

## 3.3. MOLECULAR MODELLING AND GRAPHICS

$$\begin{pmatrix} c^2\psi s^2\varphi + (1 - c^2\psi s^2\varphi)c\theta & -s\psi c\psi s\varphi(1 - c\theta) - c\psi c\varphi s\theta \\ -s\psi c\psi s\varphi(1 - c\theta) + c\psi c\varphi s\theta & s^2\psi + c^2\psi c\theta \\ -c^2\psi s\varphi c\varphi(1 - c\theta) - s\psi s\theta & s\psi c\psi c\varphi(1 - c\theta) - c\psi s\varphi s\theta \\ & -c^2\psi s\varphi c\varphi(1 - c\theta) + s\psi s\theta \\ & s\psi c\psi c\varphi(1 - c\theta) + c\psi s\varphi s\theta \\ & c^2\psi c^2\varphi + (1 - c^2\psi c^2\varphi)c\theta \end{pmatrix}$$

in which  $\cos$  and  $\sin$  are abbreviated to  $c$  and  $s$ , which is the standard form with  $l = -\cos\psi\sin\varphi$ ,  $m = \sin\psi$ ,  $n = \cos\psi\cos\varphi$ .

## 3.3.1.3.11. Other useful rotations

If rotations in display space are to be controlled by trackerball or tablet then there are two measures available, an  $x$  and a  $y$ , which can define an axis of rotation in the plane of the screen and an angle  $\theta$ . If  $x$  and  $y$  are suitably scaled coordinates of a pen on a tablet then the rotation

$$\begin{pmatrix} \frac{y^2 + x^2c}{x^2 + y^2} & \frac{-xy(1 - c)}{x^2 + y^2} & x\sqrt{x^2 + y^2} \\ \frac{-xy(1 - c)}{x^2 + y^2} & \frac{x^2 + y^2c}{x^2 + y^2} & y\sqrt{x^2 + y^2} \\ -x\sqrt{x^2 + y^2} & -y\sqrt{x^2 + y^2} & c \end{pmatrix}$$

with  $c = \sqrt{1 - (x^2 + y^2)^2}$  is about an axis in the  $xy$  plane (*i.e.* the screen face) normal to  $(x, y)$  and with  $\sin\theta = x^2 + y^2$ . Applied repetitively this gives a quadratic velocity characteristic. Similarly, if an atom at  $(x, y, z, w)$  in display space is to be brought onto the  $z$  axis by a rotation with its axis in the  $xy$  plane the necessary matrix, in homogeneous form, is

$$\begin{pmatrix} \frac{x^2z + y^2r}{x^2 + y^2} & \frac{-xy(r - z)}{x^2 + y^2} & -x & 0 \\ \frac{-xy(r - z)}{x^2 + y^2} & \frac{x^2r + y^2z}{x^2 + y^2} & -y & 0 \\ x & y & z & 0 \\ 0 & 0 & 0 & r \end{pmatrix}$$

with  $r = \sqrt{x^2 + y^2 + z^2}$ .

## 3.3.1.3.12. Symmetry

In Section 3.3.1.1.1 it was pointed out that it is usual to express coordinates for graphical purposes in Cartesian coordinates in ångström units or nanometres. Symmetry, however, is best expressed in crystallographic fractional coordinates. If a molecule, with Cartesian coordinates, is being displayed, and a symmetry-related neighbour is also to be displayed, then the data-space coordinates must be multiplied by

$$\begin{pmatrix} \mathbf{W} & \mathbf{T} \\ \mathbf{0}^T & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathcal{S} \begin{pmatrix} \mathbf{M}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{W} & -\mathbf{T} \\ \mathbf{0}^T & \mathbf{W} \end{pmatrix},$$

where

$$\begin{pmatrix} \mathbf{T} \\ \mathbf{W} \end{pmatrix}$$

are the data-space coordinates of the crystallographic origin,  $\mathbf{M}$  and  $\mathbf{M}^{-1}$  are as in Section 3.3.1.1.1 and  $\mathcal{S}$  is a crystallographic symmetry operator in homogeneous coordinates, expressed relative to the same crystallographic origin.

For example, in  $P2_1$  with the origin on the screw dyad along  $\mathbf{b}$ ,

$$\mathcal{S} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathcal{S} \begin{pmatrix} \mathbf{M}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$\mathcal{S}$  comprises a proper or improper rotational partition,  $\mathbf{S}$ , in the upper-left  $3 \times 3$  in the sense that  $\mathbf{MSM}^{-1}$  is orthogonal, and with the associated fractional lattice translation in the last column, with the last row always consisting of three zeros and 1 at the 4, 4 position. See *IT A* (1983, Chapters 5.3 and 8.1) for a fuller discussion of symmetry using augmented (*i.e.*  $4 \times 4$ ) matrices.

## 3.3.1.4. Modelling transformations

The two sections under this heading are concerned only with the graphical aspects of conformational changes. Determination of such changes is considered under Section 3.3.2.2.

## 3.3.1.4.1. Rotation about a bond

It is a common requirement in molecular modelling to be able to rotate part of a molecule relative to the remainder about a bond between two atoms.

If four atoms are bonded 1–2–3–4 then the dihedral angle in the bond 2–3 is zero if the four atoms are *cis* planar, and a rotation in the 2–3 bond is, by convention (IUPAC–IUB Commission on Biochemical Nomenclature, 1970), positive if, when looking along the 2–3 bond, the far end rotates clockwise relative to the near end. This is valid for either viewing direction. This sign convention, when applied to the  $\mathbf{R}$  matrix of Section 3.3.1.2.1, leads to the following statement.

If one of the two atoms is selected as the near atom and the direction cosines are those of the vector from the near atom to the far atom, and if the matrix is to rotate material attached to the far atom (with the reference axes fixed), then a positive rotation in the foregoing sense is generated by a positive  $\theta$ .

Rotation about a bond normally involves compounding  $\mathbf{R}$  with translations in the manner of Section 3.3.1.3.8.

## 3.3.1.4.2. Stacked transformations

A flexible molecule may require to be drawn in any of a number of conformations which are related to one another by, for example, rotations about single bonds, changes of bond angles or changes of bond lengths, all of which changes may be brought about by the application of suitable homogeneous transformations during the drawing of the molecule (Section 3.3.1.3.8). With suitable organization, this may be done without necessarily altering the coordinates of the atoms in the coordinate list, only the transformations being manipulated during drawing.

The use of transformations in the manner shown below is straightforward for simply connected structures or structures containing only rigid rings. Flexible rings may be similarly handled provided that the matrices employed are consistent with the consequential constraints as described in Section 3.3.2.2.1, though this requirement may make real-time folding of flexible rings difficult.

Any simply connected structure may be organized as a tree with a node at each branch point and with an arbitrary number of sites of

### 3. DUAL BASES IN CRYSTALLOGRAPHIC COMPUTING

conformational change between one node and the next. We shall call such sites and their associated matrices 'conformons'. The technique then depends on the stacking technique in which matrices are stored and later recovered in the reverse order of their storage.

One begins at some reference point deemed to be fixed in data space and at this point one stacks the prevailing viewing transformation. From this reference point one advances through the molecule along the structural tree and as each conformon is encountered its matrix is calculated. The product of the prevailing matrix with the conformon matrix is formed and stacked, and this product becomes the prevailing matrix. This product is constructed with the conformon matrix as a factor on the right, *i.e.* in data space as defined in Section 3.3.1.3.1, and is calculated using the coordinates of the molecule in their unmodified form, *i.e.* before any shape changes are brought about.

This progression leads eventually to an extremity of the tree. At this point drawing is commenced using the prevailing matrix and working backwards towards the fixed root, unstacking (or 'popping') a matrix as each conformon is passed until a node is reached, which, in general, will occur only part way back to the root. On reaching such a node drawing is suspended and one advances along the newly found branch as before, stacking matrices, until another extremity is reached when drawing towards the root is resumed. This alternation of stacking matrices while moving away from the root and drawing and unstacking matrices while moving towards the root is continued until the whole tree is traversed.

This process is illustrated schematically in Fig. 3.3.1.2 for a simple tree with one node, numbered 1, and three conformons at *a*, *b* and *c*. One enters the tree with a current viewing transformation *T* and progresses upwards from the fixed lower extremity. When the conformon at *a* is encountered, *T* is stacked and the product  $TM_a$  is formed. Continuing up the tree, at node 1 either branch may be chosen; we choose the left and, on reaching *b*,  $TM_a$  is stacked and  $TM_aM_b$  is formed. On reaching the tip drawing down to *b* is done with this transformation,  $TM_a$  is then unstacked and drawing continues with this matrix until node 1 is reached. The other branch is then followed to *c* whereupon  $TM_a$  is again stacked and the product  $TM_aM_c$  is formed. From the tip down as far as *c* is drawn with this matrix, whereupon  $TM_a$  is unstacked and drawing continues down to *a*, where *T* is unstacked before drawing the section nearest the root.

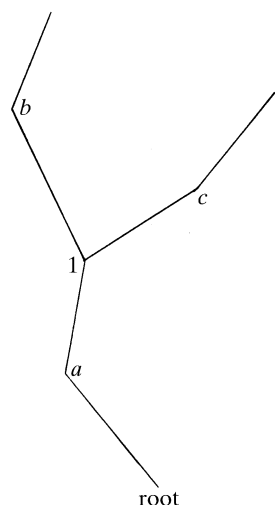


Fig. 3.3.1.2. Schematic representation of a simple branched-chain molecule with a stationary root and two extremities. The positions marked *a*, *b* and *c* are the loci of possible conformational change, here called conformons, and there is a single, numbered branch point.

With this organization the matrices associated with *b* and *c* are unaffected by changes in the conformation at *a*, notwithstanding the fact that changes at *a* alter the direction of the axis of rotation at *b* or *c*.

Two other approaches are also possible. One of these is to start at the tip of the left branch, replace the coordinates of atoms between *b* and the tip by  $M_bX$ , and later replace all coordinates between the tip and *a* by  $M_aX$ , with a similar treatment for the other branch. The advantage of this is that no storage is required for stacked matrices, but the disadvantage is that atoms near the tips of the tree have to be reprocessed for every conformon. It also modifies the stored coordinates, which may or may not be desirable.

The second alternative is to draw upwards from the root using *T* until *a* is reached, then using  $TM_a$  until *b* is reached, then using  $TM'_bM_a$  to the tip, but in this formulation  $M'_b$  must be based on the geometry that exists at *b* after the transformation  $M_a$  has been applied to this region of the molecule, *i.e.*  $M'_b$  is characteristic of the final conformation rather than the initial one.

#### 3.3.1.5. Drawing techniques

##### 3.3.1.5.1. Types of hardware

There are two main types of graphical hardware in use for interactive work, in addition to plotters used for batch work. These main types are raster and vector. In raster equipment the screen is scanned as in television, with a grid of points, called pixels, addressed sequentially as the scan proceeds. Associated with each pixel is a word of memory, usually containing something in the range of 1 to 24 bits per pixel, which controls the colour and intensity to be displayed. Many computer terminals have one bit per pixel (said to be 'single-plane' systems) and these are essentially monochrome and have no grey scale. Four-plane systems are cheap and popular and commonly provide 4-bit by 4-bit look-up tables between the pixel memory and the monitor with one such table for each of the colours red, green and blue. If these tables are each loaded identically then 16 levels of monochrome grey scale are available, but if they are loaded differently 16 different colours are available simultaneously chosen from a total of 4096 possibilities. Four-plane systems are adequate for many applications where colour is used for coding, but are inadequate if colour is intended also to provide realism, where brilliance and saturation must be varied as well as hue. For these applications eight-plane systems are commonly used which permit 256 colours chosen from 16 million using three look-up tables, though the limitations of these can also be felt and full colour is only regarded as being available in 24-plane systems.

Raster-graphics devices are ideal for drawing objects represented by opaque surfaces which can be endowed with realistic reflecting properties (Max, 1984) and they have been successfully used to give effects of transparency. They are also capable of representing shadows, though these are generally difficult to calculate (see Section 3.3.1.5.5). Many devices of this type provide vectorization, area fill and anti-aliasing. Vectorization provides automatically for the loading of relevant pixels on a straight line between specified points. Area fill automatically fills any irregular pre-defined polygon on the screen with a uniform colour with the user specifying only the colour and one point within the polygon. Anti-aliasing is the term used for a technique which softens the staircase effect that may be seen on a line which runs at a small angle to a vertical or horizontal row of pixels.

The main drawback with this type of equipment is that it is slow compared to vector machines. Only relatively simple objects can be displayed with smooth rotation in real time as transformed coordinates have to be converted to pixel addresses and the