MATH431: Quaternions

Justin Wyss-Gallifent

December 5, 2021

1 Definitions

Quaternions are essentially an extension of the complex numbers. Rather than introducing just one value whose square is -1 we introduce three.

Definition 1.0.1. We define $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{k}}$ such that

$$\hat{\boldsymbol{\imath}}^2 = \hat{\boldsymbol{\jmath}}^2 = \hat{\boldsymbol{k}}^2 = -1$$

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

1. $\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}} = +\hat{\boldsymbol{k}}$
2. $\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} = +\hat{\boldsymbol{\imath}}$
3. $\hat{\boldsymbol{k}}\hat{\boldsymbol{\imath}} = +\hat{\boldsymbol{\jmath}}$

□

Theorem 1.0.1. From these rules we get some other facts:

(a) $\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} = -\hat{\boldsymbol{k}}$

(b) $\hat{\boldsymbol{k}}\hat{\boldsymbol{\jmath}} = -\hat{\boldsymbol{\imath}}$

(c) $\hat{\boldsymbol{\imath}}\hat{\boldsymbol{k}} = -\hat{\boldsymbol{\jmath}}$

(d) $\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} = \hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}}\hat{\boldsymbol{\imath}} = \hat{\boldsymbol{k}}\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}} = -1$

(e) $\hat{\boldsymbol{k}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} = \hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}}\hat{\boldsymbol{k}} = \hat{\boldsymbol{\imath}}\hat{\boldsymbol{k}}\hat{\boldsymbol{\jmath}} = +1$

Proof. For example:

$$\begin{aligned}
\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}} &= \hat{\boldsymbol{k}} \\
\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} &= \hat{\boldsymbol{k}}\hat{\boldsymbol{k}} \\
\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} &= -1 \\
\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} &= -\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} \\
\hat{\boldsymbol{\jmath}}(-1)\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} &= -\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} \\
-(\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\jmath}})\hat{\boldsymbol{k}} &= -\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} \\
-(-1)\hat{\boldsymbol{k}} &= -\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} \\
\hat{\boldsymbol{k}} &= -\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} \\
-\hat{\boldsymbol{k}} &= \hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}}
\end{aligned}$$

□

At this point notice that for example $\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}} = -\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}}$ so multiplication as defined is not commutative when it includes $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{k}}$. As we'll see soon it's not anti-commutative either when we go full-on with quaternions.

Exercise 1.1. Using a variation on the above proof show that $\hat{\boldsymbol{k}}\hat{\boldsymbol{\jmath}} = -\hat{\boldsymbol{\imath}}$.

□

We can then define a quaternion.

Definition 1.0.2. A *quaternion* is an expression of the form:

$$q = s + a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}} \text{ with } s, a, b, c \in \mathbb{R}$$

We extend addition, subtraction and multiplication to the quaternions by obeying the above rules as well as distributivity and associativity.

The set of quaternions is denoted $\mathbb{H}$. This is for William Rowan Hamilton who first described them in 1843.

□

Example 1.1. If $q_1 = 2 + \hat{\boldsymbol{\imath}}$ and $q_2 = 3 + 4\hat{\boldsymbol{\jmath}}$ then:

$$\begin{aligned}
q_1 q_2 &= (2 + \hat{\boldsymbol{\imath}})(3 + 4\hat{\boldsymbol{\jmath}}) \\
&= 2(3 + 4\hat{\boldsymbol{\jmath}}) + \hat{\boldsymbol{\imath}}(3 + 4\hat{\boldsymbol{\jmath}}) \\
&= 6 + 8\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}} \\
&= 6 + 8\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{k}} \\
&= 6 + 3\hat{\boldsymbol{\imath}} + 8\hat{\boldsymbol{\jmath}} + 4\hat{\boldsymbol{k}}
\end{aligned}$$

Just to compare, note that:

$$\begin{aligned}
q_2 q_1 &= (3 + 4\hat{\boldsymbol{\jmath}})(2 + \hat{\boldsymbol{\imath}}) \\
&= 3(2 + \hat{\boldsymbol{\imath}}) + 4\hat{\boldsymbol{\jmath}}(2 + \hat{\boldsymbol{\imath}}) \\
&= 6 + 3\hat{\boldsymbol{\imath}} + 8\hat{\boldsymbol{\jmath}} + 4\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} \\
&= 6 + 3\hat{\boldsymbol{\imath}} + 8\hat{\boldsymbol{\jmath}} - 4\hat{\boldsymbol{k}} \\
&= 6 + 3\hat{\boldsymbol{\imath}} + 8\hat{\boldsymbol{\jmath}} - 4\hat{\boldsymbol{k}}
\end{aligned}$$

These are different!

□

Example 1.2. If $q_1 = 2 + 3\hat{\boldsymbol{\imath}} - 2\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}$ and $q_2 = 1 - 1\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + 5\hat{\boldsymbol{k}}$ then:

$$\begin{aligned}
q_1q_2 &= (2 + 3\hat{\boldsymbol{\imath}} - 2\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}})(1 - 1\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + 5\hat{\boldsymbol{k}}) \\
&= 2(1 - 1\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + 5\hat{\boldsymbol{k}}) \\
&\quad + 3\hat{\boldsymbol{\imath}}(1 - 1\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + 5\hat{\boldsymbol{k}}) \\
&\quad - 2\hat{\boldsymbol{\jmath}}(1 - 1\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + 5\hat{\boldsymbol{k}}) \\
&\quad + 1\hat{\boldsymbol{k}}(1 - 1\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + 5\hat{\boldsymbol{k}}) \\
&= 2 - 2\hat{\boldsymbol{\imath}} + 8\hat{\boldsymbol{\jmath}} + 10\hat{\boldsymbol{k}} \\
&\quad + 3\hat{\boldsymbol{\imath}} - 3\hat{\boldsymbol{\imath}}^2 + 12\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}} + 15\hat{\boldsymbol{\imath}}\hat{\boldsymbol{k}} \\
&\quad - 2\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{\imath}} - 8\hat{\boldsymbol{\jmath}}^2 - 10\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} \\
&\quad + 1\hat{\boldsymbol{k}} - 1\hat{\boldsymbol{k}}\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{k}}\hat{\boldsymbol{\jmath}} + 5\hat{\boldsymbol{k}}^2 \\
&= 2 - 2\hat{\boldsymbol{\imath}} + 8\hat{\boldsymbol{\jmath}} + 10\hat{\boldsymbol{k}} \\
&\quad + 3\hat{\boldsymbol{\imath}} - 3(-1) + 12(\hat{\boldsymbol{k}}) + 15(-\hat{\boldsymbol{\jmath}}) \\
&\quad - 2\hat{\boldsymbol{\jmath}} + 2(-\hat{\boldsymbol{k}}) - 8(-1) - 10(\hat{\boldsymbol{\imath}}) \\
&\quad + 1\hat{\boldsymbol{k}} - 1(\hat{\boldsymbol{\jmath}}) + 4(-\hat{\boldsymbol{\imath}}) + 5(-1) \\
&= 8 - 13\hat{\boldsymbol{\imath}} - 10\hat{\boldsymbol{\jmath}} + 21\hat{\boldsymbol{k}}
\end{aligned}$$

□

It's worth doing one or two of these just to settle the rules in your head.

Exercise 1.2. If $q_1 = 2 + 3\hat{\boldsymbol{\imath}} + 5\hat{\boldsymbol{k}}$ and $q_2 = 1 - 2\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$ find q_1q_2 and q_2q_1.

□

Definition 1.0.3. For a quaternion $q = s + a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}$ we have:

(a) The real, or scalar, part of q, denoted $\mathrm{Re}(q) = s$.

(b) The imaginary, or vector, part of q, denoted $\mathrm{Im}(q) = a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}$.

□

Definition 1.0.4. A *pure quaternion* (also a *vector quaternion*) is a quaternion with scalar part equal to 0.

□

Definition 1.0.5. A *scalar* (also a *scalar quaternion* or a *real quaternion*)) is a quaternion with vector part equal to 0.

□

Example 1.3. $2+3\hat{\boldsymbol{\imath}}-1\hat{\boldsymbol{\jmath}}+2\hat{\boldsymbol{k}}$ is a quaternion, $3\hat{\boldsymbol{\imath}}-1\hat{\boldsymbol{\jmath}}+2\hat{\boldsymbol{k}}$ is a pure quaternion and 7 is a scalar.

□

2 Quaternion Properties

2.1 Non-Commutativity

Theorem 2.1.1. Observe that (see examples above) in general $q_1q_2 \neq q_2q_1$ and $q_1q_2 \neq -q_2q_1$ so quaternion multiplication is neither commutative nor anti-commutative. This is not really a theorem, I just called it one so it would have an impact. There are some special cases as we will notice later.

□

2.2 Vector Connection

The use of the notation $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{k}}$ are not arbitrary, we can use vector operations like cross products and dot products on the vector part.

Example 2.1. If $\mathbf{q}_1 = 2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}}$ and $\mathbf{q}_2 = 5\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + 6\hat{\boldsymbol{k}}$ then $\mathbf{q}_1 \cdot \mathbf{q}_2 = 16$ and $\mathbf{q}_1 \times \mathbf{q}_2 = 22\hat{\boldsymbol{\imath}} - 17\hat{\boldsymbol{\jmath}} - 7\hat{\boldsymbol{k}}$. 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:

$$q = s + \mathbf{v} \text{ or } q = [s, \mathbf{v}] \text{ where } s \in \mathbb{R} \text{ and } \mathbf{v} = a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}.$$

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

Theorem 2.2.1. For quaternions $q_1 = s_1 + \mathbf{v}_1$ and $q_2 = s_2 + \mathbf{v}_2$ we have:

$$\begin{aligned} q_1q_2 &= (s_1s_2 - \mathbf{v}_1 \cdot \mathbf{v}_2) + (s_1\mathbf{v}_2 + s_2\mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2) \\ q_2q_1 &= (s_1s_2 - \mathbf{v}_2 \cdot \mathbf{v}_1) + (s_2\mathbf{v}_1 + s_1\mathbf{v}_2 + \mathbf{v}_2 \times \mathbf{v}_1) \\ &= (s_1s_2 - \mathbf{v}_1 \cdot \mathbf{v}_2) + (s_1\mathbf{v}_2 + s_2\mathbf{v}_1 - \mathbf{v}_1 \times \mathbf{v}_2) \end{aligned}$$

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

Proof. The first is just brute force calculation. The second follows from the first and from the commutativity of the dot product and the anti-commutativity of the cross product. The second line isn't really necessary, it's just there to make it obvious how the (anti-)commutativity fails. It's the cross product part which complicates the situation. □

Note 2.2.1. Note that for $q_1, q_2 \in \mathbb{H}$ we have:

$$\begin{aligned} &\mathrm{Re}(q_1q_2) = \mathrm{Re}(q_2q_1) \\ &\mathrm{Im}(q_1q_2) \neq \mathrm{Im}(q_2q_1) \qquad \text{(In General)} \end{aligned}$$

$\square$

Exercise 2.1. Given $q_1 = 2 + 1\hat{\boldsymbol{\imath}} - 2\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$ and $q_2 = 5 - 5\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 4\hat{\boldsymbol{k}}$, use the above theorem to calculate q_1q_2 and q_2q_1.

$\square$

Exercise 2.2. Do the brute brute force calculation for the above theorem.

$\square$

Exercise 2.3. Prove that for $q_1 = s_1 + \mathbf{v}_1$ and $q_2 = s_2 + \mathbf{v}_2 \in \mathbb{H}$ we have $q_1q_2 = q_2q_1$ iff $\mathbf{v}_1 \parallel \mathbf{v}_2$.

$\square$

As a special case of this we have:

Theorem 2.2.2. For pure quaternions $\mathbf{v}$ and $\mathbf{w}$ we have:

$$\begin{aligned}
\mathbf{vw} &= -\mathbf{v} \cdot \mathbf{w} + \mathbf{v} \times \mathbf{w} \\
\mathbf{wv} &= -\mathbf{w} \cdot \mathbf{v} + \mathbf{w} \times \mathbf{v} \\
&= -\mathbf{v} \cdot \mathbf{w} - \mathbf{v} \times \mathbf{w}
\end{aligned}$$

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. $\square$

Note 2.2.2. Note that for pure quaternions $\mathbf{v}, \mathbf{w}$ we have:

$$\begin{aligned}
\mathrm{Re}(\mathbf{wv}) &= \mathrm{Re}(\mathbf{vw}) \\
\mathrm{Im}(\mathbf{wv}) &= -\mathrm{Im}(\mathbf{vw})
\end{aligned}$$

$\square$

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

Theorem 2.2.3. For vectors $\mathbf{v}$ and $\mathbf{w}$ we have:

$$\begin{aligned}
\mathbf{v} \times \mathbf{w} &= +\frac{1}{2}\left(\mathbf{vw} - \mathbf{wv}\right) \\
\mathbf{v} \cdot \mathbf{w} &= -\frac{1}{2}\left(\mathbf{vw} + \mathbf{wv}\right)
\end{aligned}$$

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

Note 2.2.3. We are not suggesting that dot and cross products should be computed this way but it is good to keep in mind that they can be. Moreover it's important to note that the quaternion product can be considered the fundamental thing here from which the dot and cross products emerge.

□

2.3 Conjugation

Definition 2.3.1. The *conjugate* of a quaternion $q = s + a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}$ is denoted q^* and is defined by:

$$q^* = s - a\hat{\boldsymbol{\imath}} - b\hat{\boldsymbol{\jmath}} - c\hat{\boldsymbol{k}}$$

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:

$$(s + \mathbf{v})^* = s - \mathbf{v}$$

□

Theorem 2.3.1. If $q_1, q_2 \in \mathbb{H}$ then $(q_1 q_2)^* = q_2^* q_1^*$

Proof. Brute force. □

Exercise 2.4. Work out the brute force.

□

Exercise 2.5. Prove that for pure quaternions $\mathbf{v}$ and $\mathbf{w}$ we have $\mathbf{wv} = (\mathbf{vw})^*$.

□

Theorem 2.3.2. The conjugate of a quaternion can be expressed using addition and multiplication of quaternions. Specifically:

$$q^* = -\frac{1}{2}(q + \hat{\boldsymbol{\imath}} q \hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}} q \hat{\boldsymbol{\jmath}} + \hat{\boldsymbol{k}} q \hat{\boldsymbol{k}})$$

Proof. Brute force. Note that the same is not true in $\mathbb{C}$. In other words there is no way to express the conjugate of a complex number using addition and multiplication of complex numbers. This is not obvious. □

2.4 Norm

Definition 2.4.1. The *magnitude* (or *norm*) of a quaternion $q = s + a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}$ is:

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

Note that if $q = s + \mathbf{v}$ then $|q|^2 = s^2 + |\mathbf{v}|^2$.

□

Just like in $\mathbb{C}$ now we get:

Theorem 2.4.1. If $q = s + a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}} \in \mathbb{H}$ then

$$qq^* = q^*q = s^2 + a^2 + b^2 + c^2 = |q|^2$$

Proof. Brute force. □

Exercise 2.6. Work out the brute force.

□

Theorem 2.4.2. It follows that $|q|^2 = qq^* = q^*q$, that $|q| = \sqrt{qq^*} = \sqrt{q^*q}$.

Proof. Immediate from previous theorems. □

Exercise 2.7. Elaborate the above proof.

□

Theorem 2.4.3. The norm is multiplicative. That is, for $q_1, q_2 \in \mathbb{H}$ we have:

$$|q_1q_2| = |q_1||q_2|$$

Proof. We have:

$$\begin{aligned}
|q_1q_2| &= \sqrt{(q_1q_2)(q_1q_2)^*} \\
&= \sqrt{q_1q_2q_2^*q_1^*} \\
&= \sqrt{q_1|q_2|^2q_1^*} \\
&= |q_2|\sqrt{q_1q_1^*} \\
&= |q_2||q_1| \\
&= |q_1||q_2|
\end{aligned}$$

Notice that the same is true in $\mathbb{C}$. □

2.5 Unit Quaternions

Definition 2.5.1. A *unit quaternion* is a quaternion with norm 1.

□

Note 2.5.1. Note that for a unit quaternion we have $qq^* = q^*q = 1$.

□

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

Theorem 2.5.1. q is a unit pure quaternion iff $q^2 = -1$.

Proof. For a general quaternion $q = s + \mathbf{v} = s + a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}$ we have:

$$\begin{aligned} q^2 &= (ss - \mathbf{v}\cdot\mathbf{v}) + (s\mathbf{v} + s\mathbf{v} + \mathbf{v}\times\mathbf{v}) \\ &= (s^2 - |\mathbf{v}|^2) + (2s\mathbf{v} + 0) \\ &= (s^2 - a^2 - b^2 - c^2) + (2as\hat{\boldsymbol{\imath}} + 2bs\hat{\boldsymbol{\jmath}} + 2cs\hat{\boldsymbol{k}}) \end{aligned}$$

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:

$$\begin{aligned} s^2 - a^2 - b^2 - c^2 &= -1 \\ 2as &= 0 \\ 2bs &= 0 \\ 2cs &= 0 \end{aligned}$$

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 $|q| = 1$. $\square$

Example 2.2. If $q = \frac{2}{\sqrt{14}}\hat{\boldsymbol{\imath}} - \frac{1}{\sqrt{14}}\hat{\boldsymbol{\jmath}} + \frac{3}{\sqrt{14}}\hat{\boldsymbol{k}}$ then:

$$|q| = \sqrt{\left(\frac{2}{\sqrt{14}}\right)^2 + \left(\frac{1}{\sqrt{14}}\right)^2 + \left(\frac{3}{\sqrt{14}}\right)^2} = 1$$

Thus q is a unit pure quaternion and hence $q^2 = -1$.

$\square$

Consider what this states. If we think of unit pure quaternions as unit 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.

It turns out that this is where we start to see more similarities to $\mathbb{C}$. In $\mathbb{C}$ unit complex numbers correspond to rotations in 2D and there are a circle's worth of rotations (one per angle).

In $\mathbb{H}$ a rotation has an axis (of rotation) and each axis can be represented by a vector so it turns out that each unit pure quaternion corresponds to an axis of rotation. We'll need to go a little further in order to bring the angle of rotation into the picture, but that will happen in the next section.

2.6 Invertibility

Definition 2.6.1. A quaternion q is *invertible* if there is another quaternion, denoted q^{-1}, such that $qq^{-1} = q^{-1}q = 1$.

$\square$

Note 2.6.1. Note that for a unit quaternion we have $q^{-1} = q^*$.

$\square$

Theorem 2.6.1. All nonzero quaternions are invertible and in fact:

$$q^{-1} = \frac{q^*}{|q|^2} = \frac{s - a\hat{\boldsymbol{\imath}} - b\hat{\boldsymbol{\jmath}} - c\hat{\boldsymbol{k}}}{|q|^2}$$

Proof. Observe that:

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

and similarly for the other product. $\square$

Exercise 2.8. Calculate the inverse of $2 + 4\hat{\boldsymbol{\imath}} - 2\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$.

$\square$

Corollary 2.6.1. If $q = s + a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}$ is a unit quaternion then

$$q^{-1} = s - a\hat{\boldsymbol{\imath}} - b\hat{\boldsymbol{\jmath}} - c\hat{\boldsymbol{k}} = q^*$$

$\square$

Theorem 2.6.2. For nonzero quaternions q_1 and q_2 we have:

$$(q_1q_2)^{-1} = q_2^{-1}q_1^{-1}$$

Proof. This follows from the fact that:

$$q_1q_2q_2^{-1}q_1^{-1} = 1 \text{ and } q_2^{-1}q_1^{-1}q_1q_2 = 1$$

$\square$

2.7 Divisibility

We won't use divisibilty so we won't cover it here except to say that we need to be very careful because of the non-commutative nature of quaternion multiplication.

In other words when we write an expression like this:

$$\frac{q_1}{q_2}$$

we need to understand that we mean:

$$\frac{q_1}{q_2} = q_1 q_2^{-1}$$

This is important because we cannot arbitrarily cancel. For example observe that:

$$\frac{q_1 q_2}{q_1} = q_1 q_2 q_1^{-1}$$

which is not necessarily the same as q_2. In other words the q_1 don't cancel.

2.8 Conjugation of a Pure Quaternion is Pure

Lastly a fact that will be relevant when we discuss rotation using quaternions:

Theorem 2.8.1. If $\mathbf{v}$ is pure and $q \in \mathbb{H}$ then $q\mathbf{v}q^{-1}$ is pure.

Proof. From an earlier theorem we see that for $q_1, q_2 \in \mathbb{H}$ we have $\mathrm{Re}(q_1 q_2) = \mathrm{Re}(q_2 q_1)$. In this case:

$$\mathrm{Re}(q\mathbf{v}q^{-1}) = \mathrm{Re}(qq^{-1}\mathbf{v}) = \mathrm{Re}(\mathbf{v}) = 0$$

□

2.9 Summary

Here is a brief summary of properties for reference:

(a) $q_1 q_2 = (s_1 s_2 - \mathbf{v}_1 \cdot \mathbf{v}_2) + (s_1 \mathbf{v}_2 + s_2 \mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2)$

(b) $\mathbf{vw} = -\mathbf{v} \cdot \mathbf{w} + \mathbf{v} \times \mathbf{w}$

(c) $\mathbf{wv} = -\mathbf{v} \cdot \mathbf{w} + \mathbf{w} \times \mathbf{v} = -\mathbf{v} \cdot \mathbf{w} - \mathbf{v} \times \mathbf{w}$

(d) $\mathbf{v} \times \mathbf{w} = +\frac{1}{2}(\mathbf{vw} - \mathbf{wv})$

(e) $\mathbf{v} \cdot \mathbf{w} = -\frac{1}{2}(\mathbf{vw} + \mathbf{wv})$

(f) $qq^* = |q|^2$

(g) $|q| = \sqrt{qq^*}$

(h) $(q_1 q_2)^* = q_2^* q_1^*$

(i) q is a unit pure quaternion iff $q^2 = -1$

(j) $|q_1 q_2| = |q_1||q_2|$

(k) $q^* = -\frac{1}{2}(q + \hat{\boldsymbol{\imath}}q\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}q\hat{\boldsymbol{\jmath}} + \hat{\boldsymbol{k}}q\hat{\boldsymbol{k}})$

(l) $q^{-1} = \frac{q^*}{|q|^2}$

(m) If q is a unit quaternion then $q^{-1} = q^*$

3 Visualization of Quaternions

The only quaternions we represent graphically are pure quaternions, meaning those of the form $a\hat{\boldsymbol{\imath}} + b\hat{\boldsymbol{\jmath}} + c\hat{\boldsymbol{k}}$. We represent these either as vectors or as points, depending on how we're using them.

It's worth taking a moment to appreciate that when dealing with vectors $\mathbf{v}$ and $\mathbf{w}$ that the quaternion product:

$$\mathbf{vw} = \mathbf{v} \times \mathbf{w} - \mathbf{v} \cdot \mathbf{w}$$

captures both the dot product (in the scalar part of the result) and the cross product (in the vector part of the result).

Imagine two vectors $\mathbf{v}$ and $\mathbf{w}$. When $\mathbf{v} \perp \mathbf{w}$ the dot product is zero and the result is just $\mathbf{v} \times \mathbf{w}$ and is perpendicular to both. If we turn $\mathbf{v}$ and $\mathbf{w}$ a bit (without changing their lengths) so $\mathbf{v} \not\perp \mathbf{w}$, the resulting cross product shrinks (since $|\mathbf{v} \times \mathbf{w}| = |\mathbf{v}||\mathbf{w}| \sin\theta$) and we interpret that what we've lost from the cross product we've gained in the dot product, but as a scalar. As $\mathbf{v}$ and $\mathbf{w}$ get less perpendicular and more parallel the cross product shrinks towards zero and we gain more dot product until they're parallel, at which point the cross product part vanishes and the dot product part is everything.

4 Translations

If a point $x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}$ is to be translated in 3D space we simply add or subtract another pure quaternion.

Example 4.1. To shift $2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}}$ by 5 in the x-direction, 2 in the y-direction, and 7 in the z-direction we simply do:

$$2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}} \mapsto 2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}} + 5\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 7\hat{\boldsymbol{k}} = 7\hat{\boldsymbol{\imath}} + 5\hat{\boldsymbol{\jmath}} + 6\hat{\boldsymbol{k}}$$

$\square$

5 Rotations

5.1 About 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 5.1.1. (Rodrigues Rotation Formula)
Suppose $\hat{\mathbf{u}}$ is a unit vector and $\mathbf{v}$ is some vector. Then the result of rotating $\mathbf{v}$ around $\hat{\mathbf{u}}$ by an angle θ counterclockwise with regards to the right-hand rule equals:

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

Proof. We begin by breaking $\mathbf{v}$ into components, one perpendicular to $\hat{\mathbf{u}}$ and one parallel to $\hat{\mathbf{u}}$:

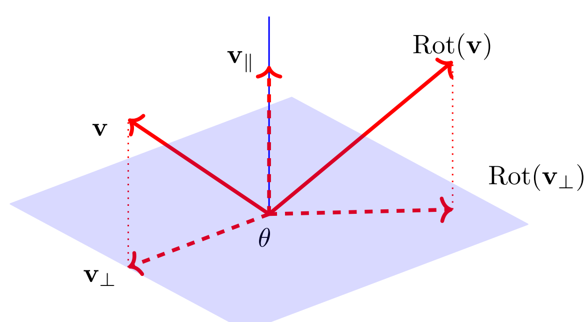
$$\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel$$

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

That is:

$$\mathrm{Rot}(\mathbf{v}) = \mathbf{v}_\parallel + \mathrm{Rot}\left(\mathbf{v}_\perp\right)$$

The reason for this is illustrated by this picture:



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

$$\mathrm{Rot}(\mathbf{v}_\perp) = (\cos\theta)\mathbf{v}_\perp + (\sin\theta)(\hat{\mathbf{u}} \times \mathbf{v}_\perp)$$

If we use this along with the facts that:

$$\begin{aligned} \mathbf{v}_\parallel &= (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} \\ \mathbf{v}_\perp &= \mathbf{v} - (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} \end{aligned}$$

So now we calculate:

$$\begin{aligned} \mathrm{Rot}(\mathbf{v}) &= \mathbf{v}_\parallel + \mathrm{Rot}\,(\mathbf{v}_\perp) \\ &= \mathbf{v}_\parallel + (\cos\theta)\mathbf{v}_\perp + (\sin\theta)(\hat{\mathbf{u}} \times \mathbf{v}_\perp) \\ &= (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} + (\cos\theta)\left(\mathbf{v} - (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}}\right) + (\sin\theta)\left(\hat{\mathbf{u}} \times (\mathbf{v} - (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}})\right) \\ &= (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} + (\cos\theta)\mathbf{v} - (\cos\theta)(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} + (\sin\theta)(\hat{\mathbf{u}} \times \mathbf{v} - \underbrace{\hat{\mathbf{u}} \times ((\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}})}_{\hat{\mathbf{u}} \times \hat{\mathbf{u}} = 0}) \\ &= (1 - \cos\theta)(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} + (\cos\theta)\mathbf{v} + (\sin\theta)(\hat{\mathbf{u}} \times \mathbf{v}) \end{aligned}$$

□

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 $\hat{\mathbf{u}}$, $\mathbf{v}$ and $\hat{\mathbf{u}} \times \mathbf{v}$.

Exercise 5.1. Use RRF to calculate the result of rotating $\mathbf{v} = 2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}}$ about $\mathbf{u} = 3\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} - 5\hat{\boldsymbol{k}}$ by $7\pi/6$ radians. Note that $\mathbf{u}$ has not been normalized so do this first.

□

Before our theorem, a few notes:

(a) Note the trig identity $\cos(2x) = \cos^2 x - \sin^2 x$.

(b) Note the trig identity $\sin(2x) = 2\sin x \cos x$.

(c) Note the trig identity $2\sin^2 x = 1 - \cos(2x)$.

(d) We have $\hat{\mathbf{u}}\mathbf{v} - \mathbf{v}\hat{\mathbf{u}} = 2(\hat{\mathbf{u}} \times \mathbf{v})$. This follows directly from $\hat{\mathbf{u}} \times \mathbf{v} = \frac{1}{2}(\hat{\mathbf{u}}\mathbf{v} - \mathbf{v}\hat{\mathbf{u}})$.

(e) We have $\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}} = -2(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} + \mathbf{v}$. This is not so obvious. We know that $\hat{\mathbf{u}} \cdot \mathbf{v} = -\frac{1}{2}(\hat{\mathbf{u}}\mathbf{v} + \mathbf{v}\hat{\mathbf{u}})$ and so $\hat{\mathbf{u}}\mathbf{v} = -2\hat{\mathbf{u}} \cdot \mathbf{v} - \mathbf{v}\hat{\mathbf{u}}$. Then:

$$\begin{aligned}
\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}} &= \left[-2\hat{\mathbf{u}} \cdot \mathbf{v} - \mathbf{v}\hat{\mathbf{u}}\right]\hat{\mathbf{u}} \\
&= -2(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} - \mathbf{v}\hat{\mathbf{u}}\hat{\mathbf{u}} \\
&= -2(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} - \mathbf{v}(\hat{\mathbf{u}} \times \hat{\mathbf{u}} - \hat{\mathbf{u}} \cdot \hat{\mathbf{u}}) \\
&= -2(\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} - \mathbf{v}(-1)
\end{aligned}$$

And now our theorem:

Theorem 5.1.2. Suppose $\hat{\mathbf{u}}$ is a unit vector and $\mathbf{v}$ is some vector. Then the result of rotating $\mathbf{v}$ around $\hat{\mathbf{u}}$ by an angle θ counterclockwise with regards to the right-hand rule can be obtained by letting:

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

and then doing:

$$\mathbf{v}_{\text{Rot}} = p\mathbf{v}p^{-1} = p\mathbf{v}p^*$$

Before embarking on the proof, note that p is a unit quaternion because $|p|^2 = \cos^2(\theta/2) + |\sin(\theta/2)\hat{\mathbf{u}}|^2 = 1 + 1 = 1$. In addition every unit unit quaternion p can be decomposed this way because for any unit quaternion $p = s + \mathbf{w}$ we can write:

$$p = s + |\mathbf{w}|\left(\frac{1}{|\mathbf{w}|}\mathbf{w}\right)$$

and then simply assign $\hat{\mathbf{u}} = \frac{1}{|\mathbf{w}|}\mathbf{w}$ and choose θ so that $\cos(\theta/2) = s$ and $\sin(\theta/2) = |\mathbf{w}|$ which is possible since $s^2 + |\mathbf{w}|^2 = |p|^2 = 1$.

Thus unit quaternions correspond to rotations where the vector part corresponds to the axis of rotation and the angle is built into the scalar part and the magnitude of the vector part. This is very important because when discussing rotations we can say that an arbitrary rotation can be performed via $\mathbf{v} \mapsto p\mathbf{v}p^*$ where p is a unit quaternion. This will arise frequently.

Proof. This is just calculation. First note that since $|p| = 1$ that $p^{-1} = p^*$.

Then consider:

$$\begin{aligned}
p\mathbf{v}p^* &= \left(\cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right)\hat{\mathbf{u}}\right)\mathbf{v}\left(\cos\left(\frac{\theta}{2}\right) - \sin\left(\frac{\theta}{2}\right)\hat{\mathbf{u}}\right) \\
&= \cos^2\left(\frac{\theta}{2}\right)\mathbf{v} + \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)\underbrace{(\hat{\mathbf{u}}\mathbf{v} - \mathbf{v}\hat{\mathbf{u}})}_{(d)} - \sin^2\left(\frac{\theta}{2}\right)\underbrace{\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}}}_{(e)} \\
&= \cos^2\left(\frac{\theta}{2}\right)\mathbf{v} + \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)2\left(\hat{\mathbf{u}}\times\mathbf{v}\right) - \sin^2\left(\frac{\theta}{2}\right)\left(-2(\hat{\mathbf{u}}\cdot\mathbf{v})\hat{\mathbf{u}} + \mathbf{v}\right) \\
&= \underbrace{\left(\cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right)\right)}_{(a)}\mathbf{v} + \underbrace{2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)}_{(b)}\left(\hat{\mathbf{u}}\times\mathbf{v}\right) + \underbrace{2\sin^2\left(\frac{\theta}{2}\right)}_{(c)}\left(\hat{\mathbf{u}}\cdot\mathbf{v}\right)\hat{\mathbf{u}} \\
&= (\cos\theta)\,\mathbf{v} + (\sin\theta)\left(\hat{\mathbf{u}}\times\mathbf{v}\right) + (1-\cos\theta)\left(\hat{\mathbf{u}}\cdot\mathbf{v}\right)\hat{\mathbf{u}} \\
&= \mathbf{v}_{\mathrm{Rot}}
\end{aligned}$$

□

This should and should not surprise you. In $\mathbb{C}$ it was multiplication by $\cos\theta + \hat{\imath}\sin\theta$ which did rotation and so this 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 p is not a pure quaternion the result of the calculation $p\mathbf{v}p^*$ where $\mathbf{v}$ is a pure quaternion results in a pure quaternion.

Example 5.1. To rotate $\mathbf{v} = 2\hat{\imath} + 1\hat{\jmath}$ by $\theta = \pi/3$ radians about $\hat{\mathbf{u}} = \frac{1}{\sqrt{2}}\hat{\jmath} + \frac{1}{\sqrt{2}}\hat{k}$ we set:

$$p = \cos(\pi/6) + \sin(\pi/6)\left(\frac{1}{\sqrt{2}}\hat{\jmath} + \frac{1}{\sqrt{2}}\hat{k}\right) = \frac{\sqrt{3}}{2} + \frac{1}{2\sqrt{2}}\hat{\jmath} + \frac{1}{2\sqrt{2}}\hat{k}$$

and then the result is:

$$\begin{aligned}
p\mathbf{v}p^* &= \left(\frac{\sqrt{3}}{2} + \frac{1}{2\sqrt{2}}\hat{\jmath} + \frac{1}{2\sqrt{2}}\hat{k}\right)(2\hat{\imath} + 1\hat{\jmath})\left(\frac{\sqrt{3}}{2} - \frac{1}{2\sqrt{2}}\hat{\jmath} - \frac{1}{2\sqrt{2}}\hat{k}\right) \\
&\approx \text{...Matlab...} \\
&\approx 0 + 0.3876\hat{\imath} + 1.9747\hat{\jmath} - 0.9747\hat{k}
\end{aligned}$$

□

Exercise 5.2. Use the above formula to rotate $\mathbf{v} = \hat{\imath}$ about $\hat{\mathbf{u}} = \hat{k}$ by $\pi/2$ radians. Does this correspond to your expectations? Hint: What should it do?

$\square$

Exercise 5.3. Use the above formula to rotate $\mathbf{v} = 1\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$ about $\mathbf{u} = 2\hat{\boldsymbol{\imath}} + 0\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$ by $2\pi/3$ radians.

$\square$

After discussions with a student here is a slightly shorter proof of the above theorem. For now I'll keep the previous one as the default one because I think there's value in seeing Rodrigues Rotation Formula.

Here is the theorem again:

Theorem 5.1.3. Suppose $\hat{\mathbf{u}}$ is a unit vector and $\mathbf{v}$ is some vector. Then the result of rotating $\mathbf{v}$ around $\hat{\mathbf{u}}$ by an angle θ counterclockwise with regards to the right-hand rule can be obtained by letting:

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

and then doing:

$$\mathbf{v}_{\text{Rot}} = p\mathbf{v}p^{-1} = p\mathbf{v}p^{*}$$

Proof. As with the start of RRF we break $\mathbf{v}$ into components parallel and perpendicular to $\mathbf{u}$, so $\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}$ and we wish to rotate the perpendicular part while keeping the parallel part fixed. See the RRF picture for clarification if needed.

Observe that the mapping is then:

$$\mathbf{v} \mapsto p(\mathbf{v}_{\parallel} + \mathbf{v}_{\perp})p^{-1} = p\mathbf{v}_{\parallel}p^{-1} + p\mathbf{v}_{\perp}p^{-1}$$

Observe then the following two things:

(a) For parallel pure quaternions $\mathbf{a}$ and $\mathbf{b}$ we have:

$$\mathbf{ab} = \mathbf{a} \times \mathbf{b} - \mathbf{a} \cdot \mathbf{b} = 0 - \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \times \mathbf{a} - \mathbf{b} \cdot \mathbf{a} = \mathbf{ba}$$

It follows that $\mathbf{v}_{\parallel}\hat{\mathbf{u}} = \hat{\mathbf{u}}\mathbf{v}_{\parallel}$ and so:

$$
\begin{aligned}
p\mathbf{v}_{\|}p^{-1} &= p\mathbf{v}_{\|}p^{*} \\
&= p\mathbf{v}_{\|}\left(\cos(\theta/2) - \hat{\mathbf{u}}\sin(\theta/2)\right) \\
&= p\left(\cos(\theta/2)\mathbf{v}_{\|} - \sin(\theta/2)\mathbf{v}_{\|}\hat{\mathbf{u}}\right) \\
&= p\left(\cos(\theta/2)\mathbf{v}_{\|} - \sin(\theta/2)\hat{\mathbf{u}}\mathbf{v}_{\|}\right) \\
&= p\left(\cos(\theta/2) - \sin(\theta/2)\hat{\mathbf{u}}\right)\mathbf{v}_{\|} \\
&= pp^{*}\mathbf{v}_{\|} \\
&= pp^{-1}\mathbf{v}_{\|} \\
&= \mathbf{v}_{\|}
\end{aligned}
$$

(b) For perpendicular pure quaternions $\mathbf{a}$ and $\mathbf{b}$ we have:

$$
\mathbf{ab} = \mathbf{a}\times\mathbf{b} - \mathbf{a}\cdot\mathbf{b} = \mathbf{a}\times\mathbf{b} - 0 = -(\mathbf{b}\times\mathbf{a}) + 0 = -(\mathbf{b}\times\mathbf{a}) + \mathbf{b}\cdot\mathbf{a} = -\mathbf{ba}
$$

It follows that $\mathbf{v}_{\perp}\hat{\mathbf{u}} = -(\hat{\mathbf{u}}\times\mathbf{v}_{\perp}) = -\hat{\mathbf{u}}\mathbf{v}_{\perp}$. and so:

$$
\begin{aligned}
p\mathbf{v}_{\perp}p^{-1} &= p\mathbf{v}_{\perp}p^{*} \\
&= p\mathbf{v}_{\perp}\left(\cos(\theta/2) - \hat{\mathbf{u}}\sin(\theta/2)\right) \\
&= p\left(\cos(\theta/2)\mathbf{v}_{\perp} - \sin(\theta/2)\mathbf{v}_{\perp}\hat{\mathbf{u}}\right) \\
&= p\left(\cos(\theta/2)\mathbf{v}_{\perp} + \sin(\theta/2)\hat{\mathbf{u}}\mathbf{v}_{\perp}\right) \\
&= p\left(\cos(\theta/2) + \sin(\theta/2)\hat{\mathbf{u}}\right)\mathbf{v}_{\perp} \\
&= \left(\cos(\theta/2) + \sin(\theta/2)\hat{\mathbf{u}}\right)\left(\cos(\theta/2) + \sin(\theta/2)\hat{\mathbf{u}}\right)\mathbf{v}_{\perp} \\
&= \left[(\cos^2(\theta/2) - \sin^2(\theta/2))\hat{\mathbf{u}}\hat{\mathbf{u}} + 2\sin(\theta/2)\cos(\theta/2)\hat{\mathbf{u}}\right]\mathbf{v}_{\perp} \\
&= \left[\cos(\theta) + \sin(\theta)\hat{\mathbf{u}}\right]\mathbf{v}_{\perp} \\
&= \cos(\theta)\mathbf{v}_{\perp} + \sin(\theta)\hat{\mathbf{u}}\mathbf{v}_{\perp} \\
&= \cos(\theta)\mathbf{v}_{\perp} + \sin(\theta)\left(\hat{\mathbf{u}}\times\mathbf{v}_{\perp}\right)
\end{aligned}
$$

Now the mapping is:

$$
\mathbf{v} \mapsto \mathbf{v}_{\|} + \cos(\theta)\mathbf{v}_{\perp} + \sin(\theta)\left(\hat{\mathbf{u}}\times\mathbf{v}_{\perp}\right)
$$

However this is exactly as desired since the parallel portion is held fixed while the perpendicular portion is rotated according to the rule from Chapter 2.

□

5.2 Exponential Form of Rotation

Recall that for complex numbers it was convenient to write rotations using exponential form:

$$z \mapsto z e^{\theta \hat{\boldsymbol{\imath}}}$$

This is because the Taylor expansion of e^z gives us:

$$e^{\theta \hat{\boldsymbol{\imath}}} = \cos\theta + \hat{\boldsymbol{\imath}} \sin\theta$$

Similarly for $\mathbb{H}$:

Definition 5.2.1. For any $q \in \mathbb{H}$ we may define the exponential function via the Taylor expansion:

$$e^q = \sum_{n=0}^{\infty} \frac{q^n}{n!}$$

which (we will not prove) converges for all q.

$\square$

It turns out that as a consequence of this we get:

$$e^{\theta \hat{\mathbf{u}}} = \cos\theta + \hat{\mathbf{u}} \sin\theta$$

And thus our rotation can be rewritten as:

$$\mathbf{v} \mapsto e^{\frac{\theta}{2}\hat{\mathbf{u}}} \mathbf{v} e^{-\frac{\theta}{2}\hat{\mathbf{u}}}$$

5.3 About 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 $\mathbf{v}_0$ with unit direction vector $\hat{\mathbf{u}}$ then rotation about this line can be done by:

$$\mathbf{v} \mapsto p(\mathbf{v} - \mathbf{v}_0)p^* + \mathbf{v}_0$$

where p is as before for rotations.

Exercise 5.4. Find the result of rotating $\mathbf{v} = 2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}$ by $\pi/4$ about the line passing through $(0, 0, 1)$ with direction $\mathbf{u} = 3\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 0\hat{\boldsymbol{k}}$.
Note that you need to make $\mathbf{u}$ into $\hat{\mathbf{u}}$.

6 Reflections

6.1 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:

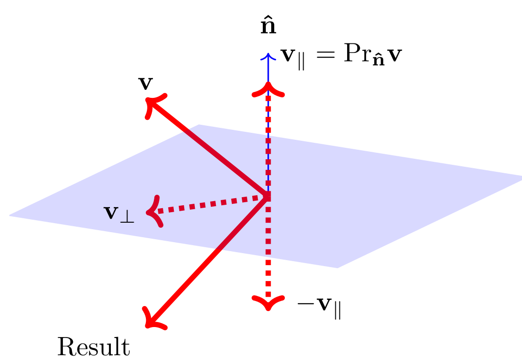
Theorem 6.1.1. Given a plane $\mathcal{P}$ through the origin represented by the unit pure quaternion (unit normal vector) $\hat{\mathbf{n}}$ the reflection of the vector $\mathbf{v}$ is given by:

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

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

$$\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel$$

Reflecting in $\mathcal{P}$ involves negating $\mathbf{v}_\parallel$ and leaving $\mathbf{v}_\perp$ alone.



We'll look at these two parts independently.

Observe that:

$$\begin{aligned}
\hat{\mathbf{n}}\mathbf{v}_\perp\hat{\mathbf{n}} &= \hat{\mathbf{n}}(\mathbf{v}_\perp \times \hat{\mathbf{n}} - \mathbf{v}_\perp \cdot \hat{\mathbf{n}}) \\
&= \hat{\mathbf{n}}(\mathbf{v}_\perp \times \hat{\mathbf{n}} - 0) \\
&= \hat{\mathbf{n}} \times (\mathbf{v}_\perp \times \hat{\mathbf{n}}) - \hat{\mathbf{n}} \cdot (\mathbf{v}_\perp \times \hat{\mathbf{n}}) \\
&= \hat{\mathbf{n}} \times (\mathbf{v}_\perp \times \hat{\mathbf{n}}) - 0) \\
&= (\hat{\mathbf{n}} \cdot \hat{\mathbf{n}})\mathbf{v}_\perp - (\hat{\mathbf{n}} \cdot \mathbf{v}_\perp)\hat{\mathbf{n}} \\
&= (1)\mathbf{v}_\perp - (0)\hat{\mathbf{n}} \\
&= \mathbf{v}_\perp
\end{aligned}$$

And observe that:

$$\begin{aligned}
\hat{\mathbf{n}}\mathbf{v}_\parallel\hat{\mathbf{n}} &= \hat{\mathbf{n}}(\mathbf{v}_\parallel \times \hat{\mathbf{n}} - \mathbf{v}_\parallel \cdot \hat{\mathbf{n}}) \\
&= \hat{\mathbf{n}}(0 - \mathbf{v}_\parallel \cdot \hat{\mathbf{n}}) \\
&= -(\mathbf{v}_\parallel \cdot \hat{\mathbf{n}})\hat{\mathbf{n}} \\
&= -\mathrm{Pr}_{\hat{\mathbf{n}}}\mathbf{v}_\parallel \\
&= -\mathbf{v}_\parallel
\end{aligned}$$

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

$$\begin{aligned}
\hat{\mathbf{n}}\mathbf{v}\hat{\mathbf{n}} &= \hat{\mathbf{n}}(\mathbf{v}_\perp + \mathbf{v}_\parallel)\hat{\mathbf{n}} \\
&= \hat{\mathbf{n}}\mathbf{v}_\perp\hat{\mathbf{n}} + \hat{\mathbf{n}}\mathbf{v}_\parallel\hat{\mathbf{n}} \\
&= \mathbf{v}_\perp - \mathbf{v}_\parallel
\end{aligned}$$

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:

$$\mathbf{v} \mapsto \left(\frac{1}{|\mathbf{n}|^2}\right)\mathbf{n}\mathbf{v}\mathbf{n} = \frac{\mathbf{n}\mathbf{v}\mathbf{n}}{\mathbf{n} \cdot \mathbf{n}}$$

Example 6.1. To reflect $\mathbf{v} = 3\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}$ in the plane through the origin with normal vector $\mathbf{n} = 2\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}}$ we calculate:

$$\begin{aligned}
3\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}} &\mapsto \left(\frac{1}{3}\right)(2\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}})(3\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}})(2\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}}) \\
&\mapsto \ldots \\
&\mapsto \left(\frac{1}{3}\right)(-1\hat{\boldsymbol{\imath}} - 19\hat{\boldsymbol{\jmath}} + 23\hat{\boldsymbol{k}})
\end{aligned}$$

$\square$

Exercise 6.1. Calculate the result of reflecting $\mathbf{v} = 3\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}$ in the plane through the origin with normal vector $\hat{\mathbf{n}} = \hat{\boldsymbol{\imath}}$. Is this what you expect?

$\square$

Exercise 6.2. Calculate the result of reflecting $\mathbf{v} = 15\hat{\boldsymbol{\imath}} + 10\hat{\boldsymbol{\jmath}} - 20\hat{\boldsymbol{k}}$ in the plane through the origin with normal vector $\mathbf{n} = 1\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$.

$\square$

6.2 In Lines Through the Origin

It's also possible in three dimensions to reflect through a line given as an axis $\hat{\mathbf{u}}$. This reflection is exactly the same as a rotation about $\hat{\mathbf{u}}$ by π radians which we can then see easily is:

$$\begin{aligned}
\mathbf{v} &\mapsto \left(\cos\frac{\pi}{2} + \sin\frac{\pi}{2}\hat{\mathbf{u}}\right)\mathbf{v}\left(\cos\frac{\pi}{2} - \sin\frac{\pi}{2}\hat{\mathbf{u}}\right) \\
&\mapsto \hat{\mathbf{u}}\mathbf{v}(-\hat{\mathbf{u}}) \\
&\mapsto -\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}}
\end{aligned}$$

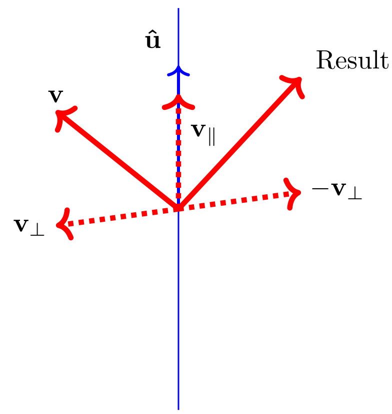
It's worth noting that we can derive the formula separately from that approach.

Theorem 6.2.1. Given a line $\mathcal{L}$ through the origin represented by the unit pure quaternion $\hat{\mathbf{u}}$ the reflection of the vector $\mathbf{v}$ is given by:

$$\mathbf{v} \mapsto -\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}}$$

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 $\mathbf{v}$ into the sum of two vectors, one perpendicular to $\hat{\mathbf{u}}$ and one parallel to (a multiple of) $\hat{\mathbf{u}}$. Here's where the proof differs. In this case reflecting in $\mathcal{L}$ involves leaving the parallel part intact and negating the perpendicular part, rather than the other way around.



For $\mathbf{v} = \mathbf{v}_{\perp} + \mathbf{v}_{\|}$ we then have:

$$\begin{aligned}
-\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}} &= -\hat{\mathbf{u}}\left(\mathbf{v}_{\perp} + \mathbf{v}_{\|}\right)\hat{\mathbf{u}} \\
&= -\left(\hat{\mathbf{u}}\mathbf{v}_{\perp}\hat{\mathbf{u}} + \hat{\mathbf{u}}\mathbf{v}_{\|}\hat{\mathbf{u}}\right) \\
&= -\left(\mathbf{v}_{\perp} - \mathbf{v}_{\|}\right) \\
&= -\mathbf{v}_{\perp} + \mathbf{v}_{\|}
\end{aligned}$$

The third equality holds using the same equations we worked out in the proof dealing with reflections in a plane.

□

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

$$\mathbf{v} \mapsto -\left(\frac{1}{|\mathbf{u}|^2}\right)\mathbf{u}\mathbf{v}\mathbf{u} = -\frac{\mathbf{u}\mathbf{v}\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$$

Exercise 6.3. Find the result when the vector $\mathbf{v} = 10\hat{\boldsymbol{\imath}} + 12\hat{\boldsymbol{\jmath}} + 8\hat{\boldsymbol{k}}$ is reflected in the axis $\hat{\mathbf{u}} = \hat{\boldsymbol{k}}$. Is the result what you expect?

Exercise 6.4. Find the result when the vector $\mathbf{v} = 10\hat{\boldsymbol{\imath}} + 12\hat{\boldsymbol{\jmath}} + 8\hat{\boldsymbol{k}}$ is reflected in the axis $\mathbf{u} = 5\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$.

6.3 In (Points Through?) The Origin

It's trivial but worth noting that reflection in the origin is simply negation of the vector:

$$\mathbf{v} \mapsto -\mathbf{v}$$

Although a bit strange, observe that we can also write this as:

$$\mathbf{v} \mapsto -1\mathbf{v}1$$

6.4 Reflections in Other Planes, Lines, and Points

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 6.5. Find the result when $\mathbf{v} = 3\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} + 10\hat{\boldsymbol{k}}$ is reflected in the plane $2x + 4y + 4z = 12$ with normal vector arising from the coefficients.

□

Exercise 6.6. Find the result when $\mathbf{v} = 3\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} + 10\hat{\boldsymbol{k}}$ is reflected in the line through $(1, 1, 2)$ with axis $\mathbf{u} = 1\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$.

□

And likewise in a point $\mathbf{v}_0$:

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

6.5 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 $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ respectively and meet at an angle of θ. Suppose we wish to reflect in $\mathcal{P}_1$ first and $\mathcal{P}_2$ second.

The axis formed by the intersection of the two has vector $\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2$ but this is probably not a unit vector. Notice that this vector follows the right-hand rule curling the fingers around the small angle between $\mathcal{P}_1$ and $\mathcal{P}_2$.

The unit vector $\hat{\mathbf{u}}$ would be:

$$\hat{\mathbf{u}} = \frac{\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2}{|\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2|}$$

and would satisfy:

$$\begin{aligned}
\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2 &= \frac{\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2}{|\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2|}|\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2| \\
&= \hat{\mathbf{u}}|\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2| \\
&= \hat{\mathbf{u}}|\hat{\mathbf{n}}_1||\hat{\mathbf{n}}_2| \sin\theta \\
&= \hat{\mathbf{u}}(1)(1)\sin\theta
\end{aligned}$$

Keeping in mind also that:

$$\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 = |\hat{\mathbf{n}}_1||\hat{\mathbf{n}}_2| \cos\theta = \cos\theta$$

The double-reflection will then be:

$$\begin{aligned}
\mathbf{v} &\mapsto \hat{\mathbf{n}}_2(\hat{\mathbf{n}}_1\mathbf{v}\hat{\mathbf{n}}_1)\hat{\mathbf{n}}_2 \\
&\mapsto (\hat{\mathbf{n}}_2\hat{\mathbf{n}}_1)\mathbf{v}(\hat{\mathbf{n}}_1\hat{\mathbf{n}}_2) \\
&\mapsto (\hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1 - \hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1)\mathbf{v}(\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2 - \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2) \\
&\mapsto (-\hat{\mathbf{u}}\sin\theta - \cos\theta)\mathbf{v}(\hat{\mathbf{u}}\sin\theta - \cos\theta) \\
&\mapsto (\cos\theta + \hat{\mathbf{u}}\sin\theta)\mathbf{v}(\cos\theta - \hat{\mathbf{u}}\sin\theta)
\end{aligned}$$

This is exactly equal to a rotation of 2θ radians about the axis $\hat{\mathbf{u}}$.

Note 6.5.1. The direction of rotation here is by the right-hand rule applied to the vector $\hat{\mathbf{u}}$. This vector arose from $\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2$ and so the direction of $\hat{\mathbf{u}}$ is such that the right-hand rule rotates $\hat{\mathbf{n}}_1$ toward $\hat{\mathbf{n}}_2$.

□

Note 6.5.2. The angle betwen $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ is not necessarily the angle between the planes. These could differ by $\pi/2$ depending on the relationship of the normal vectors to the planes.

□

7 Transformation Summary So Far

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

Rotation about a line $\hat{\mathbf{u}}$	$\mathbf{v} \mapsto p\mathbf{v}p^*$ where $p = \cos(\theta/2) + \sin(\theta/2)\hat{\mathbf{u}}$
Reflection in a plane $\hat{\mathbf{n}}$ through $\mathbf{0}$	$\mathbf{v} \mapsto \hat{\mathbf{n}}\mathbf{v}\hat{\mathbf{n}}$
Reflection in a line $\hat{\mathbf{u}}$ through $\mathbf{0}$	$\mathbf{v} \mapsto -\hat{\mathbf{u}}\mathbf{v}\hat{\mathbf{u}}$
Reflection in the origin $\mathbf{0}$	$\mathbf{v} \mapsto -1\mathbf{v}1$

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

To close this section just note that reflection in a plane is the only reflection that doesn't have a negative. Weird.

8 Transformations of Lines and Planes

8.1 Representations of Lines and Planes

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 vector $(\mathbf{v}_0, \mathbf{d}_0)$. Then the line consists of all points of the form:

$$\mathbf{v}(t) = \mathbf{v}_0 + t\mathbf{d}_0$$

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 $(\mathbf{v}_0, \mathbf{n}_0)$. Then the plane consists of all points $\mathbf{v}(t)$ satisfying:

$$\mathbf{n}_0 \cdot (\mathbf{v} - \mathbf{v}_0) = 0$$

8.2 Transformations of Lines

Theorem 8.2.1. To translate the line parametrized by $(\mathbf{v}_0, \mathbf{d}_0)$ by $\mathbf{q}$ we translate the anchor point $\mathbf{v}_0$. The direction vector doesn't change so $\mathbf{d}_0$ is not touched. That is:

$$(\mathbf{v}_0, \mathbf{d}_0) \mapsto (\mathbf{v}_0 + \mathbf{q}, \mathbf{d}_0)$$

Proof. In the statement. □

Theorem 8.2.2. To rotate the line parametrized by $(\mathbf{v}_0, \mathbf{d}_0)$ by θ radians about the axis $\hat{\mathbf{u}}$ through the origin we rotate both the anchor point and the direction vector. That is, assign $p = \cos(\theta/2) + \sin(\theta/2)\hat{\mathbf{u}}$ and map:

$$(\mathbf{v}_0, \mathbf{d}_0) \mapsto (p\mathbf{v}_0p^*, p\mathbf{d}_0p^*)$$

Proof. Observe that the original line passes through the points $\mathbf{v}_0$ and $\mathbf{v}_0 + \mathbf{d}_0$ and so the rotated line must pass through the points $p\mathbf{v}_0p^*$ and $p(\mathbf{v}_0 + \mathbf{d}_0)p^* = p\mathbf{v}_0p^* + p\mathbf{d}_0p^*$ and hence has direction vector $p\mathbf{d}_0p^*$. □

Theorem 8.2.3. To reflect the line parametrized by $(\mathbf{v}_0, \mathbf{d}_0)$ in a plane through the origin with unit normal vector $\hat{\mathbf{n}}$ we follow the same approach as rotations in other words reflect both the anchor point and the direction vector:

$$(\mathbf{v}_0, \mathbf{d}_0) \mapsto (\hat{\mathbf{n}}\mathbf{v}_0\hat{\mathbf{n}}, \hat{\mathbf{n}}\mathbf{d}_0\hat{\mathbf{n}})$$

Proof. Similar and omitted. □

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

Exercise 8.1. Find the result when the line $(2\hat{\boldsymbol{\imath}} + 0\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}, 3\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}})$ is rotated by 7.32 radians about the axis through the origin with $\mathbf{u} = 2\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}$.

□

Exercise 8.2. Find the result when the line $(2\hat{\boldsymbol{\imath}} + 0\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}, 3\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}})$ is rotated by 2.3 radians about the axis through $(10, 10, 0)$ with $\mathbf{u} = 2\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}$. Note: First translate so the axis passes through the origin, then rotate, then translate back.

□

Exercise 8.3. Find the result when the line $(2\hat{\boldsymbol{\imath}} + 0\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}, 3\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}})$ is reflected in the plane through the origin with normal vector $\mathbf{n} = 1\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$.

□

Exercise 8.4. Find the result when the line $(2\hat{\boldsymbol{\imath}} + 0\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}}, 3\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}})$ is reflected in the plane through $(4, 3, 0)$ with normal vector $\mathbf{n} = 1\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$.

□

8.3 Transformations of Planes

Theorem 8.3.1. To translate the plane parametrized by $(\mathbf{v}_0, \mathbf{n}_0)$ by $\mathbf{q}$ we translate the anchor point $\mathbf{v}_0$. The orientation doesn't change so the normal vector $\mathbf{n}_0$ is left untouched. That is:

$$(\mathbf{v}_0, \mathbf{n}_0) \mapsto (\mathbf{v}_0 + \mathbf{q}, \mathbf{n}_0)$$

Proof. In the statement. □

The proof of the result for the rotation of planes follows from the fact that conjugation by a unit quaternion fixes the dot product between vectors. In other words:

Theorem 8.3.2. For vectors $\mathbf{a}$ and $\mathbf{b}$ and for a unit quaternion p we have:

$$(p\mathbf{a}p^*) \cdot (p\mathbf{b}p^*) = \mathbf{a} \cdot \mathbf{b}$$

Proof. We have:

$$\begin{aligned}
(p\mathbf{a}p^*) \cdot (p\mathbf{b}p^*) &= -\frac{1}{2}(p\mathbf{a}p^*p\mathbf{b}p^* + p\mathbf{b}p^*p\mathbf{a}p^*) \\
&= -\frac{1}{2}(p\mathbf{a}(1)\mathbf{b}p^* + p\mathbf{b}(1)\mathbf{a}p^*) \\
&= -\frac{1}{2}(p(\mathbf{ab} + \mathbf{ba})p^*) \\
&= p(\mathbf{a} \cdot \mathbf{b})p^* \\
&= pp^*(\mathbf{a} \cdot \mathbf{b}) \\
&= 1(\mathbf{a} \cdot \mathbf{b}) \\
&= \mathbf{a} \cdot \mathbf{b}
\end{aligned}$$

□

Now then:

Theorem 8.3.3. To rotate the plane $\mathcal{P}$ parametrized by $(\mathbf{v}_0, \mathbf{n}_0)$ by θ radians about the axis $\hat{\mathbf{u}}$ through the origin yielding the new plane $\mathcal{P}'$ we rotate both the anchor point and the normal vector. That is, assign $p = \cos(\theta/2) + \sin(\theta/2)\hat{\mathbf{u}}$ and map:

$$(\mathbf{v}_0, \mathbf{n}_0) \mapsto (p\mathbf{v}_0p^*, p\mathbf{n}_0p^*)$$

Proof. We have:

$$\begin{aligned}
\mathbf{n}_0 \cdot (\mathbf{v} - \mathbf{v}_0) &= (p\mathbf{n}_0p^*) \cdot (p(\mathbf{v} - \mathbf{v}_0)p^*) \\
&= (p\mathbf{n}_0p^*) \cdot (p\mathbf{v}p^* - p\mathbf{v}_0p^*)
\end{aligned}$$

It follows that $\mathbf{v} \in \mathcal{P}$ iff $p\mathbf{v}p^* \in \mathcal{P}'$. □

Theorem 8.3.4. To reflect the plane parametrized by $(\mathbf{v}_0, \mathbf{n}_0)$ in a plane through the origin with unit normal vector $\hat{\mathbf{n}}$ we follow the same approach as rotations in other words reflect both the anchor point and the normal vector:

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

Proof. Similar and omitted. □

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

Example 8.1. Consider the plane $2x + y - z = 10$. Suppose we wish to rotate this plane by 0.5 radians about the axis $\mathbf{u} = 2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 4\hat{\boldsymbol{k}}$. We find the normal for the plane and any point on the plane:

$$\mathbf{n}_0 = 2\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}}$$
$$\mathbf{v}_0 = 0\hat{\boldsymbol{\imath}} + 10\hat{\boldsymbol{\jmath}} + 0\hat{\boldsymbol{k}}$$

We then find:

$$\hat{\mathbf{u}} = \frac{1}{\sqrt{29}}(2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 4\hat{\boldsymbol{k}})$$
$$p = \cos(0.5/2) + \frac{1}{\sqrt{29}}(2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 4\hat{\boldsymbol{k}})\sin(0.5/2)$$

Then we find the new normal and point:

$$p\mathbf{n}_0 p^* \approx 1.9371\hat{\boldsymbol{\imath}} + 0.4827\hat{\boldsymbol{\jmath}} - 1.4194\hat{\boldsymbol{k}}$$
$$p\mathbf{v}_0 p^* \approx 3.8144\hat{\boldsymbol{\imath}} + 9.1557\hat{\boldsymbol{\jmath}} + 1.2740\hat{\boldsymbol{k}}$$

Therefore the new plane has equation:

$$1.9371(x - 3.8133) + 0.4827(y - 9.1557) - 1.4194(z - 1.2740) = 0$$

$\square$

Exercise 8.5. Find the result when the plane $x+2y+z = 4$ with normal vector arising from the coefficients is rotated by 0.2 radians about the axis through the origin with $\mathbf{u} = 4\hat{\boldsymbol{\imath}} + 6\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$.
Hint: The plane can be thought of as $(\mathbf{v}_0, \hat{\mathbf{n}})$ where $\mathbf{v}_0$ is any point on the plane and $\hat{\mathbf{n}}$ is the unit vector arising from the coefficients.

$\square$

Exercise 8.6. Find the result when the plane $x + 2y + z = 4$ with normal vector arising from the coefficients is rotated by 4.3 radians about the axis through $(5, 5, 5)$ with $\mathbf{u} = 4\hat{\boldsymbol{\imath}} + 6\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$.

$\square$

Exercise 8.7. 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.

$\square$

Exercise 8.8. 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.

$\square$

Exercise 8.9. 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 $\mathbf{u} = 4\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 0\hat{\boldsymbol{k}}$.

$\square$

Exercise 8.10. 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 $\mathbf{u} = 4\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 0\hat{\boldsymbol{k}}$.

$\square$

9 Slerp

9.1 Basic Slerp

Suppose you have an object located at a certain point $\mathbf{v}_s$ and you wish to move it with a constant velocity to another point $\mathbf{v}_e$, On way to do this is simply along a straight line which we can parametrize by:

$$\mathbf{v}(t) = (1-t)\mathbf{v}_s + t\mathbf{v}_e \text{ for } 0 \leq t \leq 1$$

These points lie between $\mathbf{v}_s$ and $\mathbf{v}_e$, specifically on the plane containing the origin as well as these two points. The velocity is constant, it is simply the distance traveled since the time required is 1.

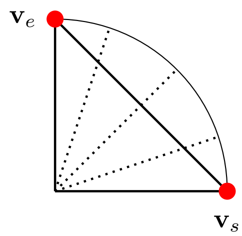
But suppose we wished to move this object along an arc. For simplicitity sake let's assume that both $\mathbf{v}_s$ and $\mathbf{v}_e$ are unit quaternions and we wish to move the object along the shortest curve on the unit sphere.

One solution might simply to normalize the above vectors:

$$\mathbf{v}(t) = \frac{(1-t)\mathbf{v}_s + t\mathbf{v}_e}{|(1-t)\mathbf{v}_s + t\mathbf{v}_e|} \text{ for } 0 \leq t \leq 1$$

While this will follow the desired route the velocity will not be constant.

This can be shown computationally but it more easily seen by this pictoral example where the straight-line route has been divided into four equal quarter lengths by the dotted line. Along the staight-line route the time required is 1/4 per quarter-length. When we normalize to get the arc, however, the four quarter-arcs are not the same length because the distances at the ends are shorter. Consequently the object speeds up as it moves towards the middle.



Instead we define the following:

Definition 9.1.1. We define *spherical linear interpolation*:

$$\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, t) = \frac{\sin(\theta(1-t))}{\sin\theta}\mathbf{v}_s + \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_e$$

where $0 < \theta < \pi$ satisfies $\cos\theta = \mathbf{v}_s \cdot \mathbf{v}_e$.

□

Theorem 9.1.1. This function has the following properties:

(a) When $t = 0$ we get $\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, 0) = \mathbf{v}_s$.

(b) When $t = 1$ we get $\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, 1) = \mathbf{v}_e$.

(c) For all t we have:

$$|\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, t)| = 1$$

(d) The speed of Slerp is constant. In fact if $\mathbf{v}_s$ and $\mathbf{v}_s$ are fixed then:

$$\left|\frac{d}{dt}\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, t)\right| = |\theta|$$

Proof. The proofs of (a) and (b) are clear when we plug in $t = 0$ and $t = 1$.

For (c) first note two facts:

(i) For unit vectors $\mathbf{v}$ and $\mathbf{w}$ we have:

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

(ii) $\sin(\theta(1-t)) = \sin(\theta - \theta t) = \sin\theta\cos(\theta t) - \sin(\theta t)\cos\theta$

From there it's just a lengthy calculation:

$$\begin{aligned}
\mathrm{slerp}(\mathbf{v}_s, \mathbf{v}_t, t) &= \frac{\sin(\theta(1-t))}{\sin\theta}\mathbf{v}_s + \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_e \\
&= \frac{1}{\sin\theta}\left[\sin(\theta - \theta t)\mathbf{v}_s + \sin(\alpha t)\mathbf{v}_e\right] \\
&= \frac{1}{\sin\theta}\left[\left(\sin\theta\cos(\theta t) - \sin(\theta t)\cos\theta\right)\mathbf{v}_s + \sin(\alpha t)\mathbf{v}_e\right] \\
\mathrm{slerp}(\mathbf{v}_s, \mathbf{v}_t, t)\sin\theta &= \left(\sin\theta\cos(\theta t) - \sin(\theta t)\cos\theta\right)\mathbf{v}_s + \sin(\alpha t)\mathbf{v}_e \\
|\mathrm{slerp}(\mathbf{v}_s, \mathbf{v}_t, t)|^2\sin^2\theta &= \sin^2\theta\cos^2(\theta t) - 2\sin\theta\cos\theta\sin(\theta t)\cos(\theta t) + \sin^2(\theta t)\cos^2\theta \\
&\quad + 2(\sin\theta\cos(\theta t) - \sin(\theta t)\cos\theta)\sin(\theta t)\cos\theta \\
&\quad + \sin^2(\theta t) \\
&= \sin^2\theta\cos^2(\theta t) - 2\sin\theta\cos\theta\sin(\theta t)\cos(\theta t) + \sin^2(\theta t)\cos^2\theta \\
&\quad + 2\sin\theta\cos\theta\sin(\theta t)\cos(\theta t) - 2\sin^2(\theta t)\cos^2\theta \\
&\quad + \sin^2(\theta t) \\
&= \sin^2\theta\cos^2(\theta t) - \sin^2(\theta t)\cos^2\theta + \sin^2(\theta t) \\
&= \sin^2(\theta t)(1 - \cos^2\theta) + \sin^2\theta\cos^2(\theta t) \\
&= \sin^2(\theta t)\sin^2\theta + \sin^2\theta\cos^2(\theta t) \\
&= \sin^2\theta \\
|\mathrm{slerp}(\mathbf{v}_s, \mathbf{v}_t, t)| &= 1
\end{aligned}$$

For (d) first note:

(i) $\cos(\theta(1-t)) = \cos(\theta - \theta t) = \cos\theta\cos(\theta t) + \sin(\theta t)\sin\theta$

From there it's just another lengthy calculation:

$$
\begin{aligned}
\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t) &= \frac{\sin(\theta(1-t))}{\sin\theta}\mathbf{v}_s + \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_e \\
\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t) &= \frac{1}{\sin\theta}\left[\sin(\theta(1-t))\mathbf{v}_s + \sin(\theta t)\mathbf{v}_e\right] \\
\frac{d}{dt}\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t) &= \frac{1}{\sin\theta}\left[-\theta\cos(\theta(1-t))\mathbf{v}_s + \theta\cos(\theta t)\mathbf{v}_e\right] \\
\left(\frac{\sin\theta}{\theta}\right)\frac{d}{dt}\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t) &= -\cos(\theta(1-t))\mathbf{v}_s + \cos(\theta t)\mathbf{v}_e \\
&= -(\cos\theta\cos(\theta t) + \sin(\theta t)\sin\theta)\mathbf{v}_s + \cos(\theta t)\mathbf{v}_e \\
\left(\frac{\sin^2\theta}{\theta^2}\right)\left|\frac{d}{dt}\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t)\right|^2 &= \cos^2\theta\cos^2(\theta t) + 2\sin\theta\cos\theta\sin(\theta t)\cos(\theta t) + \sin^2\theta\sin^2(\theta t) \\
&\quad - 2(\cos\theta\cos(\theta t) + \sin(\theta t)\sin\theta)\cos(\theta t)\cos\theta \\
&\quad + \cos^2(\theta t) \\
&= \cos^2\theta\cos^2(\theta t) + 2\sin\theta\cos\theta\sin(\theta t)\cos(\theta t) + \sin^2\theta\sin^2(\theta t) \\
&\quad - 2\cos\theta\cos^2(\theta t) - 2\sin\theta\sin(\theta t)\cos(\theta t)\cos\theta \\
&\quad + \cos^2(\theta t) \\
&= \cos^2\theta\cos^2(\theta t) + 2\sin\theta\cos\theta\sin(\theta t)\cos(\theta t) + \sin^2\theta\sin^2(\theta t) \\
&\quad - 2\cos^2\theta\cos^2(\theta t) - 2\sin\theta\cos\theta\sin(\theta t)\cos(\theta t) \\
&\quad + \cos^2(\theta t) \\
&= -\cos^2\theta\cos^2(\theta t) + \sin^2\theta\sin^2(\theta t) + \cos^2(\theta t) \\
&= \cos^2(\theta t)(1-\cos^2\theta) + \sin^2\theta\sin^2(\theta t) \\
&= \cos^2(\theta t)\sin^2\theta + \sin^2\theta\sin^2(\theta t) \\
&= \sin^2\theta \\
\left(\frac{\sin^2\theta}{\theta^2}\right)\left|\frac{d}{dt}\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t)\right|^2 &= \sin^2\theta \\
\left|\frac{d}{dt}\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t)\right|^2 &= \theta^2 \\
\left|\frac{d}{dt}\operatorname{slerp}(\mathbf{v}_s, \mathbf{v}_t, t)\right| &= |\theta|
\end{aligned}
$$

$\square$

It follows from these facts, and the fact that $\operatorname{Slerp}(\mathbf{v}_s, \mathbf{v}_e, t)$ is a a linear combination of $\mathbf{v}_s$ and $\mathbf{v}_e$, that it always lies on the plane they span, which is a plane through the origin, hence it lies on the great circle joining the two, hence it lies on the curve of shortest distance.

The reason we know this is the shortest distance (rather than going the long way around the great circle) is that the speed is θ. If it were the long way around it would need to have speed $2\pi - \theta$ in order to finish.

It travels along that great circle at constant speed, starting at $\mathbf{v}_s$ at $t = 0$ and ending at $\mathbf{v}_e$ at $t = 1$.

Example 9.1. Consider the following two unit vectors: Consider the following two unit vectors:

$$\begin{aligned}\mathbf{v}_s &= \frac{1}{\sqrt{5}}(0\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}) \\ \mathbf{v}_e &= \frac{1}{\sqrt{13}}(3\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 0\hat{\boldsymbol{k}})\end{aligned}$$

We find our angle θ via:

$$\begin{aligned}\cos\theta &= \mathbf{v}_s \cdot \mathbf{v}_e \\ \cos\theta &= \frac{2}{\sqrt{65}} \\ \theta &= \cos^{-1}(2/\sqrt{65})\end{aligned}$$

Note that by default the arccosine function traditionally returns values in the range $[0, \pi]$ so we're getting the right θ here.

The expression for Slerp is then:

$$\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, t) = \frac{\sin(\theta(1-t))}{\sin\theta}\left(\frac{1}{\sqrt{5}}(0\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}})\right) + \frac{\sin(\theta t)}{\sin\theta}\left(\frac{1}{\sqrt{13}}(3\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 0\hat{\boldsymbol{k}})\right)$$

Then for example at $t = 0.2$ we have location:

$$\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, 0.2) \approx [0.2241; 0.5513; 0.8037]$$

Exercise 9.1. Consider the following two unit vectors:

$$\begin{aligned}\mathbf{v}_s &= \frac{1}{\sqrt{14}}(2\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}) \\ \mathbf{v}_e &= \frac{1}{\sqrt{26}}(5\hat{\boldsymbol{\imath}} + 0\hat{\boldsymbol{\jmath}} + 1\hat{\boldsymbol{k}})\end{aligned}$$

(a) Write down the expression for Slerp.

(b) Find the location at $t = 0, 0.25, 0.5, 0.75, 1$.

□

Exercise 9.2. The two points $P = (1, 0, 0)$ and $Q = \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right)$ both lie on the unit sphere.

(a) Write down the expression for Slerp going from P to Q.

(b) Find the location at $t = 0, 0.25, 0.5, 0.75, 1$.

□

9.2 Adapting Slerp

If $\mathbf{v}_e$ and $\mathbf{v}_s$ are equidistant from $\mathbf{c}$ then we can adapt Slerp to provide a rotation from $\mathbf{v}_e$ to $\mathbf{v}_s$ along the great circle on the sphere of radius $R = |\mathbf{v}_s - \mathbf{c}| = |\mathbf{v}_e - \mathbf{c}|$ by translating so that $\mathbf{c}$ goes to the origin, using the Slerp formula multiplied by R but with normalized versions of our translated $\mathbf{v}_s$ and $\mathbf{v}_e$, and then translating back.

In this formula θ is the angle between the translated $\mathbf{v}_s$ and $\mathbf{v}_e$.

$$\text{NewSlerp}(\mathbf{v}_s, \mathbf{v}_e, t) = \mathbf{c} + R\left[\frac{\sin(\theta(1-t))}{\sin\theta}\frac{\mathbf{v}_s - \mathbf{c}}{R} + \frac{\sin(\theta t)}{\sin\theta}\frac{\mathbf{v}_e - \mathbf{c}}{R}\right]$$

Exercise 9.3. The two points $P = (4, 7, 0)$ and $Q = (5, 4, 2)$ are equidistant from $C = (1, 3, -1)$.

(a) Write down the expression for NewSlerp going from P to Q.

(b) Find the location at $t = 0, 0.25, 0.5, 0.75, 1$.

9.3 Exponential Form

Although we won't delve into too much detail it is worth saying a few things about exponentials here and mentioning the exponential version of Slerp. Partly we do this because it places Slerp in the context of quaternions, as the definition given above doesn't really require any quaternion manipulation and partly because the exponential version is rather elegant.

Definition 9.3.1. For any $q \in \mathbb{H}$ we may define the exponential function via the Taylor expansion:

$$e^q = \sum_{n=0}^{\infty} \frac{q^n}{n!}$$

which (we will not prove) converges for all q.

□

This definition gives rise to the inverse function, the natural logarithm of a quaternion.

Theorem 9.3.1. If u is a unit quaternion and $r \in \mathbb{R}$ then if we write:

$$u = \cos\theta + \hat{\mathbf{u}}\sin\theta$$

then we obtain:

$$u^t = \cos(\theta t) + \hat{\mathbf{u}} \sin(\theta t)$$

Proof. Details omitted, but just like with the reals the basic idea is:

$$u^t = e^{t \ln u} = \text{...apply series expansion...} = \cos(\theta t) + \hat{\mathbf{u}} \sin(\theta t)$$

□

Exercise 9.4. Suppose we have the unit quaternion:

$$u = \frac{\sqrt{3}}{2} + \frac{4}{2\sqrt{21}}\hat{\boldsymbol{\imath}} + \frac{1}{2\sqrt{21}}\hat{\boldsymbol{\jmath}} + \frac{1}{\sqrt{21}}\hat{\boldsymbol{k}}$$

(a) Write u in the form $\cos\theta + \hat{\mathbf{u}} \sin\theta$ for appropriate $\hat{\mathbf{u}}$ and θ.

(b) Find q^2, $q^{1/3}$ and q^9.

□

Corollary 9.3.1. It follows that for any quaternion q we can write $q = |q|u$ where $u = q/|q|$ and then if we write u as above then we have:

$$q^t = |q|^t(\cos(\theta t) + \hat{\mathbf{u}} \sin(\theta t))$$

Proof. In the statement. □

Exercise 9.5. Find an approximate value for $(2 + 3\hat{\boldsymbol{\imath}} - 1\hat{\boldsymbol{\jmath}} + 4\hat{\boldsymbol{k}})^{1.3}$. To do this factor out the magnitude and find an approximate value for θ, then proceed from there.

□

Under this notation we can rewrite Slerp in this particularly convenient form which doesn't require us to worry about angles or trig functions explicitly. They're taken care of by the quaternion calculation.

Theorem 9.3.2. We have the alternate definition of Slerp:

$$\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, t) = \mathbf{v}_s \left(\mathbf{v}_s^{-1}\mathbf{v}_e\right)^t$$

Proof. First since $\mathbf{v}_s$ is a pure unit quaternion its inverse equals its conjugate which equals its negative. Thus we have:

$$
\begin{aligned}
\mathbf{v}_s\left(\mathbf{v}_s^{-1}\mathbf{v}_e\right)^t &= \mathbf{v}_s\left(-\mathbf{v}_s\mathbf{v}_e\right)^t \\
&= \mathbf{v}_s\left(-(\mathbf{v}_s \times \mathbf{v}_e - \mathbf{v}_s \cdot \mathbf{v}_e)\right)^t \\
&= \mathbf{v}_s\left(-(\mathbf{v}_s \times \mathbf{v}_e - \cos\theta)\right)^t \\
&= \mathbf{v}_s\left(\cos\theta - \mathbf{v}_s \times \mathbf{v}_e\right)^t \\
&= \mathbf{v}_s\left(\cos\theta - \frac{\mathbf{v}_s \times \mathbf{v}_e}{|\mathbf{v}_s \times \mathbf{v}_e|}|\mathbf{v}_s \times \mathbf{v}_e|\right)^t \\
&= \mathbf{v}_s\left(\cos\theta - \frac{\mathbf{v}_s \times \mathbf{v}_e}{|\mathbf{v}_s||\mathbf{v}_e|\sin\theta}|\mathbf{v}_s||\mathbf{v}_e|\sin\theta\right)^t \\
&= \mathbf{v}_s\left(\cos\theta - \frac{\mathbf{v}_s \times \mathbf{v}_e}{\sin\theta}\sin\theta\right)^t \\
&= \mathbf{v}_s\left(\cos(\theta t) - \frac{\mathbf{v}_s \times \mathbf{v}_e}{\sin\theta}\sin(\theta t)\right) \\
&= \cos(\theta t)\mathbf{v}_s - \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_s(\mathbf{v}_s \times \mathbf{v}_e) \\
&= \cos(\theta t)\mathbf{v}_s - \frac{\sin(\theta t)}{\sin\theta}\left((\mathbf{v}_s \cdot \mathbf{v}_e)\mathbf{v}_s - (\mathbf{v}_s \cdot \mathbf{v}_s)\mathbf{v}_e\right) \\
&= \cos(\theta t)\mathbf{v}_s - \frac{\sin(\theta t)}{\sin\theta}\left(\cos\theta\,\mathbf{v}_s - \mathbf{v}_e\right) \\
&= \left(\cos(\theta t) - \frac{\cos\theta\sin(\theta t)}{\sin\theta}\right)\mathbf{v}_s + \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_e \\
&= \left(\frac{\sin\theta\cos(\theta t) - \cos\theta\sin(\theta t)}{\sin\theta}\right)\mathbf{v}_s + \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_e \\
&= \left(\frac{\sin(\theta - \theta t)}{\sin\theta}\right)\mathbf{v}_s + \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_e \\
&= \left(\frac{\sin(\theta(1-t))}{\sin\theta}\right)\mathbf{v}_s + \frac{\sin(\theta t)}{\sin\theta}\mathbf{v}_e
\end{aligned}
$$

□

Our previous example can then be rewritten:

Example 9.2. Consider the following two unit vectors: Consider the following two unit vectors:

$$
\begin{aligned}
\mathbf{v}_s &= \frac{1}{\sqrt{5}}(0\hat{\boldsymbol{\imath}} + 1\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}) \\
\mathbf{v}_e &= \frac{1}{\sqrt{13}}(3\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 0\hat{\boldsymbol{k}})
\end{aligned}
$$

The expression for Slerp is then:

$$\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, t) = \mathbf{v}_s(\mathbf{v}_s^{-1}\mathbf{v}_e)^t$$

Then for example at $t = 0.2$ we have location:

$$\text{Slerp}(\mathbf{v}_s, \mathbf{v}_e, 0.2) = \mathbf{v}_s(\mathbf{v}_s^{-1}\mathbf{v}_e)^{0.2} \approx [0.2241; 0.5513; 0.8037]$$

□

10 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 $p\mathbf{v}p^*$. 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.