Other taxicab geometric sets might also be explored and additional theorems established as valid or invalid for this geometry. Surely changing the distance function and changing the coordinate grid from a square configuration to a triangular configuration leads to geometries that behave very differently from Euclidean geometry.

The author expresses her appreciation to Mr. Ron Spangler, ECU, for his assistance in the preparation of the illustrations.

REFERENCES

1. Martin Gardner, Mathematical games, *Scientific American* 243 (Nov. 1980), 18–30.
2. Eugene F. Krause, Taxicab geometry, *Mathematics Teacher* 66 (1973), 695–706.
3. ——, *Taxicab Geometry*, Addison-Wesley, Menlo Park, NJ, 1975.
4. Richard Laatsch, Pyramidal sections in taxicab geometry, this MAGAZINE 55 (1982), 205–212.
5. Barbara E. Reynolds, Taxicab geometry, *Pi Mu Epsilon J.* 7 (1980), 77–88.
6. J. J. Schaffer, *Geometry of Spheres in Normed Spaces*, Marcel Dekker, Inc., New York, 1976, pp. 17–18.
7. Doris J. Schattschneider, The taxicab group, *Amer. Math. Monthly* 91 (1984), 423–428.

The Axis of a Rotation: Analysis, Algebra, Geometry

DAN KALMAN
5332 Ironwood St.
Rancho Palos Verdes, CA 90274

Introduction Recently, while working on some problems related to coordinate transformations, I happened on the following discovery:

> If the 3 by 3 matrix A represents a rotation (i.e., A is orthogonal with determinant 1), and the trace of A is $\mathrm{tr}(A)$, then for any vector $\mathbf{x}$
>
> $$A\mathbf{x} + A^T\mathbf{x} + [1 - \mathrm{tr}(A)]\mathbf{x}$$
>
> lies on the axis of the rotation.

I recognized immediately that this result must be well known (in certain circles). However, it seems to me that my route of discovery illustrates some important principles of problem solving and mathematical discovery. I present this account with the idea in mind that a suitable modification might be presented in a linear algebra course. In addition to serving as a case study in discovery, the topic is a natural application of eigenvalues and eigenvectors, and the result has an attractive simplicity. As suggested by the title, analysis, algebra, and geometry each play a role in the development to follow.

Background Before proceeding further, it will be useful to review some facts about rotation matrices and to establish the notation and nomenclature to be used. A *rotation* is a rigid transformation of real 3 dimensional space leaving the origin fixed. Such a transformation is necessarily linear, and is represented with respect to the standard basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$ by a *rotation matrix*, A. The columns of A are the images of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ under the rigid motion, and so comprise a right-handed triple of orthogonal unit vectors. Therefore, $A^TA = I$ and $\det(A) = 1$.

In general, a square matrix A that satisfies $A^TA = I$ is called orthogonal. Using the matrix product notation $\mathbf{x}^T\mathbf{y}$ and the inner product notation $\mathbf{x} \cdot \mathbf{y}$ interchangeably (for vectors $\mathbf{x}$ and $\mathbf{y}$), observe that $A\mathbf{x} \cdot A\mathbf{y} = (A\mathbf{x})^TA\mathbf{y} = \mathbf{x}^TA^TA\mathbf{y} = \mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$. Thus, orthogonal matrices preserve inner products, and in particular, preserve angles and lengths.

For the special case of a 3×3 orthogonal matrix with unit determinant, it can be shown that 1 is an eigenvalue, and that the corresponding eigenspace is one dimensional, as follows. Since the matrix preserves lengths, its eigenvalues must all have magnitude 1. The product of the eigenvalues is the determinant, and must therefore equal 1, as well. Thus, aside from the trivial case of the identity matrix, there must be a unique eigenvalue equal to 1, and a pair of complex conjugates $r = \cos\theta + i\sin\theta$ and $s = \cos\theta - i\sin\theta$. As an unrepeated eigenvalue, 1 has a one-dimensional eigenspace, as asserted.

The eigenspace for the eigenvalue 1 is a line of fixed points for the transformation. It will now be shown that the transformation is geometrically a rotation about this fixed line. Every vector may be resolved into orthogonal components parallel to the fixed line and in the plane perpendicular to the fixed line. By virtue of linearity, it suffices to show that the transformation acts as a rotation on the perpendicular plane. Accordingly, consider the special case of a vector perpendicular to the fixed line. Its image has equal length, and must also be perpendicular to the fixed line. Therefore, it is possible to rotate the vector about the fixed line to obtain the image vector. Moreover, preservation of the angle between vectors implies that any two vectors in the perpendicular plane must be rotated by the same amount. Thus, the transformation acts on the perpendicular plane as a rotation about the fixed line, as desired.

To summarize the preceding paragraphs, a 3×3 matrix represents a rotation if and only if it is orthogonal with unit determinant. A nontrivial rotation matrix possesses a one-dimensional eigenspace corresponding to the eigenvalue 1, and this eigenspace is, in fact, the axis of the rotation. With this background established, the discussion proceeds to the main topic of the paper.

Analysis Permit me to set the stage. I was interested in developing a computer program to generate the rotation matrix linking two right-handed coordinate systems in $\mathbf{R}^3$, given some information about their relative orientations. As a side topic, I wished to find the axis of the rotation, that is, to find one eigenvector corresponding to the eigenvalue 1. Let the rotation matrix A have entries a_{ij}. A solution of $(A - I)\mathbf{x} = 0$ must be orthogonal to the first two rows of $A - I$. Denoting the mth row of A by (row m), the first two rows of $A - I$ are (row 1) $-\ \mathbf{i}$ and (row 2) $-\ \mathbf{j}$. A vector orthogonal to both is obtained by taking the vector product $\mathbf{c}$. This results in

$$\mathbf{c} = (\text{row } 1) \times (\text{row } 2) - \mathbf{i} \times (\text{row } 2) + \mathbf{j} \times (\text{row } 1) + \mathbf{i} \times j.$$

Now, since A is a rotation matrix, its columns form a right handed triple. The same may be said of A^T, so the rows of A also form a right handed triple. In particular, (row 1) $\times$ (row 2) $=$ (row 3). Applying this result and simplifying the remaining three cross products leads to

$$\mathbf{c} = \begin{bmatrix} a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} - \begin{bmatrix} 0 \\ -a_{23} \\ a_{22} \end{bmatrix} + \begin{bmatrix} a_{13} \\ 0 \\ -a_{11} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{31} + a_{13} \\ a_{32} + a_{23} \\ 1 + a_{33} - a_{22} - a_{11} \end{bmatrix}.$$

Of course, the vector $\mathbf{c}$ might equal zero (if the first two rows of $A - I$ are dependent) but in this case a cross product of a different pair can be calculated. This provides enough information for a computer program, and concludes the analysis phase of discovery.

Algebra The formula for $\mathbf{c}$ derived above has too much symmetry to be left alone. Among possible rearrangements, the following pleases the eye:

$$\mathbf{c} = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} + \begin{bmatrix} a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 - a_{11} - a_{22} - a_{33} \end{bmatrix}$$

$$= \text{row } 3 + \text{column } 3 + [1 - \text{tr}(A)]\mathbf{k}.$$

Next, observe that column 3 is just $A\mathbf{k}$, and similarly, row 3 is $A^T\mathbf{k}$. Thus, we have

$$\mathbf{c} = A\mathbf{k} + A^T\mathbf{k} + [1 - \text{tr}(A)]\mathbf{k}$$

$$= \left(A + A^T + [1 - \text{tr}(A)]I\right)\mathbf{k}.$$

Is there any reason for the vector $\mathbf{k}$ to be singled out in this fashion? Surely a similar formula involving $\mathbf{i}$ or $\mathbf{j}$ must exist. It is even tempting to believe that replacing $\mathbf{k}$ with *any* vector produces a vector $\mathbf{c}$ in the eigenspace corresponding to the eigenvalue 1. How might such an assertion be proved?

The conjecture is this: for any vector $\mathbf{v}$, $(A + A^T + [1 - \text{tr}(A)]I)\mathbf{v}$ is an eigenvector with eigenvalue 1. That is,

$$A\left(A + A^T + [1 - \text{tr}(A)]I\right)\mathbf{v} = \left(A + A^T + [1 - \text{tr}(A)]I\right)\mathbf{v}.$$

To establish this for all $\mathbf{v}$ requires showing that the matrices multiplying $\mathbf{v}$ on each side of the equation are equal. Rearranging the necessary identity yields

$$A^2 + I + [1 - \text{tr}(A)]A = A + A^T + [1 - \text{tr}(A)]I$$

and hence

$$A^2 - \text{tr}(A)A + \text{tr}(A)I - A^T = 0.$$

Finally, since A is nonsingular, we may multiply both sides by A to obtain

$$A^3 - \text{tr}(A)A^2 + \text{tr}(A)A - I = 0.$$

Thus, the conjecture at hand is equivalent to a certain polynomial identity for A. This immediately suggests consideration of the characteristic polynomial of A.

Let $p(x)$ be the characteristic polynomial of A. As mentioned earlier, p has roots 1, $r = \cos\theta + i\sin\theta$, and $s = \cos\theta - i\sin\theta$. Moreover, $rs = 1$ and $r + s = 2\cos\theta$. Then, the factored form $p(x) = (x-1)(x-r)(x-s)$ may be multiplied out to give

$$\begin{aligned} p(x) &= x^3 - (1+r+s)x^2 + (1+r+s)x - 1 \\ &= x^3 - (1+2\cos\theta)x^2 + (1+2\cos\theta)x - 1 \\ &= x^3 - \mathrm{tr}(A)x^2 + \mathrm{tr}(A)x - 1. \end{aligned}$$

Now we note that $p(A) = 0$, and the desired identity is established. What was discovered through analysis has been more generally supported by algebra.

Geometry In this section a geometric explanation will be presented for the result established algebraically above. Paraphrased, this result states that for any vector v, $A\mathrm{v} + A^T\mathrm{v} + [1 - \mathrm{tr}(A)]\mathrm{v}$ lies on the axis of rotation A. Assume that A represents a rotation of space through an angle ϕ about a fixed axis. (Note here that no connection has been established between ϕ and θ at this point.) To simplify notation, identify vectors with points in space in the usual way, and perform vector operations on points accordingly. Thus, given a point R, we apply the rotation A to find $S = A(R)$ and the inverse rotation A^T to find $T = A^T(R)$. The points R, S, and T all lie on a cone whose axis is the axis of rotation, and with vertex at the origin. This situation is illustrated in FIGURE 1 as a perspective view, and in FIGURES 2 and 3 as top and side views, respectively.

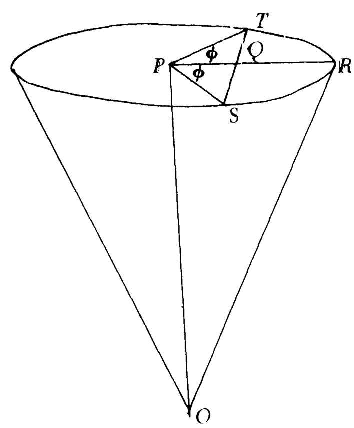

FIGURE 1

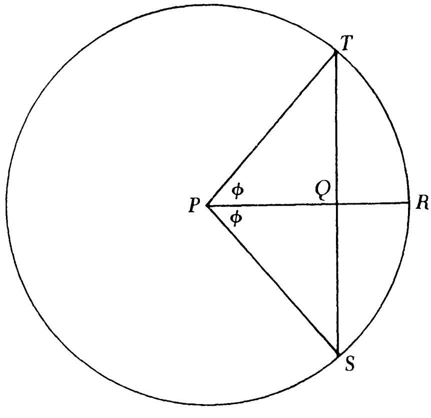

FIGURE 2

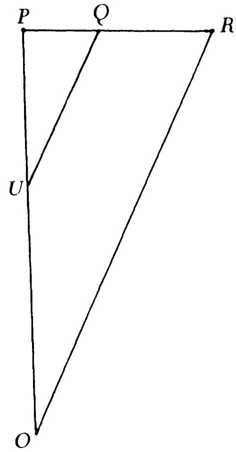

FIGURE 3

Let $Q = .5(S + T)$, the midpoint of segment ST. Clearly, some vector parallel to R can be drawn at Q so that it terminates at a point U on the axis of rotation. In fact, this vector must have length UQ and be parallel to unit vector $-R/(OR)$, hence it is given by $-(UQ/OR)R$. From FIGURE 3, the ratio UQ/OR is equal to PQ/PR. Now the vector from point Q to point U is given by the vector difference $U - Q$. Thus, with $f = PQ/PR$, we have $U - Q = -fR$ or $U = Q - fR$. To relate f to the angle ϕ, observe in FIGURE 2 that $PR = PS$ giving $f = PQ/PS = \cos\phi$. Combining these results produces $U = .5(S + T) - \cos\phi R$ as a point on the axis of rotation. Furthermore, $2U$ is also on the axis of rotation, and is given by $2U = S + T - 2\cos\phi R = (A + A^T - 2\cos\phi)R$. It remains but to show that $-2\cos\phi = 1 - \mathrm{tr}(A)$ and the geometric construction will reestablish the result of the preceding section.

Since $\mathrm{tr}(A)$ is invariant under similarity transformations, we may choose to represent the rotation relative to an orthonormal basis in which the third element lies on the axis of rotation. The corresponding matrix is easily seen to be

$$\begin{bmatrix} \cos\phi & \pm\sin\phi & 0 \\ \mp\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where the ambiguous signs depend on the direction of the rotation. Regardless, the trace is evidently $1 + 2\cos\phi$, as required.

Classroom presentations In describing this material to you, the reader, quite a bit of background has been presented or assumed: characterization of rotations as unit determinant orthogonal matrices, the Cayley-Hamilton theorem, invariance of $\mathrm{tr}(A)$ under changes of basis, etc. Depending on the background knowledge of the students involved, some modifications may be required for classroom presentation. One possible approach is to assume from the outset that A is a geometric rotation about a fixed axis. The existence of a unique one-dimensional eigenspace is then evident. If need be, the algebraic part of the discussion can be omitted in favor of passing directly from the cross product argument to the geometric construction. Indeed, one may even leave the connection between $\cos\phi$ and $\mathrm{tr}(A)$ unproved and use the geometric discussion as a plausibility argument. The most general version of the result could then be established for vectors **i** and **j** by using cross products, and extended to all vectors by linearity. This approach can also be assigned as an exercise. At the other extreme, with sufficient background, the students should be able to follow the development presented here. For these students especially, this topic provides a simple example of the interplay between various approaches to a problem, and illustrates one way that mathematical discoveries are propagated.