International Tables for Crystallography (2006). Vol. B, Section 3.3.1.3.1, pp. 367–368.

3.3.1.2.3. *Orthogonalization of impure rotations*

There are several ways of deriving a strictly orthogonal matrix from a given approximately orthogonal matrix, among them the following.

(i) The Gram–Schmidt process. This is probably the simplest and the easiest to compute. If the given matrix consists of three column vectors $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ (later referred to as primers) which are to be replaced by three column vectors $\mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{u}_3$ then the process is

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1/|\mathbf{v}_1| \\
\mathbf{u}_2 &= \mathbf{v}_2 - (\mathbf{u}_1\cdot\mathbf{v}_2)\mathbf{u}_1 \\
\mathbf{u}_2 &= \mathbf{u}_2/|\mathbf{u}_2| \\
\mathbf{u}_3 &= \mathbf{v}_3 - (\mathbf{u}_1\cdot\mathbf{v}_3)\mathbf{u}_1 - (\mathbf{u}_2\cdot\mathbf{v}_3)\mathbf{u}_2 \\
\mathbf{u}_3 &= \mathbf{u}_3/|\mathbf{u}_3|.
\end{aligned}$$

As successive vectors are established, each vector $\mathbf{v}$ has subtracted from it its components in the directions of established vectors, and the remainder is normalized. The method will fail at the normalization step if the vectors $\mathbf{v}$ are not linearly independent. Otherwise, the process may be extended to any number of dimensions.

The weakness of the method is that, though $\mathbf{u}_1$ differs from $\mathbf{v}_1$ only in scale, $\mathbf{u}_N$ may differ grossly from $\mathbf{v}_N$ as the various columns are not treated equivalently.

(ii) A preferable method which treats all vectors equivalently is to iteratively replace the matrix $\boldsymbol{M}$ by $\frac{1}{2}(\boldsymbol{M}+\boldsymbol{M}^{T-1})$.

Defining the residual matrix $\boldsymbol{E}$ as

$$\boldsymbol{E} = \boldsymbol{M}\boldsymbol{M}^T - \boldsymbol{I},$$

then on each iteration $\boldsymbol{E}$ is replaced by

$$\boldsymbol{E}^2(\boldsymbol{M}\boldsymbol{M}^T)^{-1}/4$$

and convergence necessarily ensues.

(iii) A third method resolves $\boldsymbol{M}$ into its symmetric and antisymmetric parts

$$\boldsymbol{S} = \tfrac{1}{2}(\boldsymbol{M}+\boldsymbol{M}^T), \quad \boldsymbol{A} = \tfrac{1}{2}(\boldsymbol{M}-\boldsymbol{M}^T), \quad \boldsymbol{M} = \boldsymbol{S}+\boldsymbol{A}$$

and constructs an orthogonal matrix for which only $\boldsymbol{S}$ is altered. $\boldsymbol{A}$ determines l, m, n and θ as shown in Section 3.3.1.2.1, and from these a new $\boldsymbol{S}$ may be constructed.

(iv) A fourth method is to treat the general matrix $\boldsymbol{M}$ as a combination of pure strain and pure rotation. Setting

$$\boldsymbol{M} = \boldsymbol{R}\boldsymbol{T}$$

with $\boldsymbol{R}$ orthogonal and $\boldsymbol{T}$ symmetrical gives

$$\boldsymbol{T} = (\boldsymbol{M}^T\boldsymbol{M})^{1/2}, \quad \boldsymbol{R} = \boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M})^{-1/2}.$$

The rotation so found is the one which exactly superposes those three mutually perpendicular directions which remain mutually perpendicular under the transformation $\boldsymbol{M}$.

$\boldsymbol{T}-\boldsymbol{I}$ is then the strain tensor of an unrotated body.

Writing $\boldsymbol{M} = \boldsymbol{T}\boldsymbol{R}$, $\boldsymbol{T} = (\boldsymbol{M}\boldsymbol{M}^T)^{1/2}$, $\boldsymbol{R} = (\boldsymbol{M}\boldsymbol{M}^T)^{-1/2}\boldsymbol{M}$ may also be useful, in which $\boldsymbol{T}-\boldsymbol{I}$ is the strain tensor of a rotated body. See also Section 3.3.1.2.2 (iv).

3.3.1.2.4. *Eigenvalues and eigenvectors of orthogonal matrices*

If $\boldsymbol{R}$ is the orthogonal matrix given in Section 3.3.1.2.1 in terms of the direction cosines l, m and n of the axis of rotation, then it is clear that (l, m, n) is an eigenvector of $\boldsymbol{R}$ with eigenvalue unity because

$$\boldsymbol{R}\begin{pmatrix} l \\ m \\ n \end{pmatrix} = \begin{pmatrix} l \\ m \\ n \end{pmatrix}.$$

Consideration of the determinant $|\boldsymbol{R}-\lambda\boldsymbol{I}| = 0$ shows that the sum of the three eigenvalues is $1+2\cos\theta$ and that their product is unity. Hence the three eigenvalues are 1, $e^{i\theta}$ and $e^{-i\theta}$. Since $\boldsymbol{R}$ is real, its product with any real vector is also real, yet its product with an eigenvector must, in general, be complex. Thus the eigenvectors must themselves be complex.

The remaining two eigenvectors $\mathbf{u}$ may be found using the results of Section 3.3.1.2.1 (*q.v.*) according to

$$\boldsymbol{R}\mathbf{u} = \mathbf{u} + \frac{2}{1+t^2}\{(\mathbf{r}\times\mathbf{u}) + [\mathbf{r}\times(\mathbf{r}\times\mathbf{u})]\} = \mathbf{u}e^{\pm i\theta} = \mathbf{u}\frac{1\pm it}{1\mp it},$$

which is solved by any vector of the form

$$\mathbf{u} = \mathbf{l}\times\mathbf{v} \mp i\mathbf{l}\times(\mathbf{l}\times\mathbf{v})$$

for any real vector $\mathbf{v}$, where $\mathbf{l}$ is the normalized axis vector, $\mathbf{l}t = \mathbf{r}$, $|\mathbf{l}| = 1$, $t = \tan(\theta/2)$. Eigenvectors for the two eigenvalues may have unrelated $\mathbf{v}$ vectors though the sign choices are coupled. If the vector $\mathbf{v}$ is rotated about $\mathbf{l}$ through an angle φ the corresponding vector $\mathbf{u}$ is multiplied by $e^{-i\varphi}$ and remains an eigenvector. Using superscript signs to denote the sign of θ in the eigenvalue with which each vector is associated, the matrix

$$\boldsymbol{U} = (\mathbf{l},\ \mathbf{u}^+,\ \mathbf{u}^-)$$

has the properties that

$$\boldsymbol{R}\boldsymbol{U} = \boldsymbol{U}\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\theta} & 0 \\ 0 & 0 & e^{-i\theta} \end{pmatrix}$$

and

$$\boldsymbol{U}^{*T}\boldsymbol{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2|\mathbf{l}\times\mathbf{v}^+|^2 & 0 \\ 0 & 0 & 2|\mathbf{l}\times\mathbf{v}^-|^2 \end{pmatrix}$$

which places restrictions on $\mathbf{v}$ if this is to be the identity. Note that the 23 element vanishes even in the absence of any relationship between $\mathbf{v}^+$ and $\mathbf{v}^-$.

A convenient form for $\boldsymbol{U}$, symmetrical in the elements of $\mathbf{l}$, is obtained by setting $\mathbf{v}^+ = \mathbf{v}^- = [111]$ and is

$$\boldsymbol{U} = \begin{pmatrix} l & \{(m-n) - i[l(l+m+n)-1]\}/d & \{(m-n) + i[l(l+m+n)-1]\}/d \\ m & \{(n-l) - i[m(l+m+n)-1]\}/d & \{(n-l) + i[m(l+m+n)-1]\}/d \\ n & \{(l-m) - i[n(l+m+n)-1]\}/d & \{(l-m) + i[n(l+m+n)-1]\}/d \end{pmatrix}$$

in which the normalizing denominator is given by

$$d = 2\sqrt{1 - lm - mn - nl}.$$

3.3.1.3. *Projection transformations and spaces*

In the following section we address the question of the relationship between the coordinates of a molecular model and the corresponding coordinates on the screen of the graphics device. A good introduction to this topic is given by Newman & Sproull (1973), and Foley *et al.* (1990) give a comprehensive account of the field, including recent developments, especially those arising from the development of raster-graphics technologies.

3.3.1.3.1. *Definitions*

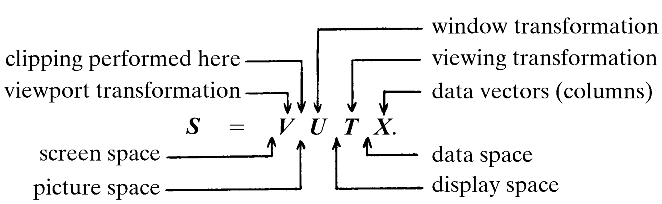
Typically, the coordinates, $\mathbf{X}$, of points in an object to be drawn are held in homogeneous Cartesian form as described in Section 3.3.1.1.2. Such coordinates are said to be in *data space* or world coordinates and this coordinate system is generally a constant aspect of the problem.

In order to view these data in convenient ways such coordinates may be subjected to a 4×4 *viewing transformation* $\boldsymbol{T}$, affecting orientation, scale *etc.*, the resulting coordinates $\boldsymbol{T}\mathbf{X}$ being then in *display space*. Here, and throughout what follows, we treat position vectors as columns with transformation matrices as factors on the left, though some writers do the reverse.

In general, only some portion of display space which lies inside a certain frustum of a pyramid is required to fall within the picture. The pyramid may be thought of as having the observer's eye at its vertex, with a rectangular base corresponding to the picture area. This volume is called a *window*. A transformation, $\boldsymbol{U}$, which takes *display-space* coordinates as input and generates vectors (X, Y, Z, W) for which X/W and $Y/W = \pm 1$ for points on the left, right, top and bottom boundaries of the window and for which Z/W takes particular values on the front and back planes of the window, is said to be a *windowing transformation*. In machines for which Z/W controls intensity depth cueing, the range of Z/W corresponding to the window is likely to be 0 to 1 rather than -1 to 1. Coordinates obtained by multiplying display-space coordinates by $\boldsymbol{U}$ are termed *picture-space* coordinates. Mathematically, $\boldsymbol{U}$ is a 4×4 matrix like any other, but functionally it is special. Declaring a transformation to be a windowing transformation implies that only resulting points having $|X|, |Y| < W$ and positive $Z < W$ are to be plotted. Machines with clipping hardware to truncate lines which run out of the picture perform clipping on the output from the windowing transformation.

Finally, the picture has to be drawn in some rectangular portion of the screen which is allocated for the purpose. Such an area is termed a *viewport* and is defined in terms of *screen coordinates* which are defined absolutely for the hardware in question as $\pm n$ for full-screen deflection, where n is declared by the manufacturer. *Screen coordinates* are obtained from picture coordinates with a *viewport transformation*, $\boldsymbol{V}$.*

To summarize, screen coordinates, $\mathbf{S}$, are given by



3.3.1.3.2. *Translation*

The transformation

$$\begin{pmatrix} N\boldsymbol{I} & \mathbf{V} \\ \mathbf{0}^T & N \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ W \end{pmatrix} = \begin{pmatrix} \mathbf{X}N + \mathbf{V}W \\ NW \end{pmatrix} \simeq \begin{pmatrix} \mathbf{X} + \mathbf{V}W/N \\ W \end{pmatrix} \simeq \begin{pmatrix} \mathbf{X}/W + \mathbf{V}/N \\ 1 \end{pmatrix}$$

evidently corresponds to the addition of the vector $\mathbf{V}W/N$ to the components of $\mathbf{X}$ or of $\mathbf{V}/N$ to the components of $\mathbf{X}/W$. ($\boldsymbol{I}$ is the identity.) Displacements may thus be affected by expressing the required displacement vector in homogeneous coordinates with any suitable choice of N (commonly, $N = W$), with $\mathbf{V}$ scaled to correspond to this choice, and loading the 4×4 transformation matrix as indicated above.

* In recent years it has become increasingly common, especially in two-dimensional work, to apply the term 'window' to what is here called a viewport, but in this chapter we use these terms in the manner described in the text.

3.3.1.3.3. *Rotation*

Rotation about the origin is achieved by

$$\begin{pmatrix} N\boldsymbol{R} & \mathbf{0} \\ \mathbf{0}^T & N \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ W \end{pmatrix} = \begin{pmatrix} N\boldsymbol{R}\mathbf{X} \\ NW \end{pmatrix} \simeq \begin{pmatrix} \boldsymbol{R}\mathbf{X} \\ W \end{pmatrix},$$

in which $\boldsymbol{R}$ is an orthogonal 3×3 matrix. $\boldsymbol{R}$ necessarily has elements not exceeding one in modulus. For machines using integer arithmetic, therefore, N would be chosen large enough (usually half the largest possible integer) for the product $N\boldsymbol{R}$ to be well represented in the available word length. Characteristically, N affects resolution but not scale.

3.3.1.3.4. *Scale*

The transformation

$$\begin{pmatrix} SN\boldsymbol{I} & \mathbf{0} \\ \mathbf{0}^T & N \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ W \end{pmatrix} = \begin{pmatrix} SN\mathbf{X} \\ NW \end{pmatrix} \simeq \begin{pmatrix} S\mathbf{X} \\ W \end{pmatrix}$$

scales the vector ($\mathbf{X}$, W) by the factor S. For integer working and $|S| < 1$, N should be set to the largest representable integer. For $|S| > 1$ the product SN should be the largest representable integer, N being reduced accordingly.

3.3.1.3.5. *Windowing and perspective*

It is necessary at this point to relate the discussion to the axial system inherent in the graphics device employed. One common system adopts X horizontal and to the right when viewing the screen, Y vertically upwards in the plane of the screen, and Z normal to X and Y with $+Z$ into the screen. This is, unfortunately, a left-handed system in that $(\mathbf{X} \times \mathbf{Y}) \cdot \mathbf{Z}$ is negative. Since it is usual in crystallographic work to use right-handed axial systems it is necessary to incorporate a matrix of the form

$$\begin{pmatrix} W & 0 & 0 & 0 \\ 0 & W & 0 & 0 \\ 0 & 0 & -W & 0 \\ 0 & 0 & 0 & W \end{pmatrix}$$

either as the left-most factor in the matrix $\boldsymbol{T}$ or as the right-most factor in the windowing transformation $\boldsymbol{U}$ (see Section 3.3.1.3.1). The latter choice is to be preferred and is adopted here. The former choice leads to complications if transformations in display space will be required. Display-space coordinates are necessarily referred to this axial system.

Let L, R, T, B, N and F be the left, right, top, bottom, near and far boundaries of the windowed volume $(L < R, T > B, N < F)$, S be the Z coordinate of the screen, and C, D and E be the coordinates of the observer's eye position, all ten of these parameters being referred to the origin of display space as origin, which may be anywhere in relation to the hardware. L, R, T and B are to be evaluated in the screen plane. All ten parameters may be referred to their own fourth coordinate, V, meaning that the point (X, Y, Z, W) in display space will be on the left boundary of the picture if $X/W = L/V$ when $Z/W = S/V$. V may be freely chosen so that all eleven quantities and all elements of $\boldsymbol{U}$ suit the word length of the machine. These relationships are illustrated in Fig. 3.3.1.1.

Since

$$(X, Y, Z, W) \simeq \left(\frac{XV}{W}, \frac{YV}{W}, \frac{ZV}{W}, V\right),$$

XV/W is a display-space coordinate on the same scale as the window parameters. This must be plotted on the screen at an X coordinate (on the scale of the window parameters) which is the weighted mean of XV/W and C, the weights being $(S - E)$ and