An arbitrary rotation without matrices

Keith Pedersen

NOTE: This paper uses the nomenclature that $\boldsymbol{w}$ is a vector and $\hat{\boldsymbol{w}}$ is a unit vector.

A right-handed (RH) rotation, by some angle ψ about some axis $\hat{\boldsymbol{x}}$, can be defined without matrices using Rodrigues' rotation formula. The victim vector $\boldsymbol{w}$ is projected into three pieces: ($\boldsymbol{w}_\parallel$) the piece parallel to the axis, ($\boldsymbol{w}_\perp$) the piece perpendicular to the axis, and ($\boldsymbol{w}'_\perp$), the perpendicular $\boldsymbol{w}_\perp$ rotated by 90°. The parallel piece is unaltered by the rotation, and the two $\boldsymbol{w}_\perp$ form a basis for the rotation;

$$\boldsymbol{w}' = R(\boldsymbol{w}) = \underbrace{(\boldsymbol{w}\cdot\hat{\boldsymbol{x}})\hat{\boldsymbol{x}}}_{\boldsymbol{w}_\parallel} + \cos\psi\underbrace{(\boldsymbol{w} - \boldsymbol{w}_\parallel)}_{\boldsymbol{w}_\perp} + \sin\psi\underbrace{(\hat{\boldsymbol{x}}\times\boldsymbol{w})}_{\boldsymbol{w}'_\perp}\,. \tag{1}$$

We can validate this scheme by showing it preserves the defining properties of rotations. First, the angle to the axis is unaltered;

$$\begin{aligned}\boldsymbol{w}'\cdot\hat{\boldsymbol{x}} &= (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})\cancelto{1}{\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{x}}} + \cos\psi(\boldsymbol{w}\cdot\hat{\boldsymbol{x}} - (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})\cancelto{1}{\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{x}}}) + \sin\psi\cancelto{0}{((\hat{\boldsymbol{x}}\times\boldsymbol{w})\cdot\hat{\boldsymbol{x}})} \\ &= \boldsymbol{w}\cdot\hat{\boldsymbol{x}}\end{aligned} \tag{2}$$

(which uses $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a}) = \boldsymbol{c}\cdot(\boldsymbol{a}\times\boldsymbol{b})$). Second, the vector's length is unaltered;

$$\begin{aligned}|\boldsymbol{w}'|^2 &= (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})^2\cancelto{1}{\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{x}}} + \cos^2\psi\,(|\boldsymbol{w}|^2 - 2(\boldsymbol{w}\cdot\hat{\boldsymbol{x}})^2 + (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})^2) + \sin^2\psi\,|\boldsymbol{w}|^2(|\boldsymbol{w}|^2 - (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})^2) \\ &= (\boldsymbol{w}\cdot\hat{\boldsymbol{x}}) + (\sin^2\psi + \cos^2\psi)(|\boldsymbol{w}|^2 - (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})^2) \\ &= |\boldsymbol{w}|^2\end{aligned} \tag{3}$$

Here, we have used the identity $|\boldsymbol{u}\times\boldsymbol{v}|^2 = |\boldsymbol{u}|^2|\boldsymbol{v}|^2 - (\boldsymbol{u}\cdot\boldsymbol{v})^2$ (from $\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$), and implicitly used the orthogonality of the each component vector

$$\boldsymbol{w}_\parallel\cdot\boldsymbol{w}_\perp = (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})^2 - (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})^2 = 0\,, \tag{4}$$

$$\boldsymbol{w}_\perp\cdot\boldsymbol{w}'_\perp = \cancelto{0}{\boldsymbol{w}\cdot(\hat{\boldsymbol{x}}\times\boldsymbol{w})} - (\boldsymbol{w}\cdot\hat{\boldsymbol{x}})\,\cancelto{0}{\hat{\boldsymbol{x}}\cdot(\hat{\boldsymbol{x}}\times\boldsymbol{w})} = 0\,. \tag{5}$$

Instead of an axis $\hat{\boldsymbol{x}}$ and angle ψ as our input degrees of freedom, we may want the rotation that takes vector $\boldsymbol{u} \to \boldsymbol{v}$. We can reuse Eq. 1, and define:

$$\cos\psi = \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{v}} \tag{6}$$

$$\sin\psi = |\boldsymbol{u} \times \boldsymbol{v}| \tag{7}$$

$$\hat{\boldsymbol{x}}_1 = \frac{\hat{\boldsymbol{u}} \times \hat{\boldsymbol{v}}}{|\hat{\boldsymbol{u}} \times \hat{\boldsymbol{v}}|} = \frac{\hat{\boldsymbol{u}} \times \hat{\boldsymbol{v}}}{\sin\psi} \tag{8}$$

However, $\hat{\boldsymbol{x}}_1$ is only one *possible* rotation axis which takes $\boldsymbol{u} \to \boldsymbol{v}$. Consider a rotation of $\psi = \pi$ about the axis bisecting the two normalized vectors

$$\hat{\boldsymbol{x}}_2 = \frac{\hat{\boldsymbol{u}} + \hat{\boldsymbol{v}}}{|\hat{\boldsymbol{u}} + \hat{\boldsymbol{v}}|} = \frac{\hat{\boldsymbol{u}} + \hat{\boldsymbol{v}}}{\sqrt{2(1 + \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{v}})}}\,. \tag{9}$$

This rotation also clearly takes $\boldsymbol{u} \to \boldsymbol{v}$. In fact, any axis which satisfies

$$\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{x}} = \hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{x}} \tag{10}$$

defines a valid rotation (since $\boldsymbol{u}$ and $\boldsymbol{v}$ will have the same latitude relative to $\hat{\boldsymbol{x}}$, and thus trace out the same circle during the rotation). Thus, "the rotation which takes $\boldsymbol{u} \to \boldsymbol{v}$" is ambiguous.

This ambiguity stems from a hidden degree of freedom. We are free to choose (x, y, z) vectors

$$\boldsymbol{u} = (0,\, 0,\, 1) \tag{11}$$

$$\boldsymbol{v} = (\cos\phi\sin\theta,\, \sin\phi\sin\theta,\, \cos\theta) \tag{12}$$

We can get from $\boldsymbol{u}$ to $\boldsymbol{v}$ in two steps: (i) A RH rotation about $\hat{\boldsymbol{z}}$ by angle ϕ. (ii) A RH rotation about $\hat{\boldsymbol{y}}'$ (the new y-axis, after the first rotation) by angle θ. This procedure uses only two of the three Euler angles required to cover SO(3) (the group of 3-dimensional rotations). To complete the coverage, we need a final RH rotation about $\hat{\boldsymbol{z}}''$ (the final z-axis, which in our case is the newly minted $\boldsymbol{v}$).

By construction, this post-rotation about $\boldsymbol{v}$ by angle ω cannot alter $\boldsymbol{v}$, so it does not spoil the original purpose of this rotation (take $\boldsymbol{u} \to \boldsymbol{v}$). Instead, ω determines what happens to *every other* vector, and does so by selecting *one* axis $\hat{\boldsymbol{x}}$ from the set which map $\boldsymbol{u} \to \boldsymbol{v}$. If we can find this $\hat{\boldsymbol{x}}$, we can describe the complete operation as a single rotation (instead of two sequential rotations). We have already found two valid $\hat{\boldsymbol{x}}$ (Eqs. 8 and 9), and they are fortuitously orthogonal, so we can use them to construct a basis that parameterizes all possible axes of rotation

$$\hat{\boldsymbol{x}} = a\,\hat{\boldsymbol{x}}_1 + b\,\hat{\boldsymbol{x}}_2 \tag{13}$$

Our task is now clear; given $\boldsymbol{u}$, $\boldsymbol{v}$ and ω, determine a and b to find $\hat{\boldsymbol{x}}$.

R_1 is the rotation about $\hat{\boldsymbol{x}}_1$ by $\theta = \arccos(\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{v}})$ and R_2 is the rotation about $\hat{\boldsymbol{x}}_2$ by ω. Applying them consecutively produces a composite rotation which cannot alter the axis of rotation;

$$\hat{\boldsymbol{x}} = R_2(R_1(\hat{\boldsymbol{x}}))\,. \tag{14}$$

Since this defining property applies to both of $\hat{\boldsymbol{x}}$'s component individually (and linearly), we can solve for a and b by using unitarity as one equation (i.e. $a^2 + b^2 = 1$), then obtain the other equation by calculating

$$\begin{aligned} a &= R_2(R_1(a\,\hat{\boldsymbol{x}}_1 + b\,\hat{\boldsymbol{x}}_2)) \cdot \hat{\boldsymbol{x}}_1 \\ &= a\,R_2(R_1(\hat{\boldsymbol{x}}_1)) \cdot \hat{\boldsymbol{x}}_1 + b\,R_2(R_1(\hat{\boldsymbol{x}}_2)) \cdot \hat{\boldsymbol{x}}_1\,. \end{aligned} \tag{15}$$

Beginning with $\hat{\boldsymbol{x}}_1$, we are lucky that R_1 does not alter it's own axis, while R_2 creates only one term parallel to $\hat{\boldsymbol{x}}_1$;

$$R_1(\hat{\boldsymbol{x}}_1) = \hat{\boldsymbol{x}}_1\ ; \tag{16}$$

$$R_2(R_1(\hat{\boldsymbol{x}}_1)) \cdot \hat{\boldsymbol{x}}_1 = (\cancelto{0}{(\hat{\boldsymbol{x}}_1 \cdot \hat{\boldsymbol{v}})}\hat{\boldsymbol{v}} + \cos\omega(\hat{\boldsymbol{x}}_1 - 0) + \sin\omega\underbrace{(\hat{\boldsymbol{v}} \times \hat{\boldsymbol{x}}_1)}_{\perp\text{ to }\hat{\boldsymbol{x}}_1}) \cdot \hat{\boldsymbol{x}}_1 = \cos\omega \tag{17}$$

The effect on $\hat{\boldsymbol{x}}_2$ is slightly more complicated;

$$\begin{aligned} R_1(\hat{\boldsymbol{x}}_2) &= (\cancelto{0}{(\hat{\boldsymbol{x}}_2 \cdot \hat{\boldsymbol{x}}_1)}\hat{\boldsymbol{x}}_1 + \cos\theta(\hat{\boldsymbol{x}}_2 - 0) + \sin\theta(\hat{\boldsymbol{x}}_1 \times \hat{\boldsymbol{x}}_2) \\ &= \frac{1}{\sqrt{2(1+\cos(\theta))}}\left(\cos\theta\,(\hat{\boldsymbol{u}} + \hat{\boldsymbol{v}}) + \sin\theta\,\frac{(\hat{\boldsymbol{u}} \times \hat{\boldsymbol{v}})}{\sin\theta} \times (\hat{\boldsymbol{u}} + \hat{\boldsymbol{v}})\right) \\ &= \frac{1}{\sqrt{2(1+\cos(\theta))}}\left(\cos\theta\,(\hat{\boldsymbol{u}} + \hat{\boldsymbol{v}}) + \hat{\boldsymbol{v}} - \hat{\boldsymbol{u}}\cos\theta + \hat{\boldsymbol{v}}\cos\theta - \hat{\boldsymbol{u}}\right) \\ &= \frac{1}{\sqrt{2(1+\cos(\theta))}}\left((2\cos\theta + 1)\hat{\boldsymbol{v}} - \hat{\boldsymbol{u}}\right) \end{aligned} \tag{18}$$

(using $(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = \boldsymbol{b}\,(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{a}\,(\boldsymbol{b} \cdot \boldsymbol{c})$). Before we put all of Eq. 18 through R_2, we should recall that the final equation only uses terms parallel to $\hat{\boldsymbol{x}}_1$. R_2 only returns terms which are (i) parallel to the incoming vector (which in this case is in the uv-plane, and thus orthogonal to $\hat{\boldsymbol{x}}_1$), (ii) parallel to $\hat{\boldsymbol{v}}$ (also in the uv plane) and (iii) perpendicular to $\hat{\boldsymbol{v}}$ (via $\hat{\boldsymbol{v}} \times \boldsymbol{w}$). Hence, only the 3$^{\text{rd}}$ piece is meaningful, and since $\hat{\boldsymbol{v}} \times \hat{\boldsymbol{v}} = 0$, we only need to give it Eq. 18's $\hat{\boldsymbol{u}}$ term;

$$
\begin{aligned}
R_2(R_1(\hat{\boldsymbol{x}}_2)) \cdot \hat{\boldsymbol{x}}_1 &= \frac{1}{\sqrt{2(1+\cos(\theta))}} R_2(-\hat{\boldsymbol{u}}) \cdot \hat{\boldsymbol{x}}_1 = \frac{1}{\sqrt{2(1+\cos(\theta))}} (\hat{\boldsymbol{v}} \times (-\hat{\boldsymbol{u}})) \cdot \hat{\boldsymbol{x}}_1 \\
&= \sin\omega \frac{\sin\theta}{\sqrt{2(1+\cos(\theta))}} \qquad (19)
\end{aligned}
$$

(using $\hat{\boldsymbol{u}} \times \hat{\boldsymbol{v}} = \sin\theta\, \hat{\boldsymbol{x}}_1$). Combining Eq. 15, 16 and 19 we get our system of equations

$$
a = a\cos\omega + b\sin\omega \frac{\sin\theta}{\sqrt{2(1+\cos(\theta))}} \qquad (20)
$$

$$
1 = a^2 + b^2 \qquad (21)
$$

Having done the hard work (see Appendix), we can plug our system of equations into Mathematica to obtain

$$
a = 2\cos(\omega/2)\frac{\sin(\theta/2)}{c(\theta)} \qquad (22)
$$

$$
b = 2\sin(\omega/2)\frac{1}{c(\theta)} \qquad (23)
$$

$$
c(\theta) = \sqrt{3 - \cos(\omega) - \cos(\theta)(1+\cos(\omega))} \qquad (24)
$$

I then used Mathematica to validate this solution by checking that $\hat{\boldsymbol{x}}$ matches the eigenvector of the composite rotation $R_2(R_1(\boldsymbol{w}))$. (up to the parity operation $\hat{\boldsymbol{x}} \to -\hat{\boldsymbol{x}}$, which is irreducibly ambiguous).

However, $c(\theta)$ is numerically unstable if used naïvely, due to the cosine cancellations. These terms should be rewritten in a form which *doubles* the precision of the floating point result

$$
1 - \cos(t) \mapsto 2\sin^2(t/2)\,. \qquad (25)
$$

This gives us

$$
c(\theta) \mapsto \sqrt{2\sin^2(\theta/2) + 2\sin^2(\omega/2) + (1 - \cos(\theta)\cos(\omega))}\,. \qquad (26)
$$

The final cancellation can be corrected using

$$
\cos(\theta)\cos(\omega) = \frac{1}{2}(\cos(\theta+\omega) + \cos(\theta-\omega))\,; \qquad (27)
$$

$$
(1 - \cos(\theta)\cos(\omega)) \mapsto \frac{1}{2}(1 - \cos(\theta+\omega)) + \frac{1}{2}(1 - \cos(\theta-\omega))\,. \qquad (28)
$$

Again using Eq. 25, we obtain the final expression

$$c(\theta) \mapsto \sqrt{2\sin^2(\theta/2) + 2\sin^2(\omega/2) + \sin^2(\theta+\omega) + \sin^2(\theta-\omega)}\,. \tag{29}$$

Given $\boldsymbol{u}$, $\boldsymbol{v}$ and ω, we now have the tools to find the axis $\hat{\boldsymbol{x}}$ about which the composite rotations occur.[1] But what is the angle ψ of rotation? We can determine ψ empirically by projecting $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ into the plane of rotation (e.g. $\hat{\boldsymbol{u}}_\perp = \hat{\boldsymbol{u}} - (\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{x}})\hat{\boldsymbol{x}}$). The rotation angle is then defined via

$$\psi = \mathrm{atan2}(\sin\psi, \cos\psi) = \mathrm{atan2}\left(\mathrm{sign}(a)|\hat{\boldsymbol{u}}_\perp \times \hat{\boldsymbol{v}}_\perp|,\ \hat{\boldsymbol{u}}_\perp \cdot \hat{\boldsymbol{v}}_\perp\right)\,. \tag{30}$$

It is best to use atan2 because it is more precise for angles near 0, $\pi/2$ and π. Note that we have to inject the *sign* of a into $\sin(\psi)$, because when $a < 0$, the RH rotation becomes larger than π, so we must instead uses a *negative* RH rotation.

I have tested an implementation of this algorithm and it works quite well (it is both length and angle preserving). I have additionally validated that rotating once about $\hat{\boldsymbol{x}}$ gives the same result as the two-step composite rotation.

Appendix: The apology

I see three steps in answering any question scientifically:

1. Write down a model which describes the system with sufficient accuracy.
2. Try and fail ad naseum until the Eureka moment allows you to *describe* the solution.
3. Find the solution.

The difference between step 2 and 3 is subtle, but important. A *description* of the solution is a complete statement of *where* the solution can be found, its GPS coordinates — it is a differential equation, an unevaluated integral, the polynomial whose roots we will find. Step 3 is the vehicle that takes us there — an algorithm, a clever change of variable, an analytic continuation. It is math, and frequently the heavy kind.

[1] There is one class of system where $\hat{\boldsymbol{x}}$ remains ambiguous; when $\boldsymbol{u}$ and $\boldsymbol{v}$ are antiparallel, $\hat{\boldsymbol{x}}_1$ and $\hat{\boldsymbol{x}}_2$ are both null. If $\boldsymbol{x}$ is supplied externally, ω rotates it around the shared uv axis, but $\boldsymbol{x}$ cannot be determined from ω alone.

In my opinion, competence and creativity in step 1 and 2 are the mark of a good scientist. This requires a vast array of knowledge and experience (e.g. you have to know what an eigenvalue is, and have previously used them to solve simple problems, before you can use them to describe the solution for some novel problem). But once we reach step 3 (calculate the eigenvalue), we should beg, borrow and steal. Of course, someone has to blaze the original trail when a new technique is found, and for that they earn their reverence in the pages of a textbook. But since they bled to build that trail, we should have the common decency to use it. Conversely, each of us should forge our own trail for the step 3 in which *we* are the experts, and then tell the world.

But we cannot be experts in the huge library of step 3's. Furthermore, since step 3's are often quite general, they lend themselves to automation. Computers are still not very good at answering *declarative* statements like "what is the volume of a sphere" (and here I don't mean using a search engine or a neural net to find a solution published by a human). However, Mathematica will swiftly provide the symbolic solution to a very *imperative* statement

$$V_{\text{sphere}} = \int_0^R \mathrm{d}r \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \, r^2 \sin(\theta) = \frac{4}{3}\pi \, R^3 \, .$$

So when I find myself with a system of equations which I can ask the computer to solve (nearly instantaneously), I use the computer — that's what computers are for.