J. fixed point theory appl. 2 (2007), 215–220

1661-7738/020215-6, *published online* 20.12.2007
DOI 10.1007/s11784-007-0042-5

Journal of Fixed Point Theory and Applications

Euler's fixed point theorem: The axis of a rotation

Bob Palais and Richard Palais

Dedicated to the memory of Leonhard Euler, "The Master of us all", on the occasion of the 300th anniversary of his birth

Abstract. We give an elementary proof of what is perhaps the earliest fixed point theorem; namely Leonhard Euler's theorem of 1775 on the existence of an axis $\mathbf{v}$ for any three-dimensional rotation $\mathbf{R}$. The proof is constructive and shows that no multiplications are required to compute $\mathbf{v}$.

Mathematics Subject Classification (2000). Primary 55M20, 01A50.

Keywords. Rotation, axis.

In what is perhaps the historically earliest fixed point theorem, Leonhard Euler [1] stated in 1775 that in three dimensions, every rotation has an axis. Euler's original formulation of the result is that if a sphere is rigidly rotated about its center then there is a diameter that remains fixed. A modern reformulation is:

Euler's Theorem. *If* $\mathbf{R}$ *is a* 3×3 *matrix satisfying* $\mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}$ *and* $\det \mathbf{R} = +1$*, then there is a non-zero vector* $\mathbf{v}$ *satisfying* $\mathbf{R}\mathbf{v} = \mathbf{v}$*.*

Counterexamples are easy to find in two or other even dimensions. Even in three dimensions, it is not immediately obvious that the composition of rotations about distinct axes is equivalent to a rotation about a single axis. This important fact has a myriad of applications in pure and applied mathematics, and as a result there are many known proofs. In this 300th anniversary year of Euler's birth, we give a constructive proof of the fixed vector $\mathbf{v}$ that appears to be new.

Proof. Our construction is this: if $\mathbf{A} := \frac{1}{2}(\mathbf{R} - \mathbf{R}^T)$ is the skew-symmetric part of the matrix $\mathbf{R}$, then the vector $\mathbf{v} := (a_{23}, a_{31}, a_{12})$ is left fixed by $\mathbf{R}$. This follows from $\mathbf{R}\mathbf{A}\mathbf{R}^T = \mathbf{A}$, which is immediate from the definition of $\mathbf{A}$. To see that $\mathbf{R}\mathbf{A}\mathbf{R}^{-1} = \mathbf{R}\mathbf{A}\mathbf{R}^T = \mathbf{A}$ is equivalent to $\mathbf{R}\mathbf{v} = \mathbf{v}$, we invoke the linear isomorphism $\mathbf{v} \mapsto \mathbf{J}_\mathbf{v}$ between $\mathbb{R}^3$ and the skew-symmetric 3×3 matrices, given explicitly by

$$\mathbf{J}_\mathbf{v} := \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}.$$

$\mathbf{J}_\mathbf{v}$ is just the matrix implementing cross-product with $\mathbf{v}$; i.e., $\mathbf{J}_\mathbf{v}(\mathbf{w}) := \mathbf{v} \times \mathbf{w}$ for all $\mathbf{w} \in \mathbb{R}^3$. An essential fact about the cross-product operation is its invariance under any such rotation matrix $\mathbf{R}$, that is, $\mathbf{R}(\mathbf{v} \times \mathbf{w}) = \mathbf{R}\mathbf{v} \times \mathbf{R}\mathbf{w}$. In terms of the map $\mathbf{J}$, this says $\mathbf{J}_{\mathbf{R}\mathbf{v}}(\mathbf{R}\mathbf{w}) = \mathbf{R}\mathbf{J}_\mathbf{v}(\mathbf{w})$ or simply $\mathbf{J}_{\mathbf{R}\mathbf{v}}\mathbf{R} = \mathbf{R}\mathbf{J}_\mathbf{v}$. Rearranging this, we arrive at $\mathbf{J}_{\mathbf{R}\mathbf{v}} = \mathbf{R}\mathbf{J}_\mathbf{v}\mathbf{R}^{-1}$. (In the language of representation theory, this says that $\mathbf{J}$ is an equivalence or "intertwining operator" between the standard representation of $\mathbf{SO}(3)$ on $\mathbb{R}^3$ and the adjoint representation on its Lie algebra, $\mathbf{so}(3)$, the skew-adjoint 3×3 matrices. Or, in more elementary terms, the matrix for cross-product with a rotated vector is obtained by conjugating (by the rotation) the matrix for cross-product with the un-rotated vector.) It is now immediate that a vector $\mathbf{v}$ is fixed by a rotation $\mathbf{R}$ if and only if the skew-adjoint matrix $\mathbf{J}_\mathbf{v}$ is fixed under conjugation by $\mathbf{R}$.

But we are not through yet! What if $\mathbf{A} = \mathbf{0}$? Such orthogonal matrices—the symmetric ones—have measure zero, and correspond to angles of rotation 0 and π. But still they are important, and we have a similar result for them. In this (non-generic) case, $\mathbf{R} = \mathbf{R}^T$ and so $\mathbf{R}^2 = \mathbf{I}$. Therefore, $\mathbf{R}(\mathbf{I} + \mathbf{R}) = \mathbf{R} + \mathbf{R}^2 = \mathbf{I} + \mathbf{R}$, so the columns of $\mathbf{I} + \mathbf{R}$ are fixed. Since $\mathbf{R}$ is proper ($\det \mathbf{R} = +1$), $\mathbf{R} \neq -\mathbf{I}$, so $\mathbf{I} + \mathbf{R}$ has a non-zero column that is fixed by $\mathbf{R}$. □

Examples. If $\mathbf{R} = \mathbf{R}_\theta = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is rotation about the z-axis by an angle θ, then $\mathbf{A} = \begin{pmatrix} 0 & \sin\theta & 0 \\ -\sin\theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ so $\mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ \sin\theta \end{pmatrix}$. If θ is not 0 or π then we are in the generic case, and we find that the z-axis is a fixed axis of the rotation. If $\theta = 0$, then $\mathbf{R} = \mathbf{I}$, so every vector is fixed, while if $\theta = \pi$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ then $\mathbf{v}' = \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix}$ (which is non-zero if we choose $x_3 \neq 0$) and again we find that the z-axis is fixed.

If $\mathbf{R} = \mathbf{R}_G = \begin{pmatrix} 2/3 & 2/3 & -1/3 \\ -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \end{pmatrix}$, then $\mathbf{A} = \begin{pmatrix} 0 & 1/2 & -1/2 \\ -1/2 & 0 & 1/2 \\ 1/2 & -1/2 & 0 \end{pmatrix}$, so $\begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$ is fixed.

If $\mathbf{R} = \mathbf{R}_N = \begin{pmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{pmatrix}$, then $\mathbf{A} = \mathbf{0}$, and $\mathbf{R} + \mathbf{I} = \begin{pmatrix} 2/3 & 2/3 & 2/3 \\ 2/3 & 2/3 & 2/3 \\ 2/3 & 2/3 & 2/3 \end{pmatrix}$, so $\begin{pmatrix} 2/3 \\ 2/3 \\ 2/3 \end{pmatrix}$ is fixed.

Remarks. 1) Our proof of Euler's theorem uses one non-trivial result, the "intrinsic" nature of the cross-product operation in a three-dimensional, oriented, real inner-product space. But this is obvious from the usual characterization of $\mathbf{u} \times \mathbf{v}$, namely that its length is the area of the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$, that it is orthogonal to $\mathbf{u}$ and $\mathbf{v}$, and that $(\mathbf{u}, \mathbf{v}, \mathbf{u} \times \mathbf{v})$ is an oriented triple.

2) In the generic case our key fixed-point fact, $\mathbf{R}\mathbf{A}\mathbf{R}^T = \mathbf{A}$, requires only the orthogonality of $\mathbf{R}$. However, if $\mathbf{R}$ is improper then $\mathbf{R}(\mathbf{v} \times \mathbf{w}) = -\mathbf{R}\mathbf{v} \times \mathbf{R}\mathbf{w}$,

so our argument shows that the vector $\mathbf{v}$ corresponding to $\mathbf{A}$ is not fixed by $\mathbf{R}$, but rather is "reversed", i.e., mapped to $-\mathbf{v}$ (but note that the axis through $\mathbf{v}$ is still fixed). In the non-generic case we are using another elementary fixed-point result; if a linear operator $\mathbf{T}$ on a vector space $\mathbf{V}$ satisfies $\mathbf{T}^2 = \mathbf{I}$, and we define $\mathbf{P} := \frac{1}{2}(\mathbf{I} + \mathbf{T})$, then $\mathbf{P}^2 = \mathbf{P}$, so $\mathbf{P}$ is the projection on the image of $\mathbf{P}$ along its kernel. But if $\mathbf{w} = \mathbf{Pv} = \frac{1}{2}(\mathbf{v} + \mathbf{Tv})$, then $\mathbf{Tw} = \frac{1}{2}(\mathbf{Tv} + \mathbf{v}) = \mathbf{w}$, so $\mathbf{P}$ projects $\mathbf{V}$ onto the fixed-point set $\mathbf{V}^+$ (the $+1$ eigenspace) of $\mathbf{T}$. And if $\mathbf{Pv} = \mathbf{0}$, then $\mathbf{Tv} = -\mathbf{v}$, so $\mathbf{P}$ projects $\mathbf{V}$ along the space $\mathbf{V}^-$ of reversed vectors (the -1 eigenspace of $\mathbf{T}$). Symmetrically, $\mathbf{I} - \mathbf{P} = \frac{1}{2}(\mathbf{I} - \mathbf{T})$ is the projection on $\mathbf{V}^-$ along $\mathbf{V}^+$. In particular $\mathbf{V} = \mathbf{V}^+ \oplus \mathbf{V}^-$, and of course when $\mathbf{T}$ is symmetric this is an orthogonal decomposition. So regardless of whether $\mathbf{R}$ is proper or improper, in the non-generic orthogonal case in which $\mathbf{R} = \mathbf{R}^T$, we have $\mathbf{R}^2 = \mathbf{I}$, so the two projections $\frac{1}{2}(\mathbf{I} + \mathbf{R})$ and $\frac{1}{2}(\mathbf{I} - \mathbf{R})$ give an orthogonal decomposition of $\mathbb{R}^3$ into fixed and reversed subspaces. We may classify four cases up to $\mathbf{O}(3)$ conjugacy by their complementary dimensions between 0 and 3. We also note that proper and improper examples are in one-to-one correspondence under composition with $-\mathbf{I}$.

3) Euler's original demonstration never invokes the distinction between proper and improper transformations, apparently leaving it implicit in the concept of a rigid motion. Moreover, Euler's construction of a fixed axis for a non-generic rotation fails (in different ways) for both the proper and improper cases, so it does require modification to be valid by modern standards. (Indeed, there have been many other purported proofs of Euler's theorem over the years that also fail for similar reasons, and this is why we have paid such detailed attention to the non-generic case above). Remarkably, Euler's demonstration **does** contain constructions foreshadowing both cases of the proof we have given here, but in his more geometric language! We address these considerations, provide a (previously unavailable) English translation of the original proof, and survey other proofs in a forthcoming paper [8].

4) We note that, using our construction, no multiplications are required to find an axis in either case. Similarly, the cosine of the angle of rotation about this axis can easily be found by taking the trace of $\mathbf{R}$. The trace is independent of basis, and in a proper orthonormal basis containing the axis, the matrix for $\mathbf{R}$ is given by the matrix $\mathbf{R}_\theta$ in the example above, whose trace is clearly $1 + 2\cos\theta$. For comparison, other approaches have been suggested to compute the axis involving from 1 to 3 cross-products or up to 18 multiplications, followed by more cross-products, normalizations and projections for up to about 40 multiplications to find the cosine of the angle (see for example [3, Chapter 12, Section 9: Finding the axis and angle]).

Final remarks suggested by the referee's comments

The fact that the axis of a generic rotation $\mathbf{R}$ is a multiple of the vector corresponding to the anti-symmetric part of $\mathbf{R}$ has been observed and used in other contexts in the past. There is also another construction that does not require any multiplications to obtain the axis. As we will outline below, all of these constructions *assume the existence of an axis*, and seem to require appeals to less elementary concepts than our approach above to turn them into proofs of Euler's theorem.

The axis-angle representation of a rotation $\mathbf{R}$ about a unit axis $\mathbf{v}$ by an angle θ is based upon the decomposition of a vector $\mathbf{x}$ into a component along $\mathbf{v}$, and a component orthogonal to $\mathbf{v}$: $\mathbf{x} = (\mathbf{x}\cdot\mathbf{v})\mathbf{v} + (\mathbf{x} - (\mathbf{x}\cdot\mathbf{v})\mathbf{v})$. Since $\mathbf{v} \times ((\mathbf{x} - (\mathbf{x}\cdot\mathbf{v})\mathbf{v}))$ is the image of $\mathbf{x} - (\mathbf{x}\cdot\mathbf{v})\mathbf{v}$ under a quarter-turn counterclockwise rotation of $\mathbf{v}^\perp$, and $\mathbf{R}\mathbf{v} = \mathbf{v}$, we can check that

$$\mathbf{R}\mathbf{x} = (\mathbf{x}\cdot\mathbf{v})\mathbf{v} + (\cos\theta)(\mathbf{x} - (\mathbf{x}\cdot\mathbf{v})\mathbf{v}) + (\sin\theta)(\mathbf{v} \times (\mathbf{x} - (\mathbf{x}\cdot\mathbf{v})\mathbf{v})).$$

Using $\mathbf{v} \times \mathbf{v} = \mathbf{0}$, this simplifies to

$$\mathbf{R}\mathbf{x} = \mathbf{x} + (\sin\theta)(\mathbf{v} \times \mathbf{x}) + (1 - \cos\theta)((\mathbf{x}\cdot\mathbf{v})\mathbf{v} - \mathbf{x}).$$

If $\mathbf{J}_\mathbf{v}$ is the matrix defined earlier by $\mathbf{J}_\mathbf{v}(\mathbf{x}) = \mathbf{v}\times\mathbf{x}$, the fact that $\mathbf{v}$ is a unit vector implies $\mathbf{J}_\mathbf{v}^2 = \mathbf{v}\mathbf{v}^T - \mathbf{I}$, or $\mathbf{J}_\mathbf{v}^2\mathbf{x} = (\mathbf{x}\cdot\mathbf{v})\mathbf{v} - \mathbf{x}$, which leads to the *Euler–Rodrigues rotation formula:*

$$\mathbf{R} = \mathbf{I} + (\sin\theta)\mathbf{J}_\mathbf{v} + (1 - \cos\theta)\mathbf{J}_\mathbf{v}^2.$$

In [1], Alperin derives this result from Lagrange's well-known triple product formula $\mathbf{p} \times (\mathbf{q} \times \mathbf{r}) = (\mathbf{p}\cdot\mathbf{r})\mathbf{q} - (\mathbf{p}\cdot\mathbf{q})\mathbf{r}$, while Fillmore [5] and Kahan [6] derive it for $\mathbf{R}(\theta) = \exp(\theta\mathbf{J}_\mathbf{v})$ using the series representation for exp and the recurrence $\mathbf{J}_\mathbf{v}^{n+2k} = (-1)^k\mathbf{J}_\mathbf{v}^{n+2k}$, $n > 0$. Fillmore takes the series as the definition of $\exp(\theta\mathbf{J}_\mathbf{v})$ while Kahan solves the initial value problem $d\mathbf{R}/d\theta = \mathbf{J}_\mathbf{v}\mathbf{R}$, $\mathbf{R}(0) = \mathbf{I}$. Kahan observes that this formula implies $\frac{1}{2}(\mathbf{R} - \mathbf{R}^T) = (\sin\theta)\mathbf{J}_\mathbf{v}$ which can be used (as we did) to recover the axis in the generic case, and points out this can fail numerically when $\sin\theta$ is small. Fillmore uses the formula to show that if we define $\mathbf{S} := \mathbf{R} + \mathbf{R}^T + (1 - \mathrm{trace}(\mathbf{R}))\mathbf{I}$ (clearly a symmetric matrix) then $\mathbf{S} = (3 - \mathrm{trace}(\mathbf{R}))\mathbf{v}\mathbf{v}^T$. This gives another method to obtain the axis without multiplication, as the right-hand side exhibits this matrix as a rank one symmetric matrix, any of whose non-zero columns are multiples of $\mathbf{v}$. Kalman [7] makes the related observation that for any $\mathbf{x}$, $\mathbf{R}(\mathbf{S}\mathbf{x}) = \mathbf{S}\mathbf{x}$. He shows that this is equivalent to $\mathbf{R}^3 - \mathrm{trace}(\mathbf{R})\mathbf{R}^2 + \mathrm{trace}(\mathbf{R})\mathbf{R} - \mathbf{I} = \mathbf{0}$, and confirms this by applying the Cayley–Hamilton theorem. Fillmore and Kalman both use the fact that the eigenvalues of $\mathbf{R}$ are $\{1, \cos\theta + i\sin\theta, \cos\theta - i\sin\theta\}$ to show $\mathrm{trace}(\mathbf{R}) = 1 + 2\cos\theta$. Note that *all of these approaches start with the assumption that* $\mathbf{R}$ *is a rotation about an axis* $\mathbf{v}$ *through an angle* θ*,* and so cannot be used as a proof of Euler's theorem without considerable modification. (Moreover they all invoke results like the Cayley–Hamilton theorem, the power series for exp, and facts about solutions of the characteristic equation of a matrix that seem less elementary than the approach we have given above.) However, granting the truth of Euler's theorem, they do give alternative derivations for our construction of the axis of a rotation in the generic case, and recognizing this, Kahan presents the classical proof based on characteristic polynomials that 1 is an eigenvalue of $\mathbf{R}$, and suggests finding the corresponding eigenvector by elimination or computing an appropriate column of $\mathrm{Adj}(\mathbf{R} - \mathbf{I})$. Fillmore also examines the improper case, and the representation of orthogonal transformations as the product of two (in the proper case) or three (in the improper case) reflections.

There have also been several algorithms proposed to compute the unit quaternions $\pm\mathbf{Q}$ corresponding to $\mathbf{R}$, [2, 10–12]. Most are based upon a modified version of the Euler–Rodrigues rotation formula, which has a nice expression in terms of the components of the unit quaternion associated with $\mathbf{R}$, $\mathbf{Q} = [q, \mathbf{q}] = [\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\mathbf{v}]$. Using $\sin\theta = 2\cos\frac{\theta}{2}\sin\frac{\theta}{2}$, we get $\sin\theta\mathbf{J}_{\mathbf{v}} = 2q\mathbf{J}_{\mathbf{q}}$, and since $1 - \cos\theta = 1 - \cos(2\frac{\theta}{2}) = 1 - (\cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}) = 2\sin^2\frac{\theta}{2}$ we get $(1-\cos\theta)\mathbf{J}_{\mathbf{v}}^2 = 2\mathbf{J}_{\mathbf{q}}^2$. This leads to the well-known quaternion-to-matrix formula

$$\mathbf{R} = \mathbf{I} + 2q\mathbf{J}_{\mathbf{q}} + 2\mathbf{J}_{\mathbf{q}}^2.$$

Because the vector part of $\mathbf{Q}$ is a unit axis of $\mathbf{R}$, scaled by $\sin\frac{\theta}{2}$ (rather than $\sin\theta$ as in our construction), matrix-to-quaternion algorithms compute the anti-symmetric part of $\mathbf{R}$, then perform square roots and division to accomplish the desired scaling. In another forthcoming paper [9], we present new and efficient algorithms for implementing and interpolating rotations based on a novel realization of a rotation and unit quaternions *without* reference to an axis.

Acknowledgments

We are indebted to the referee for pointing out to us references [1] and [5].

References

[1] R. Alperin, *The matrix of a rotation.* College Math. J. **20** (1989), 230.

[2] I. Y. Bar-Itzhack, *New method for extracting the quaternion from a rotation matrix.* AIAA J. Guidance Control Dynamics **23** (2000), 1085–1087.

[3] B. Casselman, *A Manual of Geometry and PostScript.* Cambridge Univ. Press, 2005; http://www.math.ubc.ca/∼cass/graphics/manual/pdf/ch12.pdf.

[4] L. Euler, *Formulae generales pro translatione quacunque corporum rigidorum.* Novi Commentarii Academiae Scientiarum Petropolitanae **20** (1776), 189–207. Reprinted in *Opera Omnia*, Series 2, Volume 9, 84–98, and in *Theoria motus corporum solidorum seu rigidorum*, ed. nova, 1790, 449–460.

[5] J. P. Fillmore, *A note on rotation matrices.* IEEE J. Computer Graphics Appl. **4** (1984), 30–33.

[6] W. Kahan, *Computing cross-products and rotations.* UC Berkeley Math H110 Lecture Notes; http://www.cs.berkeley.edu/∼wkahan/MathH110/Cross.pdf.

[7] D. Kalman, *The axis of a rotation: analysis, algebra, geometry.* Math. Mag. **62** (1989), 248–252.

[8] B. Palais, R. Palais and S. Rodi, *Euler's theorem on the axis of a rotation: proofs old and new.* To appear.

[9] B. Palais and R. Palais, *Reflections on rotations.* To appear.

[10] S. W. Shepperd, *Quaternion from rotation matrix.* J. Guidance Control **1** (1978), 223–224.

[11] K. Shoemake, *Quaternion calculus and fast animation.* In: Computer Animation: 3-D Motion Specification and Control, SIGGRAPH 1987 Tutorial (1987), 101–121.

[12] R. A. Spurrier, *Comment on singularity-free extraction of a quaternion from a direction-cosine matrix.* J. Spacecraft Rockets **15** (1978), 255.

Bob Palais
Department of Mathematics
The University of Utah
Salt Lake City, UT 84112, USA
e-mail: `palais@math.utah.edu`

Richard Palais
Department of Mathematics
103 MSTB University of California at Irvine
Irvine, CA 92697, USA
e-mail: `palais@uci.edu`

To access this journal online:
www.birkhauser.ch/jfpta