ROTATIONS AND QUATERNIONS

VIJAY KODIYALAM

This is a companion article for a 1 hour talk at the “One per cent” Mathematics workshop for students of classes XI and XII organised by K. N. Raghavan and Parameswaran Sankaran at IMSc, Chennai on 29th November, 2013. It covers some extensions and proofs that were omitted. Ideally it should be read having watched the video of the talk.

1. Rotations in 3-dimensional space

We will restrict attention in this article to rotations about the origin in 3-space. Such a rotation is described by an axis - which we think of as a unit vector - and an angle of rotation. We will imagine that 3-space is rotated by the given angle about the given axis, where the direction of rotation is given by the right hand rule - when the right thumb points in the direction of the axis, the curled up fingers of the right hand indicate the direction of rotation.

One of our main interests is in trying to figure out the image of a given point under a given rotation. Another is to explain the rather unexpected fact - which was demonstrated using a cube - that a composite of two rotations is always given by a single rotation and to be able to compute the final angle and axis of a composite of rotations.

It turns out that both are easily explained using the algebra of quaternions.

2. The algebra of quaternions

Consider as an analogue of complex numbers, all “numbers” of the form $w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ where w, x, y, z are real numbers and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are symbols that commute with all real numbers and satisfy

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

One checks that this is an associative multiplication when extended in a linear way to all such “numbers” which are the quaternions.

Define the norm of a quaternion $q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, again by analogy with the complex numbers, as $|q| = \sqrt{w^2 + x^2 + y^2 + z^2}$. Just as for complex numbers, the quaternionic norm is multiplicative, i.e., $|q\tilde{q}| = |q||\tilde{q}|$.

Exercise: Prove this and use it to show that if two positive integers are each a sum of 4 perfect squares, then, so is their product. Thus, to show that every

Date: November 29, 2013.

positive integer is a sum of (at most) 4 squares, it suffices to verify this for prime numbers.

Also, if we define the conjugate of $q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ as $\overline{q} = w - x\mathbf{i} - y\mathbf{j} - z\mathbf{k}$, then, $q^{-1} = \frac{\overline{q}}{|q|}$. This q^{-1} is both a left and right inverse of q - we do need to consider the possibility that these might be different, though with multiplication being associative, a simple proof - which everyone must do once - shows that if something has both a left and right inverse, then they are equal.

The quaternions form what is known as a division ring, i.e., roughly, they obey all the usual laws of arithmetic (of, say, the real or complex numbers) except that of commutativity of multiplication. Thus, they extend the number system of complex numbers - albeit, non-commutatively. It should be noted that while the real numbers commute with all quaternions, the complex numbers do not.

We will regard 3-space as contained in the 4-dimensional space of quaternions by identifying the point (x, y, z) with the "purely imaginary" quaternion $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

3. Describing rotations by quaternions

The main result is the following. Take an axis of rotation determined by a unit vector $\mathbf{u}$ and an angle of rotation θ. Associate to this the (unit) quaternion

$$q = cos(\frac{\theta}{2}) + sin(\frac{\theta}{2})\mathbf{u}.$$

For any point P in 3-space, represented by the quaternion p, the quaternion qpq^{-1} still lies in 3-space (i.e, is still purely imaginary) and represents the image of p under the rotation of θ about the axis $\mathbf{u}$.

To try and prove this, begin with the given $\mathbf{u}$ and take the plane through the origin that is perpendicular to $\mathbf{u}$. Pick any non-zero unit vector $\mathbf{v}$ in this plane and let $\mathbf{w} = \mathbf{u} \times \mathbf{v}$. Recall the 3-vector identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a}.\mathbf{c})\mathbf{b} - (\mathbf{a}.\mathbf{b})\mathbf{c}$, and use it to conclude that $\mathbf{u} = \mathbf{v} \times \mathbf{w}$ and $\mathbf{v} = \mathbf{w} \times \mathbf{u}$ - notice the similarity with the multiplication of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ in the quaternions (and the difference).

Also observe at this point that a quaternion may be regarded as a "sum" of a scalar w and a vector $\mathbf{v}$. With that, the definition of quaternion multiplication can be written in terms of dot and cross product of vectors as:

$$(w + \mathbf{v})(\tilde{w} + \tilde{\mathbf{v}}) = (w\tilde{w} - \mathbf{v}.\tilde{\mathbf{v}}) + (w\tilde{\mathbf{v}} + \tilde{w}\mathbf{v} + \mathbf{v} \times \tilde{\mathbf{v}})$$

Exercise: One of the nice things about quaternions is that -1 has infinitely many distinct quaternion square roots - in fact, one for every direction in 3-space. Show this by using the formula above to see that if $\mathbf{v}$ is a unit vector in 3-space regarded as a (purely imaginary) unit quaternion, then $\mathbf{v}^2 = -1$.

Next, observe that rotation is an operation on 3-space that commutes with scaling. What this means is that the result of multiplying a vector by a scalar and then rotating it is the same as first rotating it and then multiplying by that scalar. Also note that rotation distributes over addition. In other words, the rotation of a sum of two vectors is the sum of their rotations.

Hence, to verify that that qpq^{-1} represents the rotation of vector p, it suffices to check this when $p = \mathbf{u}, \mathbf{v}, \mathbf{w}$. For p can be written as a sum of scalings of $\mathbf{u}, \mathbf{v}, \mathbf{w}$ and both rotations and the operation $p \mapsto qpq^{-1}$ commute with scaling and distribute over addition of vectors.

Now $q = cos(\frac{\theta}{2}) + sin(\frac{\theta}{2})\mathbf{u}$ and $q^{-1} = cos(\frac{\theta}{2}) - sin(\frac{\theta}{2})\mathbf{u}$. We will now check in the 3 cases $p = \mathbf{u}, \mathbf{v}, \mathbf{w}$ that qpq^{-1} is the same as the result of rotation of p around the $\mathbf{u}$ axis by an angle θ.

Case I: Suppose that $p = \mathbf{u}$. We know that a rotation around the $\mathbf{u}$ axis should keep $\mathbf{u}$ fixed. Now consider

$$\begin{aligned} qpq^{-1} &= q\mathbf{u}q^{-1} = (cos(\frac{\theta}{2}) + sin(\frac{\theta}{2})\mathbf{u})\mathbf{u}(cos(\frac{\theta}{2}) - sin(\frac{\theta}{2})\mathbf{u}) \\ &= (-sin(\frac{\theta}{2}) + cos(\frac{\theta}{2})\mathbf{u})(cos(\frac{\theta}{2}) - sin(\frac{\theta}{2})\mathbf{u}) \\ &= (cos^2(\frac{\theta}{2}) + sin^2(\frac{\theta}{2}))\mathbf{u} = \mathbf{u} \end{aligned}$$

The last two equalities follow from the definition of quaternion multiplication in terms of dot and cross products. So for $p = \mathbf{u}$, we have what we need.

Case II: Suppose that $p = \mathbf{v}$. We first calculate:

$$\begin{aligned} qpq^{-1} &= q\mathbf{v}q^{-1} = (cos(\frac{\theta}{2}) + sin(\frac{\theta}{2})\mathbf{u})\mathbf{v}(cos(\frac{\theta}{2}) - sin(\frac{\theta}{2})\mathbf{u}) \\ &= (cos(\frac{\theta}{2})\mathbf{v} + sin(\frac{\theta}{2})\mathbf{w})(cos(\frac{\theta}{2}) - sin(\frac{\theta}{2})\mathbf{u}) \\ &= cos^2(\frac{\theta}{2})\mathbf{v} + sin(\frac{\theta}{2})cos(\frac{\theta}{2})\mathbf{w} - cos(\frac{\theta}{2})sin(\frac{\theta}{2})(\mathbf{v} \times \mathbf{u}) - sin^2(\frac{\theta}{2})(\mathbf{w} \times \mathbf{u}) \\ &= cos(\theta)\mathbf{v} - sin(\theta)\mathbf{w} \end{aligned}$$

Case III: I'll leave it to you to verify that if $p = \mathbf{w}$, then $qpq^{-1} = sin(\theta)\mathbf{v} + cos(\theta)\mathbf{w}$.

But now considering the $\mathbf{v}$ - $\mathbf{w}$ plane, a rotation of θ about the $\mathbf{u}$-axis should take v to $cos(\theta)\mathbf{v} - sin(\theta)\mathbf{w}$ and w to $sin(\theta)\mathbf{v} + cos(\theta)\mathbf{w}$. This shows that in all 3 cases rotation does exactly the same thing as qpq^{-1} - finishing the proof.

Now let us also observe the following. Given any unit quaternion q - meaning of norm 1 - it can always be written as $cos(\frac{\theta}{2}) + sin(\frac{\theta}{2})\mathbf{u}$ - actually in two ways, one obtainable from the other from changing the sign of $\mathbf{u}$ and changing the angle by $180°$.

This now shows that the composite of two rotations is a rotation (since a product of unit quaternions is a unit quaternion) and how to compute the axis and angle of the final rotation (multiply the associated unit quaternions - carefully, in the right order - and deduce the angle and axis from the result).

4. What is the Dirac belt trick about?

As I understand it, it is an "explanation" of why $\frac{\theta}{2}$ occurs in the quaternion associated to a rotation by θ rather than θ itself. It really has to so with some nice physics which I unfortunately do not know.

Consider the set of all unit quaternions. This forms the unit 3-sphere $w^2 + x^2 + y^2 + z^2 = 1$ in 4-dimensional space. On the other hand consider the "set of all rotations of 3-space". What kind of object is this ?

Under a rotation, suppose the unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ go to the vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ respectively. It is clear geometrically that $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are also mutually perpendicular unit vectors, and as we said earlier, completely determine the rotation. Thus a rotation is "the same" as the set of 3 mutually perpendicular unit vectors. We can arrange these three vectors in order as the columns of a 3×3 matrix.

So every rotation R gives a 3×3 matrix, also denoted R, whose columns are unit mutually perpendicular vectors. The matrix thus satisfies $R^T R = I$. This also implies that $RR^T = I$. Here, of course, R^T stands for the transpose matrix of R.

The thing is that the converse holds too - almost. Given any three mutually perpendicular unit vectors, is there a rotation that takes $\mathbf{i}, \mathbf{j}, \mathbf{k}$ to those ? Not quite, for geometrically, rotations preserve cross products of vectors, so if we know where $\mathbf{i}, \mathbf{j}$ go under a rotation, there is only one choice for where $\mathbf{k}$ can go - namely the cross product of the two images. This is a unit vector perpendicular to both and one of the only two such. Thus given any 3 mutually perpendicular vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$, there is a rotation that maps $\mathbf{i}, \mathbf{j}, \mathbf{k}$ to either $\mathbf{u}, \mathbf{v}, \mathbf{w}$ or to $\mathbf{u}, \mathbf{v}, -\mathbf{w}$.

What characterises images of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ under a rotation is that the scalar triple product $\mathbf{u}.(\mathbf{v} \times \mathbf{w}) = 1$ as opposed to -1 - which is also a possibility for a triple of mutually perpendicular vectors.

Since the scalar triple product is just the determinant of the associated matrix, we may conclude that the set of all rotations in 3-space is "the same" as the set of all 3×3 matrices R that satisfy $RR^T = I = R^T R$ and $det(R) = 1$.

So on the one hand, we have unit quaternions that describe rotations and on the other, we have the collection of 3×3 matrices as above. What is the connection ? The fact is that there is a natural way of getting a 3×3 matrix from a quaternion in such a way that exactly two quaternions correspond to a single matrix and these two are negatives of each other.

Dirac's belt trick has to finally do with the fact that the negative of a negative is the identity. We will stop at this for the interested reader to pursue further.

Acknowledgements: My thanks to K. N. Raghavan for a careful reading of this and comments and for preparing a companion sheet to this article with two detailed worked out examples of using quaternions to understand rotations.

The Institute of Mathematical Sciences, Chennai, India
E-mail address: vijay@imsc.res.in