Three-Dimensional Proper and Improper Rotation Matrices

1. Proper and improper rotation matrices

A real orthogonal matrix R is a matrix whose elements are real numbers and satisfies $R^{-1} = R^{\mathsf{T}}$ (or equivalently, $RR^{\mathsf{T}} = \mathbf{I}$, where $\mathbf{I}$ is the 3×3 identity matrix). Taking the determinant of the equation $RR^{\mathsf{T}} = \mathbf{I}$ and using the fact that $\det(R^{\mathsf{T}}) = \det R$, it follows that $(\det R)^2 = 1$, which implies that either $\det R = 1$ or $\det R = -1$.

A real orthogonal matrix with $\det R = 1$ provides a matrix representation of a *proper* rotation. The most general rotation matrix represents a counterclockwise rotation by an angle θ about a fixed axis that lies along the unit vector $\hat{\boldsymbol{n}}$. The rotation matrix operates on vectors to produce rotated vectors, while the coordinate axes are held fixed. In typical parlance, a rotation refers to a proper rotation. Thus, in the following sections of these notes we will often omit the adjective *proper* when referring to a proper rotation.

A real orthogonal matrix with $\det R = -1$ provides a matrix representation of an *improper* rotation. To perform an improper rotation requires mirrors. That is, the most general improper rotation matrix is a product of a proper rotation by an angle θ about some axis $\hat{\boldsymbol{n}}$ and a mirror reflection through a plane that passes through the origin and is perpendicular to $\hat{\boldsymbol{n}}$.

In these notes, we shall explore the matrix representations of three-dimensional proper and improper rotations. By determining the most general form for a three-dimensional proper and improper rotation matrix, we can then examine any 3×3 orthogonal matrix and determine the rotation and/or reflection it produces as an operator acting on vectors. If the matrix is a proper rotation, then the axis of rotation and angle of rotation can be determined. If the matrix is an improper rotation, then the reflection plane and the rotation, if any, about the normal to that plane can be determined.

2. Properties of the 3×3 rotation matrix

A rotation in the x–y plane by an angle θ measured counterclockwise from the positive x-axis is represented by the 2×2 real orthogonal matrix with determinant equal to 1,

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} .$$

If we consider this rotation as occurring in three-dimensional space, then it can be described as a counterclockwise rotation by an angle θ about the z-axis. The matrix representation of this three-dimensional rotation is given by the 3×3 real orthogonal matrix with determinant equal to 1 [cf. eq. (7.18) on p. 129 of Boas],

$$R(\boldsymbol{k},\theta) \equiv \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1}$$

where the axis of rotation and the angle of rotation are specified as arguments of R.

The most general three-dimensional rotation, denoted by $R(\hat{\boldsymbol{n}},\theta)$, can be specified by an axis of rotation, $\hat{\boldsymbol{n}}$, and a rotation angle θ.[1] Conventionally, a positive rotation angle corresponds to a counterclockwise rotation. The direction of the axis is determined by the right hand rule. Namely, curl the fingers of your right hand around the axis of rotation, where your fingers point in the θ direction. Then, your thumb points perpendicular to the plane of rotation in the direction of $\hat{\boldsymbol{n}}$. In general, rotation matrices do not commute under multiplication. However, if both rotations are taken with respect to the *same* fixed axis, then

$$R(\hat{\boldsymbol{n}},\theta_1)R(\hat{\boldsymbol{n}},\theta_2) = R(\hat{\boldsymbol{n}},\theta_1+\theta_2)\,. \tag{2}$$

Simple geometric considerations will convince you that the following relations are satisfied:

$$R(\hat{\boldsymbol{n}},\theta+2\pi k) = R(\hat{\boldsymbol{n}},\theta)\,, \qquad k = 0, \pm 1\, \pm 2\, \ldots\,, \tag{3}$$

$$[R(\hat{\boldsymbol{n}},\theta)]^{-1} = R(\hat{\boldsymbol{n}},-\theta) = R(-\hat{\boldsymbol{n}},\theta)\,. \tag{4}$$

Combining these two results, it follows that

$$R(\hat{\boldsymbol{n}},2\pi-\theta) = R(-\hat{\boldsymbol{n}},\theta)\,, \tag{5}$$

which implies that any three-dimensional rotation can be described by a counterclockwise rotation by an angle θ about an arbitrary axis $\hat{\boldsymbol{n}}$, where $0 \leq \theta \leq \pi$. However, if we substitute $\theta = \pi$ in eq. (5), we conclude that

$$R(\hat{\boldsymbol{n}},\pi) = R(-\hat{\boldsymbol{n}},\pi)\,, \tag{6}$$

which means that for the special case of $\theta = \pi$, $R(\hat{\boldsymbol{n}},\pi)$ and $R(-\hat{\boldsymbol{n}},\pi)$ represent the *same* rotation. In particular, note that

$$[R(\hat{\boldsymbol{n}},\pi)]^2 = \mathbf{I}\,. \tag{7}$$

Indeed for any choice of $\hat{\boldsymbol{n}}$, the $R(\hat{\boldsymbol{n}},\pi)$ are the only non-trivial rotation matrices whose square is equal to the identity operator. Finally, if $\theta = 0$ then $R(\hat{\boldsymbol{n}},0) = \mathbf{I}$ is the identity operator (sometimes called the trivial rotation), independently of the direction of $\hat{\boldsymbol{n}}$.

[1] There is an alternative convention for the range of possible angles θ and rotation axes $\hat{\boldsymbol{n}}$. We say that $\hat{\boldsymbol{n}} = (n_1, n_2, n_3) > 0$ if the first nonzero component of $\hat{\boldsymbol{n}}$ is positive. That is $n_3 > 0$ if $n_1 = n_2 = 0$, $n_2 > 0$ if $n_1 = 0$, and $n_1 > 0$ otherwise. Then, all possible rotation matrices $R(\hat{\boldsymbol{n}},\theta)$ correspond to $\hat{\boldsymbol{n}} > 0$ and $0 \leq \theta < 2\pi$. However, we will not employ this convention in these notes.

To learn more about the properties of a general three-dimensional rotation, consider the matrix representation $R(\hat{\boldsymbol{n}},\theta)$ with respect to the standard basis $\mathcal{B}_s = \{\boldsymbol{i}\,,\,\boldsymbol{j}\,,\,\boldsymbol{k}\}$. We can define a new coordinate system in which the unit vector $\hat{\boldsymbol{n}}$ points in the direction of the new z-axis; the corresponding new basis will be denoted by $\mathcal{B}'$. The matrix representation of the rotation with respect to $\mathcal{B}'$ is then given by $R(\boldsymbol{k},\theta)$. Using the formalism developed in the class handout, *Vector coordinates, matrix elements and changes of basis*, there exists an invertible matrix P such that

$$R(\hat{\boldsymbol{n}},\theta) = PR(\boldsymbol{k},\theta)P^{-1}\,, \tag{8}$$

where $R(\boldsymbol{k},\theta)$ is given by eq. (1). In Section 3, we will determine the matrix P, in which case eq. (14) provides an explicit form for the most general three-dimensional rotation. However, the mere existence of the matrix P in eq. (8) is sufficient to provide a simple algorithm for determining the rotation axis $\hat{\boldsymbol{n}}$ (up to an overall sign) and the rotation angle θ that characterize a general three-dimensional rotation matrix.

To determine the rotation angle θ, we note that the properties of the trace imply that $\mathrm{Tr}(PRP^{-1}) = \mathrm{Tr}(P^{-1}PR) = \mathrm{Tr}\,R$, since one can cyclically permute the matrices within the trace without modifying its value. Hence, it immediately follows from eq. (8) that

$$\mathrm{Tr}\;R(\hat{\boldsymbol{n}},\theta) = \mathrm{Tr}\;R(\boldsymbol{k},\theta) = 2\cos\theta + 1\,, \tag{9}$$

after taking the trace of eq. (1). By convention, $0 \leq \theta \leq \pi$, which implies that $\sin\theta \geq 0$. Hence, the rotation angle is uniquely determined by eq. (9) To identify $\hat{\boldsymbol{n}}$, we observe that any vector that is parallel to the axis of rotation is unaffected by the rotation itself. This last statement can be expressed as an eigenvalue equation,

$$R(\hat{\boldsymbol{n}},\theta)\hat{\boldsymbol{n}} = \hat{\boldsymbol{n}}\,. \tag{10}$$

Thus, $\hat{\boldsymbol{n}}$ is an eigenvector of $R(\hat{\boldsymbol{n}},\theta)$ corresponding to the eigenvalue 1. In particular, the eigenvalue 1 is unique for any $\theta \neq 0$, in which case $\hat{\boldsymbol{n}}$ can be determined up to an overall sign by computing the eigenvalues and the normalized eigenvectors of $R(\hat{\boldsymbol{n}},\theta)$. A simple proof of this result is given in Appendix A. Here, we shall establish this assertion by noting that the eigenvalues of any matrix are invariant with respect to a similarity transformation. Using eq. (8), it follows that the eigenvalues of $R(\hat{\boldsymbol{n}},\theta)$ are identical to the eigenvalues of $R(\boldsymbol{k},\theta)$. The latter can be obtained from the characteristic equation,

$$(1-\lambda)\left[(\cos\theta-\lambda)^2 + \sin^2\theta\right] = 0\,,$$

which simplifies to:

$$(1-\lambda)(\lambda^2 - 2\lambda\cos\theta + 1) = 0\,,$$

after using $\sin^2\theta + \cos^2\theta = 1$. Solving the quadratic equation, $\lambda^2 - 2\lambda\cos\theta + 1 = 0$, yields:

$$\lambda = \cos\theta \pm \sqrt{\cos^2\theta - 1} = \cos\theta \pm i\sqrt{1-\cos^2\theta} = \cos\theta \pm i\sin\theta = e^{\pm i\theta}\,. \tag{11}$$

It follows that the three eigenvalues of $R(\boldsymbol{k},\theta)$ are given by,

$$\lambda_1 = 1\,, \quad \lambda_2 = e^{i\theta}\,, \quad \lambda_3 = e^{-i\theta}\,, \qquad \text{for } \ 0 \leq \theta \leq \pi\,.$$

There are three distinct cases:

Case 1:	$\theta = 0$	$\lambda_1 = \lambda_2 = \lambda_3 = 1\,,$	$R(\hat{\boldsymbol{n}},0) = \mathbf{I}\,,$
Case 2:	$\theta = \pi$	$\lambda_1 = 1\,,\, \lambda_2 = \lambda_3 = -1\,,$	$R(\hat{\boldsymbol{n}},\pi)\,,$
Case 3:	$0 < \theta < \pi$	$\lambda_1 = 1, \lambda_2 = e^{i\theta}\,,\, \lambda_3 = e^{-i\theta}\,,$	$R(\hat{\boldsymbol{n}},\theta)\,,$

where the corresponding rotation matrix is indicated for each of the three cases. Indeed, for $\theta \neq 0$ the eigenvalue 1 is unique. Moreover, the other two eigenvalues are complex conjugates of each other, whose real part is equal to $\cos\theta$, which uniquely fixes the rotation angle in the convention where $0 \leq \theta \leq \pi$. Case 1 corresponds to the identity (i.e. no rotation) and Case 2 corresponds to a 180° rotation about the axis $\hat{\boldsymbol{n}}$. In Case 2, the interpretation of the the doubly degenerate eigenvalue -1 is clear. Namely, the corresponding two linearly independent eigenvectors span the plane that passes through the origin and is perpendicular to $\hat{\boldsymbol{n}}$. In particular, the two doubly degenerate eigenvectors (along with any linear combination $\vec{\boldsymbol{v}}$ of these eigenvectors that lies in the plane perpendicular to $\hat{\boldsymbol{n}}$) are inverted by the 180° rotation and hence must satisfy $R(\hat{\boldsymbol{n}},\pi)\vec{\boldsymbol{v}} = -\vec{\boldsymbol{v}}$.

Since $\hat{\boldsymbol{n}}$ is a real vector of unit length, it is determined only up to an overall sign by eq. (10) when its corresponding eigenvalue 1 is unique. This sign ambiguity is immaterial in Case 2 in light of eq. (6). The sign ambiguity in Case 3 cannot be resolved without further analysis. To make further progress, in Section 3 we shall obtain the general expression for the three dimensional rotation matrix $R(\hat{\boldsymbol{n}},\theta)$.

3. An explicit formula for the matrix elements of a general 3×3 rotation matrix

In this section, the matrix elements of $R(\hat{\boldsymbol{n}},\theta)$ will be denoted by R_{ij}. Since $R(\hat{\boldsymbol{n}},\theta)$ describes a rotation by an angle θ about an axis $\hat{\boldsymbol{n}}$, the formula for R_{ij} that we seek depends on the angle θ and on the coordinates of $\hat{\boldsymbol{n}} = (n_1\,,\, n_2\,,\, n_3)$ with respect to a fixed Cartesian coordinate system. Note that since $\hat{\boldsymbol{n}}$ is a unit vector, it follows that:

$$n_1^2 + n_2^2 + n_3^2 = 1\,. \tag{12}$$

Suppose we are given a rotation matrix $R(\hat{\boldsymbol{n}},\theta)$ and are asked to determine the axis of rotation $\hat{\boldsymbol{n}}$ and the rotation angle θ. The matrix $R(\hat{\boldsymbol{n}},\theta)$ is specified with respect to the standard basis $\mathcal{B}_s = \{\boldsymbol{i}\,,\, \boldsymbol{j}\,,\, \boldsymbol{k}\}$. We shall rotate to a new orthonormal basis,

$$\mathcal{B}' = \{\boldsymbol{i}'\,,\, \boldsymbol{j}'\,,\, \boldsymbol{k}'\}\,,$$

in which new positive z-axis points in the direction of $\hat{\boldsymbol{n}}$. That is,

$$\boldsymbol{k}' = \hat{\boldsymbol{n}} \equiv (n_1, n_2, n_3)\,, \qquad \text{where } \ n_1^2 + n_2^2 + n_3^2 = 1\,.$$

The new positive y-axis can be chosen to lie along

$$\boldsymbol{j}' = \left(\frac{-n_2}{\sqrt{n_1^2+n_2^2}}\,,\ \frac{n_1}{\sqrt{n_1^2+n_2^2}}\,,\ 0\right)\,,$$

since by construction, $\boldsymbol{j}'$ is a unit vector orthogonal to $\boldsymbol{k}'$. We complete the new right-handed coordinate system by choosing:

$$\boldsymbol{i}' = \boldsymbol{j}'\times\boldsymbol{k}' = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{-n_2}{\sqrt{n_1^2+n_2^2}} & \dfrac{n_1}{\sqrt{n_1^2+n_2^2}} & 0 \\ n_1 & n_2 & n_3 \end{vmatrix} = \left(\frac{n_3 n_1}{\sqrt{n_1^2+n_2^2}}\,,\ \frac{n_3 n_2}{\sqrt{n_1^2+n_2^2}}\,,\ -\sqrt{n_1^2+n_2^2}\right)\,.$$

Following the class handout entitled, *Vector coordinates, matrix elements and changes of basis*, we determine the matrix P whose matrix elements are defined by

$$\boldsymbol{b}'_j = \sum_{i=1}^{n} P_{ij}\hat{\boldsymbol{e}}_i\,,$$

where the $\hat{\boldsymbol{e}}_i$ are the basis vectors of $\mathcal{B}_s$ and the $\boldsymbol{b}'_j$ are the basis vectors of $\mathcal{B}'$. The columns of P are the coefficients of the expansion of the new basis vectors in terms of the old basis vectors. Thus,

$$P = \begin{pmatrix} \dfrac{n_3 n_1}{\sqrt{n_1^2+n_2^2}} & \dfrac{-n_2}{\sqrt{n_1^2+n_2^2}} & n_1 \\ \dfrac{n_3 n_2}{\sqrt{n_1^2+n_2^2}} & \dfrac{n_1}{\sqrt{n_1^2+n_2^2}} & n_2 \\ -\sqrt{n_1^2+n_2^2} & 0 & n_3 \end{pmatrix}\,. \tag{13}$$

The inverse P^{-1} is easily computed since the columns of P are orthonormal, which implies that P is an orthogonal matrix,[2] i.e. $P^{-1} = P^{\mathsf{T}}$.

According to eq. (14) of the class handout, *Vector coordinates, matrix elements and changes of basis*,

$$[R]_{\mathcal{B}'} = P^{-1}\,[R]_{\mathcal{B}_s}\,P\,. \tag{14}$$

where $[R]_{\mathcal{B}_s}$ is the matrix R with respect to the standard basis, and $[R]_{\mathcal{B}'}$ is the matrix R with respect to the new basis (in which $\hat{\boldsymbol{n}}$ points along the new positive z-axis). In particular,

$$[R]_{\mathcal{B}} = R(\hat{\boldsymbol{n}},\theta)\,, \qquad [R]_{\mathcal{B}'} = R(\boldsymbol{k},\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\,.$$

[2]In fact, P can be expressed as the product of two simple rotation matrices as shown in Appendix B.

Hence, eq. (14) yields

$$R(\hat{\boldsymbol{n}},\theta) = PR(\boldsymbol{k},\theta)P^{-1}\,, \tag{15}$$

where P is given by eq. (13) and $P^{-1} = P^{\mathsf{T}}$. Eq. (15) is a special case of a more general result [cf. eq. (67)], which is derived in Appendix B.

For ease of notation, we define

$$N_{12} \equiv \sqrt{n_1^2 + n_2^2}\,. \tag{16}$$

Note that $N_{12}^2 + n_3^2 = 1$, since $\hat{\boldsymbol{n}}$ is a unit vector. Writing out the matrices in eq. (15),

$$R(\hat{\boldsymbol{n}},\theta) = \begin{pmatrix} n_3n_1/N_{12} & -n_2/N_{12} & n_1 \\ n_3n_2/N_{12} & n_1/N_{12} & n_2 \\ -N_{12} & 0 & n_3 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} n_3n_1/N_{12} & n_3n_2/N_{12} & -N_{12} \\ -n_2/N_{12} & n_1/N_{12} & 0 \\ n_1 & n_2 & n_3 \end{pmatrix}$$

$$= \begin{pmatrix} n_3n_1/N_{12} & -n_2/N_{12} & n_1 \\ n_3n_2/N_{12} & n_1/N_{12} & n_2 \\ -N_{12} & 0 & n_3 \end{pmatrix} \begin{pmatrix} \frac{n_3n_1\cos\theta + n_2\sin\theta}{N_{12}} & \frac{n_3n_2\cos\theta - n_1\sin\theta}{N_{12}} & -N_{12}\cos\theta \\ \frac{n_3n_1\sin\theta - n_2\cos\theta}{N_{12}} & \frac{n_3n_2\sin\theta + n_1\cos\theta}{N_{12}} & -N_{12}\sin\theta \\ n_1 & n_2 & n_3 \end{pmatrix}.$$

Using $N_{12}^2 = n_1^2 + n_2^2$ and $n_3^2 = 1 - N_{12}^2$, the final matrix multiplication then yields the desired result:

$$\boxed{R(\hat{\boldsymbol{n}},\theta) = \begin{pmatrix} \cos\theta + n_1^2(1-\cos\theta) & n_1n_2(1-\cos\theta) - n_3\sin\theta & n_1n_3(1-\cos\theta) + n_2\sin\theta \\ n_1n_2(1-\cos\theta) + n_3\sin\theta & \cos\theta + n_2^2(1-\cos\theta) & n_2n_3(1-\cos\theta) - n_1\sin\theta \\ n_1n_3(1-\cos\theta) - n_2\sin\theta & n_2n_3(1-\cos\theta) + n_1\sin\theta & \cos\theta + n_3^2(1-\cos\theta) \end{pmatrix}} \tag{17}$$

Eq. (17) is called the *Rodriguez formula* for the 3×3 rotation matrix $R(\hat{\boldsymbol{n}},\theta)$.

One can easily check that eqs. (3) and (4) are satisfied. In particular, as indicated by eq. (5), the rotations $R(\hat{\boldsymbol{n}},\pi)$ and $R(-\hat{\boldsymbol{n}},\pi)$ represent the same rotation,

$$R_{ij}(\hat{\boldsymbol{n}},\pi) = \begin{pmatrix} 2n_1^2 - 1 & 2n_1n_2 & 2n_1n_3 \\ 2n_1n_2 & 2n_2^2 - 1 & 2n_2n_3 \\ 2n_1n_3 & 2n_2n_3 & 2n_3^2 - 1 \end{pmatrix} = 2n_in_j - \delta_{ij}\,, \tag{18}$$

where the Kronecker delta δ_{ij} is defined to be the matrix elements of the identity,

$$\delta_{ij} = \begin{cases} 1\,, & \text{if } i = j\,, \\ 0\,, & \text{if } i \neq j\,. \end{cases} \tag{19}$$

Finally, as expected, $R_{ij}(\hat{\boldsymbol{n}},0) = \delta_{ij}$, independently of the direction of $\hat{\boldsymbol{n}}$. I leave it as an exercise to the reader to verify explicitly that $R = R(\hat{\boldsymbol{n}},\theta)$ given in eq. (17) satisfies the conditions $RR^{\mathsf{T}} = \mathbf{I}$ and $\det R = +1$.

4. Determining the rotation axis and the rotation angle

Given a general three-dimensional rotation matrix, $R(\hat{\boldsymbol{n}},\theta)$, we can determine the angle of rotation θ and the axis of rotation $\hat{\boldsymbol{n}}$. Using eq. (17), the trace of $R(\hat{\boldsymbol{n}},\theta)$ is given by:

$$\boxed{\text{Tr } R(\hat{\boldsymbol{n}},\theta) = 1 + 2\cos\theta} \tag{20}$$

which coincides with our previous result obtained in eq. (9). Thus eq. (20) yields,

$$\boxed{\cos\theta = \tfrac{1}{2}\,(\text{Tr } R - 1) \quad \text{and} \quad \sin\theta = (1-\cos^2\theta)^{1/2} = \tfrac{1}{2}\sqrt{(3-\text{Tr } R)(1+\text{Tr } R)}} \tag{21}$$

where $\sin\theta \geq 0$ is a consequence of our convention for the range of the rotation angle, $0 \leq \theta \leq \pi$. If $\sin\theta \neq 0$, then we can immediately use eqs. (17) and (21) to obtain

$$\hat{\boldsymbol{n}} = \frac{1}{\sqrt{(3-\text{Tr } R)(1+\text{Tr } R)}}\biggl(R_{32}-R_{23}\,,\, R_{13}-R_{31}\,,\, R_{21}-R_{12}\biggr)\,, \qquad \text{Tr } R \neq -1\,,\,3\,. \tag{22}$$

The overall sign of $\hat{\boldsymbol{n}}$ is fixed by eq. (4) due to our convention in which $\sin\theta \geq 0$. If $\sin\theta = 0$, then eq. (17) implies that $R_{ij} = R_{ji}$, in which case $\hat{\boldsymbol{n}}$ cannot be determined from eq. (22). In this case, eq. (20) determines whether $\cos\theta = +1$ or $\cos\theta = -1$. If $\cos\theta = +1$, then $R_{ij} = \delta_{ij}$ and the axis $\hat{\boldsymbol{n}}$ is undefined. If $\cos\theta = -1$, then eq. (18) determines the direction of $\hat{\boldsymbol{n}}$ up to an overall sign. That is,

$$\begin{aligned} &\hat{\boldsymbol{n}} \text{ is undetermined if } \theta = 0\,, \\ &\hat{\boldsymbol{n}} = \left(\epsilon_1\sqrt{\tfrac{1}{2}(1+R_{11})}\,,\, \epsilon_2\sqrt{\tfrac{1}{2}(1+R_{22})}\,,\, \epsilon_3\sqrt{\tfrac{1}{2}(1+R_{33})}\right)\,, \qquad \text{if } \theta = \pi\,, \end{aligned} \tag{23}$$

where the individual signs $\epsilon_i = \pm 1$ are determined up to an overall sign via[3]

$$\epsilon_i\epsilon_j = \frac{R_{ij}}{\sqrt{(1+R_{ii})(1+R_{jj})}}\,, \qquad \text{for fixed } i \neq j\,,\, R_{ii} \neq -1\,,\, R_{jj} \neq -1\,. \tag{24}$$

The ambiguity of the overall sign of $\hat{\boldsymbol{n}}$ sign is not significant in this case, since $R(\hat{\boldsymbol{n}},\pi)$ and $R(-\hat{\boldsymbol{n}},\pi)$ represent the same rotation [cf. eq. (6)].

One slightly inconvenient feature of the above analysis is that the case of $\theta = \pi$ (or equivalently, $\text{Tr } R = -1$) requires a separate treatment in order to determine $\hat{\boldsymbol{n}}$. Moreover, for values of θ very close to π, the numerator and denominator of eq. (22) are very small, so that a very precise numerical evaluation of both the numerator and denominator is required to accurately determine the direction of $\hat{\boldsymbol{n}}$. Thus, we briefly mention another approach for determining $\hat{\boldsymbol{n}}$ which avoids these problems.

In this alternate approach, we define the matrix

$$S = R + R^{\mathsf{T}} + (1 - \text{Tr } R)\mathbf{I}\,. \tag{25}$$

[3] If $R_{ii} = -1$, where i is a fixed index, then $n_i = 0$, in which case the corresponding ϵ_i is irrelevant.

Then, eq. (17) yields $S_{jk} = 2(1 - \cos\theta)n_j n_k = (3 - \text{Tr } R)n_j n_k$. Hence,[4]

$$\boxed{n_j n_k = \frac{S_{jk}}{3 - \text{Tr } R}\,, \qquad \text{Tr } R \neq 3} \tag{26}$$

Note that for θ close to π (which corresponds to Tr $R \simeq -1$), neither the numerator nor the denominator of eq. (26) is particularly small, and the direction of $\hat{\boldsymbol{n}}$ can be determined numerically without significant roundoff error. To determine $\hat{\boldsymbol{n}}$ up to an overall sign, we simply set $j = k$ in eq. (26), which fixes the value of n_j^2. If $\sin\theta \neq 0$, the overall sign of $\hat{\boldsymbol{n}}$ is determined by eq. (22). If $\sin\theta = 0$ then there are two cases. For $\theta = 0$ (corresponding to the identity rotation), $S = \mathbf{0}$ and the rotation axis $\hat{\boldsymbol{n}}$ is undefined. For $\theta = \pi$, the ambiguity in the overall sign of $\hat{\boldsymbol{n}}$ is immaterial, in light of eq. (6).

In summary, eqs. (21), (22) and (23) provide a simple algorithm for determining the rotation axis $\hat{\boldsymbol{n}}$ and the rotation angle θ for any rotation matrix $R(\hat{\boldsymbol{n}}, \theta) \neq \mathbf{I}$.

5. Boas, p. 161, problem 3.11–54 revisited

Boas poses the following question in problem 3.11–54 on p. 161. Show that the matrix,

$$R = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & -1 \\ \sqrt{2} & 0 & \sqrt{2} \\ 1 & -\sqrt{2} & -1 \end{pmatrix}\,, \tag{27}$$

is orthogonal and find the rotation it produces as an operator acting on vectors. Determine the rotation axis and the angle of rotation.

After checking that $RR^{\mathsf{T}} = \mathbf{I}$ and $\det R = 1$, we can deduce the rotation angle θ and the rotation axis $\hat{\boldsymbol{n}}$ as follows. First, $\hat{\boldsymbol{n}}$ is identified as the normalized eigenvector of $R(\hat{\boldsymbol{n}}, \theta)$ corresponding to the eigenvalue $+1$. The overall sign of $\hat{\boldsymbol{n}}$ is fixed by making some conventional choice. In order to determine θ, Boas proposes the following procedure. Noting that the matrix R provides the matrix elements of the rotation operator with respect to the standard basis $\mathcal{B}_s$, we can define a new basis such that $\hat{\boldsymbol{n}}$ points along the new y-axis. Then, the matrix elements of the rotation operator with respect to this new basis has a simple form and θ can be determined by inspection. Two concrete examples of this procedure are provided on p. 156 of Boas. A detailed solution to the above problem following the procedure of Boas can be found in Section 5 of the class handout entitled, *Vector coordinates, matrix elements and changes of basis.*

Indeed, we employed a variant of this procedure in these notes to derive the explicit form for $R(\hat{\boldsymbol{n}}, \theta)$ obtained in eq. (17). However, once we have this general form, we can make use of the results of Section 4 to determine θ and $\hat{\boldsymbol{n}}$ almost by inspection. In particular, we can use eqs. (21) and (22) to determine the rotation angle θ and the rotation axis $\hat{\boldsymbol{n}}$. Since eq. (27) yields Tr $R = 0$, it follows from eq. (21) that

$$\cos\theta = -\tfrac{1}{2}\,, \qquad \sin\theta = \tfrac{1}{2}\sqrt{3}\,,$$

[4]Eq. (25) yields $\text{Tr}\, S = 3 - \text{Tr } R$. One can then use eq. (26) to verify that $\hat{\boldsymbol{n}}$ is a unit vector.

which implies that $\theta = 120°$. Next, we employ eq. (22) to obtain the axis of rotation,

$$\hat{\boldsymbol{n}} = -\frac{1}{\sqrt{3}}\left[\sqrt{2}\,\boldsymbol{i} + \boldsymbol{j}\right] . \tag{28}$$

In the solution to this problem given in Section 5 of the class handout cited above, we initially obtained $\hat{\boldsymbol{n}} = \frac{1}{\sqrt{3}}\left[\sqrt{2}\,\boldsymbol{i} + \boldsymbol{j}\right]$ and $\theta = 240°$, which corresponds to the convention defined in footnote 1 where $0 \leq \theta < 2\pi$ and $\hat{\boldsymbol{n}} > 0$. This is in contrast to eq. (28) which is based on the convention adopted in these notes where $0 \leq \theta \leq \pi$. Of course, both choices yield the same rotation matrix, in light of eq. (5).

6. Properties of the 3 × 3 improper rotation matrix

An improper rotation matrix is an orthogonal matrix, $\overline{R}$, such that $\det \overline{R} = -1$. The most general three-dimensional improper rotation, denoted by $\overline{R}(\hat{\boldsymbol{n}}, \theta)$ consists of a product of a proper rotation matrix, $R(\hat{\boldsymbol{n}}, \theta)$, and a mirror reflection through a plane normal to the unit vector $\hat{\boldsymbol{n}}$, which we denote by $\overline{R}(\hat{\boldsymbol{n}})$. In particular, the reflection plane passes through the origin and is perpendicular to $\hat{\boldsymbol{n}}$. In equations,

$$\overline{R}(\hat{\boldsymbol{n}}, \theta) \equiv R(\hat{\boldsymbol{n}}, \theta)\overline{R}(\hat{\boldsymbol{n}}) = \overline{R}(\hat{\boldsymbol{n}})R(\hat{\boldsymbol{n}}, \theta) . \tag{29}$$

Note that the improper rotation defined in eq. (29) does not depend on the order in which the proper rotation and reflection are applied. The matrix $\overline{R}(\hat{\boldsymbol{n}})$ is called a *reflection matrix*, since it is a representation of a mirror reflection through a fixed plane. In particular,

$$\overline{R}(\hat{\boldsymbol{n}}) = \overline{R}(-\hat{\boldsymbol{n}}) = \overline{R}(\hat{\boldsymbol{n}}, 0) , \tag{30}$$

after using $R(\hat{\boldsymbol{n}}, 0) = \mathbf{I}$. Thus, the overall sign of $\hat{\boldsymbol{n}}$ for a reflection matrix has no physical meaning. Note that all reflection matrices are orthogonal matrices with $\det \overline{R}(\hat{\boldsymbol{n}}) = -1$, with the property that:

$$[\overline{R}(\hat{\boldsymbol{n}})]^2 = \mathbf{I} , \tag{31}$$

or equivalently,

$$[\overline{R}(\hat{\boldsymbol{n}})]^{-1} = \overline{R}(\hat{\boldsymbol{n}}) . \tag{32}$$

In general, the product of a two proper and/or improper rotation matrices is not commutative. However, if $\hat{\boldsymbol{n}}$ is the same for both matrices, then eq. (2) implies that:[5]

$$R(\hat{\boldsymbol{n}}, \theta_1)\overline{R}(\hat{\boldsymbol{n}}, \theta_2) = \overline{R}(\hat{\boldsymbol{n}}, \theta_1)R(\hat{\boldsymbol{n}}, \theta_2) = \overline{R}(\hat{\boldsymbol{n}}, \theta_1 + \theta_2) , \tag{33}$$

$$\overline{R}(\hat{\boldsymbol{n}}, \theta_1)\overline{R}(\hat{\boldsymbol{n}}, \theta_2) = \overline{R}(\hat{\boldsymbol{n}}, \theta_1)\overline{R}(\hat{\boldsymbol{n}}, \theta_2) = R(\hat{\boldsymbol{n}}, \theta_1 + \theta_2) , \tag{34}$$

[5]Since $\det[R(\hat{\boldsymbol{n}}, \theta_1)\overline{R}(\hat{\boldsymbol{n}}, \theta_2)] = \det R(\hat{\boldsymbol{n}}, \theta_1) \det \overline{R}(\hat{\boldsymbol{n}}, \theta_2) = -1$, it follows that $R(\hat{\boldsymbol{n}}, \theta_1)\overline{R}(\hat{\boldsymbol{n}}, \theta_2)$ must be an improper rotation matrix. Likewise, $\overline{R}(\hat{\boldsymbol{n}}, \theta_1)\overline{R}(\hat{\boldsymbol{n}}, \theta_2)$ must be a proper rotation matrix. Eqs. (33) and (34) are consistent with these expectations.

after making use of eqs. (29) and (31).

The properties of the improper rotation matrices mirror those of the proper rotation matrices given in eqs. (3)–(7). Indeed the properties of the latter combined with eqs. (30) and (32) yield:

$$\overline{R}(\hat{\boldsymbol{n}},\theta+2\pi k)=\overline{R}(\hat{\boldsymbol{n}},\theta)\,,\qquad k=0,\pm 1\,\pm 2\,\ldots\,, \tag{35}$$

$$[\overline{R}(\hat{\boldsymbol{n}},\theta)]^{-1}=\overline{R}(\hat{\boldsymbol{n}},-\theta)=\overline{R}(-\hat{\boldsymbol{n}},\theta)\,. \tag{36}$$

Combining these two results, it follows that

$$\overline{R}(\hat{\boldsymbol{n}},2\pi-\theta)=\overline{R}(-\hat{\boldsymbol{n}},\theta)\,. \tag{37}$$

We shall adopt the convention (employed in Section 2) in which the angle θ is defined to lie in the interval $0\leq\theta\leq\pi$. In this convention, the overall sign of $\hat{\boldsymbol{n}}$ is meaningful when $0<\theta<\pi$.

The matrix $\overline{R}(\hat{\boldsymbol{n}},\pi)$ is special. Geometric considerations will convince you that

$$\overline{R}(\hat{\boldsymbol{n}},\pi)=R(\hat{\boldsymbol{n}},\pi)\overline{R}(\hat{\boldsymbol{n}})=\overline{R}(\hat{\boldsymbol{n}})R(\hat{\boldsymbol{n}},\pi)=-\mathbf{I}\,. \tag{38}$$

That is, $\overline{R}(\hat{\boldsymbol{n}},\pi)$ represents an *inversion*, which is a linear operator that transforms all vectors $\vec{\boldsymbol{x}}\to-\vec{\boldsymbol{x}}$. In particular, $\overline{R}(\hat{\boldsymbol{n}},\pi)$ is *independent* of the unit vector $\hat{\boldsymbol{n}}$. Eq. (38) is equivalent to the statement that an inversion is equivalent to a mirror reflection through a plane that passes through the origin and is perpendicular to an arbitrary unit vector $\hat{\boldsymbol{n}}$, followed by a proper rotation of 180° around the axis $\hat{\boldsymbol{n}}$. Sometimes, $\overline{R}(\hat{\boldsymbol{n}},\pi)$ is called a point reflection through the origin (to distinguish it from a reflection through a plane). Just like a reflection matrix, the inversion matrix satisfies

$$[\overline{R}(\hat{\boldsymbol{n}},\pi)]^2=\mathbf{I}\,. \tag{39}$$

In general, any improper 3×3 rotation matrix $\overline{R}$ with the property that $\overline{R}^2=\mathbf{I}$ is a representation of either an inversion or a reflection through a plane that passes through the origin.

Given any proper 3×3 rotation matrix $R(\hat{\boldsymbol{n}},\theta)$, the matrix $-R(\hat{\boldsymbol{n}},\theta)$ has determinant equal to -1 and therefore represents some improper rotation which can be determined as follows:

$$-R(\hat{\boldsymbol{n}},\theta)=R(\hat{\boldsymbol{n}},\theta)\overline{R}(\hat{\boldsymbol{n}},\pi)=\overline{R}(\hat{\boldsymbol{n}},\theta+\pi)=\overline{R}(-\hat{\boldsymbol{n}},\pi-\theta)\,, \tag{40}$$

after employing eqs. (38), (33) and (37). Two noteworthy consequences of eq. (40) are:

$$\overline{R}(\hat{\boldsymbol{n}},\tfrac{1}{2}\pi)=-R(-\hat{\boldsymbol{n}},\tfrac{1}{2}\pi)\,, \tag{41}$$

$$\overline{R}(\hat{\boldsymbol{n}})=\overline{R}(\hat{\boldsymbol{n}},0)=-R(\hat{\boldsymbol{n}},\pi)\,, \tag{42}$$

after using eq. (6) in the second equation above.

To learn more about the properties of a general three-dimensional improper rotation, consider the matrix representation $\overline{R}(\hat{\boldsymbol{n}},\theta)$ with respect to the standard basis

$\mathcal{B}_s = \{\boldsymbol{i}\,, \boldsymbol{j}\,, \boldsymbol{k}\}$. We can define a new coordinate system in which the unit normal to the reflection plane $\hat{\boldsymbol{n}}$ points in the direction of the new z-axis; the corresponding new basis will be denoted by $\mathcal{B}'$. The matrix representation of the improper rotation with respect to $\mathcal{B}'$ is then given by

$$\overline{R}(\boldsymbol{k},\theta) = R(\boldsymbol{k},\theta)\overline{R}(\boldsymbol{k}) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

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

Using the formalism developed in the class handout, *Vector coordinates, matrix elements and changes of basis*, there exists an invertible matrix P (which has been explicitly obtained in Section 3) such that

$$\overline{R}(\hat{\boldsymbol{n}},\theta) = P\overline{R}(\boldsymbol{k},\theta)P^{-1}\,. \tag{43}$$

The rest of the analysis mirrors the discussion of Section 2. It immediately follows that

$$\mathrm{Tr}\ \overline{R}(\hat{\boldsymbol{n}},\theta) = \mathrm{Tr}\ \overline{R}(\boldsymbol{k},\theta) = 2\cos\theta - 1\,, \tag{44}$$

after taking the trace of eq. (43). By convention, $0 \leq \theta \leq \pi$, which implies that $\sin\theta \geq 0$. Hence, the rotation angle is uniquely determined by eq. (44) To identify $\hat{\boldsymbol{n}}$ (up to an overall sign), we observe that any vector that is parallel to $\hat{\boldsymbol{n}}$ (which points along the normal to the reflection plane) is inverted. This last statement can be expressed as an eigenvalue equation,

$$\overline{R}(\hat{\boldsymbol{n}},\theta)\hat{\boldsymbol{n}} = -\hat{\boldsymbol{n}}\,. \tag{45}$$

Thus, $\hat{\boldsymbol{n}}$ is an eigenvector of $\overline{R}(\hat{\boldsymbol{n}},\theta)$ corresponding to the eigenvalue -1. In particular, the eigenvalue -1 is unique for any $\theta \neq \pi$, in which case $\hat{\boldsymbol{n}}$ can be determined up to an overall sign by computing the eigenvalues and the normalized eigenvectors of $\overline{R}(\hat{\boldsymbol{n}},\theta)$. A simple proof of this result is given in Appendix A. Here, we shall establish this assertion by noting that the eigenvalues of any matrix are invariant with respect to a similarity transformation. Using eq. (43), it follows that the eigenvalues of $\overline{R}(\hat{\boldsymbol{n}},\theta)$ are identical to the eigenvalues of $\overline{R}(\boldsymbol{k},\theta)$. The latter can be obtained from the characteristic equation,

$$-(1+\lambda)\left[(\cos\theta-\lambda)^2 + \sin^2\theta\right] = 0\,,$$

which simplifies to:

$$(1+\lambda)(\lambda^2 - 2\lambda\cos\theta + 1) = 0\,.$$

The solution to the quadratic equation, $\lambda^2 - 2\lambda\cos\theta + 1 = 0$, was given in eq. (11). It follows that the three eigenvalues of $\overline{R}(\boldsymbol{k},\theta)$ are given by,

$$\lambda_1 = -1\,,\quad \lambda_2 = e^{i\theta}\,,\quad \lambda_3 = e^{-i\theta}\,,\qquad \text{for}\ \ 0 \leq \theta \leq \pi\,.$$

There are three distinct cases:

Case 1: $\theta = 0$ $\quad\lambda_1 = \lambda_2 = \lambda_3 = -1\,,$ $\quad\overline{R}(\hat{\boldsymbol{n}},\pi) = -\mathbf{I}\,,$

Case 2: $\theta = \pi$ $\quad\lambda_1 = -1\,,\, \lambda_2 = \lambda_3 = 1\,,$ $\quad\overline{R}(\hat{\boldsymbol{n}},0) \equiv \overline{R}(\hat{\boldsymbol{n}})\,,$

Case 3: $0 < \theta < \pi$ $\quad\lambda_1 = -1, \lambda_2 = e^{i\theta}\,,\, \lambda_3 = e^{-i\theta}\,,$ $\quad\overline{R}(\hat{\boldsymbol{n}},\theta)\,,$

where the corresponding improper rotation matrix is indicated for each of the three cases. Indeed, for $\theta \neq \pi$, the eigenvalue -1 is unique. Moreover, the other two eigenvalues are complex conjugates of each other, whose real part is equal to $\cos\theta$, which uniquely fixes the rotation angle in the convention where $0 \leq \theta \leq \pi$. Case 1 corresponds to inversion and Case 2 corresponds to a mirror reflection through a plane that passes through the origin and is perpendicular to $\hat{\boldsymbol{n}}$. In Case 2, the doubly degenerate eigenvalue $+1$ is a consequence of the two linearly independent eigenvectors that span the reflection plane. In particular, any linear combination $\vec{\boldsymbol{v}}$ of these eigenvectors that lies in the reflection plane is unaffected by the reflection and thus satisfies $\overline{R}(\hat{\boldsymbol{n}})\vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}$. In contrast, the improper rotation matrices of Case 3 do not possess an eigenvalue of $+1$, since the vectors that lie in the reflection plane transform non-trivially under the proper rotation $R(\hat{\boldsymbol{n}},\theta)$.

Since $\hat{\boldsymbol{n}}$ is a real vector of unit length, it is determined only up to an overall sign by eq. (45) when its corresponding eigenvalue -1 is unique. This sign ambiguity is immaterial in Case 2 in light of eq. (30). The sign ambiguity in Case 3 cannot be resolved without further analysis. To make further progress, in Section 7 we shall obtain the general expression for the three dimensional improper rotation matrix $\overline{R}(\hat{\boldsymbol{n}},\theta)$.

7. An explicit formula for the matrix elements of a general 3×3 improper rotation matrix

In this section, the matrix elements of $\overline{R}(\hat{\boldsymbol{n}},\theta)$ will be denoted by $\overline{R}_{ij}$. The formula for $\overline{R}_{ij}$ that we seek depends on the angle θ and on the coordinates of $\hat{\boldsymbol{n}} = (n_1\,,\, n_2\,,\, n_3)$ with respect to a fixed Cartesian coordinate system. The computation mirrors the one given in Section 3. Namely, we compute:

$$\overline{R}(\hat{\boldsymbol{n}},\theta) = P\overline{R}(\boldsymbol{k},\theta)P^{-1}\,, \tag{46}$$

where P is given by eq. (13) and $P^{-1} = P^{\mathsf{T}}$. Writing out the matrices in eq. (46),

$$\overline{R}(\hat{\boldsymbol{n}},\theta) = \begin{pmatrix} n_3n_1/N_{12} & -n_2/N_{12} & n_1 \\ n_3n_2/N_{12} & n_1/N_{12} & n_2 \\ -N_{12} & 0 & n_3 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} n_3n_1/N_{12} & n_3n_2/N_{12} & -N_{12} \\ -n_2/N_{12} & n_1/N_{12} & 0 \\ n_1 & n_2 & n_3 \end{pmatrix}$$

$$= \begin{pmatrix} n_3n_1/N_{12} & -n_2/N_{12} & n_1 \\ n_3n_2/N_{12} & n_1/N_{12} & n_2 \\ -N_{12} & 0 & n_3 \end{pmatrix} \begin{pmatrix} \frac{n_3n_1\cos\theta + n_2\sin\theta}{N_{12}} & \frac{n_3n_2\cos\theta - n_1\sin\theta}{N_{12}} & -N_{12}\cos\theta \\ \frac{n_3n_1\sin\theta - n_2\cos\theta}{N_{12}} & \frac{n_3n_2\sin\theta + n_1\cos\theta}{N_{12}} & -N_{12}\sin\theta \\ -n_1 & -n_2 & -n_3 \end{pmatrix}\,,$$

where N_{12} is defined in eq. (16). Using $N_{12}^2 = n_1^2 + n_2^2$ and $n_3^2 = 1 - N_{12}^2$, the final matrix multiplication then yields the desired result:

$$\boxed{\overline{R}(\hat{\boldsymbol{n}},\theta) = \begin{pmatrix} \cos\theta - n_1^2(1+\cos\theta) & -n_1n_2(1+\cos\theta) - n_3\sin\theta & -n_1n_3(1+\cos\theta)+n_2\sin\theta \\ -n_1n_2(1+\cos\theta)+n_3\sin\theta & \cos\theta - n_2^2(1+\cos\theta) & -n_2n_3(1+\cos\theta) - n_1\sin\theta \\ -n_1n_3(1+\cos\theta) - n_2\sin\theta & -n_2n_3(1+\cos\theta)+n_1\sin\theta & \cos\theta - n_3^2(1+\cos\theta) \end{pmatrix}} \tag{47}$$

One can easily check that eqs. (35) and (36) are satisfied. In particular, as indicated by eq. (30), the improper rotations $\overline{R}(\hat{\boldsymbol{n}},0)$ and $\overline{R}(-\hat{\boldsymbol{n}},0)$ represent the same reflection matrix,[6]

$$\overline{R}_{ij}(\hat{\boldsymbol{n}},0) \equiv \overline{R}_{ij}(\hat{\boldsymbol{n}}) = \begin{pmatrix} 1-2n_1^2 & -2n_1n_2 & -2n_1n_3 \\ -2n_1n_2 & 1-2n_2^2 & -2n_2n_3 \\ -2n_1n_3 & -2n_2n_3 & 1-2n_3^2 \end{pmatrix} = \delta_{ij} - 2n_in_j\,, \tag{48}$$

where the Kronecker delta δ_{ij} is defined in eq. (19). Finally, as expected,

$$\overline{R}_{ij}(\hat{\boldsymbol{n}},\pi) = -\delta_{ij}\,,$$

independently of the direction of $\hat{\boldsymbol{n}}$. I leave it as an exercise to the reader to verify explicitly that $\overline{R} = \overline{R}(\hat{\boldsymbol{n}},\theta)$ given in eq. (47) satisfies the conditions $\overline{R}\,\overline{R}^{\mathsf{T}} = \mathbf{I}$ and $\det \overline{R} = -1$.

8. Determining the reflection plane and the rotation angle

A general three-dimensional improper rotation matrix, $\overline{R}(\hat{\boldsymbol{n}},\theta) = R(\hat{\boldsymbol{n}},\theta)\overline{R}(\hat{\boldsymbol{n}})$, is the product of a reflection and a proper rotation. The reflection $\overline{R}(\hat{\boldsymbol{n}})$ corresponds to a mirror reflection through a plane perpendicular to $\hat{\boldsymbol{n}}$ that passes through the origin. $R(\hat{\boldsymbol{n}},\theta)$ represents a proper rotation by θ that is taken around the axis $\hat{\boldsymbol{n}}$ in the counterclockwise direction.

To determine the angle of rotation θ, we compute the trace of $\overline{R}(\hat{\boldsymbol{n}},\theta)$. In particular, using eq. (47) it follows that:

$$\boxed{\mathrm{Tr}\ \overline{R}(\hat{\boldsymbol{n}},\theta) = 2\cos\theta - 1} \tag{49}$$

which coincides with our previous result obtained in eq. (44). By convention, $0 \leq \theta \leq \pi$, which implies that $\sin\theta \geq 0$. Thus, eq. (49) yields

$$\boxed{\cos\theta = \tfrac{1}{2}\left(\mathrm{Tr}\ \overline{R} + 1\right) \quad \text{and} \quad \sin\theta = (1-\cos^2\theta)^{1/2} = \tfrac{1}{2}\sqrt{(3+\mathrm{Tr}\ \overline{R})(1-\mathrm{Tr}\ \overline{R})}} \tag{50}$$

[6]Indeed, eqs. (18) and (48) are consistent with eq. (42) as expected.

If $\sin\theta \neq 0$, then we can immediately use eqs. (17) and (50) to obtain the unit normal to the reflection plane,

$$\hat{\boldsymbol{n}} = \frac{1}{\sqrt{(3+\mathrm{Tr}\,\overline{R})(1-\mathrm{Tr}\,\overline{R})}}\left(\overline{R}_{32}-\overline{R}_{23}\,,\,\overline{R}_{13}-\overline{R}_{31}\,,\,\overline{R}_{21}-\overline{R}_{12}\right), \qquad \mathrm{Tr}\,\overline{R} \neq 1\,,\,-3\,. \tag{51}$$

The overall sign of $\hat{\boldsymbol{n}}$ is fixed by eq. (36) due to our convention in which $\sin\theta \geq 0$. If $\sin\theta = 0$, then eq. (47) implies that $\overline{R}_{ij} = \overline{R}_{ji}$, in which case $\hat{\boldsymbol{n}}$ cannot be determined from eq. (22). In this case, eq. (49) determines whether $\cos\theta = +1$ or $\cos\theta = -1$. If $\cos\theta = -1$, then $\overline{R}_{ij} = -\delta_{ij}$ and the axis $\hat{\boldsymbol{n}}$ is undefined. If $\cos\theta = 1$, then eq. (48) determines the direction of $\hat{\boldsymbol{n}}$ up to an overall sign. That is,

$$\hat{\boldsymbol{n}} \text{ is undetermined if } \theta = \pi\,,$$

$$\hat{\boldsymbol{n}} = \left(\epsilon_1\sqrt{\tfrac{1}{2}(1-\overline{R}_{11})}\,,\,\epsilon_2\sqrt{\tfrac{1}{2}(1-\overline{R}_{22})}\,,\,\epsilon_3\sqrt{\tfrac{1}{2}(1-\overline{R}_{33})}\right)\,, \qquad \text{if } \theta = 0\,, \tag{52}$$

where the individual signs $\epsilon_i = \pm 1$ are determined up to an overall sign via[7]

$$\epsilon_i\epsilon_j = \frac{\overline{R}_{ij}}{\sqrt{(1-\overline{R}_{ii})(1-\overline{R}_{jj})}}\,, \qquad \text{for fixed } i \neq j\,,\, \overline{R}_{ii} \neq 1\,,\, \overline{R}_{jj} \neq 1\,. \tag{53}$$

The ambiguity of the overall sign of $\hat{\boldsymbol{n}}$ sign is not significant, since $\overline{R}(\hat{\boldsymbol{n}})$ and $\overline{R}(-\hat{\boldsymbol{n}})$ represent the same mirror reflection [cf. eq. (30)].

One can also determine the unit normal to the reflection plane $\hat{\boldsymbol{n}}$ by defining the matrix,

$$\overline{S} = \overline{R} + \overline{R}^{\mathsf{T}} - (1+\mathrm{Tr}\,\overline{R})\mathbf{I}\,.$$

Then, eq. (47) yields $S_{jk} = -2(1+\cos\theta)n_jn_k = -(3+\mathrm{Tr}\,\overline{R})n_jn_k$. Hence,

$$n_jn_k = \frac{-\overline{S}_{jk}}{3+\mathrm{Tr}\,\overline{R}}\,, \qquad \mathrm{Tr}\,\overline{R} \neq -3\,. \tag{54}$$

To determine $\hat{\boldsymbol{n}}$ up to an overall sign, we simply set $j = k$ in eq. (54), which fixes the value of n_j^2. If $\sin\theta \neq 0$, the overall sign of $\hat{\boldsymbol{n}}$ is fixed by eq. (51). If $\sin\theta = 0$ then there are two cases. For $\theta = \pi$ (corresponding to an inversion), $\overline{S} = \mathbf{0}$ and the axis $\hat{\boldsymbol{n}}$ is undefined. For $\theta = 0$, the ambiguity in the overall sign of $\hat{\boldsymbol{n}}$ is immaterial, in light of eq. (30).

Finally, we shall derive an equation for the reflection plane, which passes through the origin and is perpendicular to $\hat{\boldsymbol{n}}$. That is, the unit normal to the reflection plane, $\hat{\boldsymbol{n}} = n_1\boldsymbol{i}+n_2\boldsymbol{j}+n_2\boldsymbol{k} = (n_1,n_2,n_3)$, is a vector perpendicular to the reflection plane that passes through the origin, i.e. the point $(x_0,y_0,z_0) = (0,0,0)$. Thus, using eq. (5.10) on p. 109 of Boas, the equation of the reflection plane is given by:

$$\boxed{n_1x + n_2y + n_3z = 0} \tag{55}$$

[7] If $\overline{R}_{ii} = 1$, where i is a fixed index, then $n_i = 0$, in which case the corresponding ϵ_i is irrelevant.

Note that the equation for the reflection plane does not depend on the overall sign of $\hat{\boldsymbol{n}}$. This makes sense, as both $\hat{\boldsymbol{n}}$ and $-\hat{\boldsymbol{n}}$ are perpendicular to the reflection plane.

The equation for the reflection plane can also be derived directly as follows. In the case of $\theta = \pi$, the unit normal to the reflection plane $\hat{\boldsymbol{n}}$ is undefined so we exclude this case from further consideration. If $\theta \neq \pi$, then the reflection plane corresponding to the improper rotation $\overline{R}(\hat{\boldsymbol{n}}, \theta)$ does not depend on θ. Thus, we can take $\theta = 0$ and consider $\overline{R}(\hat{\boldsymbol{n}})$ which represents a mirror reflection through the reflection plane. Any vector $\vec{\boldsymbol{v}} = (x, y, z)$ that lies in the reflection plane is an eigenvector of $\overline{R}(\hat{\boldsymbol{n}})$ with eigenvalue $+1$, as indicated at the end of Section 6. Thus, the equation of the reflection plane is $\overline{R}(\hat{\boldsymbol{n}})\vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}$, which is explicitly given by [cf. eq. (48)]:

$$\begin{pmatrix} 1 - 2n_1^2 & -2n_1 n_2 & -2n_1 n_3 \\ -2n_1 n_2 & 1 - 2n_2^2 & -2n_2 n_3 \\ -2n_1 n_3 & 2 - n_2 n_3 & 1 - 2n_3^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} . \tag{56}$$

The matrix equation, eq. (56), is equivalent to:

$$\begin{pmatrix} n_1^2 & n_1 n_2 & n_1 n_3 \\ n_1 n_2 & n_2^2 & n_2 n_3 \\ n_1 n_3 & n_2 n_3 & n_3^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 . \tag{57}$$

The matrix in eq. (57) is a special case of the matrix treated on problem 4 of the second midterm exam, so we can simply use the results obtained in the exam solutions. Applying two elementary row operations, the matrix equation, eq. (57), can be transformed into reduced row echelon form,

$$\begin{pmatrix} n_1^2 & n_1 n_2 & n_1 n_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 .$$

The solution to this equation is all x, y and z that satisfy:

$$n_1 x + n_2 y + n_3 z = 0 ,$$

which corresponds to the equation of the reflection plane. Thus, we have reproduced eq. (55) as expected.

9. Summary: the most general 3 × 3 orthogonal matrix

Eqs. (17) and (47) provide explicit forms for the most general 3×3 orthogonal matrix, R and $\overline{R}$, with determinant $+1$ and -1, respectively. The matrix elements of the most general proper and improper 3×3 rotation matrix can be written in an elegant form using the Kronecker delta defined in eq. (19) and the Levi-Civita epsilon symbol, which is defined as:

$$\epsilon_{ijk} = \begin{cases} +1 , & \text{if } \{i, j, k\} \text{ is an even permutation of } \{1,2,3\} , \\ -1 , & \text{if } \{i, j, k\} \text{ is an odd permutation of } \{1,2,3\} , \\ \ \ 0 , & \text{otherwise.} \end{cases}$$

The matrix elements of the proper and improper rotation matrices given by eqs. (17) and (47) can then be written as:

$$R_{ij}(\hat{\boldsymbol{n}},\theta) = \cos\theta\,\delta_{ij} + (1-\cos\theta)n_i n_j - \sin\theta\,\epsilon_{ijk}n_k\,, \tag{58}$$

$$\overline{R}_{ij}(\hat{\boldsymbol{n}},\theta) = \cos\theta\,\delta_{ij} - (1+\cos\theta)n_i n_j - \sin\theta\,\epsilon_{ijk}n_k\,. \tag{59}$$

I leave it as an exercise for the reader to check that eqs. (58) and (59) are indeed equivalent to eqs. (17) and (47), respectively. Using the above forms for the proper and improper rotation matrices, many of the general properties of these matrices derived earlier in these notes can be quickly established. Remarkably, there is a very simple and elegant derivation of eqs. (58) and (59) that makes use of the properties of tensor algebra. This derivation will be provided in a separate class handout entitled, *The matrix elements of a* 3×3 *orthogonal matrix—revisited*.

Finally, we note that one can unify the formulae obtained for proper and improper rotation matrices in these notes by incorporating a factor of the determinant in the corresponding formulae. Thus, if R is a 3×3 orthogonal matrix, corresponding to *either* a proper or improper rotation, then its matrix elements are given by:

$$R_{ij}(\hat{\boldsymbol{n}},\theta) = \cos\theta\,\delta_{ij} + (\varepsilon-\cos\theta)n_i n_j - \sin\theta\,\epsilon_{ijk}n_k\,, \qquad \text{where } \varepsilon\equiv\det R\,.$$

The rotation angle θ, in a convention where $0\leq\theta\leq\pi$, is obtained from:

$$\cos\theta = \tfrac{1}{2}(\operatorname{Tr}R - \det R)\,, \tag{60}$$

and the corresponding rotation axis $\hat{\boldsymbol{n}}$ (which can be identified with the unit normal to the reflection plane if $\det R=-1$) is obtained from:

$$\hat{\boldsymbol{n}} = \frac{1}{\sqrt{(3-\operatorname{Tr}R\det R)(1+\operatorname{Tr}R\det R)}}\biggl(R_{32}-R_{23}\,,\, R_{13}-R_{31}\,,\, R_{21}-R_{12}\biggr)\,,$$
$$\text{for } \operatorname{Tr}R\det R\neq -1\,,\,3\,. \tag{61}$$

The case of $\operatorname{Tr}R\det R=-1$ corresponds to $\cos\theta=-\det R$ and

$$R_{ij} = (2n_i n_j - \delta_{ij})\det R\,,$$

from which $\hat{\boldsymbol{n}}$ can be determined only up to an overall sign via,

$$n_i n_j = \tfrac{1}{2}\left[\delta_{ij} + (\det R)R_{ij}\right]\,, \qquad \text{for } \operatorname{Tr}R\det R = -1\,.$$

The case of $\operatorname{Tr}R\det R=3$ corresponds to $R=(\det R)\mathbf{I}$ and $\cos\theta=\det R$. That is,

$$R_{ij} = (\det R)\delta_{ij}\,, \qquad\qquad \text{for } \operatorname{Tr}R\det R = 3\,,$$

in which case $\hat{\boldsymbol{n}}$ is undefined.

Appendix A: The eigenvalues of a 3×3 orthogonal matrix[8]

Given any matrix A, the eigenvalues are the solutions to the characteristic equation,

$$\det\,(A - \lambda \mathbf{I}) = 0\,. \tag{62}$$

Suppose that A is an $n \times n$ real orthogonal matrix. The eigenvalue equation for A and its complex conjugate transpose are given by:

$$A\boldsymbol{v} = \lambda \boldsymbol{v}\,, \qquad \overline{\boldsymbol{v}}^{\mathsf{T}} A = \overline{\lambda}\,\overline{\boldsymbol{v}}^{\mathsf{T}}\,.$$

Hence multiplying these two equations together yields

$$\overline{\lambda}\,\lambda\,\overline{\boldsymbol{v}}^{\mathsf{T}}\boldsymbol{v} = \overline{\boldsymbol{v}}^{\mathsf{T}} A^{\mathsf{T}} A \boldsymbol{v} = \overline{\boldsymbol{v}}^{\mathsf{T}}\boldsymbol{v}\,, \tag{63}$$

since an orthogonal matrix satisfies $A^{\mathsf{T}}A = \mathbf{I}$. Since eigenvectors must be nonzero, it follows that $\overline{\boldsymbol{v}}^{\mathsf{T}}\boldsymbol{v} \neq 0$. Hence, eq. (63) yields $|\lambda| = 1$. Thus, the eigenvalues of a real orthogonal matrix must be complex numbers of unit modulus. That is, $\lambda = e^{i\alpha}$ for some α in the interval $0 \leq \alpha < 2\pi$.

Consider the following product of matrices, where A satisfies $A^{\mathsf{T}}A = \mathbf{I}$,

$$A^{\mathsf{T}}(\mathbf{I} - A) = A^{\mathsf{T}} - \mathbf{I} = -(\mathbf{I} - A)^{\mathsf{T}}\,.$$

Taking the determinant of both sides of this equation, it follows that[9]

$$\det A\,\det(\mathbf{I} - A) = (-1)^n \det(\mathbf{I} - A)\,, \tag{64}$$

since for the $n \times n$ identity matrix, $\det(-\mathbf{I}) = (-1)^n$. For a proper odd-dimensional orthogonal matrix, we have $\det A = 1$ and $(-1)^n = -1$. Hence, eq. (64) yields[10]

$$\det(\mathbf{I} - A) = 0\,, \qquad \text{for any proper odd-dimensional orthogonal matrix } A. \tag{65}$$

Comparing with eq. (62), we conclude that $\lambda = 1$ is an eigenvalue of A.[11] Since $\det A$ is the product of its three eigenvalues and each eigenvalue is a complex number of unit modulus, it follows that the eigenvalues of any proper 3×3 orthogonal matrix must be 1, $e^{i\theta}$ and $e^{-i\theta}$ for some value of θ that lies in the interval $0 \leq \theta \leq \pi$.[12]

Next, we consider the following product of matrices, where A satisfies $A^{\mathsf{T}}A = \mathbf{I}$,

$$A^{\mathsf{T}}(\mathbf{I} + A) = A^{\mathsf{T}} + \mathbf{I} = (\mathbf{I} + A)^{\mathsf{T}}\,.$$

[8] A nice reference to the results of this appendix can be found in L. Mirsky, *An Introduction to Linear Algebra* (Dover Publications, Inc., New York, 1982).

[9] Here, we make use of the well known properties of the determinant, namely $\det(AB) = \det A \det B$ and $\det(A^{\mathsf{T}}) = \det A$.

[10] Eq. (65) is also valid for any improper even-dimensional orthogonal matrix A since in this case $\det A = -1$ and $(-1)^n = 1$.

[11] Of course, this is consistent with the result that the eigenvalues of a real orthogonal matrix are of the form $e^{i\alpha}$ for $0 \leq \alpha < 2\pi$, since the eigenvalue 1 corresponds to $\alpha = 0$.

[12] There is no loss of generality in restricting the interval of the angle to satisfy $0 \leq \theta \leq \pi$. In particular, under $\theta \to 2\pi - \theta$, the two eigenvalues $e^{i\theta}$ and $e^{-i\theta}$ are simply interchanged.

Taking the determinant of both sides of this equation, it follows that

$$\det A \det(\mathbf{I} + A) = \det(\mathbf{I} + A)\,, \tag{66}$$

For any improper orthogonal matrix, we have $\det A = -1$. Hence, eq. (66) yields

$$\det(\mathbf{I} + A) = 0\,, \qquad \text{for any improper orthogonal matrix } A.$$

Comparing with eq. (62), we conclude that $\lambda = -1$ is an eigenvalue of A. Since $\det A$ is the product of its three eigenvalues and each eigenvalue is a complex number of unit modulus, it follows that the eigenvalues of any improper 3×3 orthogonal matrix must be -1, $e^{i\theta}$ and $e^{-i\theta}$ for some value of θ that lies in the interval $0 \leq \theta \leq \pi$ (cf. footnote 12).

Appendix B: The matrix $\boldsymbol{P}$ expressed as a product of simpler rotation matrices

The matrix P obtained in eq. (13) is a real orthogonal matrix with determinant equal to 1. In particular, it is straightforward to check that

$$\hat{\boldsymbol{n}} = P\boldsymbol{k}\,,$$

which is not surprising since the matrix P was constructed so that the vector $\hat{\boldsymbol{n}}$, which is represented by $\hat{\boldsymbol{n}} = (n_1\,,\, n_2\,,\, n_3)$ with respect to the standard basis would have coordinates $(0\,,\,0\,,\,1)$ with respect to the basis $\mathcal{B}'$ [cf. eq(10) in the class handout entitled, *Vector coordinates, matrix elements and changes of basis*].

We define the angles θ' and ϕ' to be the polar and azimuthal angles of the unit vector $\hat{\boldsymbol{n}}$,

$$\hat{\boldsymbol{n}} = (\sin\theta'\cos\phi'\,,\, \sin\theta'\sin\phi'\,,\, \cos\theta')\,.$$

Then, the matrix P can be expressed as the product of two simple rotation matrices,

$$P = R(\boldsymbol{k}, \phi')R(\boldsymbol{j}, \theta')\,.$$

This is easily checked by performing the matrix multiplication indicated above.

Finally, we note that eq. (15) is a special case of a more general result,

$$R(\hat{\boldsymbol{n}}, \theta) = PR(\hat{\boldsymbol{n}}', \theta)P^{-1}\,, \qquad \text{where } \hat{\boldsymbol{n}} = P\hat{\boldsymbol{n}}'\,, \tag{67}$$

where P is a proper rotation matrix. To prove eq. (67), we first note that

$$\text{Tr } R(\hat{\boldsymbol{n}}, \theta) = \text{Tr}\big[PR(\hat{\boldsymbol{n}}', \theta)P^{-1}\big] = \text{Tr } R(\hat{\boldsymbol{n}}', \theta)\,, \tag{68}$$

using the cyclicity of the trace. It follows from eq. (20) that the angle of rotation of $R(\hat{\boldsymbol{n}}, \theta)$ and $PR(\hat{\boldsymbol{n}}', \theta)P^{-1}$ must be the same. Next, we use $R(\hat{\boldsymbol{n}}', \theta)\hat{\boldsymbol{n}}' = \hat{\boldsymbol{n}}'$ [cf. eq. (10)] to determine the axis of rotation of $PR(\hat{\boldsymbol{n}}', \theta)P^{-1}$. The eigenvalue equation,

$$PR(\hat{\boldsymbol{n}}', \theta)P^{-1}(P\hat{\boldsymbol{n}}') = PR(\hat{\boldsymbol{n}}', \theta)\hat{\boldsymbol{n}}' = P\hat{\boldsymbol{n}}'\,, \tag{69}$$

implies that $P\hat{\boldsymbol{n}}'$ is an eigenvector of $PR(\hat{\boldsymbol{n}}', \theta)P^{-1}$ with eigenvalue $+1$. Thus, the corresponding eigenvector $P\hat{\boldsymbol{n}}'$ is the axis of rotation of $PR(\hat{\boldsymbol{n}}', \theta)P^{-1}$, up to an overall sign that is not fixed by the eigenvalue equation.

The overall sign ambiguity is not relevant if $\sin\theta = 0$, which corresponds to two possible cases. If $\theta = 0$, then $R(\hat{\boldsymbol{n}}, 0) = R(\hat{\boldsymbol{n}}', 0) = \mathbf{I}$ and eq. (67) is trivially satisfied. If $\theta = \pi$, then eq. (6) implies that both signs of the unit vector parallel to the axis of rotation represent the same rotation. Thus, the ambiguity in the overall sign determination of $\hat{\boldsymbol{n}}$ is immaterial. If $\sin\theta \neq 0$ then one must determine the overall sign of $\hat{\boldsymbol{n}}$ by another argument. To check that $\hat{\boldsymbol{n}} = P\hat{\boldsymbol{n}}'$ provides the correct overall sign of $\hat{\boldsymbol{n}}$, one can make an argument based on continuity. In the limit of $\hat{\boldsymbol{n}}' = \hat{\boldsymbol{n}}$, it follows that $P = \mathbf{I}$, with an overall positive sign. But, one can continuously vary $\hat{\boldsymbol{n}}$ starting from $\hat{\boldsymbol{n}}'$ until it points in the desired direction. The overall sign must therefore remain positive, since the sign can take on only two discrete values ($+$ and $-$) and therefore cannot change continuously from one sign to another. Hence, eq. (67) is confirmed.

Alternatively, for $\sin\theta \neq 0$, one can compute $\hat{\boldsymbol{n}}$ in terms of $\hat{\boldsymbol{n}}'$ directly from eq. (22). This method is employed in Appendix C of the class handout entitled, *The Matrix Elements of a* $\mathbf{3 \times 3}$ *Orthogonal Matrix—Revisited*, using the methods of tensor algebra. As expected, one indeed finds that $\hat{\boldsymbol{n}} = P\hat{\boldsymbol{n}}'$ and the proof of eq. (67) is complete.