1i + 1j + 2k\).

To transform a plane we do exactly the same as the above with \(\hat{n}\) in place of \(L\).

**Example 4.8.** Find the result when the plane \(x + 2y + z = 4\) with normal vector arising from the coefficients is rotated about the axis through the origin with \(u = 4i + 6j + 3k\).

Hint: The plane can be thought of as \((v_0, \hat{n})\) where \(v_0\) is any point on the plane and \(\hat{n}\) is the unit vector arising from the coefficients.

**Example 4.9.** Find the result when the plane \(x + 2y + z = 4\) with normal vector arising from the coefficients is rotated about the axis through \((5, 5, 5)\) with \(u = 4i + 6j + 3k\).

**Example 4.10.** Find the result when the plane \(x - y + z = 1\) with normal vector arising from the coefficients is reflected in the plane \(x + y + 2z = 0\), with normal vector also arising from the coefficients.
Example 4.11. Find the result when the plane $x - y + z = 1$ with normal vector arising from the coefficients is reflected in the plane $x + y + 2z = 10$, with normal vector also arising from the coefficients.

Example 4.12. Find the result when the plane $x - y + z = 1$ with normal vector arising from the coefficients is reflected in the line through the origin with direction $\mathbf{u} = 4\mathbf{i} + 1\mathbf{j} + 0\mathbf{k}$.

Example 4.13. Find the result when the plane $x - y + z = 1$ with normal vector arising from the coefficients is reflected in the line through $(1, 2, 2)$ with direction $\mathbf{u} = 4\mathbf{i} + 1\mathbf{j} + 0\mathbf{k}$.

4.7 The Downsides of Quaternions

The quaternions are pretty great, but it’s worth pointing out a couple of issues. They don’t obviously generalize. For example in $\mathbb{R}^2$ there’s no obvious way to write rotation about a point using a product like $\mathbf{p}\mathbf{v}\mathbf{p}^*$. We might think of $\mathbb{C}$ as the “2D Version” of $\mathbb{H}$ but that’s not really true. With complex numbers we need exponentials, and trigonometry to represent our transformations.

In an opposite direction it’s not clear whether or how the quaternions might extend to higher dimensions.

And even in $\mathbb{H}$ (and in Calculus 3!) it’s interesting to point out that to represent a plane we use a normal vector. A quick thought indicates that this is peculiar since the normal vector is indicating the direction the plane doesn’t go, and we simply take it for granted that the plane is perpendicular. We don’t do this with lines, so why do we do this for planes? The answer is that there’s no clear way in $\mathbb{H}$ to denote a plane in an algebraic way which talks about what the plane is, rather than what it isn’t.

The cross product (which we love) only really makes sense in $\mathbb{R}^3$ (it actually makes sense in $\mathbb{R}^7$ too but that’s another story) and this is very specific.

Geometric algebra does the job of abstracting the quaternions in a way that resolves all these issues.
Index

conjugate, 6
invertible, 8
Lagrange’s Formula, 8
magnitude, 6
pure quaternion, 4
quaternion, 4
triple product expansion, 8
unit quaternion, 7