Ari Grant
Sugih Jamin
EECS 487
Winter 2010

Rotations

1. Rotation of a vector around an arbitrary axis

Given a vector $\vec{b}$, an axis $\vec{a}$, and an angle α the goal is to find a formula (or function or operator) that produces the vector $\vec{c}$ created by rotating $\vec{b}$ an angle α around $\vec{a}$ counterclockwise (in a right-handed system). The expression will be found with brute force transformations and then with clever decomposition of vectors.

1.1. **Transformations.** Assume the tails of both $\vec{a}$ and $\vec{b}$ are at the origin. Apply the matrix that rotates the axis $\vec{a}$ around the $z-$axis into the xz-plane and apply that matrix to $\vec{b}$. So

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{b}$$

where $\sin\theta = -\frac{a_y}{\sqrt{a_x^2+a_y^2}}$ and $\cos\theta = \frac{a_x}{\sqrt{a_x^2+a_y^2}}$. The sine is negative because rotation by θ must be "undone;" cosine is an even function (so the minus does nothing).

Next apply the matrix that rotates the new axis in the xz-plane around the $y-$axis to coincide with the $z-$axis, that is apply $\hat{R}_y(\phi)$.

$$\begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{b}$$

where $\sin\phi = -\frac{a_z}{\sqrt{a_x^2+a_y^2+a_z^2}}$ and $\cos\phi = \frac{\sqrt{a_x^2+a_y^2}}{\sqrt{a_x^2+a_y^2+a_z^2}}$ Notice the negative sign on the sine term as before; without it the vector rotates away from the z-axis instead of toward it.

Now apply a rotation by α around the $z-$axis and then undo the first two transformations (apply the inverses in the correct order). Thus $\vec{c} = \hat{R}_z^{-1}(\theta)\hat{R}_y^{-1}(\phi)\hat{R}_z(\alpha)\hat{R}_y(\phi)\hat{R}_z(\theta)\vec{b}$. Remember that $\hat{R}(-\beta) = \hat{R}^{-1}(\beta)$.

$$\vec{c} = \begin{pmatrix} c\theta & s\theta & 0 \\ -s\theta & c\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\phi & 0 & s\phi \\ 0 & 1 & 0 \\ -s\phi & 0 & c\phi \end{pmatrix}$$

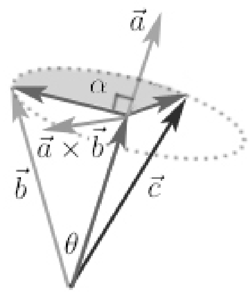
$$\cdot \begin{pmatrix} c\alpha & -s\alpha & 0 \\ s\alpha & c\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\phi & 0 & -s\phi \\ 0 & 1 & 0 \\ s\phi & 0 & c\phi \end{pmatrix} \begin{pmatrix} c\theta & -s\theta & 0 \\ s\theta & c\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{b}$$

where c is cosine and s is sine. Multiplying this out and grunting gives:

$$\vec{c} = \begin{pmatrix} (1-\cos\alpha)\frac{a_xb_x+a_yb_y+a_zb_z}{a_x^2+a_y^2+a_z^2}a_x + b_x\cos\alpha + \frac{a_yb_z-b_ya_z}{\sqrt{a_x^2+a_y^2+a_z^2}}\sin\alpha \\ (1-\cos\alpha)\frac{a_xb_x+a_yb_y+a_zb_z}{a_x^2+a_y^2+a_z^2}a_y + b_y\cos\alpha + \frac{a_zb_x-b_za_x}{\sqrt{a_x^2+a_y^2+a_z^2}}\sin\alpha \\ (1-\cos\alpha)\frac{a_xb_x+a_yb_y+a_zb_z}{a_x^2+a_y^2+a_z^2}a_z + b_z\cos\alpha + \frac{a_xb_y-b_xa_y}{\sqrt{a_x^2+a_y^2+a_z^2}}\sin\alpha \end{pmatrix}.$$

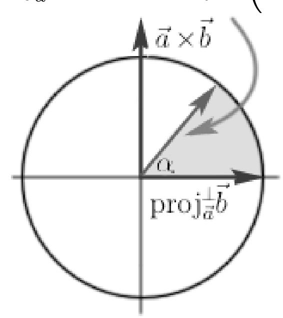
1.2. **Clever Projection.** When rotating $\vec{b}$ around $\vec{a}$ decompose $\vec{b}$ into two pieces, one that stays the same ($\text{proj}_{\vec{a}}\vec{b}$) and one that rotates ($\text{proj}^{\perp}_{\vec{a}}\vec{b}$).

Let θ be the angle between $\vec{a}$ and $\vec{b}$. Let $\vec{c}$ be $\vec{b}$ rotated around $\vec{a}$ clockwise by α.



Notice that $\vec{b}$ and $\vec{c}$ have the same projection onto $\vec{a}$ and only the orthogonal component is rotated.

$$\vec{c} = \text{proj}_{\vec{a}}\vec{b} + \text{rotateBy}\alpha\left(\text{proj}^{\perp}_{\vec{a}}\vec{b}\right)$$



Looking at the same picture from above shows that

$$\text{rotateBy}\alpha\left(\text{proj}^{\perp}_{\vec{a}}\vec{b}\right) = \cos(\alpha)\vec{u} + \sin(\alpha)\vec{v}$$

assuming $\vec{u}$ is parallel to $\text{proj}^{\perp}_{\vec{a}}\vec{b}$, $\vec{v}$ is parallel to $\vec{a}\times\vec{b}$, and $||\vec{u}|| = ||\vec{v}|| = ||\text{proj}^{\perp}_{\vec{a}}\vec{b}||$.[1]

A simple check shows that $||\vec{a}\times\vec{b}|| = ||\vec{a}||\,||\vec{b}||\sin\theta$ and $||\text{proj}^{\perp}_{\vec{a}}\vec{b}|| = ||\vec{b}||\sin\theta$ and thus

$$\text{rotateBy}\alpha\left(\text{proj}^{\perp}_{\vec{a}}\vec{b}\right) = (\text{proj}^{\perp}_{\vec{a}}\vec{b})\cos\alpha + \frac{\vec{a}\times\vec{b}}{||\vec{a}||}\sin\alpha.$$

[1]The vector $\vec{r} = \vec{u}\cos\alpha + \vec{v}\sin\alpha$ is a linear combination of $\vec{u}$ and $\vec{v}$ and thus is in the same plane as the two. Assume $\vec{u}\cdot\vec{v} = 0$ and that $||\vec{u}|| = ||\vec{v}||$. The angle between $\vec{r}$ and $\vec{u}$ is $\arccos(\vec{r}\cdot\vec{u}) = \alpha$. Then $\vec{r}$ is $\vec{u}$ rotated by α iff $||\vec{r}|| = ||\vec{u}||$ which is shown by $||\vec{r}||^2 = ||\vec{u}\cos\alpha + \vec{v}\sin\alpha||^2 = (\vec{u}\cos\alpha + \vec{v}\sin\alpha)\cdot(\vec{u}\cos\alpha + \vec{v}\sin\alpha) = ||\vec{u}||^2\cos^2\alpha + ||\vec{v}||^2\sin^2\alpha + 2\vec{u}\cdot\vec{v}\sin\alpha\cos\alpha = ||\vec{u}||^2$.

Putting it all together gives the vector $\vec{c}$ as $\vec{b}$ rotated clockwise around $\vec{a}$ by α as

$$\begin{aligned}
\vec{c} &= \mathrm{proj}_{\vec{a}}\vec{b} + (\mathrm{proj}^{\perp}_{\vec{a}}\vec{b})\cos\alpha + \frac{\vec{a}\times\vec{b}}{||\vec{a}||}\sin\alpha \\
&= \mathrm{proj}_{\vec{a}}\vec{b} + (\vec{b} - \mathrm{proj}_{\vec{a}}\vec{b})\cos\alpha + \frac{\vec{a}\times\vec{b}}{||\vec{a}||}\sin\alpha \\
\vec{c} &= (1-\cos\alpha)\frac{(\vec{a}\cdot\vec{b})}{||\vec{a}||^2}\vec{a} + \vec{b}\cos\alpha + \frac{\vec{a}\times\vec{b}}{||\vec{a}||}\sin\alpha.
\end{aligned}$$

This is Rodrigues' Rotation Formula (the same expression derived with matrices above).

$$\vec{c} = (1-\cos\alpha)(\hat{a}\cdot\vec{b})\hat{a} + \vec{b}\cos\alpha + (\hat{a}\times\vec{b})\sin\alpha \tag{1.1}$$

2. Complex numbers as rotations in the plane

2.1. **Euler's Formula.** A vector in the plane $\vec{v} = \langle a, b\rangle$ may be written in polar coordinates as $\vec{v} = \langle r, \theta\rangle$ where $r = \sqrt{a^2+b^2}$ and $\tan\theta = \frac{b}{a}$. Then the vector can be written as $\vec{v} = r\cos\theta\,\hat{x} + r\sin\theta\,\hat{y}$ with $r\in\mathbb{R}$ and $\theta\in[0,2\pi)$.

Recall the series expansions of $e^x, \sin x$, and $\cos x$ (valid for all x).

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + \cdots \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + \frac{(-1)^n}{(2n)!}x^{2n} + \cdots
\end{aligned}$$

Use these three to calculate $e^{\mathbf{i}\theta}$ where $\mathbf{i}^2 = -1$, the usual imaginary number.[2]

$$\begin{aligned}
e^{\mathbf{i}\theta} &= 1 + (\mathbf{i}\theta) + \frac{(\mathbf{i}\theta)^2}{2!} + \frac{(\mathbf{i}\theta)^3}{3!} + \cdots + \frac{(\mathbf{i}\theta)^n}{n!} + \cdots \\
&= 1 + \mathbf{i}\theta - \frac{\theta^2}{2!} - \mathbf{i}\frac{\theta^3}{3!} + \frac{\theta^4}{4!}\cdots + \frac{(i\theta)^n}{n!} \\
&= \left(1 - \frac{\theta^2}{2!} + \theta^4 4! - \cdots\right) + \mathbf{i}\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right)
\end{aligned}$$

The result is Euler's Formula

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

which gives rise to $re^{\mathbf{i}\theta} = r\cos\theta + \mathbf{i}r\sin\theta$.

[2] There are many issues in doing this, for example convergence of complex series and the ability to reorder terms (absolute convergence) but for simplicity these will be ignored.

2.2. **Complex Multiplication as Rotation.** Graphing complex numbers can be done in the usual 2-dimensional Cartesian plane with the x-axis representing the real part and the y-axis the imaginary, i.e., the complex plane. Then any vector $\vec{v} = \langle a, b\rangle$ can be written $\vec{v} = a + \mathbf{i}b = re^{\mathbf{i}\theta}$ where $r^2 = a^2 + b^2$ and $\tan\theta = \frac{b}{a}$.

What can be said about multiplication by a complex number with modulus 1? The modulus of a complex number $z = a + \mathbf{i}b$ is defined as $||z|| = \sqrt{a^2 + b^2}$. A complex number with modulus 1 has $a^2 + b^2 = r^2 = 1$. Let $z = a + b\mathbf{i}$ be a complex number with modulus 1. Then $z = e^{\mathbf{i}\phi}$ for some ϕ. Pre-multiplying some complex number $w = c + \mathbf{i}d = re^{\mathbf{i}\theta}$ by z gives:

$$zw = (a + bi)(c + di) = e^{\mathbf{i}\phi}re^{\mathbf{i}\theta} = re^{\mathbf{i}(\theta+\phi)}$$

The resultant number (vector) has the same length as w but polar angle ϕ more than it. Thus the result is w rotated by ϕ! This is a very important fact: multiplying by a complex number is equivalent to a rotation in the plane!

2.3. **Rotation as an Operator.** Again recall the Rodrigues Rotation Formula.

$$R(\vec{b}, \alpha, \vec{a}) = (1 - \cos\alpha)(\hat{a}\cdot\vec{b})\hat{a} + \vec{b}\cos\alpha + (\hat{a}\times\vec{b})\sin\alpha$$

where $R(\vec{b}, \alpha, \vec{a})$ denotes rotation of $\vec{b}$ by α around $\vec{a}$. Given the theory of transformations it would be preferable to write this as an operation, that is find $R(\alpha, \vec{a})$ such that

$$R(\vec{b}, \alpha, \vec{a}) = R(\alpha, \vec{a})\vec{b}.$$

To start, write

$$R(\vec{b}, \alpha, \vec{a}) = (1 - \cos\alpha)\hat{a}(\hat{a}^T\vec{b}) + \cos\alpha\vec{b} + \sin\alpha(\hat{a}\times\vec{b}).$$

All that remains is to find some operator $\hat{a}^*$ such that $\hat{a}^*\vec{b} = \hat{a}\times\vec{b}$.

Examining the form of $\hat{a}\times\vec{b}$.

$$\hat{a}\times\vec{b} = \begin{pmatrix} a_y b_z - b_y a_z \\ a_z b_x - b_z a_x \\ a_x b_y - b_x a_y \end{pmatrix} = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix}\begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix}$$

This shows that

$$\hat{a}^* = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix}$$

and $\hat{a}^*\vec{b} = \hat{a}\times\vec{b}$. The object (matrix) $\hat{a}^*$ is called the dual matrix of vector $\hat{a}$.[3]

Using $\hat{a}^*$ to rewrite $\hat{a}\times\vec{b}$ gives

$$\begin{aligned} R(\vec{b}, \alpha, \hat{a}) &= (1 - \cos\alpha)\hat{a}\hat{a}^T\vec{b} + \cos\alpha\vec{b} + \sin\alpha(\hat{a}\times\vec{b}) \\ &= (1 - \cos\alpha)\hat{a}\hat{a}^T\vec{b} + \cos\alpha\vec{b} + \sin\alpha\hat{a}^*\vec{b} \\ &= \left[(1 - \cos\alpha)\hat{a}\hat{a}^T + \cos\alpha + \sin\alpha\hat{a}^*\right]\vec{b} \end{aligned}$$

[3]The * was intentionally chosen because $\hat{a}^*$ is a mathematical object, called the dual of $\hat{a}$, where * is an operator called the Hodge dual or Hodge star operator.

and as noted above, this is the desired rotation operator

$$R(\alpha, \hat{a}) = (1 - \cos\alpha)\hat{a}\hat{a}^T + \cos\alpha + \sin\alpha\hat{a}^*$$

but it can be further simplified by examining $(\hat{a}^*)^2$.

$$(\hat{a}^*)^2 = \hat{a}^*\hat{a}^* = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix} \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -a_y^2 - a_z^2 & a_x a_y & a_x a_z \\ a_x a_y & -a_x^2 - a_z^2 & a_y a_z \\ a_x a_z & a_y a_z & -a_x^2 - a_y^2 \end{pmatrix}.$$

Since $\hat{a}$ is a unit vector $||\hat{a}||^2 = 1 = a_x^2 + a_y^2 + a_z^2$ and

$$(\hat{a}^*)^2 = \begin{pmatrix} a_x^2 - 1 & a_x a_y & a_x a_z \\ a_x a_y & a_y^2 - 1 & a_y a_z \\ a_x a_z & a_y a_z & a_z^2 - 1 \end{pmatrix} = \begin{pmatrix} a_x^2 & a_x a_y & a_x a_z \\ a_x a_y & a_y^2 & a_y a_z \\ a_x a_z & a_y a_z & a_z^2 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \begin{pmatrix} a_x & a_y & a_z \end{pmatrix} - \mathbf{I} = \hat{a}\hat{a}^T - \mathbf{I}.$$

Finally, the operator can be written in a succinct form.

$$\begin{aligned} R(\alpha, \hat{a}) &= (1 - \cos\alpha)\left(\mathbf{I} + (\hat{a}^*)^2\right) + \cos\alpha + \sin\alpha\hat{a}^* \\ &= \mathbf{I} - \mathbf{I}\cos\alpha + (\hat{a}^*)^2(1 - \cos\alpha) + \cos\alpha + \hat{a}^*\sin\alpha \\ R(\alpha, \hat{a}) &= \mathbf{I} + (\hat{a}^*)^2(1 - \cos\alpha) + \hat{a}^*\sin\alpha \end{aligned}$$

Recall now that to rotate $\vec{b}$ by α around $\hat{a}$, do $R(\alpha, \hat{a})\vec{b}$.

The above expression for the rotation operator shows that $(\hat{a}^*)^2\vec{b}$ and $\hat{a}^*\vec{b}$ form an orthogonal basis in the plane orthogonal to $\hat{a}$. The vector $\hat{a}^*\vec{b} = \hat{a} \times \vec{b}$ is trivially perpendicular to $\hat{a}$ and so a quick examination of $(\hat{a}^*)^2\vec{b}$ is necessary.

$$(\hat{a}^*)^2\vec{b} = (\hat{a}^*\hat{a}^*)\vec{b} = \hat{a}^*(\hat{a}^*\vec{b}) = \hat{a}^*(\hat{a} \times \vec{b}) = \hat{a} \times \hat{a} \times \vec{b}$$

Recall the vector identity $\vec{a} \times \vec{b} \times \vec{c} = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$.

$$\begin{aligned} (\hat{a}^*)^2\vec{b} &= \hat{a}(\hat{a} \cdot \vec{b}) - \vec{b}(\hat{a} \cdot \hat{a}) = \hat{a}(\hat{a} \cdot \vec{b}) - \vec{b}||\hat{a}||^2 = \hat{a}(\hat{a} \cdot \vec{b}) - \vec{b} \\ &= -\left(\vec{b} - \hat{a}(\hat{a} \cdot \vec{b})\right) = -\left(\vec{b} - \mathrm{proj}_{\hat{a}}\vec{b}\right) \\ (\hat{a}^*)^2\vec{b} &= -\mathrm{proj}_{\hat{a}}^{\perp}\vec{b} \end{aligned}$$

Thus $(\hat{a}^*)^2\vec{b}$ is just the negative of the orthogonal component of $\vec{b}$ projected onto $\hat{a}$, as shown in the geometry explored earlier.

2.4. **Rotation Operator by Symmetry with Complex Numbers.** As shown above

$$(\hat{a}^*)^2 = \hat{a}\hat{a}^T - \mathbf{I} = \begin{pmatrix} a_x^2 - 1 & a_x a_y & a_x a_z \\ a_x a_y & a_y^2 - 1 & a_y a_z \\ a_x a_z & a_y a_z & a_z^2 - 1 \end{pmatrix}.$$

Now examine the next power of $\hat{a}^*$.

$$(\hat{a}^*)^3 = \hat{a}^*(\hat{a}\hat{a}^T - \mathbf{I}) = \hat{a}^*\hat{a}\hat{a}^T - \hat{a}^*\mathbf{I} = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix} \begin{pmatrix} a_x^2 & a_x a_y & a_x a_z \\ a_x a_y & a_y^2 & a_y a_z \\ a_x a_z & a_y a_z & a_z^2 \end{pmatrix} - \hat{a}^*$$

$$= \begin{pmatrix} -a_x a_y a_z + a_x a_y a_z & -a_y^2 a_z + a_y^2 a_z & -a_y a_z^2 + a_y a_z^2 \\ a_x^2 a_z - a_x^2 a_z & a_x a_y a_z - a_x a_y a_z & a_x a_z^2 - a_x a_z^2 \\ -a_x^2 a_y + a_x^2 a_y & -a_x a_y^2 + a_x a_y^2 & -a_x a_y a_z + a_x a_y a_z \end{pmatrix} - \hat{a}^*$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \hat{a}^* = -\hat{a}^*$$

Thus the powers of $\hat{a}^*$ are as follows.

$$\begin{aligned} (\hat{a}^*)^0 &= \mathbf{I} \\ (\hat{a}^*)^1 &= \hat{a}^* \\ (\hat{a}^*)^2 &= \hat{a}\hat{a}^T - \mathbf{I} \\ (\hat{a}^*)^3 &= -\hat{a}^* \\ (\hat{a}^*)^4 &= \mathbf{I} - \hat{a}\hat{a}^T \\ (\hat{a}^*)^5 &= \hat{a}^* \\ &\vdots \end{aligned}$$

The powers cycle over every 4 powers and exhibit some similarity to i.

$$\begin{aligned} i^0 &= 1 \\ i^1 &= i \\ i^2 &= -1 \\ i^3 &= -i \\ i^4 &= 1 \\ i^5 &= i \\ &\vdots \end{aligned}$$

Using the above to compute $e^{a^*\alpha}$ with a power series expansion similar to $e^{i\alpha}$:[4]

$$\begin{aligned}
e^{\hat{a}^*\alpha} &= \mathbf{I} + \frac{(\hat{a}^*)\alpha}{1!} + \frac{(\hat{a}^*)^2\alpha^2}{2!} + \frac{(\hat{a}^*)^3\alpha^3}{3!} + \frac{(\hat{a}^*)^4\alpha^4}{4!} + \cdots \\
&= \mathbf{I} + \frac{(\hat{a}^*)\alpha}{1!} + \frac{(\hat{a}^*)^2\alpha^2}{2!} + \frac{-(\hat{a}^*)\alpha^3}{3!} + \frac{-(\hat{a}^*)^2\alpha^4}{4!} + \cdots \\
&= \mathbf{I} + (\hat{a}^*)^2\left(\frac{\alpha^2}{2!} - \frac{\alpha^4}{4!} + \cdots\right) + (\hat{a}^*)\left(\frac{\alpha}{1!} - \frac{\alpha^3}{3!} + \cdots\right) \\
&= \mathbf{I} + (\hat{a}^*)^2\left(1 - 1 - \left[-\frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} - \cdots\right]\right) + (\hat{a}^*)\sin\alpha \\
&= \mathbf{I} + (\hat{a}^*)^2\left(1 - \left[1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} - \cdots\right]\right) + (\hat{a}^*)\sin\alpha \\
&= \mathbf{I} + (\hat{a}^*)^2\left(1 - \cos\alpha\right) + (\hat{a}^*)\sin\alpha \\
&= R(\alpha, \hat{a})
\end{aligned}$$

Thus multiplying a vector by the above results in a rotation. That is $e^{\hat{a}^*\alpha}\vec{b}$ is the vector $\vec{b}$ rotated around $\hat{a}$ by α.

Unfortunately in practice computing $e^{\hat{a}^*\alpha}$ can only be done with a truncated Taylor series and is thus not optimal. What follows in the next sections is some higher theory that allows rotation with a few adds/multiplies instead of the construction of an exponential. The rotation that will be derived will have an equivalent form of $e^{\hat{a}\alpha}\vec{b}e^{-\hat{a}\alpha}$ which is somewhat similar to the above. However the two exponentials cannot be cancelled and behave quite differently than the type of object just derived. Welcome to hypercomplex numbers.

3. Hypercomplex Numbers

Sir William Rowan Hamilton, a mathematician from the nineteenth century, was aware that complex numbers were equivalent to two-dimensional rotations (or orientations) and sought to develop a similar algebra for 3-dimensional rotations.

His first thought was to extend the complex numbers by adding the element $\mathbf{j}$ where $\mathbf{j}^2 = -1$ and $\mathbf{j}$ was linearly independent from $\mathbf{i}$ and 1, thus a whole new element. Then any number could be expressed in the form $k = a + b\mathbf{i} + c\mathbf{j}$ and a theory or rotations might follow. Unfortunately such a system cannot exist. To show this, ask what is $\mathbf{ij}$?

The product of $\mathbf{i}$ and $\mathbf{j}$ must be some number in the new system and so $\mathbf{ij} = a + b\mathbf{i} + c\mathbf{j}$ for some $a, b, c \in \mathbb{R}$. Start by pre-multiplying both sides by $\mathbf{i}$.

$$\begin{aligned}
\mathbf{ij} &= a + b\mathbf{i} + c\mathbf{j} \\
\mathbf{i}^2\mathbf{j} &= \mathbf{i}a + b\mathbf{i}^2 + c\mathbf{ij} \\
-\mathbf{j} &= \mathbf{i}a - b + c\mathbf{ij}
\end{aligned}$$

This can be written as $\mathbf{ij} = \frac{1}{c}(b - \mathbf{i}a - \mathbf{j})$. Equate this with the original expression for $\mathbf{ij}$.

$$\mathbf{ij} = \frac{1}{c}(b - \mathbf{i}a - \mathbf{j}) = a + b\mathbf{i} + c\mathbf{j}$$

[4]Again ignoring all the issues of convergence expressed previously.

In order for this to be true there are three constraints.

$$\frac{b}{c} = a$$
$$-\frac{a}{c} = b$$
$$-\frac{1}{c} = c$$

The first two cause no issue; however the third states that $c^2 = -1$ which cannot be true because $a, b, c \in \mathbb{R}$. The only way for this to work is to add yet another independent element $\mathbf{k}$ where $\mathbf{ij} = \mathbf{k}$.

The next task is to identify the product $\mathbf{ji}$. It may seem that this value is also $\mathbf{k}$, so $\mathbf{ij} = \mathbf{ji} = \mathbf{k}$, but this is contradictory.[5] The only other possibility is that $\mathbf{ji} = -\mathbf{k}$ (proof omitted). Then $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$ and so $\mathbf{k}^2 = -1$, another independent root of 1.

These are the hypercomplex numbers, numbers of the form $a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ with $a, b, c, d \in \mathbb{R}$ and $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$ and $\mathbf{ij} = \mathbf{k}$. The set of all such numbers is denoted by $\mathbb{H}$ in honor of Hamilton (who discovered such objects in 1843). The term hypercomplex number is not often used any more; the preferred term is quaternion.

4. Quaternions

For the mathematically minded: the set of quaternions $\mathbb{H}$ is a finite-dimensional associative algebra (roughly a set of numbers where addition, subtraction, multiplication, and division can be performed). Note that there are only three such objects $\mathbb{R}, \mathbb{C}$, and $\mathbb{H}$.

4.1. **Basic Properties of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and Notation.** In the previous section three rules were put in place $\mathbf{i}^2 = \mathbf{j}^2 = -1$ and $\mathbf{ij} = \mathbf{k}$. From these three rules alone (and the properties of $\mathbb{R}$) the properties of quaternions will be explored. Note that it has already been shown that $\mathbf{k}^2 = -1$ and that $\mathbf{ij} = -\mathbf{ji}$.

Note that because $\mathbf{ij} \neq \mathbf{ji}$, do not assume commutativity (of multiplication) anywhere except with scalars.[6] This is the one property of $\mathbb{H}$ that separates it from $\mathbb{R}$ and $\mathbb{C}$. Thus be very careful to consistently left- or right-multiply and not mix and match.

For example, right-multiplying $\mathbf{ji}$ by $\mathbf{i}$:

$$\mathbf{ji} = -\mathbf{k}$$
$$\mathbf{ji}^2 = -\mathbf{ki}$$
$$-\mathbf{j} = -\mathbf{ki}$$
$$\mathbf{j} = \mathbf{ki}$$

[5]If $\mathbf{ij} = \mathbf{ji} = \mathbf{k}$ then $\mathbf{k}^2 = (\mathbf{ij})^2 = (\mathbf{ij})(\mathbf{ij}) = (\mathbf{ij})(\mathbf{ji}) = \mathbf{i}(\mathbf{jj})\mathbf{i} = \mathbf{i}(-1)\mathbf{i} = -\mathbf{i}^2 = -(-1) = 1$. The fact that $\mathbf{k}^2 = 1$ may not seem bad but it is, because $1 + \mathbf{k} \neq 0$ and $1 - k \neq 0$ (since $\mathbf{k} \neq \pm 1$) but

$$(1 - \mathbf{k})(1 + \mathbf{k}) = 1 - \mathbf{k}^2 = 1 - 1 = 0.$$

Thus there are two nonzero numbers whose product is zero and this is incompatible with the reals.

[6]For a scalar a, $a\mathbf{i} = \mathbf{i}a$, $a\mathbf{j} = \mathbf{j}a$, $a\mathbf{k} = \mathbf{k}a$, $a\mathbf{ij} = \mathbf{i}a\mathbf{j} = \mathbf{ij}a$, etc.

Do the same with $\mathbf{ij} = \mathbf{k}$. Then $\mathbf{i}^2\mathbf{j} = \mathbf{ik}$ and $-\mathbf{j} = \mathbf{ik}$.

$$\mathbf{j} = -\mathbf{ik}$$

Continuing similarly with all pairs yields the following.

$$\begin{aligned} \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 &= -1 \\ \mathbf{ij} = -\mathbf{ji} &= \mathbf{k} \\ \mathbf{jk} = -\mathbf{kj} &= \mathbf{i} \\ \mathbf{ki} = -\mathbf{ik} &= \mathbf{j} \end{aligned}$$

Notice the similarity of the last three lines to the standard Cartesian basis of $\mathbb{R}^3$ and the cross product.

$$\begin{aligned} \hat{\imath} \times \hat{\jmath} = -\hat{\jmath} \times \hat{\imath} &= \hat{k} \\ \hat{\jmath} \times \hat{k} = -\hat{k} \times \hat{\jmath} &= \hat{\imath} \\ \hat{k} \times \hat{\imath} = -\hat{\imath} \times \hat{k} &= \hat{\jmath} \end{aligned}$$

This is reason enough to see that quaternions and vectors have deep connections. Note that if left-handed bases is preferred, define $\mathbf{ij} = -\mathbf{k}$ *from the start* but this is far from standard.

With this similarity it is often convenient to think of a quaternion $w = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ as $w = a + b\hat{\imath} + c\hat{\jmath} + d\hat{k}$. Then write $w = a + \vec{v}$ where $\vec{v} = \langle b, c, d\rangle$ or just $w = (a, \vec{v})$. All of these notations are standard, common, convenient, and identical. At first the notation $a + \vec{v}$ may seem uncomfortable but since one is a scalar and the other is a vector it is not ambiguous as will rightly become apparent.

4.2. **General Properties and Terminology.** The task at hand is now to study how to operate on quaternions. For the following discussion allow $w, z \in \mathbb{H}$ where $w = a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ and $z = e + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$ with $a, b, c, d, e, f, g, h \in \mathbb{R}$.

For convenience also let $w = s_w + \vec{v}_w = (s_w, \vec{v}_w)$ and $z = s_z + \vec{v}_z = (s_z, \vec{v}_z)$. That is s_w is the scalar part of w, listed as a above, and $\vec{v}_w$ is the vector part, listed as $\langle b, c, d\rangle$.

4.2.1. *Addition of Quaternions.* Simply add w and z and group the like components.

$$\begin{aligned} w + z &= (a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) + (e + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}) \\ &= (a + e) + (b + f)\mathbf{i} + (c + g)\mathbf{j} + (d + h)\mathbf{k} \\ w + z &= s_w + s_z + \vec{v}_w + \vec{v}_z \\ w + z &= (s_w + s_z, \vec{v}_w + \vec{v}_z) \end{aligned}$$

Thus when adding two quaternions simply add the scalar parts and then add the vector parts.

4.3. **Multiplication of a Quaternion and a Scalar.** Given a scalar $d \in \mathbb{R}$ it is easy to show that $dw = wd = (ds_w, d\vec{v}_w)$. That is multiplication by a scalar just multiplies through to each component.

4.4. **Subtraction of Quaternions.** Subtraction is straightforward by noting that $z - w = z + (-1)w = (s_z, \vec{v}_z) + (-1)(s_w, \vec{v}_w) = (s_z, \vec{v}_z) + (-s_w, -\vec{v}_w) = (s_z - s_w, \vec{v}_z - \vec{v}_w)$.

4.5. **Multiplication of Quaternions.** Multiply w and z and group the like components.

$$\begin{aligned} wz &= (a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})(e + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}) \\ &= ae + af\mathbf{i} + ag\mathbf{j} + ah\mathbf{k} + be\mathbf{i} + bf\mathbf{i}^2 + bg\mathbf{ij} + bh\mathbf{ik} \\ &\quad + ce\mathbf{j} + cf\mathbf{ji} + cg\mathbf{j}^2 + ch\mathbf{jk} + de\mathbf{k} + df\mathbf{ki} + dg\mathbf{kj} + dh\mathbf{k}^2 \end{aligned}$$

Recall that $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$, $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$, $\mathbf{jk} = -\mathbf{kj} = \mathbf{i}$, and $\mathbf{ki} = -\mathbf{ik} = \mathbf{j}$.

$$\begin{aligned} wz &= ae + af\mathbf{i} + ag\mathbf{j} + ah\mathbf{k} + be\mathbf{i} - bf + bg\mathbf{k} - bh\mathbf{j} \\ &\quad + ce\mathbf{j} - cf\mathbf{k} - cg + ch\mathbf{i} + de\mathbf{k} + df\mathbf{j} - dg\mathbf{i} - dh \\ &= (ae - bf - cg - dh) + (af + be + ch - dg)\mathbf{i} \\ &\quad + (ag - bh + ce + df)\mathbf{j} + (ah + bg - cf + de)\mathbf{k} \\ &= (ae) - (bf + cg + dh) + a(f\mathbf{i} + g\mathbf{j} + h\mathbf{k}) + e(b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \\ &\quad + [(cg - dg)\mathbf{i} + (df - bh)\mathbf{j} + (bg - cf)\mathbf{k}] \\ wz &= (s_w s_z - \vec{v}_w \cdot \vec{v}_z, s_w \vec{v}_z + s_z \vec{v}_w + \vec{v}_w \times \vec{v}_z) \end{aligned}$$

Thus the product of two quaternions is

$$wz = s_w s_z - \vec{v}_w \cdot \vec{v}_z + s_w \vec{v}_z + s_z \vec{v}_w + \vec{v}_w \times \vec{v}_z.$$

For two **pure quaternions**, zero scalar component, this forms a new type of vector multiplication.

$$\vec{u}\vec{v} = (0, \vec{u})(0, \vec{v}) = (-\vec{u} \cdot \vec{v}, \vec{u} \times \vec{v}) = -\vec{u} \cdot \vec{v} + \vec{u} \times \vec{v}$$

This also produces two useful identities (which should only be taken in the context of quaternions).

$$\begin{aligned} \vec{u}\vec{v} + \vec{v}\vec{u} &= -2\vec{u} \cdot \vec{v} \\ \vec{u}\vec{v} - \vec{v}\vec{u} &= 2\vec{u} \times \vec{v} \end{aligned}$$

4.6. **Hypercomplex Conjugate of a Quaternion.** The hypercomplex conjugate of a quaternion w, denoted by w^* or $\overline{w}$, is defined as

$$w^* \overset{\text{def}}{=} (s_w, -\vec{v}_w).$$

4.7. **Norm of a Quaternion.** The norm, magnitude, or length of a quaternion w is defined as

$$||w|| \overset{\text{def}}{=} \sqrt{a^2 + b^2 + c^2 + d^2} = \sqrt{s_w^2 + \vec{v}_w \cdot \vec{v}_w} = \sqrt{ww^*}.$$

A quaternion with unit norm (magnitude 1) is called a **unit quaternion**.

4.8. **Dot Product of Quaternions.** The dot, scalar, or inner product is defined as

$$w \cdot z \overset{\text{def}}{=} ae + bf + cg + dh = s_w s_z + \vec{v}_w \cdot \vec{v}_z.$$

This gives two more identities (the second of which is familiar from vectors).

$$w \cdot z = wz^* + w^*z$$

$$||w|| = \sqrt{w \cdot w}$$

4.9. **Angle Between Quaternions.** The angle θ between two quaternions w and z is defined by

$$\cos\theta \stackrel{\text{def}}{=} \frac{w \cdot z}{||w||\,||z||}.$$

4.10. **Inverse of a Quaternion.** For a quaternion w, $ww^* = ||w||^2$ thus $\frac{ww^*}{||w||^2} = 1$. Hence multiplying w by $\frac{1}{||w||^2}w^*$ yields 1. The inverse of w is the quaternion w^{-1} such that $ww^{-1} = w^{-1}w = 1$. Thus, for $||w|| \neq 0$ (which is true for all quaternions except the zero quaternion)

$$w^{-1} = \frac{1}{||w||^2}w^*.$$

For a unit quaternion $w^{-1} = w^*$.

5. Unit Quaternions

For the following let w and z be defined as above but require them to both have unit norm. So $||w||^2 = a^2 + b^2 + c^2 + d^2 = 1$ and similarly for z.

Since a is a real number, $-1 \leq a \leq 1$. Without loss of generality create a variable θ such that $a = \cos\theta$.[7]

Thus a unit quaternion w may be written $w = \cos\theta + \vec{v}$ where $\vec{v} = \langle b, c, d\rangle$. Then $||w||^2 = \cos\theta\cos\theta + \vec{v}\cdot\vec{v} = \cos^2\theta + ||\vec{v}||^2$. Because $||w||^2 = 1$ it follows that $||\vec{v}||^2 = 1 - \cos^2\theta = \sin^2\theta$.

Now assume that $\sin\theta \neq 0$, in which case the quaternion would be $w = 1$ anyway. Then write $\vec{v} = \sin\theta\vec{u}$ where $\vec{u} = \left\langle \frac{b}{\sin\theta}, \frac{c}{\sin\theta}, \frac{d}{\sin\theta}\right\rangle$ and then $||\vec{u}|| = 1$ since $||\vec{v}|| = \sin\theta$.

Thus any unit quaternion w may be written as $w = \cos\theta + \hat{u}\sin\theta$ where $\hat{u}$ is a unit vector (in 3-space) and θ is in $[0, 2\pi)$.[8] Notice that this appears as an analog to Euler's Formula. This urges the inspection of $e^{\hat{u}\theta}$. The second power of $\hat{u}$ (via the new quaternion-based vector multiplication) is $\hat{u}\hat{u} = -\hat{u}\cdot\hat{u} + \hat{u}\times\hat{u} = -||\hat{u}||^2 = -1$. Then using a series expansion as before

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

Like complex numbers, any quaternion (not necessarily normal) can be written as

$$w = re^{\hat{u}\theta}.$$

5.1. **Powers of Quaternions.** With the above expression for the exponential of $\hat{u}\theta$ general exponentials and logarithms can be found.

$$w^t = (re^{\vec{u}\theta})^t = r^t e^{\vec{u}t\theta} = r^t[\cos(t\theta) + \hat{u}\sin(t\theta)]$$

$$\ln(w) = \ln(r[\cos\theta + \hat{u}\sin\theta]) = \ln(re^{\hat{u}\theta}) = \ln r + \hat{u}\theta.$$

However quaternions do not commute and so many identities involving exponentials and logarithms fail.

[7] This change of variable is well-defined due to the range of a.

[8] Writing a unit quaternion as $w = (\cos\theta, \hat{u}\sin\theta)$ is called **versor form**.

5.2. **Unit Quaternions as Rotations.** When writing $w = \cos\theta + \hat{u}\sin\theta$ the convention is to replace θ by $\frac{\theta}{2}$. Notice that in doing so none of the argument above is changed. Hence any **unit** quaternion may be written in the form

$$w = \cos\frac{\theta}{2} + \hat{u}\sin\frac{\theta}{2}.$$

The reason for doing this will become clear shortly.

A very natural construction to study at this point would be qpq^{-1} for two quaternions p and q. If quaternions commuted then this would be p but since it is not, such an object needs closer inspection. Let $q = \left(\cos\frac{\alpha}{2} + \hat{a}\sin\frac{\alpha}{2}\right)$ be a unit quaternion and examine the special case where p is a pure quaternion (not necessarily a unit quaternion), that is $p = (0, \vec{b}) = \vec{b}$, a standard 3-dimensional vector.

$$\begin{aligned}
qpq^{-1} &= qpq^* \\
&= \left(\cos\frac{\alpha}{2} + \hat{a}\sin\frac{\alpha}{2}\right)\vec{b}\left(\cos\frac{\alpha}{2} + \hat{a}\sin\frac{\alpha}{2}\right)^* \\
&= \left(\cos\frac{\alpha}{2} + \hat{a}\sin\frac{\alpha}{2}\right)\vec{b}\left(\cos\frac{\alpha}{2} - \hat{a}\sin\frac{\alpha}{2}\right) \\
&= \left(\vec{b}\cos\frac{\alpha}{2} + \hat{a}\vec{b}\sin\frac{\alpha}{2}\right)\left(\cos\frac{\alpha}{2} - \hat{a}\sin\frac{\alpha}{2}\right) \\
&= \vec{b}\cos^2\frac{\alpha}{2} - \vec{b}\hat{a}\cos\frac{\alpha}{2}\sin\frac{\alpha}{2} + \hat{a}\vec{b}\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} - \hat{a}\vec{b}\hat{a}\sin^2\frac{\alpha}{2} \\
&= \vec{b}\cos^2\frac{\alpha}{2} + (\hat{a}\vec{b} - \vec{b}\hat{a})\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} - \hat{a}\vec{b}\hat{a}\sin^2\frac{\alpha}{2} \\
&= \vec{b}\cos^2\frac{\alpha}{2} + 2(\hat{a}\times\vec{b})\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} - \left(\vec{b}(\hat{a}\cdot\hat{a}) - 2\hat{a}(\hat{a}\cdot\vec{b})\right)\sin^2\frac{\alpha}{2} \\
&= \vec{b}\left(\cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2}\right) + (\hat{a}\times\vec{b})2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} + \hat{a}(\hat{a}\cdot\vec{b})\left(2\sin^2\frac{\alpha}{2}\right) \\
&= \vec{b}\cos\alpha + (\hat{a}\times\vec{b})\sin\alpha + \hat{a}(\hat{a}\cdot\vec{b})(1-\cos\alpha) \\
qpq^{-1} &= (1-\cos\alpha)(\hat{a}\cdot\vec{b})\hat{a} + \vec{b}\cos\alpha + (\hat{a}\times\vec{b})\sin\alpha
\end{aligned}$$

This is Rodrigues' Rotation Formula.

Thus given an axis $\hat{a}$, angle α ,and point $\vec{b}$, construct the two quaternions $q = \left(\cos\frac{\alpha}{2}, \hat{a}\sin\frac{\alpha}{2}\right)$ and $p = (0, \vec{b})$. Then qpq^{-1} is $\vec{b}$ rotated around $\hat{a}$ by α.

To rotate this point again around a new axis $\hat{n}$ by an angle ϕ, construct the quaternion $w = \left(\cos\frac{\phi}{2} + \hat{n}\sin\frac{\phi}{2}\right)$ and then compute $w(qpq^{-1})w^{-1} = (wq)p(q^{-1}w^{-1}) = (wq)p(wq)^{-1}$. Thus the quaternion representing the total rotation is just wq which can be precomputed and then inverted once for the entire cumulative rotation.

5.3. **Quaternions and Axis/Angle Rotation.** Given an axis $\hat{a} = \langle x, y, z\rangle$ and an angle θ the general matrix for rotation around $\hat{a}$ by θ is

$$\begin{pmatrix}
(1-c)x^2 + c & (1-c)xy + sz & (1-c)xz - sy & 0 \\
(1-c)xy - sz & (1-c)y^2 + c & (1-c)yz + sx & 0 \\
(1-c)xz + sy & (1-c)yz - sx & (1-c)z^2 + c & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}$$

where $c = \cos\theta$ and $s = \sin\theta$.

For a unit quaternion $q = (s, v_x, v_y, v_z) = \left(\cos\frac{\theta}{2}, x\sin\frac{\theta}{2}, y\sin\frac{\theta}{2}, z\sin\frac{\theta}{2}\right)$ the corresponding matrix would be

$$\begin{pmatrix} s^2+v_x^2-v_y^2-v_z^2 & 2v_xv_y+2sv_z & 2v_xv_z-2sv_y & 0 \\ 2v_xv_y-2sv_z & s^2-v_x^2+v_y^2-v_z^2 & 2v_yv_z+2sv_x & 0 \\ 2v_xv_z+2sv_y & 2v_yv_z-2sv_x & s^2-v_x^2-v_y^2+v_z^2 & 0 \\ 0 & 0 & 0 & s^2+v_x^2+v_y^2+v_z^2 \end{pmatrix}.$$

The two matrices are identical. To demonstrate this the first two elements in the top row will be checked; the other elements would follow similarly.

The upper-leftmost element:

$$\begin{aligned} s^2+v_x^2-v_y^2-v_z^2 &= \cos^2\frac{\theta}{2} + x^2\sin^2\frac{\theta}{2} - y^2\sin^2\frac{\theta}{2} - z^2\sin^2\frac{\theta}{2} \\ &= \cos^2\frac{\theta}{2} + (x^2-y^2-z^2)\sin^2\frac{\theta}{2} \\ &= \cos^2\frac{\theta}{2} + (2x^2-x^2-y^2-z^2)\sin^2\frac{\theta}{2} \\ &= \cos^2\frac{\theta}{2} + (2x^2-1)\sin^2\frac{\theta}{2} \\ &= \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} + 2x^2\sin^2\frac{\theta}{2} \\ &= \cos\theta + 2x^2\left(\frac{1-\cos\theta}{2}\right) \\ &= \cos\theta + x^2(1-\cos\theta) \end{aligned}$$

The second element in the top row:

$$\begin{aligned} 2v_xv_y+2sv_z &= 2x\sin\frac{\theta}{2}y\sin\frac{\theta}{2} + 2\cos\frac{\theta}{2}z\sin\frac{\theta}{2} \\ &= 2xy\sin^2\frac{\theta}{2} + z\sin\theta = (1-\cos\theta)xy + z\sin\theta \end{aligned}$$

5.4. **Axis/Angle Rotation and Quaternions.** Given a general rotation matrix

$$\begin{pmatrix} a & b & c & 0 \\ d & e & f & 0 \\ g & h & i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

it is desirable to construct the equivalent quaternion $q = (s, x, y, z)$. The first element s will be done as follows. Note from the previous section that

$$\begin{aligned} a+e+i+1 = {} & (s^2+v_x^2-v_y^2-v_z^2) + (s^2-v_x^2+v_y^2-v_z^2) \\ & + (s^2-v_x^2-v_y^2+v_z^2) + (s^2+v_x^2+v_y^2+v_z^2) = 4s^2. \end{aligned}$$

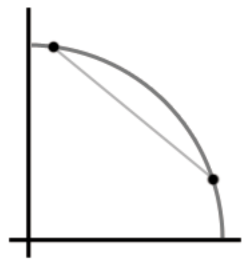
Therefore $s = \frac{1}{4}(a+e+i+1)$. Now find x, y, and z. Notice that $f - h = (2v_yv_z + 2sv_x) - (2v_yv_z - 2sv_x) = 2sv_x$. Therefore $v_x = \frac{1}{2s}(f-h) = \frac{2}{a+e+i+1}(f-h)$. Similarly it follows that $v_y = \frac{2}{a+e+i+1}(g-c)$ and $v_z = \frac{2}{a+e+i+1}(b-d)$.

6. Slerp

6.1. **Lerp.** Quaternions are ideal for many reasons. The fact that rotations can be combined with simple multiplication is one of the most quoted benefits. Another is that interpolation is simple and has many desirable properties. Given two quaternions q and p it is often necessary to interpolate between the two over time. The simplest approach would be LERP (linear interpolation).

$$r(t) = (1-t)q + tp$$

This is not desirable. Look at the image below (a cross section of the 4-dimensional sphere on which quaternions live).

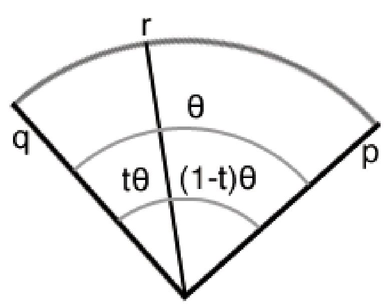


When linear interpolation is used the interpolated quaternion will duck below the unit 4-sphere causing undesirable behavior. The resultant quaternion could be normalized each moment but that would cause the interpolation to speed up in the middle. The solution to this problem is SLERP or spherical linear interpolation.

6.2. **Slerp.** Any interpolation of two quaternions p and q would have the form

$$r(t) = a(t)q + b(t)p.$$

Now find the desired functions $a(t)$ and $b(t)$. Recall that $q \cdot p = \cos\theta$; the dot product of two unit quaternions is the angle between them. The following figure may help visualize the angles.



Take the above equation for $r(t)$ and take the dot product of q with both sides.

$$q \cdot r(t) = q \cdot [a(t)q + b(t)p]$$

$$\cos(t\theta) = a(t) + b(t)\cos(\theta) \tag{6.1}$$

Then do the same with p.

$$p \cdot r(t) = p \cdot [a(t)q + b(t)p]$$
$$\cos[(1-t)\theta] = a(t)\cos(\theta) + b(t) \tag{6.2}$$

Now solve 6.2 for $b(t)$, which is $b(t) = \cos[(1-t)\theta] - a(t)\cos(\theta)$, and substitute it into 6.1.

$$\begin{aligned}
\cos(t\theta) &= a(t) + b(t)\cos(\theta) \\
&= a(t) + [\cos[(1-t)\theta] - a(t)\cos(\theta)]\cos\theta \\
&= a(t)\left[1 - \cos^2\theta\right] + \cos(\theta - t\theta)\cos\theta \\
&= a(t)\sin^2\theta + [\cos\theta\cos(t\theta) + \sin\theta\sin(t\theta)]\cos\theta \\
\cos(t\theta)[1-\cos^2\theta] &= a(t)\sin^2\theta + \sin\theta\cos\theta\sin(t\theta) \\
\cos(t\theta)\sin^2\theta &= a(t)\sin^2\theta + \sin\theta\cos\theta\sin(t\theta) \\
\cos(t\theta)\sin\theta &= a(t)\sin\theta + \cos\theta\sin(t\theta) \\
a(t)\sin\theta &= \cos(t\theta)\sin\theta - \cos\theta\sin(t\theta) \\
a(t)\sin\theta &= \sin(\theta - t\theta) \\
a(t) &= \frac{\sin[(1-t)\theta]}{\sin\theta}
\end{aligned}$$

Similar manipulation can be used to solve for $b(t)$.

$$b(t) = \frac{\sin(t\theta)}{\sin(\theta)}$$

Thus putting all the above together yields the expression for Slerp.

$$\mathrm{Slerp}(q,p;t) = \frac{\sin[(1-t)\theta]q + \sin(t\theta)p}{\sin(\theta)}$$

6.3. **Slerp Again.** Another way to derive Slerp is to interpolate between two orthogonal quaternions.

$$\mathrm{Slerp}(q,p;t) = a\cos(t\theta) + b\sin(t\theta)$$

after finding good choices for a and b. There is no reason not to chose $a = q$, so do so. Then the choice for b must be orthogonal to q. Just like with vectors, an orthogonal quaternion can be constructed from the orthogonal projection of p onto q. That is, set $b = p - q\cos\theta$, which is not necessarily normalized, so let $b = \frac{p - q\cos\theta}{||p - q\cos\theta||}$.

The denominator of b can be simplified.

$$\begin{aligned}
||p - q\cos\theta||^2 &= (p - q\cos\theta)(p - q\cos\theta)^* \\
&= (p - q\cos\theta)(p^* - q^*\cos\theta) \\
&= pp^* + qq^*\cos^2\theta - (pq^* + qp^*)\cos\theta
\end{aligned}$$

Recall that $pq^* + qp^* = 2\cos\theta$.

$$||p - q\cos\theta||^2 = 1 + \cos^2\theta - 2\cos^2\theta = 1 - \cos^2\theta = \sin^2\theta \tag{6.3}$$

Thus Slerp can be found as follows.

$$\begin{aligned}
\mathrm{Slerp}(q,p;t) &= q\cos(t\theta) + \frac{(p - q\cos\theta)}{\sin\theta}\sin(t\theta) \\
&= \frac{q\cos(t\theta)\sin\theta}{\sin\theta} + \frac{p\sin(t\theta) - q\cos\theta\sin(t\theta)}{\sin\theta} \\
&= \frac{q\cos(t\theta)\sin\theta + p\sin(t\theta) - q\cos\theta\sin(t\theta)}{\sin\theta} \\
&= \frac{q[\cos(t\theta)\sin\theta - \cos\theta\sin(t\theta)] + p\sin(t\theta)}{\sin\theta} \\
&= \frac{q\sin[\theta - t\theta] + p\sin(t\theta)}{\sin\theta} = \frac{q\sin[(1-t)\theta] + p\sin(t\theta)}{\sin\theta}
\end{aligned}$$

6.4. **Normalized Slerp.** Recall the previous expression for Slerp.

$$\mathrm{Slerp}(q,p;t) = a\cos(t\theta) + b\sin(t\theta)$$

where $a \cdot b = 0$ (they are orthogonal). Showing that Slerp is normalized follows.

$$\begin{aligned}
||\mathrm{Slerp}(q,p;t)||^2 &= [a\cos(t\theta) + b\sin(t\theta)][a\cos(t\theta) + b\sin(t\theta)]^* \\
&= [a\cos(t\theta) + b\sin(t\theta)][a^*\cos(t\theta) + b^*\sin(t\theta)] \\
&= aa^*\cos^2(t\theta) + bb^*\sin^2(t\theta) + (ab^* + ba^*)\sin(t\theta)\cos(t\theta) \\
&= \cos^2(t\theta) + \sin^2(t\theta) + (0)\sin(t\theta)\cos(t\theta) = 1
\end{aligned}$$

6.5. **Properties of Slerp.** There are many equivalent forms of Slerp (proof omitted).

$$\begin{aligned}
\mathrm{Slerp}(p,q;t) &= p(p^{-1}q)^t \\
&= q(q^{-1}p)^{1-t} \\
&= (pq^{-1})^{1-t}q \\
&= (qp^{-1})^t p \\
&= \frac{q\sin[(1-t)\theta] + p\sin(t\theta)}{\sin\theta}
\end{aligned}$$

6.5.1. *Angular Velocity of Slerp.* In the second derivation of Slerp above it was shown that

$$\mathrm{Slerp}(p,q;t) = a\cos(t\theta) + b\sin(t\theta)$$

for two time-independent quaternions p and q and with $\cos\theta = p \cdot q$.

$$\begin{aligned}
\frac{d}{dt}\mathrm{Slerp}(p,q;t) &= \frac{d}{dt}\left(a\cos(t\theta) + b\sin(t\theta)\right) \\
&= -a\theta\sin(t\theta) + b\theta\cos(t\theta) \\
\frac{d^2}{dt^2}\mathrm{Slerp}(p,q;t) &= \frac{d}{dt}\left(-a\theta\sin(t\theta) + b\theta\cos(t\theta)\right) \\
&= -a\theta^2\cos(t\theta) - b\theta^2\sin(t\theta) \\
&= -\theta^2[a\cos(t\theta) + b\sin(t\theta)] \\
&= -\theta^2\mathrm{Slerp}(p,q;t)
\end{aligned}$$

Since $\theta^2 \geq 0$, $-\theta^2 \leq 0$. Thus $\frac{d^2}{dt^2}\text{Slerp}(p, q; t) = \gamma\text{Slerp}(p, q; t)$ where $\gamma \in \mathbb{R}$ and $\gamma < 0$. A system where the second derivative (acceleration) is antiparallel to the position is a circle with constant velocity.[9]

Another way to see this is that the angle Ω between $\text{Slerp}(p, q; t)$ and $p(t)$ is $t\theta$. Thus $\frac{d\Omega}{dt} = \theta$ and so the angular velocity of the interpolation is constant.

Thus on the unit 4-sphere Slerp has constant angular velocity (hence is over a circular arc) and is normalized (magnitude is one throughout). These two features show that the interpolation is a great arc.[10]

The only issue left is that given two points on a sphere there are two great arcs connecting them. One is a half circle or smaller and the other is a half circle or larger. Slerp interpolates over the shorter of the two arcs but the proof is omitted.

Recap: Slerp interpolates between two quaternions, thus two orientations, over the shortest path possible in 4-space with constant angular velocity. It is a degree-one interpolation (in spherical geometry) between two quaternions and can be extended to produce higher degree interpolations as is done with splines. For instance, given quaternions p, q, r, construct

$$\text{QuadSlerp}(p, q, r; t) = \text{Slerp}\left(\text{Slerp}(p, q; t), \text{Slerp}(q, r; t); t\right)$$

and build up higher oder interpolations as needed. Note that QuadSlerp is normalized but that it does not have constant angular velocity since the two interpolated quaternions are time-dependent.[11]

[9] $f''(t) + \gamma f(t) = 0$ has the solutions $f(t) = re^{\pm \mathbf{i} t\theta}$ which are circles with constant angular velocity θ.

[10] A closed great arc (called a great circle) is a path along a sphere that splits it into two equal halves; airplanes try to fly as close to great arcs as possible since they minimize distance.

[11] This is not unexpected or unfortunate. A line has constant linear velocity, but a parabola does not.