NASA TECHNICAL NOTE

NASA TN D-4696

NASA TN D-4696

A VECTOR APPROACH TO THE ALGEBRA OF ROTATIONS WITH APPLICATIONS

by Paul B. Davenport

Goddard Space Flight Center

Greenbelt, Md.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • AUGUST 1968

NASA TN D-4696

A VECTOR APPROACH TO THE ALGEBRA OF ROTATIONS WITH APPLICATIONS

By Paul B. Davenport

Goddard Space Flight Center
Greenbelt, Md.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

ABSTRACT

A new type of vector multiplication defines a group (whose elements are vectors) isomorphic to the group of rotations. This allows a vector representation of rotations which has many advantages over the usual matrix or Eulerian angles approach—e.g., this vector representation avoids the need for trigonometric relationships and requires only three independent parameters. Several applications show the simplicity of this vector representation—in particular, an analytic solution to a least-squares rotation problem. The differential equations defining the motion of a rigid body are obtained in terms of a vector differential equation.

CONTENTS

A VECTOR APPROACH TO THE ALGEBRA OF ROTATIONS WITH APPLICATIONS

by

Paul B. Davenport

Goddard Space Flight Center

INTRODUCTION

The algebra of rotations is generally approached through matrix algebra, where the matrix of the rotation is defined by the product of simple rotation matrices. Each of the simple rotations is about a coordinate axis and is uniquely determined by an angle. The most general rotation is uniquely defined by three angles (Eulerian angles), provided that the axis of each rotation is known. The convention of defining Eulerian angles varies considerably in the literature, so that, when only three angles are known, the product matrix is ambiguous. Furthermore, the angular or matrix approach to rotations requires the evaluation of trigonometric functions and therefore the use of tables or a computer. The approach here is to define a group (whose elements are three-dimensional vectors) isomorphic to the group of rotation matrices. The algebra of rotations can then be represented by the algebra of vectors (dot and cross product plus a new vector product which will be defined later), and does not require the evaluation of trigonometric functions.

A comment on notation: matrices and vectors are denoted by capital English letters and their elements by small English letters with subscripts. All vectors are considered as column vectors; a superscript T denotes the transpose of a vector or matrix. A superscript -1 denotes the inverse matrix. Primes denote the transform of a vector. Vector notation is used merely to simplify the algebraic relationships existing among the *components* of various vectors, and is not intended to suggest a physical interpretation. Even though two vectors may represent the same physical quantity, they are not regarded as equal unless they are expressed in the same coordinate system. Similarly, the notation $U = V \times W$ may be used (even though V and W are expressed in different coordinate systems) merely to indicate that the components of U are formed from the components of V and W according to the standard equations for cross products. It will be assumed throughout that all coordinate systems are right-handed orthonormal systems.

VECTORS DEFINING ROTATIONS AND THEIR ALGEBRA

If R is the matrix of a rotation, then References 1 and 2 show that R may be expressed by

$$R_x(\theta) = \cos\theta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (1-\cos\theta) \begin{pmatrix} x_1^2 & x_1 x_2 & x_1 x_3 \\ x_1 x_2 & x_2^2 & x_2 x_3 \\ x_1 x_3 & x_2 x_3 & x_3^2 \end{pmatrix} + \sin\theta \begin{pmatrix} 0 & x_3 & -x_2 \\ -x_3 & 0 & x_1 \\ x_2 & -x_1 & 0 \end{pmatrix}, \quad (1)$$

where $X = (x_1, x_2, x_3)^T$ is a unit vector defining the axis of rotation and θ is the angle of rotation. Except when otherwise stated, it is assumed that X has been selected so that $0 \le \theta \le \pi$; this is always possible, since $R_x(\theta) = R_{-x}(-\theta)$ and $R_x(\pi + \theta) = R_{-x}(\pi - \theta)$.

The following expression for R may also be found in Reference 1:

$$R = \left(B^T\right)^{-1} B, \tag{2}$$

where

$$B = \begin{pmatrix} 1 & x_3 \tan\frac{\theta}{2} & -x_2 \tan\frac{\theta}{2} \\ -x_3 \tan\frac{\theta}{2} & 1 & x_1 \tan\frac{\theta}{2} \\ x_2 \tan\frac{\theta}{2} & -x_1 \tan\frac{\theta}{2} & 1 \end{pmatrix} \tag{3}$$

In order to eliminate the angle in Equations 1 and 3, we assign a length (which is a function of θ) to X, in such a way that the sine and cosine of θ can be determined by this length. Many such functions exist; we choose $\tan\theta/2$ and $\sin\theta/2$ as examples. Thus, let

$$Y = \tan\frac{\theta}{2} X, \qquad Z = \sin\frac{\theta}{2} X, \qquad \text{where} \qquad 0 \le \theta \le \pi;$$

by standard trigonometric identities we obtain:

$$\sin\frac{\theta}{2} = \frac{\sqrt{Y^2}}{\sqrt{1 + Y^2}} = \sqrt{Z^2},$$

$$\cos\frac{\theta}{2} = \frac{1}{\sqrt{1 + Y^2}} = \sqrt{1 - Z^2},$$

$$\sin\theta = \frac{2\sqrt{Y^2}}{1 + Y^2} = 2\sqrt{Z^2\left(1 - Z^2\right)},$$

$$\cos\theta = \frac{1 - Y^2}{1 + Y^2} = 1 - 2Z^2.$$

Hence the definition of R as given by Equation 1, using Y and Z, alternatively, becomes

$$R = \frac{1}{1 + Y^2}\left[\left(1 - Y^2\right)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 2\begin{pmatrix} y_1^2 & y_1 y_2 & y_1 y_3 \\ y_1 y_2 & y_2^2 & y_2 y_3 \\ y_1 y_3 & y_2 y_3 & y_3^2 \end{pmatrix} + 2\begin{pmatrix} 0 & y_3 & -y_2 \\ -y_3 & 0 & y_1 \\ y_2 & -y_1 & 0 \end{pmatrix}\right]. \tag{4}$$

and

$$R = \left(1 - 2Z^2\right)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 2\begin{pmatrix} z_1^2 & z_1 z_2 & z_1 z_3 \\ z_1 z_2 & z_2^2 & z_2 z_3 \\ z_1 z_3 & z_2 z_3 & z_3^2 \end{pmatrix} + 2\sqrt{1 - Z^2}\begin{pmatrix} 0 & z_3 & -z_2 \\ -z_3 & 0 & z_1 \\ z_2 & -z_1 & 0 \end{pmatrix} . \qquad (5)$$

If B is expressed as a function of Y, Equation 3 becomes

$$B = \begin{pmatrix} 1 & y_3 & -y_2 \\ -y_3 & 1 & y_1 \\ y_2 & -y_1 & 1 \end{pmatrix} . \qquad (6)$$

Thus, every vector Y defines a rotation matrix by Equation 4, and every vector Z with $Z^2 \leq 1$ defines a rotation matrix by Equation 5. Hence Equations 4 and 5 define mappings that map sets of vectors into the set of rotations. We have yet to verify that the mapping is "onto," i.e., that for every rotation matrix there exist vectors Y and Z such that Equations 4 and 5 define the matrix of the rotation.

Let R be the matrix (with elements r_{ij}) of a rotation. Separating R into symmetric and skew-symmetric parts gives

$$R = \frac{1}{2}\left(R + R^T\right) + \frac{1}{2}\left(R - R^T\right) .$$

Comparing this with Equation 5 gives the following relations:

$$3 - 4Z^2 = \sigma , \qquad (7)$$

(where σ is the trace of R)

$$2\sqrt{1 - Z^2}\, Z = \frac{1}{2}\begin{pmatrix} r_{23} - r_{32} \\ r_{31} - r_{13} \\ r_{12} - r_{21} \end{pmatrix} , \qquad (8)$$

$$\left.\begin{aligned} 2z_2 z_3 &= \frac{1}{2}\left(r_{32} + r_{23}\right) , \\ 2z_1 z_3 &= \frac{1}{2}\left(r_{13} + r_{31}\right) , \\ 2z_1 z_2 &= \frac{1}{2}\left(r_{21} + r_{12}\right) , \end{aligned}\right\} \qquad (9)$$

$$1 - 2\left(z_2^2 + z_3^2\right) = r_{11} ,$$

$$1 - 2\left(z_1^2 + z_3^2\right) = r_{22} ,$$

$$1 - 2\left(z_1^2 + z_2^2\right) = r_{33} .$$

These last three equations imply

$$\left.\begin{aligned} z_1^2 &= \left(1 + r_{11} - r_{22} - r_{33}\right)/4 , \\ z_2^2 &= \left(1 + r_{22} - r_{11} - r_{33}\right)/4 , \\ z_3^2 &= \left(1 + r_{33} - r_{11} - r_{22}\right)/4 . \end{aligned}\right\} \tag{10}$$

Thus, to express R in the form of Equation 5 demands that Z satisfy nine conditions; by using the well-known properties of rotation matrices, it is easy to verify that these nine conditions are consistent.

Equations 7 and 8 give a convenient expression for Z:

$$Z^2 = \frac{3 - \sigma}{4} ,$$

$$Z = \frac{1}{2\sqrt{1 + \sigma}} \begin{pmatrix} r_{23} - r_{32} \\ r_{31} - r_{13} \\ r_{12} - r_{21} \end{pmatrix} , \tag{11}$$

provided that $Z^2 \neq 1$. If $Z^2 = 1$, then Equations 9 and 10 may be used to determine Z. In this case, however, Z is not unique; if it satisfies Equations 9 and 10, so does -Z. In practice, this ambiguity is of no consequence; because, if $Z^2 = 1$, Z and -Z yield the same matrix.

From Equation 1, we find that $\sigma = 1 + 2\cos\theta$, which by Equation 11 implies that $Z^2 \leq 1$.

Hence the mapping defined by Equation 5 is a mapping of three-dimensional vectors over the field of real numbers whose length is less than or equal to unity (denote this set of vectors by ζ) *onto* the group of rotation matrices. The mapping is also one-to-one except when $Z^2 = 1 (\sigma = -1)$.

The vector Y may be obtained, like the vector Z, directly from the matrix R or indirectly from Z itself, by the relationship

$$Y = \frac{1}{\sqrt{1 - Z^2}} Z . \tag{12}$$

Thus,

$$Y = \frac{1}{1+\sigma}\begin{pmatrix} r_{23} - r_{32} \\ r_{31} - r_{13} \\ r_{12} - r_{21} \end{pmatrix} . \tag{13}$$

The vector Y is undefined when $\sigma = -1$. This singularity may be removed if we agree to allow vectors of infinite magnitude whose direction is given by a unit vector, X, say. For such vectors, Equation 4 becomes

$$R = -I + 2XX^T . \tag{14}$$

With this convention, Y may be obtained from the rotation matrix R with trace -1, if it is noted that X = Z in this special case. Hence, X may be obtained from Equations 9 and 10.

Let η denote the set of all real three-dimensional vectors augmented by the vectors of infinite magnitude discussed above. Then Equations 4 and 14 define a mapping from η onto the group of rotation matrices. Thus, either of the two sets ζ or η may be used to parametrize the group of rotations. However, when vectors are used for this purpose it should be emphasized that they correspond to transformations. Hence, it is their algebraic properties as transformations that concern us and not their properties as vectors (indeed, ζ is not even a vector space). It does not matter that two elements of ζ or η may be equal in the sense of equality of transformations but not equal in the normal sense of vector equality. On the other hand, the abundance of vector operations (and their algebraic properties) that are normally employed to simplify relationships between components of vectors can also be applied to vectors corresponding to a rotation. What makes the vector parametrization appealing is the fact that the algebra of vectors as rotations can be expressed in terms of standard vector notation. In order to distinguish vectors used to parametrize rotations from ordinary vectors transformed by rotations, we introduce a new type of vector, which is obtained by defining a new equivalence relation among real three-dimensional vectors. Definition: Given a set δ of real three-dimensional vectors and a mapping τ that maps δ onto the group of rotation matrices; then two elements of δ are said to be equivalent if they map into the same rotation matrix. For the two sets ζ and η defined above, vector equivalence is the same as vector equality except for vectors mapping into rotation matrices with trace of -1.

Clearly, vector equivalence as defined above is an equivalence relation and separates the sets ζ and η into disjoint classes. These equivalence classes are merely the inverse images under τ of the rotation matrices. We shall call these equivalence classes "rotation vectors" and denote them in the same manner as ordinary vectors. The algebra usually associated with vectors is also applicable to rotation vectors. The well-known operations of scalar multiplication, vector addition, and dot and cross products are performed on rotation vectors in the classical manner, and the ordinary symbolism is used to denote these operation. When there is more than one vector in the equivalence class, the desired result of an operation can be that obtained by using either vector. However, the same vector must be used throughout any one expression.

Given two rotation vectors Y_1 and Y_2, Equation 4 defines two rotation matrices, say R_1 and R_2, respectively. It is well known that $R = R_2 R_1$ is also the matrix of a rotation. Thus, by a previous discussion there exists a rotation vector Y that defines R. This vector can be obtained from Equation 13 by forming the product $R_2 R_1$ in terms of Y_1 and Y_2, which gives the remarkably simple expression

$$Y = \frac{1}{1 - Y_1 \cdot Y_2} \left(Y_1 + Y_2 + Y_1 \times Y_2\right) . \tag{15}$$

If Z_1 and Z_2 are members of ζ defining R_1 and R_2, respectively, then the $Z \epsilon \zeta$ defining R, such that $R = R_2 R_1$, may be obtained in a similar manner, i.e., by comparing the trace and skew-symmetric part of $R_2 R_1$ with the trace σ and skew-symmetric part of R, respectively, as given by Equation 5. This gives

$$\sigma = 3 - 4Z^2 = 3 - 4Z_0^2 ,$$

$$\sqrt{1 - Z^2}\, Z = \left(\sqrt{1 - Z_1^2}\sqrt{1 - Z_2^2} - Z_1 \cdot Z_2\right) Z_0 ,$$

where

$$Z_0 = \sqrt{1 - Z_2^2}\, Z_1 + \sqrt{1 - Z_1^2}\, Z_2 + Z_1 \times Z_2 .$$

Thus

$$1 - Z^2 = 1 - Z_0^2 = \left[\sqrt{1 - Z_1^2}\sqrt{1 - Z_2^2} - Z_1 \cdot Z_2\right]^2 ,$$

and

$$\frac{\sqrt{1 - Z_1^2}\sqrt{1 - Z_2^2} - Z_1 \cdot Z_2}{\sqrt{1 - Z^2}} = \pm 1 ,$$

where the sign depends only on the numerator, since $\sqrt{1 - Z^2} \geq 0$ by definition, i.e., $\sqrt{1 - Z^2} = \cos \theta/2$ where $0 \leq \theta \leq \pi$. Hence, using the signum function, $\mathrm{sgn}(x)$, which is defined as $+1$, -1 according as x is positive or negative, respectively (here zero is considered positive or negative), leads to

$$Z = \left[\mathrm{sgn}\left(\sqrt{1 - Z_1^2}\sqrt{1 - Z_2^2} - Z_1 \cdot Z_2\right)\right] Z_0 . \tag{16}$$

The above argument does not guarantee that the Z as given by Equation 16 corresponds to $R = R_2 R_1$ when $Z^2 = 1$. However, it is straightforward though tedious to verify that Equation 16 does indeed give the Z defining $R = R_2 R_1$ for all Z_1 and Z_2 contained in ζ corresponding to R_1 and R_2, respectively.

Actually, Equation 16 may be obtained more readily by combining Equations 12 and 15 and the identity

$$Z = \frac{1}{\sqrt{1 + Y^2}} Y, \tag{17}$$

which is a useful formula since it is also valid in the limit as Y^2 approaches infinity.

Equations 15 and 16 suggest a new type of vector product. Definition: Let δ be a set of vectors, let $*$ be a binary operation on δ, and let τ be a mapping of δ onto the group of rotation matrices. Then $*$ is said to be a *rotation product* if it is preserved by τ, i.e., if $\tau(V * W) = \tau(V)\,\tau(W)$, for every V and W in δ. Thus for Z_1, Z_2 in ζ, the product $Z = Z_2 * Z_1$ given by Equation 16 is a rotation product. In a similar vein, for Y_1, Y_2 in η, Equation 15 defines a rotation product, except when $Y_1 \cdot Y_2 = 1$, or when either of the rotation vectors has infinite magnitude. In these exceptional cases, we define the product $Y_2 * Y_1$ by forming Z_1 and Z_2 by Equation 17, $Z_2 * Z_1$ by Equation 16, and then $Y_2 * Y_1$ by Equation 12.

Let η' and ζ' denote the two sets of rotation vectors (equivalence classes) defined by vector equivalence and the sets η and ζ, respectively. Then each of the two sets η' and ζ', together with its rotation product, forms a group. Clearly, η' is closed under the rotation product. Also, since Equation 17 implies that $Z = Z_2 * Z_1$ has length less than or equal to unity, ζ' is closed. The associativity of the rotation product is easily verified. The identity of each set is the null rotation vector, and the inverse of V (in either set) is $-V$. Furthermore, the respective mappings of the two sets defined herein are isomorphisms onto the group of rotation matrices. Thus I (the identity rotation matrix) is the image of the null rotation vector; R^{-1} is the image of $-V$ if R is the image of V. Actually, a one-to-one relationship between vectors and rotations may be defined without the concept of equivalence class. For example, when $\sigma = -1$, additional conditions may be imposed on Z to ensure uniqueness. The choice of additional constraints, however, depends on the preference of the user (a situation analogous to the many definitions of Eulerian angles). Thus, we introduced vector equivalence to emphasize that for the purpose of representing a rotation it does not matter which vector one selects out of an equivalence class for a 180-degree rotation.

The group of rotation transformations can now be represented by either of the two groups, ζ' or η'. This representation has certain advantages over the usual matrix representation, stemming from the fact that the rotation is defined by three independent parameters, without recourse to trigonometry (the matrix approach requires nine elements; or the evaluation of six trigonometric functions plus two matrix multiplications). In many applications the vector representation requires fewer calculations than the matrix representation. For example, matrix multiplication requires 27 scalar multiplications plus 18 additions; the rotation product of Equation 15 requires only 13 multiplications and 10 additions.

The choice of vector representation depends on the application. The formulas associated with the Y rotation vector are generally quite simple, requiring no radicals or trigonometric functions. Thus this vector representation is a valuable tool for hand calculations or for deriving theoretical results, but has the inconvenience of becoming infinite for all 180-degree rotations. The Z-vector representation, on the other hand, is always finite, always defined (uniquely, except for its sign at 180-degree rotations), and the rotation product is valid for all rotations; but the existence of the radical is inconvenient for hand calculations. The vector representation $W = \tan \theta/4X$ may prove to be useful since it combines some of the assets of both the Y and Z. In this case we have

$$R = \frac{1}{\left(1 + W^2\right)^2} \left[\left(W^4 - 6W^2 + 1\right) I + 8WW^T + 4\left(1 - W^2\right) \begin{pmatrix} 0 & w_3 & -w_2 \\ -w_3 & 0 & w_1 \\ w_2 & -w_1 & 0 \end{pmatrix} \right],$$

$$Y = \frac{2}{1 - W^2} W, \qquad Z = \frac{2}{1 + W^2} W,$$

$$W = \frac{1}{1 + \sqrt{1 + Y^2}} Y = \frac{1}{1 + \sqrt{1 - Z^2}} Z.$$

COORDINATES OF A ROTATED VECTOR

A rotation matrix is frequently used to find the coordinates of a vector after rotation. The vector representation yields a useful formula for this application. Let V' be the image of V under the rotation and R the matrix of the rotation (i.e., $V' = RV$). If Y and Z are the rotation vectors defining R, then Equations 4, 5, and 14 give V' as

$$V' = YV = \begin{cases} \frac{1}{1 + Y^2} \left[\left(1 - Y^2\right) V + 2(V \cdot Y) Y + 2V \times Y\right], & Y^2 < \infty \\ -V + 2(V \cdot X) X & Y^2 = \infty \end{cases}$$

$$V' = ZV = \left(1 - 2Z^2\right) V + 2(V \cdot Z) Z + 2\sqrt{1 - Z^2}\, V \times Z.$$

Note that the symbolism YV denotes the image of V under the rotation corresponding to Y. Thus the above equations define a vector multiplication in the same sense that RV defines a matrix multiplication.

ROTATIONS DETERMINED BY A VECTOR AND ITS IMAGE

In many instances we are given two vectors (V and V') and wish to determine the matrix R such that $V' = RV$. If R is to be a rotation matrix, then we must have $V'^2 = V^2$. For simplicity, we assume that $V'^2 = V^2 = 1$. The rotation taking V into V' is not unique; however, the vector representation of the "shortest path" rotation is immediately obvious—the axis of rotation is collinear with

$V' \times V$ and the angle of rotation is $\cos^{-1} V \cdot V'$. Thus

$$Y = \frac{1}{1 + V \cdot V'} V' \times V, \qquad V \cdot V' \neq -1,$$

$$Z = \frac{1}{\sqrt{2(1 + V \cdot V')}} V' \times V, \qquad V \cdot V' \neq -1.$$

If $V \cdot V' = -1$, then Z is any vector satisfying the two conditions $Z^2 = 1$ and $Z \cdot V = 0$; Y has infinite magnitude with direction defined by Z.

ROTATIONS DETERMINED BY TWO VECTORS AND THEIR IMAGES

A common practice in orbit theory is to construct a rotation that takes the xy-plane into the plane of the orbit and takes the x-axis into the direction of perigee. This is a particular example of the following problem: Given V_1, V_2 (V_1 and V_2 noncollinear), V_1', and V_2' such that $V_1'^2 = V_1^2$, $V_2'^2 = V_2^2$, and $V_1' \cdot V_2' = V_1 \cdot V_2$, find the rotation that takes V_1 into V_1' and V_2 into V_2'. Here again, the vector representation of the rotation yields a simple solution.

If $V' = RV$, then Equation 2 implies that $B^T V' = BV$. From Equation 6 and by matrix multiplication, this condition can be written as

$$V' + Y \times V' = V - Y \times V,$$

or

$$V - V' = Y \times (V + V'),$$

where Y is the vector representation of R.

Thus, the conditions $V_1' = RV_1$ and $V_2' = RV_2$ may be expressed by

$$V_1 - V_1' = Y \times (V_1 + V_1') \qquad \text{and} \qquad V_2 - V_2' = Y \times (V_2 + V_2').$$

Let

$$A_i = V_i - V_i', \qquad B_i = V_i + V_i', \qquad (i = 1, 2).$$

Then the condition equations become

$$A_i = Y \times B_i, \qquad (i = 1, 2),$$

where $A_i \cdot B_i = 0$ and $A_1 \cdot B_2 = -A_2 \cdot B_1$. It is immediately obvious that $Y \cdot A_i = 0$. If each side of the first equation is cross-multiplied by the vector A_2, we obtain

$$A_2 \times A_1 = A_2 \times (Y \times B_1) = (A_2 \cdot B_1)\,Y - (A_2 \cdot Y)\,B_1 = (A_2 \cdot B_1)\,Y ;$$

thus, if $A_2 \cdot B_1 = -A_1 \cdot B_2 \neq 0$, we obtain a simple expression for Y, namely,

$$Y = \frac{1}{V_1 \cdot V_2' - V_2 \cdot V_1'} (V_1 - V_1') \times (V_2 - V_2') . \tag{18}$$

A more general expression may be obtained as follows. Cross-multiply each side of the ith equation by B_i; this gives

$$B_i \times A_i = B_i^2 Y - (B_i \cdot Y)\,B_i , \qquad (i = 1, 2) . \tag{19}$$

Therefore,

$$B_2 \cdot (B_1 \times A_1) = B_1^2 B_2 \cdot Y - (B_1 \cdot B_2)\,B_1 \cdot Y ,$$

$$B_1 \cdot (B_2 \times A_2) = B_2^2 B_1 \cdot Y - (B_1 \cdot B_2)\,B_2 \cdot Y ,$$

and solving these two linear equations for $B_1 \cdot Y$ and $B_2 \cdot Y$ yields

$$B_1 \cdot Y = \frac{(B_1 \times B_2) \cdot \left[B_1^2 A_2 - (B_1 \cdot B_2)\,A_1\right]}{(B_1 \times B_2)^2} ,$$

$$B_2 \cdot Y = \frac{(B_1 \times B_2) \cdot \left[(B_1 \cdot B_2)\,A_2 - B_2^2 A_1\right]}{(B_1 \times B_2)^2} .$$

Substitution of these last expressions in Equation 19 gives the solution for Y, provided that $B_1 \times B_2 \neq 0$.

If $V_1' = V_1$ and/or $V_2' = V_2$, the Z rotation vector is obtained from Y by Equations 17 and 19 unless $B_1 \times B_2 = 0$; otherwise, Equations 17 and 18 provide a much simpler expression for Z, namely,

$$Z = \frac{\operatorname{sgn}(V_1 \cdot V_2' - V_2 \cdot V_1')}{\sqrt{(V_1 \cdot V_2' - V_2 \cdot V_1')^2 + \left[(V_1 - V_1') \times (V_2 - V_2')\right]^2}} (V_1 - V_1') \times (V_2 - V_2')$$

When $V_i' = V_i$ and $B_1 \times B_2 = 0$, then $Z = V_i$. The above expressions provide another way to make a vector representation of a rotation from the matrix of the rotation R: choose two independent vectors (V_1 and V_2) and set $V_1' = RV_1$, $V_2' = RV_2$.

Given V_1, V_2, V_1' and V_2' such that $V_1^2 = V_1'^2$, $V_2^2 = V_2'^2$ and $V_1 \cdot V_2 = V_1' \cdot V_2'$, the matrix of the rotation, R, that takes V_1 into V_1' and V_2 into V_2' can be determined from the vector representation obtained above and Equation 4 or 5. A more direct approach is to construct two right-handed orthonormal coordinate systems; one from vectors V_1 and V_2, and the other from vectors V_1' and V_2'. The rotation taking the first system into the second will then take V_i into V_i'. The matrix of this rotation can be easily written as the product of two matrices, each of whose rows or columns are formed from the components of the constructed axes relative to some underlying coordinate system.

$$
\begin{aligned}
U_1 &= V_1/|V_1| \, , & U_1' &= V_1'/|V_1'| \, , \\
W &= \left[V_2 - \frac{(V_1 \cdot V_2)}{V_1^2} V_1\right] , & W' &= \left[V_2' - \frac{(V_1' \cdot V_2')}{V_1'^2} V_1'\right] , \\
U_2 &= W/|W| \, , & U_2' &= W'/|W'| \, , \\
U_3 &= U_1 \times U_2 \, , & U_3' &= U_1' \times U_2' \, , \\
R_1 &= (U_1, U_2, U_3) \, , & R_2 &= (U_1', U_2', U_3') \, ;
\end{aligned}
$$

then

$$R = R_2 R_1^{-1} \, .$$

A LEAST-SQUARES ROTATION—DETERMINATION OF ATTITUDE*

Consider a vehicle with a local coordinate system that has been rotated from a fixed coordinate system. If the vehicle has equipment that can find the direction (relative to the local coordinate system) of a point whose direction relative to the fixed coordinate system is known, then the rotation relating the local and fixed systems can be determined by two or more observations of such points.† If the directions were exact, then the rotation relating the two coordinate systems could be obtained from any two noncollinear observations as described in the previous section. In practice, however, the directions are not exact and in such cases we generally seek the "least-squares" solution.

For the case at hand, we seek a rotation matrix R such that the scalar function

$$\phi(R) = \sum_{i=1}^{n} (W_i - RV_i)^2$$

*This section is a solution to Problem 65-1 by Grace Wahba in Reference 3.

†The Orbiting Astronomical Observatory has this capability, where the points are known stars.

is a minimum. Here, V_i and W_i are vectors defining a point relative to the fixed and local coordinate systems, respectively. V_i and W_i need not be unit vectors; their length may indicate the reliability of the measurement. The matrix R—thus also the function $\phi(R)$—has only three independent parameters. The rotation that makes $\phi(R)$ a minimum is a solution of the three equations obtained by setting the partial derivatives of ϕ with respect to each independent parameter to zero (assuming that the derivatives do exist). The use of Eulerian angles as independent parameters leads to three condition equations that are complicated and tedious to derive and must be dealt with as three scalar equations. On the other hand, the condition equations in terms of a vector representation of R can be expressed as a single vector equation, easily derived and solvable by vector and matrix algebra.

Since R is a rotation matrix, $\phi(R)$ may be written as

$$\phi(R) = \sum_{i=1}^{n} \left[V_i^2 + W_i^2 - 2W_i \cdot (RV_i)\right] ,$$

which as a function of $Y \epsilon \eta$ gives

$$\phi(R) = \sum_{i=1}^{n} \left\{V_i^2 + W_i^2 - \frac{2}{1+Y^2}\left[(1-Y^2)\, V_i \cdot W_i + 2(V_i \cdot Y)(W_i \cdot Y) + 2(W_i \times V_i) \cdot Y\right]\right\} , \tag{20a}$$

when $Y^2 < \infty$, and

$$\phi(R) = \sum_{i=1}^{n} \left\{V_i^2 + W_i^2 - 2\left[-V_i \cdot W_i + 2(V_i \cdot X)(W_i \cdot X)\right]\right\} , \tag{20b}$$

(where $X^2 = 1$) when $Y^2 = \infty$.

In the first case,

$$\frac{\partial \phi}{\partial y_j} = \frac{-4}{(1+Y^2)^2} \sum_{i=1}^{n} \left\{2\left[(V_i \times W_i) \cdot Y - (V_i \cdot Y)(W_i \cdot Y) - V_i \cdot W_i\right] y_j \right.$$

$$\left. + (1+Y^2)\left[(V_i \cdot Y)\, w_{ij} + (W_i \cdot Y)\, v_{ij} - (V_i \times W_i)_j\right]\right\} ,$$

where the subscript j refers to the jth component of the vector. Thus a necessary condition for ϕ to have a minimum (at least among the rotations of less than 180-degrees) at some Y is given by the vector equation

$$2\left[\sum(W_i \times V_i) \cdot Y + (V_i \cdot Y)(W_i \cdot Y) + V_i \cdot W_i\right] Y = (1 + Y^2)\sum \left[(V_i \cdot Y) W_i + (W_i \cdot Y) V_i + W_i \times V_i\right] . \quad (21)$$

(For convenience, the range of summation is omitted hereinafter; the letter i will denote the summation index.)

Taking the dot product of each side with Y gives

$$2\sum(V_i \cdot Y)(W_i \cdot Y) = \sum\left[(W_i \times V_i) \cdot Y + 2V_i \cdot W_i\right] Y^2 - \sum (W_i \times V_i) \cdot Y .$$

When this last expression is substituted in the vector equation and the factor $1 + Y^2$ is divided out, we obtain

$$\sum\left\{\left[(W_i \times V_i) \cdot Y + 2V_i \cdot W_i\right] Y - \left[(V_i \cdot Y) W_i + (W_i \cdot Y) V_i + W_i \times V_i\right]\right\} = 0 . \quad (22)$$

The same substitution in Equation 20a gives

$$\phi(R) = \sum \left(V_i^2 + W_i^2 - 2V_i \cdot W_i\right) - 2Y \cdot \sum W_i \times V_i , \quad (23)$$

when Y satisfies Equation 22. Hence, to minimize ϕ we must take the solution of Equation 22 that makes $Y \cdot \Sigma W_i \times V_i$ a maximum.

Equation 22 may be written in matrix notation as follows:

$$\left(A^T YI + B\right) Y = A , \quad (24)$$

where I is the identity matrix, A is the vector $\Sigma W_i \times V_i$, and B is a symmetric matrix, with elements

$$b_{jk} = -\sum\left(v_{ij} w_{ik} + v_{ik} w_{ij}\right) , \qquad j \neq k ,$$

$$b_{jj} = 2\sum\left(V_i \cdot W_i - v_{ij} w_{ij}\right) .$$

(To see this most easily, express Equations 22 and 24 in component form.)

If $A = 0$, then $Y = 0$ is the desired solution (there may be other solutions if $\det B = 0$, but the value of ϕ will be the same for all solutions). If $A \neq 0$, then the solutions may be obtained as outlined below. Multiplying each side of the matrix equation by the adjoint of $A^T YI + B$ gives

$$\left[\det\left(A^T YI + B\right)\right] Y = \left[\operatorname{adj}\left(A^T YI + B\right)\right] A ;$$

multiplication of this equation by A^T yields the scalar equation

$$\left[\det\left(A^T YI + B\right)\right] A^T Y = A^T \left[\operatorname{adj}\left(A^T YI + B\right)\right] A .$$

Denote the scalar $A^T Y$ by λ; denote the scalar functions $\det(\lambda I + B)$ and $A^T\left[\operatorname{adj}(\lambda I + B)\right] A$ by $f(\lambda)$ and $g(\lambda)$, respectively. Then the above scalar equation may be written as

$$\lambda f(\lambda) - g(\lambda) = 0 .$$

Note that the left-hand side, say $h(\lambda)$, is a fourth-degree polynomial in λ, and that $-f(-\lambda)$ is the characteristic polynomial of the symmetric matrix B.

The solutions to Equation 24 are obtained by determining the zeros of $h(\lambda)$ and solving the resulting linear equations. However, in the discussion following Equation 23, it was determined that the maximum value of $Y \cdot \Sigma\left(W_i \times V_i\right)$ leads to the minimum value of $\phi(Y)$. The maximum value of $Y \cdot \Sigma\left(W_i \times V_i\right)$ is the largest zero of $h(\lambda)$; denote this zero by λ_0.

Since $h(\lambda)$ is a fourth-degree polynomial, λ_0 may be obtained analytically; however, a numerical iterative solution is probably more practical, since the zeros of $h(\lambda)$ are easily bounded. Note that $\phi(Y)$ is a non-negative function; therefore, Equation 23 implies that

$$\lambda_0 \leq \frac{1}{2} \sum_{i=1}^{n} \left(V_i - W_i\right)^2 ,$$

which provides an upper bound.

Since B is symmetric, there exists an orthogonal matrix P, say, such that $P^{-1} BP$ is diagonal. Let Y' and U be vectors such that $Y = PY'$ and $A = PU$. Then, in terms of Y' and U, Equation 24 becomes

$$\left[\left(U^T Y'\right) I + B\right] PY' = PU ,$$

and a premultiplication by P^{-1} gives

$$\left[\left(U^T Y'\right) I + D\right] Y' = U .$$

where D is a diagonal matrix whose entries are merely the eigenvalues of B. (Without loss of generality, we may assume the eigenvalues to be arranged in increasing order, say $\lambda_1 \leq \lambda_2 \leq \lambda_3$.) Multiplying each side of the last equation by $\mathrm{adj}\left[\left(U^T Y'\right) I + D\right]$ and then by U^T yields the scalar equation

$$\det\left(U^T Y' I + D\right) U^T Y' = U^T \mathrm{adj}\left(U^T Y' I + D\right) U .$$

However, $U^T Y' = A^T Y = \lambda$ and D is diagonal, so this last equation may be written in the form

$$\lambda(\lambda + \lambda_1)(\lambda + \lambda_2)(\lambda + \lambda_3) = u_1^2(\lambda + \lambda_2)(\lambda + \lambda_3) + u_2^2(\lambda + \lambda_1)(\lambda + \lambda_3) + u_3^2(\lambda + \lambda_1)(\lambda + \lambda_2) .$$

The left-hand side of the above equation is $\lambda f(\lambda)$ and the right-hand side is $g(\lambda)$. We may now easily determine the sign of $g(\lambda)$ when $\lambda = -\lambda_3, -\lambda_2, -\lambda_1$, respectively (assuming $\lambda_1 \leq \lambda_2 \leq \lambda_3$), and thus also the sign of $h(\lambda) = \lambda f(\lambda) - g(\lambda)$ at these points. From these considerations, it can be shown that λ_0 is at least as great as $-\lambda_1$, and also that the largest zero of $g(\lambda)$ (provided that $g(\lambda)$ is not identically zero, i.e., $A \neq 0$) is less than or equal to $-\lambda_1$. Hence, a lower bound of our desired zero λ_0 is $-\lambda_1$, which can be determined only by solving a cubic. But the largest zero of $g(\lambda)$ is also a lower bound, and this can be found by solving a quadratic.

Thus far, we have obtained the minimum of $\phi(R)$ for rotations of less than 180 degrees. To obtain the minimum among all 180-degree rotations, we use the method of Lagrange for solving extremum problems with a constraint. The necessary conditions become

$$-4\sum(V_i \cdot X) w_{ij} + (W_i \cdot X) v_{ij} + 2\mu x_j = 0, \qquad j = 1, 2, 3 , \tag{25}$$

where μ is Lagrange's multiplier and X must be such that $X^2 = 1$. In terms of our previous notation, the above conditions may be collected into the single vector equation

$$\left[\left(\frac{\mu}{2} - 2\sum V_i \cdot W_i\right) I + B\right] X = 0 ,$$

where B is the symmetric matrix introduced in Equation 24. This last equation can be satisfied if and only if μ is a root of the cubic equation

$$\det\left[\left(\frac{\mu}{2} - 2\sum V_i \cdot W_i\right) I + B\right] = 0 .$$

But the roots of this equation are given by

$$\frac{\mu_k}{2} - 2\sum V_i \cdot W_i = -\lambda_k , \qquad k = 1, 2, 3 ,$$

where the λ_k are eigenvalues of B. The condition equations thus become

$$\left(B - \lambda_k I\right) X = 0 , \qquad X^2 = 1 .$$

and the solutions are the unit eigenvectors of B. The value of $\phi(R)$ at each of these solutions may be obtained by multiplying the jth equation of Equations 25 by x_j and adding the three equations together. This yields

$$\mu = 4\sum (V_i \cdot X)(W_i \cdot X) ,$$

which, when substituted in the expression for $\phi(R)$ (Equation 20b), gives

$$\phi(R) = \sum \left(V_i^2 + W_i^2 + 2V_i \cdot W_i\right) - \mu_k ,$$

$$= \sum (V_i - W_i)^2 + 2\lambda_k .$$

Thus the eigenvector (with unit length) corresponding to the minimum eigenvalue λ_1 gives the minimum of $\phi(R)$ for all 180-degree rotations.

Equation 23 gives the minimum of $\phi(R)$ for rotations other than 180 degrees:

$$\phi(R) = \sum (V_i - W_i)^2 - 2\lambda_0 ;$$

it was also shown that $\lambda_0 \geq -\lambda_1$. Hence, the rotation giving the minimum $\phi(R)$ among all rotations is

$$Y = \left(\lambda_0 I + B\right)^{-1} A , \qquad \text{when} \qquad f\left(\lambda_0\right) \neq 0 ,$$

or by a 180-degree rotation with axis of rotation X defined by

$$(\lambda_0 I + B) X = 0 , \qquad \text{and} \qquad X^2 = 1 ,$$

when $f(\lambda_0) = 0$, where λ_0 is the largest zero of $h(\lambda)$.

To avoid inverting a near-singular matrix and dealing with large values of the components of Y when $f(\lambda_0)$ is near zero, obtain the Z vector representation of R from Equation 17 and from the relationship

$$(\lambda_0 I + B)^{-1} = \frac{\operatorname{adj}(\lambda_0 I + B)}{\det(\lambda_0 I + B)} = \frac{\operatorname{adj}(\lambda_0 I + B)}{f(\lambda_0)} .$$

This gives

$$Z = \frac{\operatorname{sgn}\left[f(\lambda_0)\right]}{\sqrt{\left[f(\lambda_0)\right]^2 + A^T \left[\operatorname{adj}(\lambda_0 I + B)\right]^2 A}} \left[\operatorname{adj}(\lambda_0 I + B)\right] A , \qquad \text{when} \qquad f(\lambda_0) \neq 0 ,$$

and

$$(\lambda_0 I + B) Z = 0 \qquad \text{and} \qquad Z^2 = 1 , \qquad \text{when} \qquad f(\lambda_0) = 0 ,$$

where

$$\operatorname{adj}(\lambda I + B) = \lambda^2 I + \lambda \left[\operatorname{Tr}(B) I - B\right] + \operatorname{adj}(B) ,$$

$$f(\lambda) = \lambda^3 + \operatorname{Tr}(B)\lambda^2 + \operatorname{Tr}\left[\operatorname{adj}(B)\right] \lambda + \det(B) ,$$

$$g(\lambda) = A^T \left[\operatorname{adj}(\lambda I + B)\right] A ,$$

and $\operatorname{Tr}(B)$ denotes the trace of B. These last equations are easily verified by direct calculations.

If $n = 2$, $V_1^2 = V_2^2$, and $W_1^2 = W_2^2$, then the following simple procedure gives the least-squares solution. Let

$$U_1 = \frac{V_1 - V_2}{|V_1 - V_2|} , \qquad U_1' = \frac{W_1 - W_2}{|W_1 - W_2|} ,$$

$$U_2 = \frac{V_1 + V_2}{|V_1 + V_2|} , \qquad U_2' = \frac{W_1 + W_2}{|W_1 + W_2|} ,$$

and use the techniques of the preceding section to obtain the rotation that takes each U_i into U_i'. Such a rotation takes the plane determined by V_1 and V_2 into the plane determined by W_1 and W_2 and ensures that $W_1 \cdot RV_1 = W_2 \cdot RV_2$. That such a rotation is the least-squares rotation can be verified by expressing the V_i and W_i as functions of U_i and U_i'. It is then clear that $\phi(R)$ is a minimum when $U_i' = RU_1$.

We have also been very successful in solving Equation 21 in the cases where $n > 2$, by successive substitutions, i.e., using the iteration

$$Y_{j+1} = \frac{1 + Y_j^2}{2\sum_{i=1}^{n} \left[(W_i \times V_i) \cdot Y_j + (V_i \cdot Y_j)(W_i \cdot Y_j) + V_i \cdot W_i\right]} \sum_{i=1}^{n} \left[(V_i \cdot Y_j) W_i + (W_i \cdot Y_j) V_i + W_i \times V_i\right] .$$

Here, Y_j is the jth approximation to Y, and Y_0 is obtained as above using two of the V_i and their corresponding images W_i. In fact, in one case studied, the procedure converged even when the angle of rotation was as large as 179 degrees.

RELATIONS BETWEEN EULERIAN ANGLES, MATRIX OF ROTATION, AND VECTOR APPROACH

Since Eulerian angles have a wide usage (especially when the angles correspond to yaw, pitch, and roll) it may be convenient or necessary to transform the matrix or vector parametrization of a rotation into Eulerian angles. To do this, we must establish a convention or positive sense of rotation. Here, we assume that the matrices of the simple rotations about each of the coordinate axes are given by

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \qquad R_2(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix},$$

$$R_3(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $R_i(\theta)$ indicates a rotation of θ about the ith axis. The matrix of any rotation R can then be written as

$$R = R_k(\theta_3) R_j(\theta_2) R_i(\theta_1), \qquad \begin{cases} i = 1, 2, 3, \\ j = 1, 2, 3, \\ k = 1, 2, 3, \\ i \neq j \neq k . \end{cases}$$

If i, j, and k are distinct, then direct calculation shows that

$$\sin\theta_2 = \delta_{ijk} r_{ki}, \qquad \cos\theta_2 \sin\theta_1 = -\delta_{ijk} r_{kj},$$

$$\cos\theta_2 \cos\theta_1 = r_{kk}, \qquad \cos\theta_2 \sin\theta_3 = -\delta_{ijk} r_{ji},$$

$$\cos\theta_2 \cos\theta_3 = r_{ii},$$

where $\delta_{ijk} = 1$ if ijk is a cyclic permutation of 123, and $\delta_{ijk} = -1$ otherwise. Thus, if R is given, the angles are defined as follows:

$$\sin\theta_2 = \delta_{ijk} r_{ki},$$

$$\cos\theta_2 = \pm\sqrt{r_{ii}^2 + r_{ji}^2} = \pm\sqrt{r_{kk}^2 + r_{kj}^2};$$

if $\cos\theta_2 \neq 0$,

$$\sin\theta_1 = -\delta_{ijk} r_{kj}/\cos\theta_2, \qquad \cos\theta_1 = r_{kk}/\cos\theta_2,$$

$$\sin\theta_3 = -\delta_{ijk} r_{ji}/\cos\theta_2, \qquad \cos\theta_3 = r_{ii}/\cos\theta_2;$$

if $\cos\theta_2 = 0$, θ_1 and θ_3 are subject only to the conditions

$$\sin\left(\theta_3 \pm \delta_{ijk}\theta_1\right) = \delta_{ijk} r_{ij},$$

$$\cos\left(\theta_3 \pm \delta_{ijk}\theta_1\right) = \mp\delta_{ijk} r_{ik},$$

where the upper signs are taken if $\sin\theta_2 = 1$ and the lower signs when $\sin\theta_2 = -1$.

The factorization is not unique even when $\cos\theta_2 \neq 0$, since either choice of sign for $\cos\theta_2$ produces the same product matrix R.

To factorize R in the form

$$R = R_i\left(\theta_3\right) R_j\left(\theta_2\right) R_i\left(\theta_1\right)$$

where the first and last factors are of the same form, let

$$\sin\theta_2 = \pm\sqrt{r_{ij}^2 + r_{ik}^2} = \pm\sqrt{r_{ji}^2 + r_{ki}^2},$$

$$\cos\theta_2 = r_{ii};$$

if $\sin\theta_2 \neq 0$,

$$\sin\theta_1 = r_{ij}/\sin\theta_2 , \qquad \cos\theta_1 = \delta_{ji}\, r_{ik}/\sin\theta_2 ,$$

$$\sin\theta_3 = r_{ji}/\sin\theta_2 , \qquad \cos\theta_3 = -\delta_{ji}\, r_{ki}/\sin\theta_2 ,$$

where $\delta_{ji} = 1$ if ji is in natural cyclic order, and $\delta_{ji} = -1$ otherwise. If $\sin\theta_2 = 0$, θ_1 and θ_3 are subject only to the conditions

$$\sin(\theta_3 \pm \theta_1) = \delta_{ji}\, r_{kj} ,$$

$$\cos(\theta_3 \pm \theta_1) = \pm r_{kk} ,$$

where the plus sign is taken if $r_{ii} > 0$ and the minus sign if $r_{ii} < 0$.

Given Eulerian angles θ_1, θ_2, θ_3 as defined above, we can obtain the vector representation by forming the product matrix R, and then using the techniques of the previous sections. A more direct approach is to use Equations 15 or 16 twice, where

$$Y_i = \tan\frac{\theta}{2} E_i , \qquad Z_i = \sin\frac{\theta}{2} E_i , \qquad -\pi < \theta \leq \pi ,$$

are the vectors corresponding to $R_i(\theta)$, and E_i is the coordinate axis about which the rotation is taken. (If $\pi < \theta < 2\pi$, then the negatives of the above expressions for the range $-\pi < \theta < 0$ must be used.)

To obtain the Eulerian angles from the vector Z, use the expressions (derived from Equation 5):

$$r_{ij} = 2\left[z_i z_j \pm z_k \sqrt{1 - Z^2}\right] , \qquad i \neq j ,$$

$$r_{ii} = 1 - 2Z^2 + 2z_i^2 ,$$

where the upper sign is used when ij is in natural cyclic order and the lower sign otherwise. The angles are then obtained from the appropriate formulas above.

EQUATIONS OF MOTION OF A RIGID BODY

If $R(t)$ is the matrix of a rotation that defines the orientation of a coordinate system (attached to a moving rigid body) relative to a fixed coordinate system, then R satisfies the matrix differential equation

$$\dot{R}(t) = \Omega(t)\, R(t) , \qquad R(0) = I , \tag{26}$$

where Ω is a skew-symmetric matrix such that $\Omega V = V \times \omega(t)$ for all vectors V, and ω is the angular velocity vector (Reference 4). This equation, in the form

$$\dot{R}R^{-1} = \Omega ,$$

with $R = R_k(\theta_3) R_j(\theta_2) R_i(\theta_1)$, gives

$$\Omega = \dot{R}R^{-1} = R_k R_j \dot{R}_i R_i^{-1} R_j^{-1} R_k^{-1} + R_k \dot{R}_j R_j^{-1} R_k^{-1} + \dot{R}_k R_k^{-1} . \tag{27}$$

Although this is a matrix equation, it represents only three independent component equations, since each of the product matrices on the right is skew-symmetric. These three independent equations can be collected into a single vector equation by means of the well-known isomorphism between 3×3 skew-symmetric matrices and three-dimensional vectors,

$$S(V) = \begin{pmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{pmatrix} \leftrightarrow \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = V . \tag{28}$$

It is easy to verify that if R is a rotation matrix, then $RS(V)R^{-1} \leftrightarrow RV$. Also, from the definitions of R_ℓ, it is easy to show that $\dot{R}_\ell(\theta) R_\ell^{-1}(\theta) \leftrightarrow \dot{\theta} E_\ell$, where E_ℓ is the coordinate axis of rotation (this is also valid for rotations about any fixed line). For these reasons, and because $\Omega \leftrightarrow \omega$, Equation 27 is equivalent to

$$\begin{aligned} \omega &= \dot{\theta}_1 R_k(\theta_3) R_j(\theta_2) E_i + \dot{\theta}_2 R_k(\theta_3) E_j + \dot{\theta}_3 E_k \\ &= A\dot{\theta} , \qquad \theta(0) = 0 , \end{aligned} \tag{29}$$

where A is a matrix whose columns are the vectors $R_k R_j E_i$, $R_k E_j$, and E_k, respectively, and $\dot{\theta} = (\dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3)^T$.

The matrix differential equation has no singularities, but requires the integration of nine scalar functions. Equation 29, on the other hand, only involves three scalar functions, but the matrix A is singular when $\cos\theta_2 = 0$ for $k \neq i$, or when $\sin\theta_2 = 0$ for $k = i$. Thus, no set of Eulerian angles can be chosen so that $\dot{\theta}$ will be defined for all rotations. In fact, any set of Eulerian angles gives singularities for rotations as small as 90 degrees.

To obtain the equations of motion expressed in terms of the Y vector, we merely differentiate Equation 13 and make the proper substitutions using Equations 4 and 13 and the matrix differential equation. This gives

$$\dot{Y} = \frac{1}{2}\left[(\omega \cdot Y) Y + Y \times \omega + \omega\right] , \qquad Y(0) = 0 . \tag{30}$$

To solve for ω, cross-multiply each side by Y, and subtract the resulting equation from the original. Thus,

$$\frac{1}{2}\left(1+Y^2\right)\omega = \dot{Y} - Y \times \dot{Y} .$$

Differential Equation 30 has no singularities, but from the definition of Y we know that solutions involving 180-degree rotations will diverge to infinity. However, for many applications (involving only moderate displacements of the moving frame) this presents no difficulties. The differential equations in terms of the Z rotation vector can be obtained as they were in terms of the Y rotation vector. A more direct approach is to use the identity

$$Z = \frac{1}{\sqrt{1+Y^2}} Y , \tag{31}$$

whence,

$$\dot{Z} = \frac{1}{\sqrt{1+Y^2}} \dot{Y} - \frac{Y \cdot \dot{Y}}{\left(1+Y^2\right)\sqrt{1+Y^2}} Y . \tag{32}$$

Equation 30 gives

$$Y \cdot \dot{Y} = \frac{1}{2}\left(1+Y^2\right)\omega \cdot Y . \tag{33}$$

Combining Equations 30 through 33 gives

$$\dot{Z} = \frac{1}{2}\left[Z \times \omega + \sqrt{1-Z^2}\,\omega\right], \qquad Z(0) = 0 . \tag{34}$$

Conversely, let $Z(t)$ be a differentiable vector function such that over some interval (say, $t_0 < t < t_1$) $Z^2 < 1$ and Z satisfies Equation 34. Then $Z(t)$ defines a rotation matrix $R(t)$ by Equation 5, and $\dot{R}$ is given by

$$\dot{R} = -4Z \cdot \dot{Z}I + 2\left(Z\dot{Z}^T + \dot{Z}Z^T\right) + 2\sqrt{1-Z^2}\,S(\dot{Z}) - 2Z \cdot \dot{Z}\left(1-Z^2\right)^{-1/2} S(Z) ,$$

where $S(V)$ denotes the skew-symmetric matrix formed from the vector V by Equation 28. Taking the dot product of each side of Equation 34 with Z leads to

$$Z \cdot \dot{Z} = \frac{1}{2}\frac{d\left(Z^2\right)}{dt} = \frac{1}{2}\omega \cdot Z\sqrt{1-Z^2} ,$$

or

$$Z \cdot \dot{Z}(1 - Z^2)^{-1/2} = \frac{1}{2}\omega \cdot Z, \tag{35}$$

since $Z^2 < 1$. When Equations 34 and 35 are substituted in the expression for $\dot{R}$, then

$$\dot{R} = \sqrt{1 - Z^2}\left[Z\omega^T + \omega Z^T + S\left(Z \times \omega + \sqrt{1 - Z^2}\,\omega\right) - 2\omega \cdot ZI\right] + Z(Z \times \omega)^T + (Z \times \omega)Z^T - \omega \cdot ZS(Z),$$

and direct calculation will verify that $\dot{R} = \Omega R$ for $t_0 < t < t_1$. In fact, $\dot{R} = \Omega R$ even if $Z^2 = 1$, provided that Z has a derivative at this point satisfying Equation 34 and that Equation 35 is valid in the limit as Z^2 approaches unity. For example, if ω is a non-zero constant vector, then

$$Z = \sin\left(\frac{|\omega|\, t}{2}\right)\frac{\omega}{|\omega|}$$

is a function satisfying Equation 34 for $-\pi \leq |\omega|\, t \leq \pi$; indeed, the matrix R defined by Z satisfies Equation 26 over the same interval.

Thus, both Equations 30 and 34 define the motion of a rigid body over a wider range of allowable orientations than Euler's equations (Equation 29) and require only the integration of three scalar equations (which do not contain trigonometric functions). Furthermore, the results of the integration (especially Z) can be used directly. There is no need to generate the matrix; for example, we can obtain the coordinates of a vector relative to the body system and Y or Z as shown under "Coordinates of a Rotated Vector." If Z_1 defines the orientation of the body at time t_1 relative to the body system at $t = 0$, and Z_2 defines the orientation at time t_2 relative to the body system at $t = t_1$, then the rotation product $Z = Z_2^* Z_1$ gives the orientation at time t_2 relative to the body system at $t = 0$ for all Z_1 and Z_2. Eulerian angles can also be obtained as described in the previous section.

As with most differential equations, one would have to devote considerable time to Equations 30 and 34 in order to describe completely the properties of the solutions. The equations have one property in common that is useful for approximating the solutions for small increments of time. If $Y(0) = Z(0) = 0$, then the nth derivative of Y or Z at $t = 0$ contains the term $\omega^{(n-1)}(0)$—the $(n-1)$th derivative of ω at $t = 0$; for the first two derivatives this is the only term. Thus,

$$Y(h) = \frac{1}{2}\int_0^h \omega\, dt + U,$$

$$Z(h) = \frac{1}{2}\int_0^h \omega\, dt + V,$$

where U and V are of the order of h^3. Hence, to the second order, both solutions may be approximated by the integral of the angular velocity.

CONCLUSION

The significant advantages of the vector approach to rotations, as presented here, over other parametrizations is that the vector parameters can be obtained with ease from basic data; we need not transform them to a new set in order to perform the algebra of rotations (the product of rotations and the product of a vector by a rotation). We need not evaluate the trigonometric functions; there are no longer the singularities that existed when we tried to write a vector in polar form or factor a rotation matrix into simple rotations. To illustrate these last remarks, we cite one final important application of our vector approach to rotations.

In orbit theory, it is customary to obtain the components of the position and velocity vectors by rotating a coordinate system in which the direction cosines of the angular momentum vector are given by $E_3 = (0, 0, 1)^T$ into a fixed system in which the direction cosines of this vector are also known, say $H = (h_1, h_2, h_3)^T$. We usually assume that H is of the form $H = (\sin i \sin \Omega, -\sin i \cos \Omega, \cos i)^T$; thus the rotation is normally given by $R = R_3(-\Omega) R_1(-i)$. Unfortunately, this technique produces a singularity even in the trivial case $i = 0$ (no rotation required). On the other hand, we can refer to the section "Rotations Determined by Two Vectors and Their Images," obtaining

$$Z = \frac{1}{\sqrt{2(1 + h_3)}} H \times E_3 = \begin{pmatrix} -\sin \frac{i}{2} \cos \Omega \\ -\sin \frac{i}{2} \sin \Omega \\ 0 \end{pmatrix}$$

as the rotation vector taking E_3 into H. Thus, the two parameters z_1 and z_2 define the rotation uniquely for all i except $i = \pi$.

ACKNOWLEDGMENT

The author wishes to express his sincere thanks to Mr. Richard desJardins for his painstaking review of the manuscript and for his many helpful suggestions concerning the manner of presentation.

Goddard Space Flight Center
National Aeronautics and Space Administration
Greenbelt, Maryland, April 11, 1968
311-07-21-01-51

REFERENCES

1. Frazer, R. A., Duncan, W. J., and Collar, A. R., "Elementary Matrices," New York: Macmillan, 1947.

2. Mayer, A., "Rotations and Their Algebra," SIAM Review, 2(2):77-122, April 1960.

3. Wahba, G., "A Least Squares Estimate of Satellite Attitude," Problem 65-1, SIAM Review, 7(3):409, July 1965.

4. Stuelpnagel, J., "On The Parametrization of The Three-Dimensional Rotation Group," SIAM Review, 6(4):422-430, Oct. 1964.