Why is angular velocity a vector?

Andrew Guy*

In the Engineering Tripos IB "Vector Calculus & PDEs" course, students are asked why angular velocity is a vector despite finite rotations being obviously not. The lectures teach that it is a "pseudo-vector" due to it requiring a "sense" to work properly (the right-hand rule). Despite having studied vector calculus and having a good understanding of physical processes, students arrive at Part II modules (such as 3M1: Mathematical Methods) *not actually knowing what a vector is.*

In supervisions, I tend to explain it as "angular velocity is a vector because infinitesimal rotations are, and we can get to these from suitably small chunks of time".

Finite rotations as not-vectors

Encode a finite rotation as a 'vector' (it's not, we're just borrowing the colloquial word for a tuple) $v = u\theta$: let the axis u be a unit vector and the angle θ, and we'll use a right-handed convention. By considering the effect of the rotation on a vector in the axis direction, and on those in the plane perpendicular to it, you should be able to conclude that we may produce a rotation matrix $Q = f(v)$ as:

$$Q = (\cos\theta)I + (\sin\theta)[u]_\times + (1 - \cos\theta)uu^\intercal \tag{1}$$

where the notation $[u]_\times$ represents the "cross product matrix":

$$[u]_\times = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} \tag{2}$$

i.e., for any vectors u, v we have $[u]_\times v = u \times v$.

Infinitesimal rotations

Considering a small rotation we have $\cos\theta \approx 1$ and $\sin\theta \approx \theta$. So we are left with:

$$Q = I + \theta[u]_\times \tag{3}$$

and it may be helpful to think of the angle as $\theta\epsilon$, where ϵ is a nilpotent element such that $\epsilon^2 = 0$ — when we multiply things this will behave like we are used to when multiplying infinitesimals, where we argue that we may always make cross terms negligibly small by choosing a small enough length scale. This deals with the first potential point of confusion: that whilst we can choose arbitrarily large numbers to live within these matrices (and our "infinitesimal rotation vectors"

*aag43 (at) cam (dot) ac (dot) uk

will span the entirety of $\mathbb{R}^3$), the "order of realisation" matters. Since we realise the angular velocity vectors first, no matter how big they are we can always choose a small enough time step that our approximation holds to any desired tolerance[1].

Let's create two infinitesimal rotation matrices, $P = f(u\theta)$ and $Q = f(v\phi)$, and multiply them together, remembering that $\theta\phi = 0$:

$$\begin{aligned} PQ &= \begin{bmatrix} 1 & -u_z\theta & u_y\theta \\ u_z\theta & 1 & -u_x\theta \\ -u_y\theta & u_x\theta & 1 \end{bmatrix} \begin{bmatrix} 1 & -v_z\phi & v_y\phi \\ v_z\phi & 1 & -v_x\phi \\ -v_y\phi & v_x\phi & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -v_z\phi - u_z\theta & v_y\phi + u_y\theta \\ u_z\theta + v_z\phi & 1 & -v_x\phi - u_x\theta \\ -u_y\theta - v_y\phi & u_x\theta + v_x\phi & 1 \end{bmatrix} \end{aligned} \tag{4}$$

and the other way:

$$\begin{aligned} QP &= \begin{bmatrix} 1 & -v_z\phi & v_y\phi \\ v_z\phi & 1 & -v_x\phi \\ -v_y\phi & v_x\phi & 1 \end{bmatrix} \begin{bmatrix} 1 & -u_z\theta & u_y\theta \\ u_z\theta & 1 & -u_x\theta \\ -u_y\theta & u_x\theta & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -u_z\theta - v_z\phi & u_y\theta + v_y\phi \\ v_z\phi + u_z\theta & 1 & -u_x\theta - v_x\phi \\ -v_y\phi - u_y\theta & v_x\phi + u_x\theta & 1 \end{bmatrix} \end{aligned} \tag{5}$$

So we see that not only does this commute ($PQ = QP$) but more importantly, we have

$$f(u\theta)f(v\phi) = f(u\theta + v\phi) \tag{6}$$

and that f does not do anything fancy with the magnitude and unit axis of the rotation — we assign a unique infinitesimal rotation vector to a unique infinitesimal rotation matrix simply by placing its components in the off-diagonals in a skew-symmetric way.

What are vectors, anyway?

We start by defining a group as a set G and an operation $\circ\colon G \times G \to G$ such that

- The operation is closed in G: $(\forall a, b \in G) a \circ b \in G$.
- The operation is associative: $(\forall a, b, c \in G) a \circ (b \circ c) = (a \circ b) \circ c$.
- There is an identity element $e \in G$ such that $a \circ e = e \circ a = a$.
- Each element has an inverse: $(\forall a \in G)(\exists a' \in G) a \circ a' = a' \circ a = e$.

A vector space is a set of vectors V over a field F. It has an operation which behaves like a group (so for $\mathbb{R}^3$, the identity is 0, the operation is addition, inverses are $v' = -v$), but it also requires the following:

- The group operation is commutative: $(\forall a, b \in G) a \circ b = b \circ a$. (A group whose operation commutes is referred to as Abelian. Fortunately for us, addition in $\mathbb{R}^3$ commutes.)

[1] Essentially, it's a loaded game: you have to commit to a number before I do, so I always win. And unlike in the playground, you can't respond with "No times infinity plus one".

It also allows scalar multiplication, which sends $F \times V \to V$. The rules for this make it a module over F:

- Scalar and field multiplication are compatible: $(\forall a, b \in F)(\forall v \in V)(ab)v = a(bv)$.
- Scalar identity behaves as expected: $(\forall v \in V)1_F v = v$, where 1_F is the multiplicative identity of the field (for our purposes, the field is $\mathbb{R}$, and the identity is 1).
- Scalar multiplication distributes over the vector group operation: $(\forall a \in F)(\forall u, v \in V)a(u \circ v) = au \circ av$.
- Scalar addition distributes over scalar-vector multiplication: $(\forall a, b \in F)(\forall v \in V)(a + b)v = av \circ bv$.

Clearly, we have $\mathbb{R}^3$ being a vector space. You are all happy with vectors behaving like this, even if you had never considered the algebraic structure of what it was underneath: it is intuitive.

Infinitesimal rotation matrices form a vector space

Now consider the space of all infinitesimal rotation matrices, with the group operation of matrix multiplication. We have already seen that these matrices commute, we already know that matrix multiplication associates, and obviously the identity element is I. It is not difficult to show that

$$f(v)f(-v) = I = f(0) = f(v - v), \tag{7}$$

i.e., not only does the space of infinitesimal rotations form a group, but we also have a function f which preserves the group structure.

We take a little more care to define scalar multiplication correctly, but we can see that multiplying the off-diagonal elements behaves as required: if we multiplied the diagonal elements we would get a matrix which is not of the form $I + [v]_\times$, and therefore not an element of the space. In this way, we find that, if:

$$\alpha P = \alpha f(v) = \begin{bmatrix} 1 & -\alpha v_z & \alpha v_y \\ \alpha v_z & 1 & -\alpha v_x \\ -\alpha v_y & \alpha v_x & 1 \end{bmatrix} = f(\alpha v) \tag{8}$$

and we can trivially check that this satisfies the laws of compatibility, identity, and distribution (remembering that our equivalent of addition is matrix multiplication). Not only that, but we may also perform these operations on our original $\mathbb{R}^3$ vectors as well, then transform by f and get the same result — f also preserves the laws of scalar multiplication!

Time for some terminology. We can say that these two spaces are homomorphic (transliterated: “of the same form”, but this is supposedly a mistranslation of the true meaning, “of similar form”); f is therefore a vector space homomorphism. Even better — we can invert the mapping by repopulating our vector in $\mathbb{R}^3$ from the off-diagonal elements, so we may call them isomorphic (of equal form). A function which maps elements one-to-one (injective, injectivity, an injection) and hits each element of the target space (surjective) has an inverse in the sense you are used to and is called a bijection.

In any case, if the form in which it acts on things looks like a vector space and you can map it isomorphically to another representation, then the other way of representing it is uncontroversially a vector space.

Angular velocity and infinitesimal rotations

Now we are ready to convince ourselves that angular velocities are actually vectors, and we'll clarify how the 'pseudovector' and 'sense' terminology fits in. Recall that an angular velocity is the rate of rotation around an axis:

$$\omega = \frac{\mathrm{d}\theta}{\mathrm{d}t} u \tag{9}$$

and since $\mathrm{d}\theta/\mathrm{d}t$ may be arbitrarily large, $\omega \in \mathbb{R}^3$. Generate an infinitesimal rotation by multiplying by some δt to get $\omega \delta t = \delta\theta$. Now, using our rules from earlier we may add two angular velocities by pulling out factors of δt and show that we may also pull the δt from our composed matrices. This of course also works with scalar multiplication.

So what about the term 'pseudovector'? Essentially, it comes down to the way in which we have defined f. If we had used a left-handed coordinate system, or added in a coordinate permutation along the way we could have defined our mapping between vector and matrix elements differently, but it wouldn't have affected the the algebraic equivalence. So this is what the lecturers are referring to when they talk about the 'sense' in which the rotation is defined impacting the definition of angular velocity as a vector, because there's enough ambiguity in there that you need to be consistent for them to act meaningfully on things.