Constructing Rotation Matrices using Power Series

David Eberly, Geometric Tools, Redmond WA 98052
https://www.geometrictools.com/

Created: December 11, 2007
Last Modified: December 29, 2007

Contents

1 Introduction

Although we tend to work with rotation matrices in two or three dimensions, sometimes the question arises about how to generate rotation matrices in arbitrary dimensions. This document describes a method for computing rotation matrices using power series of matrices. The approach is one you see in an undergraduate mathematics course on solving systems of linear differential equations with constant coefficients; for example, see [Bra83, Chapter 3].

1.1 Power Series of Functions

The natural exponential function is introduced in Calculus, namely, $\exp(x) = e^x$, where the base is (approximately) $e \doteq 2.718281828$. The function may be written as a power series

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^k}{k!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{1}$$

The power series is known to converge for any real number x. The formula extends to complex numbers $z = x + iy$, where x and y are real numbers and $i = \sqrt{-1}$,

$$e^z = e^{x+iy} = e^x e^{iy} = \exp(x)\left(\cos(y) + i\sin(y)\right) \tag{2}$$

The term $\exp(x)$ may be written as a power series using Equation (1). The trigonometric terms also have power series representations,

$$\sin(y) = y - \frac{y^3}{3!} + \frac{y^5}{5!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k+1}}{(2k+1)!} \tag{3}$$

$$\cos(y) = 1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k}}{(2k)!} \tag{4}$$

The latter two power series also converge for any real number y.

1.2 Power Series Involving Matrices

The power series representations extend to functions whose inputs are square matrices rather than scalars, taking us into the realm of matrix algebra; for example, see [HJ85]. That is, $\exp(M)$, $\cos(M)$, and $\sin(M)$ are power series of the square matrix M, and they converge for all M.

In particular, we are interested in $\exp(M)$ for square matrix M. A note of caution is in order. The scalar-valued exponential function $\exp(x)$ has various properties that do not immediately extend to the matrix-valued function. For example, we know that

$$\exp(a+b) = e^{a+b} = e^a e^b = e^b e^a = e^{b+a} = \exp(b+a) \tag{5}$$

for any scalars a and b. This formula does not apply to all pairs of matrices A and B. The problem is that matrix multiplication is not commutative, so reversing the order of terms $\exp(A)$ and $\exp(B)$ in the products generally produces different values. It is true, though that

$$\exp(A+B) = \exp(A)\exp(B), \quad \text{when } AB = BA \tag{6}$$

That is, as long as A and B commute in the multiplication, the usual property of power-of-sum-equals-product-of-powers applies.

The power series for $\exp(M)$ of a square matrix M is formally defined as

$$\exp(M) = I + M + \frac{M^2}{2!} + \cdots + \frac{M^k}{k!} + \cdots = \sum_{k=0}^{\infty} \frac{M^k}{k!} \tag{7}$$

and converges for all M. The first term, I, is the identity matrix of the same size as M. The immediate question is how one goes about computing the power series. That is one of the goals of this document and is described next.

1.3 Reduction of the Matrix Power Series

Suppose that the $n \times n$ matrix M is diagonalizable; that is, suppose we may factor $M = PDP^{-1}$, where $D = \text{Diag}(d_1, \ldots, d_n)$ is a diagonal matrix and P is an invertible matrix with inverse P^{-1}. The matrix powers are easily shown to be $M^k = PD^kP^{-1}$, where $D^k = \text{Diag}(d_1^k, \ldots, d_n^k)$. The power series factors as

$$\begin{aligned}
\exp(M) &= \textstyle\sum_{k=0}^{\infty} \frac{M^k}{k!} = \sum_{k=0}^{\infty} P \frac{D^k}{k!} P^{-1} \\
&= P \left(\textstyle\sum_{k=0}^{\infty} \frac{D^k}{k!} \right) P^{-1} \\
&= P \ \text{Diag} \left(\textstyle\sum_{k=0}^{\infty} \frac{d_1^k}{k!}, \ldots, \sum_{k=0}^{\infty} \frac{d_n^k}{k!} \right) P^{-1} \\
&= P \ \text{Diag} \left(e^{d_1}, \ldots, e^{d_n} \right) P^{-1}
\end{aligned} \tag{8}$$

The expression is valid regardless of whether P and D have real-valued or complex-valued entries. For the purpose of numerical computation, when D has a real-valued entry, d_j, then $\exp(d_j)$ may be computed using any standard mathematics library. Naturally, the computation is only an approximation. When D has a complex-valued entry, $d_j = x_j + iy_j$, then Equation (2) may be used to obtain $\exp(d_j) = \exp(x_j)(\cos(y_j) + i\sin(y_j))$. All of $\exp(x_j)$, $\cos(y_j)$, and $\sin(y_j)$ may be computed using any standard mathematics library.

The problem, though, is that not all matrices M are diagonalizable. A special case that arises often is a real-valued symmetric matrix M. Such matrices have only real-valued eigenvalues and a complete basis of orthonormal eigenvectors. The eigenvalues are the diagonal entries of D and the corresponding eigenvectors become the columns of P. The orthonormality of the eigenvectors guarantees that P is orthogonal, so $P^{-1} = P^{\mathsf{T}}$ (the transpose is the inverse). A more general result is discussed in [HS74, Chapter 6], which applies to all matrices. A brief summary is provided here.

A real-valued matrix S that is diagonalizable is said to be *semisimple*. A real-valued matrix N is *nilpotent* if there exists a power $p \geq 1$ for which $N^p = 0$. Generally, a real-valued matrix M may be uniquely decomposed as $M = S + N$, where S is real-valued and semisimple, N is real-valued and nilpotent, and $SN = NS$. Because S is diagonalizable, we may factor it is $S = PDP^{-1}$, where $D = \text{Diag}(d_1, \ldots, d_n)$ is a diagonal matrix and P is invertible. What this means regarding the exponential function is

$$\exp(M) = \exp(S + N) = \exp(S)\exp(N) = P \ \text{Diag} \left(e^{d_1}, \ldots, e^{d_n} \right) P^{-1} \sum_{k=0}^{p-1} \frac{N^k}{k!} \tag{9}$$

The key here is that the power series for $\exp(N)$ is really a finite sum because N is nilpotent. Equation (9) shows us how to compute $\exp(M)$ for any matrix M; in particular, the equation may be implemented on a computer. Notice that in the special case of a symmetric matrix M, it must be that $M = S$ and $N = 0$.

The discussion here allows for complex numbers. In particular, P and D might have complex-valued entries. Instead, we may factor the $n \times n$ matrix $S = QEQ^{-1}$ as follows. Let S have r real-valued eigenvalues λ_1 through λ_r and c real-valued eigenvalues $\alpha_1 + i\beta_1$ through $\alpha_c + i\beta_c$, where $n = r + c$. Because S is real-valued, the complex-valued eigenvalues appear in pairs, $\alpha_j + i\beta_j$ and $\alpha_j - i\beta_j$; thus, c is even, say, $c = 2m$. We may construct a basis of vectors for $\mathbb{R}^n$ that become the columns of the matrix Q and for which the matrix E has the block-diagonal form

$$E = \begin{bmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_r & & & & \\ & & & \begin{bmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{bmatrix} & & & \\ & & & & \cdots & & \\ & & & & & \begin{bmatrix} \alpha_m & \beta_m \\ -\beta_m & \alpha_m \end{bmatrix} \end{bmatrix} \tag{10}$$

Specifically, let $\mathbf{U}_i$ be a real-valued eigenvector corresponding to the real-valued eigenvalue λ_i for $1 \le i \le r$. Let $\mathbf{V}_j = \mathbf{X}_j + i\mathbf{Y}_j$ be a complex-valued eigenvector corresponding to the complex-valued eigenvalue $\alpha_j + i\beta_j$ for $1 \le j \le c$. The numbers α and β are real-valued and the vectors $\mathbf{X}$ and $\mathbf{Y}$ have real-valued components. Because S is real-valued, when μ is a complex-valued eigenvalue with eigenvector $\mathbf{V}$, the conjugate $\bar{\mu}$ has eigenvector $\bar{\mathbf{V}}$, which is the componentwise conjugation of $\mathbf{V}$. The matrix P has columns that are the eigenvectors,

$$P = \begin{bmatrix} \mathbf{U}_1 & \cdots & \mathbf{U}_r & \mathbf{V}_1 & \bar{\mathbf{V}}_1 & \cdots & \mathbf{V}_m & \bar{\mathbf{V}}_m \end{bmatrix} \tag{11}$$

To verify the factorization,

$$\begin{aligned} SP &= S \begin{bmatrix} \mathbf{U}_1 & \cdots & \mathbf{U}_r & \mathbf{V}_1 & \bar{\mathbf{V}}_1 & \cdots & \mathbf{V}_m & \bar{\mathbf{V}}_m \end{bmatrix} \\ &= \begin{bmatrix} S\mathbf{U}_1 & \cdots & S\mathbf{U}_r & S\mathbf{V}_1 & S\bar{\mathbf{V}}_1 & \cdots & S\mathbf{V}_m & S\bar{\mathbf{V}}_m \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{U}_1 & \cdots & \lambda_r\mathbf{U}_r & (\alpha_1 + i\beta_1)\mathbf{V}_1 & (\alpha_1 - i\beta_1)\bar{\mathbf{V}}_1 & \cdots & (\alpha_m + i\beta_m)\mathbf{V}_m & (\alpha_m - i\beta_m)\bar{\mathbf{V}}_m \end{bmatrix} \\ &= PD \end{aligned} \tag{12}$$

where D is the diagonal matrix of eigenvalues. Applying the inverse of P to both sides of the equation produces $S = PDP^{-1}$.

Let $\mu = \alpha + i\beta$ be a complex-valued eigenvalue of S. Let $\mathbf{V} = \mathbf{X} + i\mathbf{Y}$ be a complex-valued eigenvector for μ. We may write $S\mathbf{V} = (\alpha + i\beta)\mathbf{V}$ in terms of real parts and imaginary parts,

$$S\mathbf{X} + iS\mathbf{Y} = (\alpha\mathbf{X} - \beta\mathbf{Y}) + i(\beta\mathbf{X} + \alpha\mathbf{Y}) \tag{13}$$

Equating real parts and imaginary parts,

$$S\mathbf{X} = \alpha\mathbf{X} - \beta\mathbf{Y}, \quad S\mathbf{Y} = \beta\mathbf{X} + \alpha\mathbf{Y} \tag{14}$$

Writing this in block-matrix form,

$$S \begin{bmatrix} \mathbf{X} & \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{X} & \mathbf{Y} \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \tag{15}$$

Eigenvectors corresponding to two distinct eigenvalues are linearly independent. To see this, let e_1 and e_2 be distinct eigenvalues with corresponding eigenvectors $\mathbf{W}_1$ and $\mathbf{W}_2$, respectively. To demonstrate linear independence, we must show that

$$\mathbf{0} = c_1\mathbf{W}_1 + c_2\mathbf{W}_2 \tag{16}$$

implies $c_1 = c_2 = 0$. Applying S to Equation (16) leads to

$$\mathbf{0} = c_1 S\mathbf{W}_1 + c_2 S\mathbf{W}_2 = c_1 e_1 \mathbf{W}_1 + c_2 e_2 \mathbf{W}_2 \tag{17}$$

Multiplying Equation (16) by e_1 and subtracting from Equation (17) produces

$$\mathbf{0} = c_2(e_2 - e_1)\mathbf{W}_2 \tag{18}$$

The eigenvalues are distinct, so $e_2 - e_1 \neq 0$. Eigenvectors are nonzero, so the only possibility to satisfy Equation (18) is $c_2 = 0$, which in turn implies $\mathbf{0} = c_1\mathbf{W}_1$ and $c_1 = 0$.

This result implies that $\mathbf{V}$ and $\bar{\mathbf{V}}$ are linearly independent vectors, and consequently $\mathbf{X} = (\mathbf{V} + \bar{\mathbf{V}})/2$ and $\mathbf{Y} = (\mathbf{V} - \bar{\mathbf{V}})/(2i)$ are linearly independent vectors. In the matrix P of Equation (11), we may replace each pair of complex-valued eigenvectors $\mathbf{V}$ and $\bar{\mathbf{V}}$ by the real-valued vectors $\mathbf{X}$ and $\mathbf{Y}$ to obtain a real-valued and invertible matrix

$$Q = \begin{bmatrix} \mathbf{U}_1 & \cdots & \mathbf{U}_r & \mathbf{X}_1 & \mathbf{Y}_1 & \cdots & \mathbf{X}_m & \mathbf{Y}_m \end{bmatrix} \tag{19}$$

Using Equation (15), it may be verified that $S = QEQ^{-1}$.

The exponential of E has the block-diagonal form

$$\exp(E) = \begin{bmatrix} e^{\lambda_1} & & & & & & \\ & \ddots & & & & & \\ & & e^{\lambda_r} & & & & \\ & & & e^{\alpha_1}\begin{bmatrix} \cos\beta_1 & \sin\beta_1 \\ -\sin\beta_1 & \cos\beta_1 \end{bmatrix} & & & \\ & & & & \cdots & & \\ & & & & & e^{\alpha_m}\begin{bmatrix} \cos\beta_m & \sin\beta_m \\ -\sin\beta_m & \cos\beta_m \end{bmatrix} \end{bmatrix} \tag{20}$$

This follows from

$$E_k = \begin{bmatrix} \alpha_k & \beta_k \\ -\beta_k & \alpha_k \end{bmatrix} = \alpha_k I + \beta_k J \tag{21}$$

where

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad IJ = JI \tag{22}$$

Because I and J commute in their product,

$$\exp(E_k) = \exp(\alpha_k I + \beta_k J) = \exp(\alpha_k I)\exp(\beta_k J) = e^{\alpha_k} \begin{bmatrix} \cos\beta_k & \sin\beta_k \\ -\sin\beta_k & \cos\beta_k \end{bmatrix} \tag{23}$$

The matrix

$$\exp(M) = Q\exp(E)Q^{-1}\sum_{k=0}^{p-1}\frac{N^k}{k!} \tag{24}$$

is therefore computed using only real-valued matrices. The heart of the construction relies on factoring $M = S + N$, computing the eigenvalues of S, and computing an orthogonal basis for $\mathbb{R}^n$ from S.

1.4 The Cayley-Hamilton Theorem

This is also a useful result that allows a reduction of the power series for $\exp(M)$ to a finite sum. The eigenvalues of an $n \times n$ matrix M are the roots to the *characteristic polynomial*

$$p(t) = \det(M - tI) = p_0 + p_1 t + \cdots + p_n t^n = \sum_{k=0}^{n} p_k t^k \tag{25}$$

where I is the identity matrix. The degree of p is n and the coefficients are p_0 through $p_n = (-1)^n$. The *characteristic equation* is $p(t) = 0$. The *Cayley-Hamilton Theorem* states that if you formally substitute M into the characteristic equation, you obtain equality,

$$0 = p(M) = p_0 I + p_1 M + \cdots + M^n = \sum_{k=0}^{n} p_k M^k \tag{26}$$

Multiplying this equation by powers of M and reducing everytime the largest power is a multiple of n allows a reduction of the power series for $\exp(M)$ to

$$\exp(M) = \sum_{k=0}^{\infty}\frac{M^k}{k!} = \sum_{k=0}^{n-1} c_k M^k \tag{27}$$

for some coefficients c_0 through c_{n-1}. These coefficients are necessarily functions of the characteristic coefficients p_j for $0 \le j \le n$ but the relationships are at first glance complicated to determine in closed form. This leads us into the topic of *ordinary difference equations*.

1.5 Ordinary Difference Equations

This is a topic whose theory parallels that of ordinary differential equations. It is a large topic that is not discussed in full detail here; a summary of it is provided in [Ebe03, Appendix D]. Generally, one has a sequence of numbers $\{x_k\}_{k=0}^{\infty}$ and a relationship that determines a term in the sequence from the previous n terms:

$$x_{k+n} = f(x_k, \cdots, x_{k+n-1}), \quad k \ge 0 \tag{28}$$

This is an *explicit equation* for x_{k+n}. The equation is said to have degree $n \geq 1$. The *initial conditions* are user-specified x_0 through x_{n-1}. An *implicit equation* is

$$F(x_k, \cdots, x_{k+n-1}, x_{k+n}) = 0, \quad k \geq 0 \tag{29}$$

and are generally more difficult to solve because the x_{k+n} term might occur in such a manner as to prevent solving for it explicitly.

The subtopic of interest here is how to solve ordinary difference equations that are linear and have constant coefficients. The general equation of this form is written with all terms on one side of the equation as if it were implicit, but this is still an explicit equation because we may solve for x_{k+n},

$$p_n x_{k+n} + p_{n-1} x_{k+n-1} + \ldots p_1 x_{k+1} + p_0 x_k = 0 \tag{30}$$

where p_0 through p_n are constants with $p_n \neq 0$. Notice that the coefficients may be used to form the characteristic polynomial described in the previous section, $p(t) = p_0 + p_1 t + \cdots + p_n t^n$ with $p_n = (-1)^n$.

The method of solution for the linear equation with constant coefficients involves computing the roots of the characteristic equation. Let t_1 through t_ℓ be the distinct roots of $p(t) = 0$. Let each root t_j have multiplicity m_j; necessarily, $n = m_1 + \cdots + m_\ell$. The Fundamental Theorem of Algebra states that the polynomial factors into a product,

$$p(t) = \prod_{j=1}^{\ell} (t - t_j)^{m_j} \tag{31}$$

Equation (30) has m_j linearly independent solutions corresponding to t_j, namely,

$$x_k = t_j^k, \quad x_k = k t_j^k, \quad , \cdots, \quad x_k = k^{m_j - 1} t_j^k \tag{32}$$

This is the case even when t_j is a complex-valued root. The general solution to Equation (30) is

$$x_k = \sum_{j=1}^{\ell} \left(\sum_{s=0}^{m_j - 1} c_{js} k^s \right) t_j^k, \quad k \geq n \tag{33}$$

where the n coefficients c_{js} are determined from the initial conditions which are the user-specified values x_0 through x_{n-1}.

As it turns out, the finite difference equations and constructions apply equally as well when the x_k are matrices. In this case, the c_{js} of Equation (33) are themselves matrices.

1.6 Generating Rotation Matrices from Skew-Symmetric Matrices

Equation (6) shows how to exponentiate a sum of matrices. As noted, $\exp(A + B) = \exp(A)\exp(B)$ as long as $AB = BA$. In particular, it is the case that $I = \exp(0) = \exp(A + (-A)) = \exp(A)\exp(-A)$, where I is the identity matrix. Thus, we have the formula for inversion of an exponential of a matrix,

$$\exp(A)^{-1} = \exp(-A) \tag{34}$$

The transpose of a sum of matrices is the sum of the transposes of the matrices, a property that also extends to the exponential power series. A consequence is

$$\exp(A)^{\mathsf{T}} = \exp\left(A^{\mathsf{T}}\right) \tag{35}$$

Now consider a skew-symmetric matrix S. Such a matrix has the property $S^\mathsf{T} = -S$. Define $R = \exp(S)$. Using Equations (34) and (35),

$$R^\mathsf{T} = \exp(S)^\mathsf{T} = \exp\left(S^\mathsf{T}\right) = \exp(-S) = \exp(S)^{-1} = R^{-1} \tag{36}$$

Because the inverse and transpose of R are the same matrix, R must be an orthogonal matrix. Although in the realm of advanced mathematics, it may be shown that in fact R is a rotation matrix. The essence of the argument is that the space of orthogonal matrices has two connected components, one for which the determinant is 1 (rotation matrices) and one for which the determinant is -1 (reflection matrices). Each component is path connected. The curve of orthogonal matrices $\exp(tS)$ for $t \in [0, 1]$ is a path connecting I (the case $t = 0$) and $R = \exp(S)$ (the case $t = 1$), so R and I must have the same determinant, which is 1, and R is therefore a rotation matrix.

2 Rotations Matrices in 2D

The general skew-symmetric matrix in 2D is

$$S = \theta \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{37}$$

where θ is any real-valued number. The corresponding rotation matrix is $R = \exp(S)$. The characteristic equation for S is

$$0 = p(t) = \det(S - tI) = t^2 + \theta^2 \tag{38}$$

The Cayley-Hamilton Theorem guarantees that

$$S^2 + \theta^2 I = 0 \tag{39}$$

so that $S^2 = -\theta^2 I$. Higher powers of S are $S^3 = -\theta^2 S$, $S^4 = -\theta^2 S^2 = \theta^4 I$, and so on. Substituting these into the power series for $\exp(S)$ and grouping together the terms involving I and S produces

$$\begin{aligned} R &= \exp(S) \\ &= I + S + \tfrac{S^2}{2!} + \tfrac{S^3}{3!} + \tfrac{S^4}{4!} + \tfrac{S^5}{5!} + \cdots \\ &= I + S - \tfrac{\theta^2}{2!} I - \tfrac{\theta^2}{3!} S + \tfrac{\theta^4}{4!} I + \tfrac{\theta^5}{5!} S - \cdots \\ &= \left(1 - \tfrac{\theta^2}{2!} + \tfrac{\theta^4}{4!} - \cdots\right) I + \left(1 - \tfrac{\theta^2}{3!} + \tfrac{\theta^4}{5!} - \cdots\right) S \\ &= \cos(\theta) I + \left(\tfrac{\sin(\theta)}{\theta}\right) S \\ &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \end{aligned} \tag{40}$$

The rotation matrix may be computed using the eigendecomposition method of Equation (8). The eigenvalues of S are $\pm i\theta$, where $i = \sqrt{-1}$. Corresponding eigenvectors are $(\pm i, 1)$. The matrices D, P, and P^{-1} in the

decomposition are

$$D = \begin{bmatrix} i\theta & 0 \\ 0 & -i\theta \end{bmatrix}, \quad P = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}, \quad P^{-1} = \frac{1}{2i} \begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix} \tag{41}$$

It is easily verified that $S = PDP^{-1}$. The rotation matrix is computed to be

$$\begin{aligned} R &= P \exp(D) P^{-1} \\ &= \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} \tfrac{1}{2i} \begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix} \\ &= \begin{bmatrix} \frac{e^{i\theta}+e^{-i\theta}}{2} & -\frac{e^{i\theta}-e^{-i\theta}}{2i} \\ \frac{e^{i\theta}-e^{-i\theta}}{2i} & \frac{e^{i\theta}+e^{-i\theta}}{2} \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \end{aligned} \tag{42}$$

where we have used the identities $e^{\pm i\theta} = \cos(\theta) \pm i \sin(\theta)$.

The rotation matrix may also be computed using ordinary finite differences. The linear difference equation is

$$\begin{aligned} &X_{k+2} + \theta^2 X_k = 0, \quad k \geq 0 \\ &X_0 = I, \quad X_1 = S \end{aligned} \tag{43}$$

We know from the construction that $S^k = X_k$. Equation (33) is

$$X_k = (i\theta)^k C_0 + (-i\theta)^k C_1 \tag{44}$$

for unknown coefficient matrices C_0 and C_1. The initial conditions imply

$$\begin{aligned} I &= C_0 + C_1 \\ S &= i\theta C_0 - i\theta C_1 \end{aligned} \tag{45}$$

and have the solution

$$\begin{aligned} C_0 &= \tfrac{I-(i/\theta)S}{2} \\ C_1 &= \tfrac{I+(i/\theta)S}{2} \end{aligned} \tag{46}$$

The solution to the linear difference equation is

$$S^k = X_k = (i\theta)^k \frac{I - (i/\theta)S}{2} + (-i\theta)^k \frac{I + (i/\theta)S}{2} \tag{47}$$

When k is even, say, $k = 2p$, Equation (47) reduces to

$$S^{2p} = (-1)^p \theta^{2p} I, \quad p \geq 1 \tag{48}$$

When k is odd, say, $k = 2p + 1$, Equation (47) reduces to

$$S^{2p+1} = (-1)^p \theta^{2p} S, \quad p \geq 1 \tag{49}$$

But these powers of S are exactly what we computed manually and substituted into the power series for $\exp(S)$ to produce Equation (40).

3 Rotations Matrices in 3D

The general skew-symmetric matrix in 3D is

$$S = \theta \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix} \tag{50}$$

where θ, a, b, and c are any real-valued numbers with $a^2 + b^2 + c^2 = 1$. The corresponding rotation matrix is $R = \exp(S)$. The characteristic equation for S is

$$0 = p(t) = \det(S - tI) = -t^3 - \theta^2 t \tag{51}$$

The Cayley-Hamilton Theorem guarantees that

$$-S^3 - \theta^2 S = 0 \tag{52}$$

so that $S^3 = -\theta^2 S$. Higher powers of S are $S^4 = -\theta^2 S^2$, $S^5 = -\theta^2 S^3 = \theta^4 S$, $S^6 = -\theta^2 S^2$, and so on. Substituting these into the power series for $\exp(S)$ and grouping together the terms involving I, S, and S^2 produces

$$\begin{aligned} R &= \exp(S) \\ &= I + S + \tfrac{S^2}{2!} + \tfrac{S^3}{3!} + \tfrac{S^4}{4!} + \tfrac{S^5}{5!} + \tfrac{S^6}{6!} + \cdots \\ &= I + S + \tfrac{1}{2!} S^2 - \tfrac{\theta^2}{3!} S - \tfrac{\theta^2}{4!} S^2 + \tfrac{\theta^4}{5!} S + \tfrac{\theta^4}{6!} S^2 - \cdots \\ &= I + \left(1 - \tfrac{\theta^2}{3!} + \tfrac{\theta^4}{5!} - \cdots\right) S + \left(\tfrac{1}{2!} - \tfrac{\theta^2}{4!} + \tfrac{\theta^4}{6!} - \cdots\right) S^2 \\ &= I + \left(\tfrac{\sin(\theta)}{\theta}\right) S + \left(\tfrac{1-\cos(\theta)}{\theta^2}\right) S^2 \end{aligned} \tag{53}$$

If we define $\hat{S} = S/\theta$, then

$$R = I + \sin(\theta)\hat{S} + (1 - \cos(\theta))\hat{S}^2 \tag{54}$$

which is *Rodrigues' Formula* for a rotation matrix. The angle of rotation is θ and the axis of rotation has unit-length direction $(-c, b, -a)$.

The rotation matrix may also be computed using ordinary finite differences. The linear difference equation is

$$\begin{aligned} &X_{k+3} + \theta^2 X_k = 0, \quad k \geq 0 \\ &X_0 = I, \quad X_1 = S, \quad X_2 = S^2 \end{aligned} \tag{55}$$

We know from the construction that $S^k = X_k$. The roots of the characteristic equation are 1 and $\pm i\theta$. Equation (33) is

$$S^k = X_k = 1^k C_0 + (i\theta)^k C_1 + (-i\theta)^k C_2 \tag{56}$$

for unknown coefficient matrices C_0, C_1, and C_2. The initial conditions imply

$$\begin{aligned} I &= C_0 + C_1 + C_2 \\ S &= C_0 + i\theta C_1 - i\theta C_2 \\ S^2 &= C_0 - \theta^2 C_1 - \theta^2 C_2 \end{aligned} \tag{57}$$

and have the solution

$$\begin{aligned} C_0 &= \tfrac{S^2+\theta^2 I}{1+\theta^2} \\ C_1 &= \tfrac{(I-C_0)-(i/\theta)(S-C_0)}{2} \\ C_2 &= \tfrac{(I-C_0)+(i/\theta)(S-C_0)}{2} \end{aligned} \tag{58}$$

When k is odd, say, $k = 2p+1$, Equation (56) reduces to the following using Equation (58)

$$S^{2p+1} = (-1)^p \theta^{2p} S, \quad p \geq 1 \tag{59}$$

When k is even, say, $k = 2p+2$, Equation (56) reduces to the following using Equation (58)

$$S^{2p+2} = (-1)^p \theta^{2p} S^2, \quad p \geq 1 \tag{60}$$

But these powers of S are exactly what we computed manually and substituted into the power series for $\exp(S)$ to produce Equation (53).

4 Rotations Matrices in 4D

The general skew-symmetric matrix in 4D is

$$S = \theta \begin{bmatrix} 0 & a & b & d \\ -a & 0 & c & e \\ -b & -c & 0 & f \\ -d & -e & -f & 0 \end{bmatrix} \tag{61}$$

where θ, a, b, c, d, e, and f are any real-valued numbers with $a^2+b^2+c^2+d^2+e^2+f^2 = 1$. The corresponding rotation matrix is $R = \exp(S)$. The characteristic equation for S is

$$0 = p(t) = \det(S - tI) = t^4 + \theta^2 t^2 + \theta^4 (af - be + cd)^2 = t^4 + \theta^2 t^2 + \theta^4 \delta^2 \tag{62}$$

where $\delta = af - be + cd$. The Cayley-Hamilton Theorem guarantees that

$$S^4 + \theta^2 S^2 + \theta^4 \delta^2 = 0 \tag{63}$$

Computing closed-form expressions for powers of S is simple, much like previous dimensions, when $\delta = 0$. It is more difficult when $\delta \neq 0$. Let us consider the cases separately. In both cases, we assume that $\theta \neq 0$, otherwise $S = 0$ and the rotation matrix is the identity matrix.

4.1 The Case $\delta = 0$

Suppose that $\delta = af - be + cd = 0$; then $S^4 + \theta^2 S^2 = 0$. The closed-form expressions for powers of S are easy to generate. Specifically,

$$\begin{aligned} S^{2p+2} &= (-1)^p \theta^{2p} S^2, \quad p \geq 1 \\ S^{2p+3} &= (-1)^p \theta^{2p} S^3, \quad p \geq 1 \end{aligned} \tag{64}$$

This leads to the power series reduction

$$R = \exp(S) = I + S + \left(\frac{1 - \cos(\theta)}{\theta^2} \right) S^2 + \left(\frac{\theta - \sin(\theta)}{\theta^3} \right) S^3 \tag{65}$$

It is possible for further reduction depending on the entries of S. For example, if $d = e = f = 0$, we effectively have the 3D rotation in (x, y, z) embedded in 4D; the w-component is unchanged by the rotation. The matrix S really satisfies $S^3 + \theta^2 S = 0$, in which case we may replace $S^3 = -\theta^2 S$ in Equation (65) and obtain Equation (53). The fact that S is a "root" of a 3rd-degree polynomial when the characteristic polynomial is 4th-degree is related to the linear algebraic concept of *minimal polynomial*.

Another reduction occurs, for example, when $b = c = d = e = f = 0$. The matrix S satisfies $S^2 + \theta^2 I = 0$ and Equation (65) reduces to Equation (40).

4.2 The Case $\delta \neq 0$

Suppose that $\delta = af - be + cd \neq 0$. The characteristic polynomial $t^4 + \theta^2 t^2 + \theta^4 \delta^2$ is a quadratic polynomial in t^2. We may use the quadratic formula to obtain its roots,

$$t^2 = \frac{-\theta^2 \pm \sqrt{\theta^4 - 4\theta^4 \delta^2}}{2} = -\theta^2 \left(\frac{1 \pm \sqrt{1 - 4\delta^2}}{2} \right) \tag{66}$$

Notice that

$$\begin{aligned} 1 - 4\delta^2 &= (a^2 + b^2 + c^2 + d^2 + e^2 + f^2) - 4(af - be + cd)^2 \\ &= \left[(a - f)^2 + (b + e)^2 + (c - d)^2 \right] \left[(a + f)^2 + (b - e)^2 + (c + d)^2 \right] \\ &= \rho^2 \geq 0 \end{aligned} \tag{67}$$

where

$$\rho = \sqrt{\left[(a - f)^2 + (b + e)^2 + (c - d)^2 \right] \left[(a + f)^2 + (b - e)^2 + (c + d)^2 \right]} \tag{68}$$

Thus, the right-hand side of Equation (66) is real-valued. Moreover, $1 - 4\delta^2 < 1$, so the right-hand side of Equation (66) consists of two *negative* real numbers. Taking the square roots produces the t-roots for the characteristic equation, all having zero real part,

$$t = \pm i\theta \sqrt{\frac{1 - \rho}{2}} =: \pm i\alpha, \quad \pm i\theta \sqrt{\frac{1 + \rho}{2}} =: \pm i\beta \tag{69}$$

The values α and β are defined by these equations, both values being positive real numbers with $\alpha \leq \beta$.

4.2.1 The Case $\rho > 0$

Using the linear difference equation approach, the powers of S are

$$S^k = (i\alpha)^k C_0 + (-i\alpha)^k C_1 + (i\beta)^k C_2 + (-i\beta)^k C_3 \tag{70}$$

where the C-matrices are determined by the initial conditions

$$\begin{aligned} I &= C_0 + C_1 + C_2 + C_3 \\ S &= i\alpha C_0 - i\alpha C_1 + i\beta C_2 - i\beta C_3 \\ S^2 &= -\alpha^2 C_0 - \alpha^2 C_1 - \beta^2 C_2 - \beta^2 C_3 \\ S^3 &= -i\alpha^3 C_0 + i\alpha^3 C_1 - i\beta^3 C_2 + i\beta^3 C_3 \end{aligned} \tag{71}$$

The solution is

$$\begin{aligned} C_0 &= \tfrac{\alpha\beta^3 I - i\beta^3 S + \alpha\beta S^2 - i\beta S^3}{2\alpha\beta(\beta^2-\alpha^2)} \\ C_1 &= \tfrac{\alpha\beta^3 I + i\beta^3 S + \alpha\beta S^2 + i\beta S^3}{2\alpha\beta(\beta^2-\alpha^2)} \\ C_2 &= \tfrac{-\alpha^3\beta I + i\alpha^3 S - \alpha\beta S^2 + i\alpha S^3}{2\alpha\beta(\beta^2-\alpha^2)} \\ C_3 &= \tfrac{-\alpha^3\beta I - i\alpha^3 S - \alpha\beta S^2 - i\alpha S^3}{2\alpha\beta(\beta^2-\alpha^2)} \end{aligned} \tag{72}$$

We may use Equation (70) to compute four consecutive powers of S, namely,

$$\begin{aligned} S^{4p} &= \tfrac{\alpha^{4p}(S^2+\beta^2 I) - \beta^{4p}(S^2+\alpha^2 I)}{\beta^2-\alpha^2} \\ S^{4p+1} &= \tfrac{\alpha^{4p}(S^3+\beta^2 S) - \beta^{4p}(S^3+\alpha^2 S)}{\beta^2-\alpha^2} \\ S^{4p+2} &= \tfrac{-\alpha^{4p+2}(S^2+\beta^2 I) - \beta^{4p+2}(S^2+\alpha^2 I)}{\beta^2-\alpha^2} \\ S^{4p+3} &= \tfrac{-\alpha^{4p+2}(S^3+\beta^2 S) - \beta^{4p+2}(S^3+\alpha^2 S)}{\beta^2-\alpha^2} \end{aligned} \tag{73}$$

for $p \geq 0$. The exponential of S factors to

$$\begin{aligned} \exp(S) &= \textstyle\sum_{k=0}^{\infty} \frac{S^k}{k!} \\ &= \textstyle\sum_{p=0}^{\infty} \frac{S^{4p}}{(4p)!} + \sum_{p=0}^{\infty} \frac{S^{4p+1}}{(4p+1)!} + \sum_{p=0}^{\infty} \frac{S^{4p+2}}{(4p+2)!} + \sum_{p=0}^{\infty} \frac{S^{4p+3}}{(4p+3)!} \\ &= \tfrac{1}{\beta^2-\alpha^2}\Big[\Big(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p}}{(4p)!}\Big)(S^2+\beta^2 I) - \Big(\sum_{p=0}^{\infty} \frac{\beta^{4p}}{(4p)!}\Big)(S^2+\alpha^2 I) + \\ &\quad \Big(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p}}{(4p+1)!}\Big)(S^3+\beta^2 S) - \Big(\sum_{p=0}^{\infty} \frac{\beta^{4p}}{(4p+1)!}\Big)(S^3+\alpha^2 S) - \\ &\quad \Big(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p+2}}{(4p+2)!}\Big)(S^2+\beta^2 I) + \Big(\sum_{p=0}^{\infty} \frac{\beta^{4p+2}}{(4p+2)!}\Big)(S^2+\alpha^2 I) - \\ &\quad \Big(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p+2}}{(4p+3)!}\Big)(S^3+\beta^2 S) + \Big(\sum_{p=0}^{\infty} \frac{\beta^{4p+2}}{(4p+3)!}\Big)(S^3+\alpha^2 S)\Big] \end{aligned} \tag{74}$$

This expansion appears to be quite complicated, but the following identities allow us to simplify the expression. Each equation below defines the function $f_j(x)$,

$$\begin{aligned} f_0(x) &= \textstyle\sum_{p=0}^{\infty} \frac{x^{4p}}{(4p)!} &&= \tfrac{1}{4}\left(e^x + e^{-x}\right) + \tfrac{1}{2}\cos(x) \\ f_1(x) &= \textstyle\sum_{p=0}^{\infty} \frac{x^{4p+1}}{(4p+1)!} &&= \tfrac{1}{4}\left(e^x - e^{-x}\right) + \tfrac{1}{2}\sin(x) \\ f_2(x) &= \textstyle\sum_{p=0}^{\infty} \frac{x^{4p+2}}{(4p+2)!} &&= \tfrac{1}{4}\left(e^x + e^{-x}\right) - \tfrac{1}{2}\cos(x) \\ f_3(x) &= \textstyle\sum_{p=0}^{\infty} \frac{x^{4p+3}}{(4p+3)!} &&= \tfrac{1}{4}\left(e^x - e^{-x}\right) - \tfrac{1}{2}\sin(x) \end{aligned} \tag{75}$$

Notice that $f_0(x) = f_1'(x)$, $f_1(x) = f_2'(x)$, $f_2(x) = f_3'(x)$, and $f_3(x) = f_0'(x)$. It is also easy to verify that $\exp(x) = f_0(x) + f_1(x) + f_2(x) + f_3(x)$. Equation (74) becomes

$$\begin{aligned} R = \exp(S) &= \tfrac{1}{\beta^2-\alpha^2}\left[f_0(\alpha)(S^2+\beta^2 I) - f_0(\beta)(S^2+\alpha^2 I) + \right. \\ &\quad \tfrac{f_1(\alpha)}{\alpha}(S^3+\beta^2 S) - \tfrac{f_1(\beta)}{\beta}(S^3+\alpha^2 S) - \\ &\quad f_2(\alpha)(S^2+\beta^2 I) + f_2(\beta)(S^2+\alpha^2 I) - \\ &\quad \left.\tfrac{f_3(\alpha)}{\alpha}(S^3+\beta^2 S) + \tfrac{f_3(\beta)}{\beta}(S^3+\alpha^2 S)\right] \\ &= \left(\tfrac{\beta^2\cos(\alpha)-\alpha^2\cos(\beta)}{\beta^2-\alpha^2}\right) I + \left(\tfrac{\beta^2\sin(\alpha)/\alpha-\alpha^2\sin(\beta)/\beta}{\beta^2-\alpha^2}\right) S + \\ &\quad \left(\tfrac{\cos(\alpha)-\cos(\beta)}{\beta^2-\alpha^2}\right) S^2 + \left(\tfrac{\sin(\alpha)/\alpha-\sin(\beta)/\beta}{\beta^2-\alpha^2}\right) S^3 \end{aligned} \tag{76}$$

4.2.2 The Case $\rho = 0$

Let $\rho = 0$. Based on Equation (67), the only time this can happen is when $(a, b, c) = \pm(f, -e, d)$. This condition implies $a^2 + b^2 + c^2 = 1/2$ and $\delta^2 = 1/4$. The t-roots are $\pm i\theta/\sqrt{2}$, each one having multiplicity 2. Define $\alpha = \theta/\sqrt{2}$. Using the linear difference equation approach, the powers of S are

$$S^k = (i\alpha)^k C_0 + (-i\alpha)^k C_1 + k(i\alpha)^k C_2 + k(-i\alpha)^k C_3 \tag{77}$$

where the C-matrices are determined by the initial conditions

$$\begin{aligned} I &= C_0 + C_1 \\ S &= i\alpha(C_0 - C_1 + C_2 - C_3) \\ S^2 &= -\alpha^2(C_0 + C_1 + 2C_2 + 2C_3) \\ S^3 &= -i\alpha^3(C_0 - C_1 + 3C_2 - 3C_3) \end{aligned} \tag{78}$$

The solution is

$$\begin{aligned} C_0 &= \tfrac{I}{2} - \tfrac{3iS}{2\sqrt{2}\,\theta} - \tfrac{iS^3}{\sqrt{2}\,\theta^3} \\ C_1 &= \tfrac{I}{2} + \tfrac{3iS}{2\sqrt{2}\,\theta} + \tfrac{iS^3}{\sqrt{2}\,\theta^3} \\ C_2 &= -\tfrac{I}{4} + \tfrac{iS}{2\sqrt{2}\,\theta} - \tfrac{S^2}{2\theta^2} + \tfrac{iS^3}{\sqrt{2}\,\theta^3} \\ C_3 &= -\tfrac{I}{4} - \tfrac{iS}{2\sqrt{2}\,\theta} - \tfrac{S^2}{2\theta^2} - \tfrac{iS^3}{\sqrt{2}\,\theta^3} \end{aligned} \tag{79}$$

We may use Equation (77) to compute four consecutive powers of S, namely,

$$\begin{aligned} S^{4p} &= \alpha^{4p} I - 4p\alpha^{4p}\left(\tfrac{I}{2} + \tfrac{S^2}{2\alpha^2}\right) \\ S^{4p+1} &= \alpha^{4p+1}\tfrac{S}{\alpha} - 4p\alpha^{4p+1}\left(\tfrac{S}{2\alpha} + \tfrac{S^3}{2\alpha^3}\right) \\ S^{4p+2} &= \alpha^{4p+2}\tfrac{S^2}{\alpha^2} + 4p\alpha^{4p+2}\left(\tfrac{I}{2} + \tfrac{S^2}{2\alpha^2}\right) \\ S^{4p+3} &= \alpha^{4p+3}\tfrac{S^3}{\alpha^3} + 4p\alpha^{4p+3}\left(\tfrac{S}{2\alpha} + \tfrac{S^3}{2\alpha^3}\right) \end{aligned} \tag{80}$$

for $p \geq 0$. Define $\hat{S} = S/\alpha$. The exponential of S factors to

$$\begin{aligned}
\exp(S) &= \textstyle\sum_{k=0}^{\infty} \frac{S^k}{k!} \\
&= \textstyle\sum_{p=0}^{\infty} \frac{S^{4p}}{(4p)!} + \sum_{p=0}^{\infty} \frac{S^{4p+1}}{(4p+1)!} + \sum_{p=0}^{\infty} \frac{S^{4p+2}}{(4p+2)!} + \sum_{p=0}^{\infty} \frac{S^{4p+3}}{(4p+3)!} \\
&= \left(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p}}{(4p)!}\right) I - \left(\textstyle\sum_{p=0}^{\infty} \frac{4p\alpha^{4p}}{(4p)!}\right) \left(\frac{I}{2} + \frac{\hat{S}^2}{2}\right) + \\
&\quad \left(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p+1}}{(4p+1)!}\right) \hat{S} - \left(\textstyle\sum_{p=0}^{\infty} \frac{4p\alpha^{4p+1}}{(4p+1)!}\right) \left(\frac{\hat{S}}{2} + \frac{\hat{S}^3}{2}\right) + \\
&\quad \left(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p+2}}{(4p+2)!}\right) \hat{S}^2 + \left(\textstyle\sum_{p=0}^{\infty} \frac{4p\alpha^{4p+2}}{(4p+2)!}\right) \left(\frac{I}{2} + \frac{\hat{S}^2}{2}\right) + \\
&\quad \left(\textstyle\sum_{p=0}^{\infty} \frac{\alpha^{4p+3}}{(4p+3)!}\right) \hat{S}^3 + \left(\textstyle\sum_{p=0}^{\infty} \frac{4p\alpha^{4p+3}}{(4p+3)!}\right) \left(\frac{\hat{S}}{2} + \frac{\hat{S}^3}{2}\right)
\end{aligned} \tag{81}$$

We may simplify this using the definitions for $f_j(x)$ and the following identities,

$$\begin{aligned}
g_0(x) &= \textstyle\sum_{p=0}^{\infty} \frac{4px^{4p}}{(4p)!} &&= xf_3(x) \\
g_1(x) &= \textstyle\sum_{p=0}^{\infty} \frac{4px^{4p+1}}{(4p+1)!} &&= xf_0(x) - f_1(x) \\
g_2(x) &= \textstyle\sum_{p=0}^{\infty} \frac{4px^{4p+2}}{(4p+2)!} &&= xf_1(x) - 2f_2(x) \\
g_3(x) &= \textstyle\sum_{p=0}^{\infty} \frac{4px^{4p+3}}{(4p+3)!} &&= xf_2(x) - 3f_3(x)
\end{aligned} \tag{82}$$

Notice that $g_0(x) = g_1'(x)$, $g_1(x) = g_2'(x)$, and $g_2(x) = g_3'(x)$. Equation (81) becomes

$$\begin{aligned}
R = \exp(S) &= \left(f_0(\alpha) - \tfrac{g_0(\alpha)}{2} + \tfrac{g_2(\alpha)}{2}\right) I + \left(f_1(\alpha) - \tfrac{g_1(\alpha)}{2} + \tfrac{g_3(\alpha)}{2}\right) \hat{S} + \\
&\quad \left(f_2(\alpha) - \tfrac{g_0(\alpha)}{2} + \tfrac{g_2(\alpha)}{2}\right) \hat{S}^2 + \left(f_3(\alpha) - \tfrac{g_1(\alpha)}{2} + \tfrac{g_3(\alpha)}{2}\right) \hat{S}^3 \\
&= \left(\tfrac{2\cos(\alpha) + \alpha\sin(\alpha)}{2}\right) I + \left(\tfrac{3\sin(\alpha) - \alpha\cos(\alpha)}{2\alpha}\right) S + \\
&\quad \left(\tfrac{\sin(\alpha)}{2\alpha}\right) S^2 + \left(\tfrac{\sin(\alpha) - \alpha\cos(\alpha)}{2\alpha^3}\right) S^3
\end{aligned} \tag{83}$$

4.3 Summary of the Formulas

We start with the skew-symmetric matrix

$$S = \theta \begin{bmatrix} 0 & a & b & d \\ -a & 0 & c & e \\ -b & -c & 0 & f \\ -d & -e & -f & 0 \end{bmatrix} \tag{84}$$

where $a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1$. Define

$$\delta = af - be + cd, \quad \rho = \sqrt{1 - 4\delta^2}, \quad \alpha = \theta\sqrt{\frac{1-\rho}{2}}, \quad \beta = \theta\sqrt{\frac{1+\rho}{2}} \tag{85}$$

Equation (76) really is the most general formula for the rotation matrix R corresponding to a skew-symmetric

matrix S. To repeat that formula,

$$\begin{aligned} R &= \left(\frac{\beta^2\cos(\alpha)-\alpha^2\cos(\beta)}{\beta^2-\alpha^2}\right)I + \left(\frac{\beta^2\sin(\alpha)/\alpha-\alpha^2\sin(\beta)/\beta}{\beta^2-\alpha^2}\right)S \\ &\quad + \left(\frac{\cos(\alpha)-\cos(\beta)}{\beta^2-\alpha^2}\right)S^2 + \left(\frac{\sin(\alpha)/\alpha-\sin(\beta)/\beta}{\beta^2-\alpha^2}\right)S^3 \end{aligned} \tag{86}$$

This equation was constructed for when $\beta > \alpha > 0$. However, when $\alpha = 0$, Equation(86) reduces to Equation (65), where it is necessary that $\beta = 1$,

$$R = I + S + \left(\frac{1-\cos(\beta)}{\beta^2}\right)S^2 + \left(\frac{\beta-\sin(\beta)}{\beta^3}\right)S^3 \tag{87}$$

The evaluation of $\sin(\alpha)/\alpha$ at $\alpha = 0$ is done in the limiting sense, $\lim_{\alpha\to 0}\sin(\alpha)/\alpha = 1$. Equation (83) was for the case $\beta = \alpha$ but may be obtained also from Equation (86) in the limiting sense as $\beta \to \alpha$. l'Hôpital's Rule may be applied to the coefficients of I, S, S^2, and S^3 to obtain the coefficients in Equation (83),

$$\begin{aligned} R &= \left(\frac{2\cos(\alpha)+\alpha\sin(\alpha)}{2}\right)I + \left(\frac{3\sin(\alpha)-\alpha\cos(\alpha)}{2\alpha}\right)S \\ &\quad + \left(\frac{\sin(\alpha)}{2\alpha}\right)S^2 + \left(\frac{\sin(\alpha)-\alpha\cos(\alpha)}{2\alpha^3}\right)S^3 \end{aligned} \tag{88}$$

4.4 Source Code for Random Generation

Here is some sample code to illustrate Equation (86) when $\beta \neq \alpha$.

```
void RandomRotation4D_BetaNotEqualAlpha ()
{
    double a = Mathd::SymmetricRandom();  // number is in [-1,1)
    double b = Mathd::SymmetricRandom();  // number is in [-1,1)
    double c = Mathd::SymmetricRandom();  // number is in [-1,1)
    double d = Mathd::SymmetricRandom();  // number is in [-1,1)
    double e = Mathd::SymmetricRandom();  // number is in [-1,1)
    double f = Mathd::SymmetricRandom();  // number is in [-1,1)
    Matrix4d S
    (
        0.0, a,   b,   d,
       -a,  0.0, c,   e,
       -b,  -c,  0.0, f,
       -d,  -e,  -f,  0.0
     );

    double theta = Mathd::Sqrt(a*a + b*b + c*c + d*d + e*e + f*f);
    double invLength = 1.0/theta;
    a *= invLength;
    b *= invLength;
    c *= invLength;
    d *= invLength;
    e *= invLength;
    f *= invLength;

    double delta = a*f - b*e + c*d;
    double delta2 = delta*delta;
    double p = Mathd::Sqrt(1.0 - 4.0*delta2);
    double alpha = theta*Mathd::Sqrt(0.5*(1.0 - p));
    double beta = theta*Mathd::Sqrt(0.5*(1.0 + p));
    double invDenom = 1.0/(beta*beta - alpha*alpha);
    double cosa = Mathd::Cos(alpha);
    double sina = Mathd::Sin(alpha);
    double cosb = Mathd::Cos(beta);
    double sinb = Mathd::Sin(beta);

    double k0 = (beta*beta*cosa - alpha*alpha*cosb)*invDenom;
```

```
    double k1 = (beta*beta*sina/alpha - alpha*alpha*sinb/beta)*invDenom;
    double k2 = (cosa - cosb)*invDenom;
    double k3 = (sina/alpha - sinb/beta)*invDenom;

    Matrix4d I = Matrix4d::IDENTITY;
    Matrix4d S2 = S*S;
    Matrix4d S3 = S*S2;
    Matrix4d R = k0*I + k1*S + k2*S2 + k3*S3;  // The random rotation matrix.

    // Sanity checks.
    Matrix4d S4 = S*S3;
    double theta2 = theta*theta;
    double theta4 = theta2*theta2;
    Matrix4d zero0 = S4 + theta2*S2 + (theta4*delta2)*I;  // = the zero matrix
    double one = R.Determinant();  // = 1
    Matrix4d zero1 = R.TransposeTimes(R) - I;  // = the zero matrix
}
```

Here is some sample code to illustrate the case when $\beta = \alpha$.

```
void RandomRotation4D_BetaEqualAlpha ()
{
    double a = Mathd::SymmetricRandom();  // number is in [-1,1)
    double b = Mathd::SymmetricRandom();  // number is in [-1,1)
    double c = Mathd::SymmetricRandom();  // number is in [-1,1)
    double mult = Mathd::Sqrt(0.5)/Mathd::Sqrt(a*a + b*b + c*c);
    a *= mult;
    b *= mult;
    c *= mult;
    double d = c;
    double e = -b;
    double f = a;
    double theta = Mathd::SymmetricRandom();
    Matrix4d S
    (
        0.0, a,   b,   d,
       -a,  0.0, c,   e,
       -b,  -c,  0.0, f,
       -d,  -e,  -f,  0.0
     );
    S *= theta;

    double lensqr = a*a + b*b + c*c + d*d + e*e + f*f;  // = 1
    double delta = a*f - b*e + c*d;  // = 1/2
    double delta2 = delta*delta;  // = 1/4
    double discr = 1.0 - 4.0*delta2;  // = 0
    double p = Mathd::Sqrt(Mathd::FAbs(discr)); // = 0
    double alpha = theta*Mathd::Sqrt(0.5);
    double cosa = Mathd::Cos(alpha);
    double sina = Mathd::Sin(alpha);

    double k0 = (2.0*cosa + alpha*sina)/2.0;
    double k1 = (3.0*sina - alpha*cosa)/(2.0*alpha);
    double k2 = sina/(2.0*alpha);
    double k3 = (sina - alpha*cosa)/(2.0*alpha*alpha*alpha);

    Matrix4d I = Matrix4d::IDENTITY;
    Matrix4d S2 = S*S;
    Matrix4d S3 = S*S2;
    Matrix4d R = k0*I + k1*S + k2*S2 + k3*S3;  // The random rotation matrix.

    // Sanity checks.
    Matrix4d S4 = S*S3;
    double theta2 = theta*theta;
    double theta4 = theta2*theta2;
    Matrix4d zero0 = S4 + theta2*S2 + (theta4*delta2)*I;  // = the zero matrix
    double one = R.Determinant();  // = 1
    Matrix4d zero1 = R.TransposeTimes(R) - I;  // = the zero matrix
}
```

5 Fixed Points, Invariant Spaces, and Direct Sums

A rotation in $\mathbb{R}^2$ is sometimes described as "rotation about a point", where the point is the origin–a 0-dimensional quantity living in a 2-dimensional space. A rotation in $\mathbb{R}^3$ is sometimes described as "rotation about an axis", where the axis is a line–a 1-dimensional quantity in a 3-dimensional space. Sometimes you will hear someone attempt to extend these geometric descriptions to a rotation in $\mathbb{R}^4$ by saying that you have a "rotation about a plane", the plane being a 2-dimensional quantity living in a 4-dimensional space. This description is not meaningful, and in fact is not technically correct. It is better to describe the rotations in terms of *fixed points* and *invariant spaces*.

In the following discussion, I assume that a rotation (rotation matrix) is not the identity (identity matrix).

5.1 Fixed Points

Let $\mathbf{F}(\mathbf{X})$ be a vector field; that is, $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$, which states that $\mathbf{F}$ is a function that maps n-tuples to n-tuples.

If $\mathbf{F}$ is a linear function, then it may be represented by a matrix M; that is, $\mathbf{F}(\mathbf{X}) = M\mathbf{X}$. This representation is assumed to be with respect to the standard Euclidean basis for $\mathbb{R}^n$, which is the set of vectors $\{\mathbf{e}_k\}_{k=1}^n$ for which $\mathbf{e}_k$ has a 1 in component k of the vector and all other components are 0. In $\mathbb{R}^2$, the basis is $\{(1,0),(0,1)\}$. In $\mathbb{R}^3$, the basis is $\{(1,0,0),(0,1,0),(0,0,1)\}$. In $\mathbb{R}^4$, the basis is $\{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}$. Other matrix representations are possible when you choose different bases.

A point $\mathbf{P}$ is said to be a *fixed point* for the function when $\mathbf{F}(\mathbf{P}) = \mathbf{P}$. For a linear function, $M\mathbf{P} = \mathbf{P}$.

The only fixed point for rotation in $\mathbb{R}^2$ is the origin $\mathbf{0}$. If R is the 2×2 matrix that represents the rotation, then $R\mathbf{0} = \mathbf{0}$. For all vectors $\mathbf{X} \neq \mathbf{0}$, $R\mathbf{X} \neq \mathbf{X}$.

A rotation in $\mathbb{R}^3$ has infinitely many fixed points, all lying on the axis of rotation. If the axis is the line containing the origin and having direction $\mathbf{U}$, then a point on the axis is $t\mathbf{U}$ for some scalar t. If R is the 3×3 matrix that represents the rotation about the axis, then $R(t\mathbf{U}) = t\mathbf{U}$.

5.2 Invariant Spaces

Let $V \subseteq \mathbb{R}^n$ be a vector subspace. The trivial subspaces are $\{\mathbf{0}\}$, the set consisting of only the origin, and $\mathbb{R}^n$, the entire set of n-tuples. In 2D, the nontrivial subspaces are lines containing the origin. In 3D, the nontrivial subspaces are lines and planes containing the origin.

Let $\mathbf{F}(\mathbf{X})$ be a linear function from $\mathbb{R}^n$ to $\mathbb{R}^n$. Define $\mathbf{F}(V)$ to be the set of all n-tuples that are mapped to by n-tuples in V; that is, $\mathbf{Y} \in \mathbf{F}(V)$ means that $\mathbf{Y} = \mathbf{F}(\mathbf{X})$ for some $\mathbf{X} \in V$.

V is said to be an *invariant space* relative to $\mathbf{F}$ when $\mathbf{F}(V) \subseteq V$; that is, any vector in V is transformed by $\mathbf{F}$ to a vector in V. It is not required that a nonzero vector in V be mapped to itself, so an invariant space might not have any nonzero fixed points.

The trivial subspaces are always invariant under linear transformations. If $\mathbf{F}(\mathbf{X}) = M\mathbf{X}$ for an $n \times n$ matrix M, then $M\mathbf{0} = \mathbf{0}$, in which case $\mathbf{0}$ is a fixed point and the set $\{\mathbf{0}\}$ is invariant. Every vector $\mathbf{X} \in \mathbb{R}^n$

is mapped to a vector $M\mathbf{X} \in \mathbb{R}^n$, so $\mathbb{R}^n$ is invariant. The more interesting question is whether there are nontrivial invariant subspaces.

The only invariant subspaces for a rotation in 2D are the trivial ones. A line containing the origin is a nontrivial subspace, but a rotation causes a line to rotate to some other line, so the original line is not invariant.

A rotation in 3D has two nontrivial invariant subspaces, the axis of rotation and the plane perpendicular to it. For example, consider the canonical rotation matrix

$$R = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{89}$$

The axis of rotation is the z-axis. Points in the xy-plane are rotated to other points in the xy-plane. Thus, the z-axis is an invariant subspace and the xy-plane is an invariant subspace.

5.3 Direct Sums

Let V_1 through V_m be nontrivial subspaces of $\mathbb{R}^n$. If every vector $\mathbf{X} \in \mathbb{R}^n$ can be represented uniquely as

$$\mathbf{X} = \sum_{i=1}^{m} \mathbf{X}_i, \quad \mathbf{X}_i \in V_i \tag{90}$$

then $\mathbb{R}^n$ is said to be a *direct sum* of the subspaces, denoted by

$$\mathbb{R}^n = V_1 \oplus \cdots \oplus V_m = \bigoplus_{i=1}^{m} V_i \tag{91}$$

For example, if V_2 is the set of points in the xy-plane and V_2 is the set of points on the z-axis, then $\mathbb{R}^3 = V_1 \oplus V_2$. A vector is represented as $(x, y, z) = (x, y, 0) + (0, 0, z)$, where $(x, y, 0) \in V_1$ and $(0, 0, z) \in V_2$.

Let V_1 through V_m be nontrivial subspaces of $\mathbb{R}^n$. Let $\mathbf{F}_i : V_i \to V_i$ be linear transformations on the subspaces; that is, V_i is invariant with respect to $\mathbf{F}_i$. Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation for which $\mathbf{F}(\mathbf{X}) = \mathbf{F}_i(\mathbf{X})$ whenever $\mathbf{X} \in V_i$. $\mathbf{F}$ is said to be a *direct sum* of the linear functions, denoted by

$$\mathbf{F} = \mathbf{F}_1 \oplus \cdots \oplus \mathbf{F}_m = \bigoplus_{i=1}^{m} \mathbf{F}_i \tag{92}$$

Let the dimension of V_i be denoted by $d_i = \dim(V_i)$. Necessarily, $n = d_1 + \cdots d_m$. If $\mathbf{F}_i(\mathbf{X}) = M_i\mathbf{X}$ for matrices M_i, where M_i is a $d_i \times d_i$ matrix, then $\mathbf{F}(\mathbf{X}) = M\mathbf{X}$ where M is an $n \times n$ matrix and $M = \mathrm{Diag}(M_1, \ldots, M_m)$.

Using our example of a 3D rotation about the z-axis, V_1 is the set of points in the xy-plane, V_2 is the set of points on the z-axis, and $\mathbb{R}^3 = V_1 \oplus V_2$. We may choose

$$M_1 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 \end{bmatrix} \tag{93}$$

in which case the 3D rotation matrix R is

$$R = M_1 \oplus M_2 = \text{Diag}(M_1, M_2) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{94}$$

More generally, let the axis of rotation have unit-length direction $\mathbf{C}$. Let $\mathbf{A}$ and $\mathbf{B}$ be vectors that span the plane perpendicular to the axis, but choose them so that $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ is an orthonormal set (vectors are unit-length and mutually perpendicular). The invariant subspaces are

$$V_1 = \{x\mathbf{A} + y\mathbf{B} : (x, y) \in \mathbb{R}^2\}, \;\; V_2 = \{z\mathbf{C} : z \in \mathbb{R}\} \tag{95}$$

and $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ is a basis for $\mathbb{R}^3$. The coordinates for this basis are $\mathbf{X} = [x\ y\ z]^\mathsf{T}$ and the rotation matrix relative to this basis is R of Equation (94). The rotated vector is $\mathbf{X}' = R\mathbf{X}$. To see the representation in Cartesian coordinates, define the matrix $P = [\mathbf{A}\ \mathbf{B}\ \mathbf{C}]$ whose columns are the specified vectors. Any vector $\mathbf{Y} \in \mathbb{R}^3$ may be represented as

$$\mathbf{Y} = P\mathbf{X} \tag{96}$$

for some vector $\mathbf{X}$. The inverse of P is just its transpose, so $\mathbf{X} = P^\mathsf{T}\mathbf{Y}$. The rotated vector is

$$\mathbf{Y}' = P\mathbf{X}' = PR\mathbf{X} = PRP^\mathsf{T}\mathbf{Y} \tag{97}$$

Thus, the rotation matrix in Cartesian coordinates is

$$R' = PRP^\mathsf{T} \tag{98}$$

The question now is how do these concepts extend to 4D rotations. The next section describes this.

5.4 Rotations in 4D

We found that the skew-symmetric matrix S of Equation (61) has eigenvalues $\pm i\alpha\theta$ and $\pm i\beta\theta$, where α and β are real numbers with $0 \le \alpha \le \beta$. The factorization is

$$S = P \begin{bmatrix} 0 & \alpha\theta & 0 & 0 \\ -\alpha\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta\theta \\ 0 & 0 & -\beta\theta & 0 \end{bmatrix} P^\mathsf{T} \tag{99}$$

for some orthogonal matrix P. The corresponding rotation matrix is

$$R = P \begin{bmatrix} \cos(\alpha\theta) & \sin(\alpha\theta) & 0 & 0 \\ -\sin(\alpha\theta) & \cos(\alpha\theta) & 0 & 0 \\ 0 & 0 & \cos(\beta\theta) & \sin(\beta\theta) \\ 0 & 0 & -\sin(\beta\theta) & \cos(\beta\theta) \end{bmatrix} P^\mathsf{T} \tag{100}$$

The problem is to determine what is P. This requires determining the invariant subspaces of $\mathbb{R}^4$ relative to the rotation R or, equivalently, relative to the skew-symmetric S. There will be two nontrivial invariant subspaces, each of dimension 2. Let $V_1 = \text{span}(\mathbf{U}_1, \mathbf{U}_2)$, which will be the invariant subspace that corresponds to α, and $V_2 = \text{span}(\mathbf{U}_3, \mathbf{U}_4)$, which be the invariant subspace that corresponds to β, where the four spanning vectors form an orthonormal set. P may be constructed as the orthogonal matrix whose columns are these four vectors,

$$P = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 & \mathbf{U}_3 & \mathbf{U}_4 \end{bmatrix} \tag{101}$$

Recall that $\alpha = \sqrt{(1-\rho)/2}$ and $\beta = \sqrt{(1+\rho)/2}$ for some $\rho \in [0,1]$. Also, $\delta = af - be + cd$, $\delta^2 = \alpha^2\beta^2$, and $1 - 4\delta^2 = \rho^2$. Finally, we have always assumed that $a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1$.

5.4.1 The Case $(d, e, f) = (0, 0, 0)$

The condition $(d, e, f) = (0, 0, 0)$ reduces S to the 3D case that we already understand. The skew-symmetric matrix is

$$S = \begin{bmatrix} 0 & a & b & 0 \\ -a & 0 & c & 0 \\ -b & -c & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{102}$$

Observe that $\delta = 0$, $\alpha = 0$, $\beta = 1$, and $a^2 + b^2 + c^2 = 1$. By observation,

$$\mathbf{U}_1 = \begin{bmatrix} c \\ -b \\ a \\ 0 \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{103}$$

It is easily verified that $S\mathbf{U}_1 = -0\mathbf{U}_2 = \mathbf{0}$ and $S\mathbf{U}_2 = 0\mathbf{U}_1 = \mathbf{0}$.

When $(b, c) \neq (0, 0)$, we may choose

$$\mathbf{U}_3 = \frac{1}{\sqrt{b^2 + c^2}} \begin{bmatrix} b \\ c \\ 0 \\ 0 \end{bmatrix} \tag{104}$$

This unit-length vector is perpendicular to $\mathbf{U}_1$ and $\mathbf{U}_2$. Let $\mathbf{e}_k$ be the vector with a 1 in component k and 0 in all other components. We may then compute

$$\mathbf{U}_4 = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ c & -b & a \\ \frac{b}{\sqrt{b^2+c^2}} & \frac{c}{\sqrt{b^2+c^2}} & 0 \end{bmatrix} + 0\mathbf{e}_4 = \frac{1}{\sqrt{b^2 + c^2}} \begin{bmatrix} -ac \\ ab \\ b^2 + c^2 \\ 0 \end{bmatrix} \tag{105}$$

When $(b, c) = (0, 0)$, it must be that $a = 1$, in which case we may choose

$$\mathbf{U}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{U}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{106}$$

so that P is the identity matrix.

In either case for (b, c), it is easily verified that $S\mathbf{U}_3 = -1\mathbf{U}_4$ and $S\mathbf{U}_4 = 1\mathbf{U}_3$.

5.4.2 The Case $(a, b, c) = (0, 0, 0)$

The condition $(a, b, c) = (0, 0, 0)$ implies that the skew-symmetric matrix is

$$S = \begin{bmatrix} 0 & 0 & 0 & d \\ 0 & 0 & 0 & e \\ 0 & 0 & 0 & f \\ -d & -e & -f & 0 \end{bmatrix} \tag{107}$$

Observe that $\delta = 0$, $\alpha = 0$, $\beta = 1$, and $d^2 + e^2 + f^2 = 1$. By observation,

$$\mathbf{U}_3 = \begin{bmatrix} d \\ e \\ f \\ 0 \end{bmatrix}, \quad \mathbf{U}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{108}$$

It is easily verified that $S\mathbf{U}_3 = -1\mathbf{U}_4$ and $S\mathbf{U}_4 = 1\mathbf{U}_3$.

When $(d, e) \neq (0, 0)$, we may choose

$$\mathbf{U}_1 = \frac{1}{\sqrt{d^2 + e^2}} \begin{bmatrix} e \\ -d \\ 0 \\ 0 \end{bmatrix} \tag{109}$$

This unit-length vector is perpendicular to $\mathbf{U}_3$ and $\mathbf{U}_4$. Let $\mathbf{e}_k$ be the vector with a 1 in component k and 0 in all other components. We may then compute

$$\mathbf{U}_2 = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ d & e & f \\ \frac{e}{\sqrt{d^2+e^2}} & \frac{-d}{\sqrt{d^2+e^2}} & 0 \end{bmatrix} + 0\mathbf{e}_4 = \frac{1}{\sqrt{d^2 + e^2}} \begin{bmatrix} df \\ ef \\ -d^2 - e^2 \\ 0 \end{bmatrix} \tag{110}$$

When $(d, e) = (0, 0)$, it must be that $f = 1$, in which case we may choose

$$\mathbf{U}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{111}$$

so that P is the identity matrix.

5.4.3 The Case $(d, e, f) \neq 0$, $(a, b, c) \neq 0$, and $\delta = 0$

Consider the case $0 = \delta = af - be + cd = (c, -b, a) \cdot (d, e, f)$; then $\alpha = 0$ and $\beta = 1$. Let $\mathbf{e}_k$ be the vector with a 1 in component k and 0 in all other components. We may formally construct an eigenvector for $\beta = 1$ using the determinant,

$$\begin{aligned} \mathbf{V} &= \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ -i & a & b & d \\ -a & -i & c & e \\ -b & -c & -i & f \end{bmatrix} \\ &= [(-d) + (bf + ae)i]\mathbf{e}_1 + [(-e) + (cf - ad)i]\mathbf{e}_2 + \\ &\quad [(-f) - (bd + ce)i]\mathbf{e}_3 + [(a^2 + b^2 + c^2 - 1)i]\mathbf{e}_4 \end{aligned} \tag{112}$$

Using the results of Section 1.3, $\mathbf{V} = \mathbf{X} + i\mathbf{Y}$ with

$$\mathbf{X} = \begin{bmatrix} -d \\ -e \\ -f \\ 0 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} bf + ae \\ cf - ad \\ -bd - ce \\ a^2 + b^2 + c^2 - 1 \end{bmatrix} \tag{113}$$

The vectors are orthogonal and have the same length. To see this, we know that $S\mathbf{X} = -\beta\theta\mathbf{Y}$ and $S\mathbf{Y} = \beta\theta\mathbf{X}$. First,

$$\mathbf{X}^\mathsf{T} S\mathbf{X} = -\theta\mathbf{X}^\mathsf{T}\mathbf{Y} \tag{114}$$

The left-hand side is a scalar, so its transpose is just the scalar itself,

$$\mathbf{X}^\mathsf{T} S\mathbf{X} = (\mathbf{X}^\mathsf{T} S\mathbf{X})^\mathsf{T} = \mathbf{X}^\mathsf{T} S^\mathsf{T}\mathbf{X} = -\mathbf{X}^\mathsf{T} S\mathbf{X} \tag{115}$$

The only way a scalar and its negative can be the same is when the scalar is zero. Thus, $\mathbf{X}^\mathsf{T} S\mathbf{X} = 0$, which implies $\mathbf{X}^\mathsf{T}\mathbf{Y} = 0$, so the vectors are orthogonal. Second,

$$\theta\mathbf{X}^\mathsf{T}\mathbf{X} = \mathbf{X}^\mathsf{T} S\mathbf{Y} = (\mathbf{X}^\mathsf{T} S\mathbf{Y})^\mathsf{T} = \mathbf{Y}^\mathsf{T} S^\mathsf{T}\mathbf{X} = -\mathbf{Y}^\mathsf{T} S\mathbf{X} = \theta\mathbf{Y}^\mathsf{T}\mathbf{Y} \tag{116}$$

which implies $|\mathbf{X}| = |\mathbf{Y}|$.

Two real-valued eigenvectors for $\pm i$ that are orthogonal and unit-length are the normalized $-\mathbf{X}$ and $-\mathbf{Y}$,

$$\mathbf{U}_3 = \frac{1}{\sqrt{d^2+e^2+f^2}} \begin{bmatrix} d \\ e \\ f \\ 0 \end{bmatrix}, \quad \mathbf{U}_4 = \frac{1}{\sqrt{d^2+e^2+f^2}} \begin{bmatrix} -ae-bf \\ ad-cf \\ bd+ce \\ d^2+e^2+f^2 \end{bmatrix} \tag{117}$$

where we have used $a^2+b^2+c^2+d^2+e^2+f^2=1$.

The condition $\delta = 0$ allows us to choose

$$\mathbf{U}_1 = \frac{1}{\sqrt{a^2+b^2+c^2}} \begin{bmatrix} c \\ -b \\ a \\ 0 \end{bmatrix} \tag{118}$$

It is easily verified that $\mathbf{U}_1$ is perpendicular to both $\mathbf{U}_3$ and $\mathbf{U}_4$. We may choose the final vector using

$$\begin{aligned} \mathbf{U}_2 &= \tfrac{-1}{(d^2+e^2+f^2)\sqrt{a^2+b^2+c^2}} \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ d & e & f & 0 \\ -ae-bf & ad-cf & bd+ce & d^2+e^2+f^2 \\ c & -b & a & 0 \end{bmatrix} \\ &= \tfrac{1}{\sqrt{a^2+b^2+c^2}} \begin{bmatrix} ae+bf \\ -ad+cf \\ -bd-ce \\ a^2+b^2+c^2 \end{bmatrix} \end{aligned} \tag{119}$$

5.4.4 The Case $(d,e,f) \neq 0$, $(a,b,c) \neq 0$, and $\delta \neq 0$

In our construction when the rotation is not the identity matrix, we know that $\beta > 0$. The eigenvalues $i\beta\theta$ and $-i\beta\theta$ each have 1-dimensionsal eigenspaces. The eigenvectors for $i\beta\theta$ must satisfy the equation $(S - i\beta\theta I)\mathbf{V} = \mathbf{0}$. Because the eigenspace is 1-dimensional, the matrix $S - i\beta\theta I$ has rank 3. Equivalently, the matrix $S/\theta - i\beta I$ has rank 3. Let $\mathbf{e}_k$ be the vector with a 1 in component k and 0 in all other components.

We may formally construct an eigenvector using the determinant,

$$\begin{aligned}
\mathbf{V} &= \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ -i\beta & a & b & d \\ -a & -i\beta & c & e \\ -b & -c & -i\beta & f \end{bmatrix} \\
&= [(c\delta - d\beta^2) + \beta(bf + ae)i]\mathbf{e}_1 + [(-b\delta - e\beta^2) + \beta(cf - ad)i]\mathbf{e}_2 + \\
&\quad [(a\delta - f\beta^2) - \beta(bd + ce)i]\mathbf{e}_3 + [\beta(a^2 + b^2 + c^2 - \beta^2)i]\mathbf{e}_4
\end{aligned} \tag{120}$$

Using the results of Section 1.3, $\mathbf{V} = \mathbf{X} + i\mathbf{Y}$ with

$$\mathbf{X} = \begin{bmatrix} c\delta - d\beta^2 \\ -b\delta - e\beta^2 \\ a\delta - f\beta^2 \\ 0 \end{bmatrix}, \quad \mathbf{Y} = \beta \begin{bmatrix} bf + ae \\ cf - ad \\ -bd - ce \\ a^2 + b^2 + c^2 - \beta^2 \end{bmatrix} \tag{121}$$

The vectors are orthogonal and have the same length. To see this, we know that $S\mathbf{X} = -\beta\theta\mathbf{Y}$ and $S\mathbf{Y} = \beta\theta\mathbf{X}$. First,

$$\mathbf{X}^\mathsf{T} S\mathbf{X} = -\beta\theta\mathbf{X}^\mathsf{T}\mathbf{Y} \tag{122}$$

The left-hand side is a scalar, so its transpose is just the scalar itself,

$$\mathbf{X}^\mathsf{T} S\mathbf{X} = (\mathbf{X}^\mathsf{T} S\mathbf{X})^\mathsf{T} = \mathbf{X}^\mathsf{T} S^\mathsf{T}\mathbf{X} = -\mathbf{X}^\mathsf{T} S\mathbf{X} \tag{123}$$

The only way a scalar and its negative can be the same is when the scalar is zero. Thus, $\mathbf{X}^\mathsf{T} S\mathbf{X} = 0$, which implies $\mathbf{X}^\mathsf{T}\mathbf{Y} = 0$, so the vectors are orthogonal. Second,

$$\beta\theta\mathbf{X}^\mathsf{T}\mathbf{X} = \mathbf{X}^\mathsf{T} S\mathbf{Y} = (\mathbf{X}^\mathsf{T} S\mathbf{Y})^\mathsf{T} = \mathbf{Y}^\mathsf{T} S^\mathsf{T}\mathbf{X} = -\mathbf{Y}^\mathsf{T} S\mathbf{X} = \beta\theta\mathbf{Y}^\mathsf{T}\mathbf{Y} \tag{124}$$

which implies $|\mathbf{X}| = |\mathbf{Y}|$. We may choose

$$\mathbf{U}_3 = \frac{\mathbf{X}}{|\mathbf{X}|}, \quad \mathbf{U}_4 = \frac{\mathbf{Y}}{|\mathbf{Y}|} \tag{125}$$

We will construct vectors $\mathbf{U}_1$ and $\mathbf{U}_2$ for α next.

The Case $\alpha < \beta$. The same construction applies to α as it did to β, producing

$$\mathbf{X}' = \begin{bmatrix} c\delta - d\alpha^2 \\ -b\delta - e\alpha^2 \\ a\delta - f\alpha^2 \\ 0 \end{bmatrix}, \quad \mathbf{Y}' = \alpha \begin{bmatrix} bf + ae \\ cf - ad \\ -bd - ce \\ a^2 + b^2 + c^2 - \alpha^2 \end{bmatrix} \tag{126}$$

and

$$\mathbf{U}_1 = \frac{\mathbf{X}'}{|\mathbf{X}'|}, \quad \mathbf{U}_2 = \frac{\mathbf{Y}'}{|\mathbf{Y}'|} \tag{127}$$

The Case $\alpha = \beta$. When $\alpha = \beta = 1/\sqrt{2}$, it must be that $\rho = 0$, $\delta = \sigma/2$ for $\sigma \in \{-1, +1\}$, $(d, e, f) = \sigma(c, -b, a)$, and $a^2 + b^2 + c^2 = 1/2 = d^2 + e^2 + f^2$. The skew-symmetric matrix is

$$S = \theta \begin{bmatrix} 0 & a & b & \sigma c \\ -a & 0 & c & -\sigma b \\ -b & -c & 0 & \sigma a \\ -\sigma c & \sigma b & -\sigma a & 0 \end{bmatrix} \tag{128}$$

Observe that $S^2 + (\theta^2/2)I = 0$. Define $\hat{S} = S/\theta$. When solving for eigenvectors for $i\alpha = i/\sqrt{2}$, we solve $(\hat{S} - (i/\sqrt{2})I)\mathbf{V} = \mathbf{0}$. Because the eigenspace is 2-dimensional, the first two rows of $\hat{S} - (i/\sqrt{2})I$ are linearly independent and must be orthogonal to the eigenvectors of $i/\sqrt{2}$. However, this means that the first two rows are in the orthogonal complement of the eigenspace for $i/\sqrt{2}$, so these rows are linearly independent eigenvectors for $-i/\sqrt{2}$.

Consider the first row of $\hat{S} - (i/\sqrt{2})I$. The real and imaginary parts are the candidates we want for producing $\mathbf{U}_3$ and $\mathbf{U}_4$. Specifically, the normalized real and imaginary part are

$$\mathbf{U}_3 = \sqrt{2} \begin{bmatrix} 0 \\ a \\ b \\ \sigma c \end{bmatrix}, \quad \mathbf{U}_4 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{129}$$

A similar argument applied to $(-b, -c, -i/\sqrt{2}, \sigma a)$ leads to

$$\mathbf{U}_1 = \sqrt{2} \begin{bmatrix} -b \\ -c \\ 0 \\ \sigma a \end{bmatrix}, \quad \mathbf{U}_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} \tag{130}$$

5.4.5 A Summary of the Factorizations

In all cases, we have assumed $a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1$ so that $\hat{S} \neq 0$. The factorization is

$$\hat{S} = P \begin{bmatrix} 0 & \alpha & 0 & 0 \\ -\alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \\ 0 & 0 & -\beta & 0 \end{bmatrix} P^{\mathsf{T}} \tag{131}$$

The case $(d,e,f)=(0,0,0)$ and $(b,c)\neq(0,0)$:

$$\hat{S}=\begin{bmatrix} 0 & a & b & 0 \\ -a & 0 & c & 0 \\ -b & -c & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\quad P=\begin{bmatrix} c & 0 & \frac{b}{\sqrt{b^2+c^2}} & \frac{-ac}{\sqrt{b^2+c^2}} \\ -b & 0 & \frac{c}{\sqrt{b^2+c^2}} & \frac{ab}{\sqrt{b^2+c^2}} \\ a & 0 & 0 & \sqrt{b^2+c^2} \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{132}$$

The case $(d,e,f)=(0,0,0)$ and $(b,c)=(0,0)$:

$$\hat{S}=\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\quad P=\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{133}$$

The case $(a,b,c)=(0,0,0)$ and $(d,e)\neq(0,0)$:

$$\hat{S}=\begin{bmatrix} 0 & 0 & 0 & d \\ 0 & 0 & 0 & e \\ 0 & 0 & 0 & f \\ -d & -e & -f & 0 \end{bmatrix},\quad P=\begin{bmatrix} \frac{e}{\sqrt{d^2+e^2}} & \frac{df}{\sqrt{d^2+e^2}} & d & 0 \\ \frac{-d}{\sqrt{d^2+e^2}} & \frac{ef}{\sqrt{d^2+e^2}} & e & 0 \\ 0 & -\sqrt{d^2+e^2} & f & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{134}$$

The case $(a,b,c)=(0,0,0)$ and $(d,e)=(0,0)$:

$$\hat{S}=\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix},\quad P=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{135}$$

The case $(a,b,c)\neq(0,0,0)$, $(d,e,f)\neq(0,0,0)$, and $\delta=0$:

$$P=\begin{bmatrix} \frac{c}{\sqrt{a^2+b^2+c^2}} & \frac{ae+bf}{\sqrt{a^2+b^2+c^2}} & \frac{d}{\sqrt{d^2+e^2+f^2}} & \frac{-ae-bf}{\sqrt{d^2+e^2+f^2}} \\ \frac{-b}{\sqrt{a^2+b^2+c^2}} & \frac{-ad+cf}{\sqrt{a^2+b^2+c^2}} & \frac{e}{\sqrt{d^2+e^2+f^2}} & \frac{ad-cf}{\sqrt{d^2+e^2+f^2}} \\ \frac{a}{\sqrt{a^2+b^2+c^2}} & \frac{-bd-ce}{\sqrt{a^2+b^2+c^2}} & \frac{f}{\sqrt{d^2+e^2+f^2}} & \frac{bd+ce}{\sqrt{d^2+e^2+f^2}} \\ 0 & \sqrt{a^2+b^2+c^2} & 0 & \sqrt{d^2+e^2+f^2} \end{bmatrix} \tag{136}$$

The case $(a,b,c) \neq (0,0,0)$, $(d,e,f) \neq (0,0,0)$, $\delta = 0$, $\alpha < \beta$:

$$P = \begin{bmatrix} \frac{c\delta - d\alpha^2}{\ell} & \frac{\alpha(bf+ae)}{\ell} & \frac{c\delta - d\beta^2}{m} & \frac{\beta(bf+ae)}{m} \\ \frac{-b\delta - e\alpha^2}{\ell} & \frac{\alpha(cf-ad)}{\ell} & \frac{-b\delta - e\beta^2}{m} & \frac{\beta(cf-ad)}{m} \\ \frac{a\delta - f\alpha^2}{\ell} & \frac{\alpha(-bd-ce)}{\ell} & \frac{a\delta - f\beta^2}{m} & \frac{\beta(-bd-ce)}{m} \\ 0 & \frac{\alpha(a^2+b^2+c^2-\alpha^2)}{\ell} & 0 & \frac{\beta(a^2+b^2+c^2-\beta^2)}{m} \end{bmatrix} \tag{137}$$

where $\ell = \sqrt{(c\delta - d\alpha^2)^2 + (b\delta + e\alpha^2)^2 + (a\delta - f\alpha^2)^2}$ and $m = \sqrt{(c\delta - d\beta^2)^2 + (b\delta + e\beta^2)^2 + (a\delta - f\beta^2)^2}$.

The case $(a,b,c) \neq (0,0,0)$, $(d,e,f) \neq (0,0,0)$, $\delta = 0$, $\alpha = \beta$:

$$\hat{S} = \begin{bmatrix} 0 & a & b & \sigma c \\ -a & 0 & c & -\sigma b \\ -b & -c & 0 & \sigma a \\ -\sigma c & \sigma b & -\sigma a & 0 \end{bmatrix} \tag{138}$$

and

$$P = \begin{bmatrix} -b\sqrt{2} & 0 & 0 & -1 \\ -c\sqrt{2} & 0 & a\sqrt{2} & 0 \\ 0 & -1 & b\sqrt{2} & 0 \\ \sigma a\sqrt{2} & 0 & \sigma c\sqrt{2} & 0 \end{bmatrix} \tag{139}$$

References

[Bra83] Martin Braun. *Differential Equations and Their Applications.* Springer-Verlag, New York, NY, 3rd edition, 1983.

[Ebe03] David Eberly. *Game Physics.* Morgan Kaufmann, San Francisco, CA, 2003.

[HJ85] Roger A. Horn and Charles R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, England, 1985.

[HS74] Morris W. Hirsch and Stephen Smale. *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, Orlando, FL, 1974.