X-541-71-274

NASA TM X- 65688

SLEWING ABOUT NON-ORTHOGONAL AXES

PAUL B. DAVENPORT

JULY 1971

GSFC

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

X-541-71-274
PREPRINT
AAS No. 71-389

SLEWING ABOUT NON-ORTHOGONAL AXES

Paul B. Davenport
Goddard Space Flight Center
Greenbelt, Maryland

July 1971

AAS/AIAA
ASTRODYNAMICS SPECIALISTS CONFERENCE 1971
FT. LAUDERDALE, FLORIDA
AUGUST 17-19, 1971

SLEWING ABOUT NON-ORTHOGONAL AXES

Paul B. Davenport
Goddard Space Flight Center
Greenbelt, Maryland

ABSTRACT

Given a desired three-axis re-orientation, it is well known that there are (in general) twenty-four possible three-legged slews about orthogonal axes which will perform the re-orientation. The present paper determines whether the maneuver may be accomplished with three slews about arbitrary (not necessarily orthogonal) axes. All of the possible three-legged slews are exhibited by a closed analytical procedure which requires no assumptions concerning the degree of non-orthogonality nor any approximations. The classical orthogonal case (Euler angles) is then an immediate consequence as a special case of the generalized solution. The development also includes the solutions of single-axis re-orientations by two successive rotations about arbitrary fixed lines.

SLEWING ABOUT NON-ORTHOGONAL AXES

INTRODUCTION

The representation of rotations via successive simple rotations about coordinate axes is probably the oldest and most frequent method employed. The reasons for this are threefold: (1) They are easily visualized, (2) The general case is obtained by three simple steps rather than a single complex one, and (3) They have been adequate descriptors for most problems in numerous fields. An arbitrary rotation can be interpreted as three simple rotations regardless of the actual rotation and the control or causation of rotation can be performed by mounting the sensors and/or actuators mutually perpendicular. Two-axis systems include office chairs, the crane, and two gimbal systems for pointing guns, telescopes, etc. Three-axis applications range from amusement rides to the precise control of ships, airplanes, and spacecraft.

Recent requirements for precision and reliability, however, have strained the computational luxury of traditional orthogonal axes. The imperfections of mounting instruments, stresses, thermal bending, etc. may cause misalignments of the axes. In addition to the misalignment problem, some systems are purposely designed to be non-orthogonal for reliability and/or precision, e.g., instruments mounted on four non-parallel faces of a regular octahedron (the orthogonal case corresponds to mountings on a regular hexahedron – cube).

Here, we will take a general approach and show under what conditions rotations can be described by one, two, and three successive rotations about fixed directed lines of arbitrary orientation. In addition, procedures are presented for the determination of the angle of each individual rotation required to effect the total rotation. These angles will be referred to as generalized Euler angles. The development will follow what is believed to be the historical development of the classical Euler angles, i.e., the properties of a single axis rotation will be used to derive the characteristics of two successive rotations about fixed lines. This, in turn, forms a basis for the three axes development. In each instance the special situation of coordinate-axis rotations will be examined so that the results of the general case may be correlated to a familiar one.

The symbolism will be that of matrix algebra with a mixture of vector analysis notation. Capital English letters will be used to denote matrices. A vector, as used here, is a 3×1 matrix and no distinction will be made in notation between vectors and matrices except that the latter part of the alphabet (starting with T) will be reserved for vectors. It is assumed that all matrices (vectors) are relative to some underlying orthonormal coordinate system unless otherwise stated. The cross product $T = U \times V$ is used in the sense that the components of T are formed from the components of U and V in the classical manner. Note that the coordinate system is not necessarily right-handed. A superscript T will denote transpose and the inner product of two vectors U and V expressed as $U^T V = V^T U$ (which is equal to the dot product $U \cdot V$ when the coordinate system is orthonormal). Inverses will be described by the superscript -1. For a rotation matrix R, $R^{-1} = R^T$.

ROTATIONS ABOUT A FIXED LINE

Given a fixed directed line defined by the unit vector U, then a rotation about this line is a function only of the angle, θ, of rotation. The matrix of the rotation may be expressed as

$$R(\theta) = \cos\theta\, I + (1 - \cos\theta)\, U\, U^T - \sin\theta\, \tilde{U}, \tag{1}$$

where $\tilde{U}$ denotes the skew-symmetric matrix formed from the components of U such that for any vector V, $\tilde{U}V = U \times V$. I is the 3×3 identity matrix. The physical interpretation of eq. (1) is that a right-handed coordinate system is rotated (alias) by the right hand rule (when the right hand thumb points along the directed line the fingers point toward the positive motion) or that the vector is rotated (alibi) counterclockwise as viewed from the origin. The matrices of the rotations about the coordinate axes are given by:

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

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

If Y is any vector then its image Z = RY under the rotation is obtained from eq. (1) as

$$Z = \cos\theta\, Y + (1 - \cos\theta)\, U \cdot Y\, U - \sin\theta\, U \times Y. \tag{2}$$

The following relationships are then an immediate consequence: $Z \cdot Z = Y \cdot Y$, $Z \cdot U = Y \cdot U$, and $R(\theta)U = R^{-1}(\theta)U = U$. The first expression is true for any rotation, the second holds for any rotation about U, and the last is true only for vectors collinear to U unless R = I (Euler's theorem). If Y and Z are given non-zero vectors and satisfy the above conditions then the angle of rotation which takes Y into Z may be determined by taking the dot product of eq. (2) by Y and $U \times Y$ (provided $U \times Y \neq 0$) respectively which gives

$$\cos\theta = \frac{Y \cdot Z - (U \cdot Y)^2}{Y \cdot Y - (U \cdot Y)^2}, \quad \sin\theta = \frac{U \cdot (Z \times Y)}{Y \cdot Y - (U \cdot Y)^2}.$$

If $U \times Y = 0$ then θ is arbitrary, conversely if eq. (2) holds for arbitrary θ then $U \times Y = 0$.

In order to define angles uniquely, we will make extensive use of the function $\lambda = \mathrm{Tan}^{-1}(a,b)$ which signifies the inverse tangent of a/b where the angle is selected so that: (1) sgn(a) (sign of a) equals $\mathrm{sgn}(\sin\theta)$, (2) sgn (b) = $\mathrm{sgn}(\cos\theta)$, and (3) $-\pi < \theta \leq \pi$. If a = b = 0 then λ is undefined (arbitrary). We refer to the angle in the above definition as the "proper quadrant" inverse tangent. Thus, the "smallest" rotation satisfying eq. (2) is

$$\theta = \mathrm{Tan}^{-1}\,[U \cdot (Z \times Y),\ Y \cdot Z - (U \cdot Y)^2],$$

all other solutions (θ') are given by $\theta' = \theta + 2k\pi$ where k is an integer.

ROTATIONS ABOUT TWO FIXED LINES

In the preceding section it was shown that a non-zero vector Y can be rotated into a given vector Z by a rotation about a single fixed axis U if and only if $Z \cdot U = Y \cdot U$ and $Z \cdot Z = Y \cdot Y$. We now consider under what conditions a given non-zero vector can be rotated into another given vector via two successive single-axis rotations.

Let U_i and U_j be any two unit vectors such that $U_i \times U_j \neq 0$, and

$$\begin{aligned} R_i(\alpha) &= \cos\alpha\, I + (1 - \cos\alpha)\, U_i U_i^T - \sin\alpha\, \tilde{U}_i, \\ R_j(\beta) &= \cos\beta\, I + (1 - \cos\beta)\, U_j U_j^T - \sin\beta\, \tilde{U}_j. \end{aligned} \tag{3}$$

The question to be answered then is: Given two vectors Y and Z, under what conditions does $Z = R_j(\beta) R_i(\alpha) Y$? Since rotations do not alter length we know a priori that one condition is $Z \cdot Z = Y \cdot Y$. Hereafter, we will always assume that this is the case.

Now $Z = R_j(\beta) R_i(\alpha) Y$ if and only if $R_j^{-1}(\beta) Z = R_i(\alpha) Y$, which when expanded using eq. (3) yields

$$\begin{aligned} &(Z - U_j \cdot Z\, U_j) \cos\beta + U_j \times Z \sin\beta + U_j \cdot Z\, U_j = \\ &(Y - U_i \cdot Y\, U_i) \cos\alpha - U_i \times Y \sin\alpha + U_i \cdot Y\, U_i. \end{aligned}$$

By taking the dot product of both sides by U_i, U_j, and $U_i \times U_j$ respectively we obtain the following three scalar equations:

$$[U_i \cdot Z - (U_i \cdot U_j)(U_j \cdot Z)] \cos\beta + U_i \cdot (U_j \times Z) \sin\beta + (U_i \cdot U_j)(U_j \cdot Z) = U_i \cdot Y,$$

$$U_j \cdot Z = [U_j \cdot Y - (U_i \cdot U_j)(U_i \cdot Y)] \cos\alpha - U_j \cdot (U_i \times Y) \sin\alpha + (U_i \cdot U_j)(U_i \cdot Y),$$

$$(U_i \times U_j) \cdot Z \cos\beta + (U_i \times U_j) \cdot (U_j \times Z) \sin\beta = (U_i \times U_j) \cdot Y \cos\alpha - (U_i \times U_j) \cdot (U_i \times Y) \sin\alpha.$$

If we let $a = U_i \cdot U_j$ and

$$w_i = U_i \cdot Y - a\, U_j \cdot Z, \qquad x_i = U_i \cdot Z - a\, U_j \cdot Z,$$

$$w_j = U_j \cdot Y - a\, U_i \cdot Y, \qquad x_j = U_j \cdot Z - a\, U_i \cdot Y,$$

$$w_k = (U_i \times U_j) \cdot Y, \qquad x_k = (U_i \times U_j) \cdot Z,$$

then the above equations simplify to

$$x_i \cos\beta + x_k \sin\beta = w_i \tag{4a}$$

$$x_j = w_j \cos\alpha + w_k \sin\alpha \tag{4b}$$

$$x_k \cos\beta - x_i \sin\beta = w_k \cos\alpha - w_j \sin\alpha \tag{4c}$$

Note that if $a = 0$ the w's and x's are merely the components of Y and Z respectively relative to the basis U_i, U_j, and $U_i \times U_j$.

Introducing the variables g and λ defined by

$$g = w_k \cos\alpha - w_j \sin\alpha,$$

$$\lambda = \mathrm{Tan}^{-1}(x_i, x_k),$$

then equations (4a) and (4c) can be written as

$$\sqrt{x_i^2 + x_k^2}\, \sin(\beta + \lambda) = w_i,$$

$$\sqrt{x_i^2 + x_k^2}\, \cos(\beta + \lambda) = g.$$

These last two equations have a common solution if $x_i^2 + x_k^2 = g^2 + w_i^2$. Thus, if $x_i^2 + x_k^2 - w_i^2 \geq 0$ there exist g and β which satisfy both of the above equations. The solutions are discrete except when $x_i = x_k = 0$ in which case β is arbitrary.

Similarly, equations (4b) and (4c) can be expressed as

$$\sqrt{w_j^2 + w_k^2}\ \sin(\alpha + \mu) = x_j,$$

$$\sqrt{w_j^2 + w_k^2}\ \cos(\alpha + \mu) = g,$$

where $\mu = \text{Tan}^{-1}(w_j, w_k)$. These last two equations have a common solution if $w_j^2 + w_k^2 = x_j^2 + g^2$.

The two sets of solutions above will exist and be consistent provided that $w_j^2 + w_k^2 - x_j^2 = x_i^2 + x_k^2 - w_i^2 \geq 0$. In this case, the equations (4) have a common solution. In fact, there are at least two solutions given by:

$$g = \pm\sqrt{w_j^2 + w_k^2 - x_j^2} = \pm\sqrt{x_i^2 + x_k^2 - w_i^2},$$

$$\begin{pmatrix} x_i & x_k \\ x_k & -x_i \end{pmatrix} \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix} = \begin{pmatrix} w_i \\ g \end{pmatrix}, \quad \begin{pmatrix} w_j & w_k \\ w_k & -w_j \end{pmatrix} \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} = \begin{pmatrix} x_j \\ g \end{pmatrix}$$

In terms of the proper quadrant inverse tangent

$$\alpha = \text{Tan}^{-1}\ [(w_k\ x_j - g\ w_j),\ (w_j\ x_j + g\ w_k)],$$

$$\beta = \text{Tan}^{-1}\ [(w_i\ x_k - g\ x_i),\ (w_i\ x_i + g\ x_k)].$$

The consistency of the solution above depended upon the requirement that $w_j^2 + w_k^2 - x_j^2 = x_i^2 + x_k^2 - w_i^2$ which is equivalent to the condition $W^T W = X^T X$ where $W = (w_i, w_j, w_k)^T$ and $X = (x_i, x_j, x_k)^T$. We shall now show that, indeed, this is the case. Obviously the condition holds when the two axes U_i and U_j are orthonormal for then the expressions $W^T W$ and $X^T X$ represent the square of the length of Y and Z respectively (assumed equal). In the general case let T and V be vectors of components of Y and Z relative to the basis U_i, U_j, and $U_i \times U_j$, e.g., $t_1 = U_i \cdot Y$, $t_2 = U_j \cdot Y$, and $t_3 = (U_i \times U_j) \cdot Y$. The above consistency condition is then equivalent to $T^T T - 2at_1 t_2 = V^T V - 2av_1 v_2$, where $a = U_i \cdot U_j$ as before. It is possible to show that $T^T T - 2at_1 t_2 = Y \cdot Y(1 - a^2)$ by constructing an orthonormal basis (Gram-Schmidt Process) from U_i, U_j, and $U_i \times U_j$ and expressing the square of the length as a function of these components. A similar condition can be proved analogously for Z and V. This demonstrates that the condition for a solution to (4a) and (4c) is the same as that for (4b) and (4c) and when the solutions exist they are consistent. The above construction may also be used to verify that if one angle is arbitrary the other cannot be. For both angles to be arbitrary requires W and X to be null vectors which in turn implies that Y and Z be null (the components relative to the constructed basis are all zero).

The condition for the existence of the solution ($w_j^2 + w_k^2 - x_j^2 \geq 0$) depends on the axes of rotation (U_i and U_j) and very little more can be said about existence until the axes (or conditions on them) are specified. If the axes are a subset of a complete orthonormal set (e.g., the coordinate axes) then $U_i \cdot U_j = 0$ and the w's and x's are merely permutations of the coordinates of Y and Z respectively. In this case, the subscripts i, j, and k can be considered as members of the set of integers $\{1,2,3\}$ and the requirement that U_i and U_j be distinct can be stated as $i \neq j$. There are six permutations of the three axes taken two at a time. If all six possible choices are to be examined it is convenient to use the notation $k = i'j'$ meaning k not equal to i or j. This definition plus the "permutation" or "epsilon" symbol ϵ_{ij} ($\epsilon_{ij} = 0$ if $i = j$, $\epsilon_{ij} = 1$ if the ordered pair i, j is a cyclic permutation, and $\epsilon_{ij} = -1$ if the ordered pair i,j is a non-cyclic permutation) provide useful tools when the rotational sequence is arbitrary.

Let the ordered vectors U_1, U_2, U_3 form an orthonormal triad ($U_k = \epsilon_{ij} U_i \times U_j$ where $k = i'j'$) then all possible "two-legged" slews taking Y into Z can be expressed as: For i, j = 1,2,3 such that $i \neq j$ let $k = i'j'$,

$$g = \pm \sqrt{y_j^2 + y_k^2 - z_j^2} = \pm \sqrt{z_i^2 + z_k^2 - y_i^2},$$

$$\alpha = \mathrm{Tan}^{-1} \left[(\epsilon_{ij}\, y_k\, z_j - g\, y_j),\ (y_j\, z_j + \epsilon_{ij}\, g\, y_k) \right],$$

$$\beta = \mathrm{Tan}^{-1} \left[(\epsilon_{ij}\, y_i\, z_k - g\, z_i),\ (y_i\, z_i + \epsilon_{ij}\, g\, z_k) \right],$$

provided, of course, that the square root is real.

It can be shown for the orthogonal case that at least four of the permutations have solutions for any Y and Z (except for the trivial case Z = Y) where $Y \cdot Y = Z \cdot Z$ and depending on Y and Z there may be four, five, or six permutation solutions (excluding multiples of 2π there are two numerical solutions for each sequence). This stems from the useful fact that if the sequence (i,j) does not have a solution then the three sequences (j,i), (j,k), and (k,i) do. The non-existence implies $z_j^2 - y_j^2 > y_k^2 \geq 0$ and $z_j^2 - y_k^2 > y_j^2 \geq 0$ which in turn imply the solutions of the three sequences listed. Examples: (1) $Y = (0,0,1)^T$ and Z arbitrary has only four permutation solutions, if $z_1 = 0$ then there are five, (2) $Y = (1,1,0)^T$ and $Z = (0,1,1)^T$ has six (twelve numerical) solutions.

THREE ROTATIONS ABOUT FIXED LINES

We have now established that rotations of the form $R = R_j(\beta)R_i(\alpha)$ (where $R_i(\alpha)$ and $R_j(\beta)$ are rotations about fixed axes U_i and U_j) are specified to within a positive or negative square root by the requirement that Z = RY. If one desires to rotate two vectors simultaneously into two other vectors by rotations about fixed axes then it is clear that at least three such rotations will be required in general. Since the time of Euler it has been known that relative to coordinate axes three such rotations (whose angles bear his name) completely define any rotation. In the literature, a particular (but not unique) sequence is commonly selected; this has caused a great deal of confusion since there are actually twelve possible sequences. To make matters worse some authors use left-handed systems while most use right-handed. Some use clockwise and others counter-clockwise to define the positive direction and some fail to define from where the motion is viewed (from the origin or looking toward the origin). Although a particular sequence is easier to visualize (such as the yaw, pitch, and roll of a ship or airplane) the algebra of the general case is only slightly more complicated. Because of the increasing applications of non-orthogonal slewing axes and/or the requirement to examine all sequences (so as to select the "best") we will take a general algebraic approach and let the reader "visualize" the results to suit his own application.

It was implied above, but not explicitly stated, that two vectors and their images completely define a rotation. Although this is probably a widely accepted fact (at least in the orthogonal case) we will pursue it further for the sake of completeness, and since our development depends upon the result. Because of the orthogonality condition and the requirement that the determinant be plus one, the matrix of an arbitrary rotation may be written as $R = (U_1, U_2, U_3)$ where the U_i are unit vectors juxtaposed and $U_3 = U_1 \times U_2$. Hence, the vectors U_1 and U_2 define the rotation uniquely, they are the images of $(1,0,0)^T$ and $(0,1,0)^T$ respectively. Given any two vectors Y_1 and Y_2 such that $Y_1 \times Y_2 \neq 0$ and given their images Z_1 and Z_2 under a rotation ($Z_i \cdot Z_j = Y_i \cdot Y_j$, $i,j = 1,2$ since rotations preserve distance and angles) then the matrix of that rotation may be determined as follows: Construct two orthonormal systems, one from the vectors Y_1 and Y_2, and the other from the vectors Z_1 and Z_2 (both of the same handedness-right or left). The rotation taking the first system into the second will then take Y_i into Z_i, $i = 1,2$. Explicitly, normalize Y_1 and denote this as U_1. Let $W = Y_2 - (Y_1 \cdot Y_2 / Y_1 \cdot Y_1)Y_1$, normalize W and denote this as U_2. Let $U_3 = U_1 \times U_2$, and $R_1 = (U_1, U_2, U_3)$. Derive the rotation matrix R_2 as above replacing Y_i with Z_i. The rotation matrix $R = R_2 R_1^{-1}$ then satisfies $Z_i = RY_i$, $i = 1,2$. The uniqueness of R follows from Euler's theorem (a nontrivial rotation has only one real eigenvector).

From the discussion above we have ascertained that the practical question – Under what conditions can two vectors, a coordinate system, or a rigid body be rotated to some desired orientation via three fixed-axis rotations? – can be formulated algebraically as follows: Under what conditions can an arbitrary rotational matrix R be represented as $R = R_n(\psi)R_m(\theta)R_\ell(\phi)$ where R_i $(i = \ell, m, n)$ denotes a rotation about a fixed unit vector U_i? It is assumed that m is different from n and ℓ (otherwise the product degenerates into only two fixed-axis rotations) but that n and ℓ may be equal.

Assume $R = R_n(\psi)R_m(\theta)R_\ell(\phi)$ with the conditions above, then since $R_i(\lambda)U_i = R_i^{-1}(\lambda)U_i = U_i$ $(i = \ell, m, n$ and $\lambda = \phi, \theta, \psi)$ we obtain the following results:

$$R\, U_\ell = R_n(\psi)\, R_m(\theta)\, U_\ell, \tag{5a}$$

$$R^{-1}\, U_n = R_\ell^{-1}(\phi)\, R_m^{-1}(\theta)\, U_n, \tag{5b}$$

$$R\, R_\ell^{-1}(\phi)\, U_m = R_n(\psi)\, U_m, \tag{5c}$$

$$R^{-1}\, R_n(\psi)\, U_m = R_\ell^{-1}(\phi)\, U_m. \tag{5d}$$

Each of these equations has the property of eliminating one of the three generalized Euler angles. If the U_i are coordinate axes then the left side of eq. (5a) is merely the ℓ^{th} ($\ell = 1,2,3$) column of R and the left side of eq. (5b) is the n^{th} row of R represented as a column vector (the n^{th} column of R^T).

Conversely, assume that equations (5a) and (5b) hold and set $Y_1 = U_\ell$, $Z_2 = U_n$, $Z_1 = R_n(\psi)R_m(\theta)U_\ell$, and $Y_2 = R_\ell^{-1}(\phi)R_m^{-1}(\theta)U_n$. Then $Z_2 \cdot Z_2 = Y_2 \cdot Y_2 = Z_1 \cdot Z_1 = Y_1 \cdot Y_1 = 1$ and $Z_1 \cdot Z_2 = Y_1 \cdot Y_2 = U_n^T R_m(\theta)U_\ell$. Therefore, if $Y_1 \times Y_2 \neq 0$ ($R_m(\theta)U_\ell \neq \pm U_n$) it follows (by a previous argument) that the R satisfying $Z_i = RY_i$ ($i = 1,2$) (equivalent to equations (5a) and (5b)) is unique. This means that the simultaneous solutions of equations (5a) and (5b) provide the solution to $R = R_n(\psi)R_m(\theta)R_\ell(\phi)$ except in the isolated cases where $R_m(\theta)U_\ell$ is collinear with U_n (e.g., when $\ell = n$ and $\theta = 0$). This is indeed felicitous since both equations are of the same general form which was solved in the previous section. We merely have to demonstrate that the two solutions are consistent under the present assumptions.

Given U_1, U_2, U_3, R, and a particular rotational sequence, then the solution of eq. (5a) is obtained directly by the techniques of the previous section by letting $Y = U_\ell$, $Z = RU_\ell$, $i = m$, $j = n$, $\theta = \alpha$, and $\psi = \beta$. Eq. (5b) is solved similarly. Before doing this, however, it is convenient to introduce some intermediate quantities which occur often and which help to simplify the final result.

Let

$$V_1 = U_2 \times U_3,\ V_2 = U_3 \times U_1,\ V_3 = U_1 \times U_2,$$

$$a_{ij} = a_{ji} = U_i \cdot U_j,\ (i,\ j = 1,\ 2,\ 3)$$

$$f = U_1 \cdot V_1 = U_2 \cdot V_2 = U_3 \cdot V_3,$$

and when R is obtained

$$b_{ij} = U_i \cdot (R\, U_j) = U_j \cdot (R^{-1}\, U_i),$$

$$c_{ij} = V_i \cdot (R\, U_j), \quad (i,\ j = 1,\ 2,\ 3)$$

$$d_{ij} = U_i \cdot (R\, V_j) = V_j \cdot (R^{-1}\, U_i)$$

Define the matrices $M = (U_1, U_2, U_3)$ and $N = (V_1, V_2, V_3)$ (juxtaposition of vectors) then the above definitions may be expressed as:

$$A = M^T M,\ B = M^T R\, M,\ C = N^T R\, M,\ D = M^T R\, V.$$

The matrices C and D may also be written in terms of A and B, i.e.,

$$C = d\,(M)\ A^{-1}\ B,\ D = d\,(M)\ B\ A^{-1},$$

where d(M) denotes the determinant of M. If the U_i are the coordinate axes then $A = N = M = I$, $B = C = D = R$, and $f = 1$.

Returning to the solution of eq. (5a), the proper substitutions give

$$\begin{aligned} w_m &= a_{m\ell} - a_{mn}\, b_{n\ell}, & x_m &= b_{m\ell} - a_{mn}\, b_{n\ell}, \\ w_n &= a_{n\ell} - a_{mn}\, a_{\ell m}, & x_n &= b_{n\ell} - a_{mn}\, a_{\ell m}, \\ w_k &= \epsilon_{n\ell}\, f, & x_k &= \epsilon_{mn}\, c_{k\ell}, \end{aligned} \tag{6}$$

where $k = m'n'$. If

$$h = w_n^2 + w_k^2 - x_n^2 = x_m^2 + x_k^2 - w_m^2 \geq 0, \tag{7}$$

(the equality always holds) then $g = \pm\sqrt{h}$ and

$$\theta = \mathrm{Tan}^{-1}\ [(w_k\, x_n - g\, w_n),\ (w_n\, x_n + g\, w_k)], \tag{8a}$$

$$\psi = \mathrm{Tan}^{-1}\ [(w_m\, x_k - g\, x_m),\ (w_m\, x_m + g\, x_k)]. \tag{8b}$$

Also,

$$g = w_k \cos\theta - w_n \sin\theta,$$

$$= x_k \cos\psi - x_m \sin\psi.$$

To solve (5b), let $Y = U_n$, $Z = R^{-1}U_n$, $i = m$, $j = \ell$, $\alpha = -\theta$, and $\beta = -\phi$. Then

$$\begin{aligned} w_i &= a_{mn} - a_{\ell m}\, b_{n\ell}, & x_i &= b_{nm} - a_{\ell m}\, b_{n\ell}, \\ w_\ell &= a_{\ell n} - a_{\ell m}\, a_{mn} = w_n, & x_\ell &= b_{n\ell} - a_{\ell m}\, a_{mn} = x_n, \\ w &= \epsilon_{\ell n}\, f = -w_k, & x_p &= \epsilon_{m\ell}\, d_{np}, \end{aligned} \tag{9}$$

where $p = \ell'm'$ (the subscript i is for identification only having no numerical significance, i.e., $w_i \neq w_m$). The condition for a solution is

$$w_\ell^2 + w^2 - x_\ell^2 = x_i^2 + x_p^2 - w_i^2 \geq 0 \tag{10}$$

(again the equality is assured by the initial conditions), but the first expression is identical to h defined above in the solution of (5a). Thus, if (5a) has a solution so does (5b) and vice versa. Furthermore, the values of θ from (5b) can be brought into agreement with those already obtained from (5a) by proper selection of the sign of the square root of h. This gives

$$g' = w \cos\theta - w_\ell \sin(-\theta),$$

$$= x_p \cos\phi - x_i \sin(-\phi),$$

$$= -(w_k \cos\theta - w_\ell \sin\theta) = -g,$$

and

$$= x_i \cos\phi + x_p \sin(-\phi) = w_i.$$

It then follows that

$$\phi = \mathrm{Tan}^{-1}\left[-(w_i\, x_p + g\, x_i),\ (w_i\, x_i - g\, x_p)\right]. \tag{11}$$

In summary, if h, as defined above, is non-negative then (5a) and (5b) have a common solution given by the above equations. This implies that R can be factored as $R = R_n(\psi)R_m(\theta)R_\ell(\phi)$ provided $R_m(\theta)U_\ell \neq \pm U_n$.

At least one of the arguments of the proper quadrant inverse tangent in (8a) is different from zero (either w_n or w_k not zero) which means the generalized Euler angle θ always has well-defined discrete values when the solution exists. If this solution is such that $R_m(\theta)U_\ell = \pm U_n$ then (5a) and (5b) both reduce to $RU_\ell = \pm U_n$ regardless of ψ or ϕ. Conversely, if $RU_\ell = \pm U_n$ then (5a) and (5b) degenerate into the single condition $R_m(\theta)U_\ell = \pm U_n$. Hence, we have ascertained that the common solution of (5a) and (5b) gives the desired factorizations of R except when $RU_\ell = \pm U_n$ in which case ψ and ϕ are arbitrary. In general, there are two discrete solutions (neglecting multiples of 2π) corresponding to the two square roots of h.

When the axes of rotation are the coordinate axes then the expressions above can be greatly simplified. In this special case,

$$h = r_{m\ell}^2 + r_{k\ell}^2 = r_{nm}^2 + r_{np}^2 = 1 - r_{n\ell}^2$$

$$g = \pm\sqrt{h},$$

$$\phi = \mathrm{Tan}^{-1}(\mp r_{nm}, \pm \epsilon_{\ell m} r_{np}),$$

$$\theta = \mathrm{Tan}^{-1}[(\epsilon_{n\ell} r_{n\ell} - \delta_{n\ell} g), (\delta_{n\ell} r_{n\ell} + \epsilon_{n\ell} g),$$

$$\psi = \mathrm{Tan}^{-1}(\mp r_{m\ell}, \pm \epsilon_{mn} r_{k\ell}),$$

$k = m'n'$, $p = \ell'm'$, and $\delta_{n\ell}$ is the Kronecker Delta ($\delta_{n\ell} = 1$ if $n = \ell$ and $\delta_{n\ell} = 0$ if $n \neq \ell$). There are twelve possible sequences (six for $n = \ell$ and six where

$n = \ell'm'$) and each has two solutions except when $h = 0$. In this latter case, ϕ and ψ are undefined and $\theta = 0$ if $n = \ell$ or $\theta = \pm\pi/2$ if $n = \ell'm'$. This corresponds to the condition $RU_\ell = \pm U_n$.

The degenerate case ($RU_\ell = \pm U_n$) can be handled by the use of (5c). In fact, it can be demonstrated that the simultaneous solution of (5b) and (5c) provides a solution to the factorization of R into three fixed-axis rotations whenever a solution exists. The argument is analogous to that used for (5a) and (5b); the consequence that $R = R_r(\psi)R_m(\theta)R_\ell(\phi)$ if (5b) and (5c) hold requires only that $U_n \neq \pm U_m$ which was assumed to be the case. Eq. (5c) provides the angle ψ as a function of ϕ which can be evaluated after ϕ is determined from (5b) or assigned an arbitrary value when (5b) is degenerate. This functional relationship may be derived by employing familiar techniques; expand (5c) using the expression for a rotation about a line and then take appropriate dot products of each side. The result is:

$$(1 - a_{nm}^2)\sin\psi = -(U_n \times U_m)\cdot[(Z_m - a_{\ell m} Z_\ell)\cos\phi + Z_\ell \times Z_m \sin\phi + a_{\ell m} Z_\ell],$$

$$(1 - a_{nm}^2)\cos\psi = U_m \cdot [(Z_m - a_{\ell m} Z_\ell)\cos\phi + Z_\ell \times Z_m \sin\phi + a_{\ell m} Z_\ell - a_{nm} U_n],$$

where $Z_i = RU_i$ and $a_{ij} = U_i \cdot U_j$ as before. Introducing the matrix $E = N^TRN$ ($e_{ij} = V_i \cdot RV_j$) where the V_i and N were defined previously, the relationship between ϕ and ψ may be described as follows:

$$\psi = \mathrm{Tan}^{-1}\,[y(\phi),\ z(\phi)], \tag{12}$$

where

$$y(\phi) = \epsilon_{mn}\,[(c_{km} - a_{\ell m} c_{k\ell})\cos\phi + \epsilon_{\ell m} e_{kp} \sin\phi + a_{\ell m} c_{k\ell}],$$

$$z(\phi) = (b_{mm} - a_{\ell m} b_{m\ell})\cos\phi + \epsilon_{\ell m} d_{mp} \sin\phi + a_{\ell m} b_{m\ell} - a_{nm}^2,$$

and all other quantities have been defined heretofore. The transpose of the matrix E is equal to the adjoint of the B matrix. In the classical situation of rotations about coordinate axes this last formula reduces to

$$\psi = \mathrm{Tan}^{-1} [\epsilon_{mn} (r_{km} \cos \phi - \epsilon_{\ell m} r_{kp} \sin \phi), (r_{mm} \cos \phi + \epsilon_{\ell m} r_{mp} \sin \phi).$$

Although the derivation, including existence proofs, of the formulas contained herein has been somewhat lengthy and tedious, the application of the results is straightforward. The matrices M,N, and A are constant depending only on the fixed axes of rotation (if only a single fixed sequence is being considered many elements need not be evaluated). The matrices B, C, D, and E are easily determined by matrix multiplication once R is known. The inequality (7) or (10) establishes the existence or non-existence of the solution once the parameters in (6) have been determined. Eq. (8a) provides the well defined angles for the middle rotation. The angles for the first rotation are evaluated from (11) then the angles for the last rotation may be obtained by (12) or alternately by (8b) if the first angle is well defined.

BIBLIOGRAPHY

Paul B. Davenport, "Mathematical Analysis For The Orientation and Control Of The Orbiting Astronomical Observatory Satellite," NASA TN D-1668, 1963.

,"A Vector Approach to The Algebra of Rotations With Applications," NASA TN D-4696, 1968.

R. A. Frazer, W. J. Duncan, and A. R. Collar, "Elementary Matrices," University Press, Cambridge, 1957, pp. 246-256.

J. W. Gibbs, "Vector Analysis," Yale University Press, New Haven, 1958, pp. 332-371.

Carl Grubin, "Vector Representation of Rigid Body Rotation," Amer., J. Phys., 30 (1962), pp. 416-417.

G. A. Korn and T. M. Korn, "Mathematical Handbook for Scientists and Engineers," McGraw-Hill, New York, 1968, pp. 471-479.

Arthur Mayer, "Rotations and Their Algebra," SIAM Review, 2(1960), pp. 77-122.

Peter Rastall, "Rotations and Lorentz Transformations," Nuclear Phys., 57 (1964), pp. 191-199.

John Stuelpnagel, "On The Parametrization of the Three-Dimension Rotation Group," SIAM Review, 6(1964), pp. 422-430.

J. L. Synge, "On Classical Dynamics," Encyclopedia of Physics, Vol. III/1, pp. 17-25.

E. T. Whittaker, "A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, Dover Publications, New York, 1944, pp. 8-13.