3. DUAL BASES IN CRYSTALLOGRAPHIC COMPUTING

3.3.1.2.2. *Measurement of rotations and strains from coordinates*

Given the coordinates of a molecular fragment it is often a requirement to relate the fragment to its image in some standard orientation by a transformation which may be required to be a pure rotation, or may be required to be a combination of rotation and strain. Of the methods reviewed in this section all except (iv) are concerned with pure rotation, ignoring any strain that may be present, and give the best rigid-body superposition. In all these methods, unless inhomogeneous strain is being considered, the best possible superposition is obtained if the centroids of the two images are first brought into coincidence by translation and treated as the origin.

Methods (i) to (v) seek transformations which perform the superposition and impose on these, in various ways, the requirements of orthogonality for the rotational part. All these methods therefore need some defence against indeterminacy that arises in the general transformation if one or both of the fragments is planar, and, if improper rotations are to be excluded, need a defence against these also if the fragment and its image are of opposite chirality. Methods (vi) and (vii) pay no attention to the general transformation and work with variables which are intrinsically rotational in character, and always produce an orthogonal transformation with positive determinant, with no degeneracy arising from planar fragments which need no special attention. Even collinear atoms cause no problem, the superposition being performed correctly but with an arbitrary rotation about the length of the line being present in the result. These methods are therefore to be preferred over the earlier ones unless the purpose of the operation is to detect differences of chirality, although this, too, can be detected with a simple test.

In this review we adopt the same notation for all the methods which, unavoidably, means that symbols are used in ways which differ from the original publications. We use the symbol x for the vector set which is to be rotated and X for the vector set whose orientation is not to be altered, and write the residuals as

$$e_{IA} = D_{Ij}x_{jA} - X_{IA}$$

and, by choice of origin,

$$W_a x_{Ia} = W_a X_{Ia} = 0_I$$

for weights W. The quadratic residual to be minimized is

$$E = W_a e_{ia} e_{ia}$$

and we define the matrix $M_{IJ} = W_a x_{Ia} X_{Ja}$ and use l for the direction cosines of the rotation axis.

(i) McLachlan's first method (McLachlan, 1972, 1982) is iterative and conceptually the simplest. It sets

$$D_{IJ} = A_{Ik} R_{kJ}$$

in which A and R are both orthogonal with R being a current estimate of D and A being an adjustment which, at the beginning of each cycle, has a zero angle associated with it. One iterative cycle estimates a non-trivial A, after which the product AR replaces R.

$$A_{IJ} = (1 - \cos\theta)l_I l_J + \delta_{IJ}\cos\theta - \varepsilon_{IJk}l_k \sin\theta$$

and

$$\left(\frac{\partial A_{IJ}}{\partial\theta}\right)_{\theta=0} = -\varepsilon_{IJk}l_k,$$

therefore

$$\left(\frac{\partial E}{\partial\theta}\right)_{\theta=0} = 2W_a\left(\frac{\partial A_{ij}}{\partial\theta}\right)_{\theta=0} R_{jk}x_{ka}(A_{il}R_{lm}x_{ma} - X_{ia})$$

$$= 2\varepsilon_{ijl}R_{jk}M_{ki}l_l.$$

For this to vanish for all possible rotation axes l the vector

$$g_L = \varepsilon_{ijL}R_{jk}M_{ki}$$

must vanish, *i.e.* at the end of the iteration R must be such that the matrix

$$N_{JI} = R_{Jk}M_{kI}$$

is symmetrical. The vector g represents the couple exerted on the rotating body by forces $2W_A(R_{Ij}x_{jA} - X_{IA})$ acting at the atoms. Choosing

$$l_L = g_L/|g|$$

gives the greatest $|\partial E/\partial\theta|_{\theta=0}$ and $(\partial E/\partial\theta)$ vanishes when

$$\tan\theta = \frac{\varepsilon_{ijk}N_{ji}l_k}{N_{pq}(l_p l_q - \delta_{pq})}$$

in which N is constructed from the current R matrix. A is then constructed from l and this θ and AR replaces R. The process is iterative because a couple about some new axis can appear when rotation about g eliminates the couple about g.

Note that for each rotation axis l there are two values of θ, differing by π, which reduce $|g|$ to zero, corresponding to maximum and minimum values of E. The minimum is that which makes

$$\frac{\partial^2 E}{\partial\theta^2} = 2(\text{tr } N - l_i N_{ij}l_j)$$

positive. Adding π to θ alters R and N and negates this quantity.

Note, too, that the process is essentially characterized as that which makes the product RM symmetrical with R orthogonal. We return to this point in (iii).

(ii) Kabsch's method (Kabsch, 1976, 1978) minimizes E with respect to the nine elements of D, subject to the six constraints

$$D_{kI}D_{kJ} - \delta_{IJ} = 0_{IJ},$$

by using an auxiliary function

$$F = L_{ij}(D_{ki}D_{kj} - \delta_{ij})$$

in which L is symmetric containing six Lagrange multipliers. The Lagrangian function

$$G = E + F$$

then has minima with respect to the elements of D at locations which are dependent, *inter alia*, on the elements of L. By suitably choosing L a minimum of G may be brought into coincidence with the constrained minimum of E. A minimum of G occurs where

$$\frac{\partial G}{\partial D_{IJ}} = 2D_{Ik}(S_{Jk} + L_{Jk}) - 2M_{JI} = 0_{IJ}$$

and the 9×9 matrix

$$\frac{\partial^2 G}{\partial D_{MK}\partial D_{IJ}} = 2\delta_{MI}(S_{JK} + L_{JK})$$

is positive definite, block diagonal, and has

$$S_{JK} = W_a x_{Ja} x_{Ka}$$

which is symmetrical. Thus L must be chosen so as to make the symmetric matrix $(S + L)$ such that

$$D(S + L)^T = M^T$$

with D orthogonal, or $RN = M^T$ with R replacing D since we are now confined to the orthogonal case, and N is symmetric and positive definite.

(iii) Comparison of the Kabsch and McLachlan methods. Using the initials of these authors as subscripts, we have seen that the Kabsch solution involves solving

$$R_{\text{WK}}N_{\text{WK}} = M^T$$

for an orthogonal matrix R_{WK} given that N_{WK} is symmetrical and positive definite. Thus

$$MM^T = N_{\text{WK}}^T R_{\text{WK}}^T R_{\text{WK}} N_{\text{WK}} = N_{\text{WK}}^2$$

and

$$R_{\text{WK}} = M^T (MM^T)^{-1/2}.$$

By comparison, the McLachlan treatment leads to an orthogonal R matrix satisfying

$$R_{\text{ADM}} = N_{\text{ADM}} M^{-1}$$

in which N_{ADM} is also symmetric and positive definite, which similarly leads to

$$R_{\text{ADM}} = (M^T M)^{1/2} M^{-1}.$$

These seemingly different expressions for R_{WK} and R_{ADM} are, in fact, equal, as the following shows

$$R_{\text{WK}} = R_{\text{ADM}} R_{\text{ADM}}^{-1} R_{\text{WK}} = R_{\text{ADM}} M N_{\text{ADM}}^{-1} M^T N_{\text{WK}}^{-1},$$

therefore

$$R_{\text{WK}}^T R_{\text{WK}} = I$$
$$= N_{\text{WK}}^{-1} M N_{\text{ADM}}^{-1} M^T R_{\text{ADM}}^T R_{\text{ADM}} M N_{\text{ADM}}^{-1} M^T N_{\text{WK}}^{-1}.$$

Multiplying on both sides by N_{WK} gives

$$N_{\text{WK}}^2 = (M N_{\text{ADM}}^{-1} M^T)^2,$$

and since both N matrices are positive definite

$$N_{\text{WK}} = M N_{\text{ADM}}^{-1} M^T$$

and conversely

$$N_{\text{ADM}} = M^T N_{\text{WK}}^{-1} M,$$

therefore

$$R_{\text{WK}} = M^T M^{T-1} N_{\text{ADM}} M^{-1} = R_{\text{ADM}}.$$

(iv) Diamond's first method. This method (Diamond, 1976a) differs from the previous ones in that the transformation D is allowed to be a general transformation which is then factorized into the product of an orthogonal matrix R and a symmetrical matrix T. The transformation of x to fit X is thus interpreted as the combination of homogeneous strain and pure rotation in which x is subjected to strain and the result is rotated.

$$D = RT$$
$$T^2 = D^T D$$
$$T = (D^T D)^{1/2}$$
$$R = D(D^T D)^{-1/2}.$$

Furthermore, the solution for D is

$$D = M^T S^{-1}$$

(in the notation of Kabsch), so that

$$R = M^T S^{-1} (S^{-1} M M^T S^{-1})^{-1/2}$$

which may be compared with the results of the previous paragraph.

Although this R matrix by itself (*i.e.* applied without T) does not produce the best rotational superposition (*i.e.* smallest E), it is the one which *exactly* superposes the only three vectors in x whose mutual dispositions are conserved, on their equivalents in X, so that the rotation so found is arguably the best defined one.

Alternatives based on $D = TR$, $D^{-1} = RT$, $D^{-1} = TR$ are all easily developed, and these ideas are developed by Diamond (1976a) to include non-homogeneous strains also.

(v) McLachlan's second method. This method (McLachlan, 1979) is based on the properties of the 6×6 matrix

$$\begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix}$$

and is immune to singularity of M. If $\mathbf{p}$ and $\mathbf{q}$ are three-dimensional vectors such that $(\mathbf{p}^T, \mathbf{q}^T)$ is an eigenvector of this matrix then

$$\begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} = \begin{pmatrix} M\mathbf{q} \\ M^T\mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{p}\lambda \\ \mathbf{q}\lambda \end{pmatrix}.$$

If $\mathbf{q}$ is negated the second equality is maintained provided λ is also negated. Therefore an orthogonal 6×6 matrix

$$\begin{pmatrix} H & H \\ K & -K \end{pmatrix}$$

(consisting of 3×3 partitions) exists for which

$$\begin{pmatrix} H^T & K^T \\ H^T & -K^T \end{pmatrix} \begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix} \begin{pmatrix} H & H \\ K & -K \end{pmatrix} = \begin{pmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{pmatrix}$$

in which Λ is diagonal and contains non-negative eigenvalues. The reverse transformation shows that

$$M = 2H\Lambda K^T$$

and multiplying the eigenvectors together gives

$$H^T H = K^T K = \tfrac{1}{2}I = HH^T = KK^T.$$

Therefore

$$2KH^T M = 4KH^T H\Lambda K^T = 2K\Lambda K^T,$$

but $2KH^T$ is orthogonal and $2K\Lambda K^T$ is symmetrical, therefore [by paragraphs (i) and (iii) above] $2KH^T$ is the required rotation. Similarly, forming

$$M^T = 2K\Lambda H^T$$
$$2M^T H\Lambda^{-1} H^T = 4K\Lambda H^T H\Lambda^{-1} H^T = 2KH^T$$

corresponds to the Kabsch formulation [paragraphs (ii) and (iii)] since $2H\Lambda^{-1} H^T$ is symmetrical and the same rotation, $2KH^T$, appears.

Note that the determinant of the orthogonal matrix so found is twice the product of the determinants of H and of K, and since the positive eigenvalues are collected into Λ it follows that the sign of the determinant of M is the same as the sign of the determinant of the resulting orthogonal matrix. If this is negative it means that the best superposition is obtained if one vector set is inverted and that x and X are of opposite chirality.

Expanding the expression for E, the weighted sum of squares of errors, for an orthogonal transformation shows that this is least when the trace of the product RM is greatest. In this treatment

$$\text{tr}(RM) = \text{tr}(2KH^T \cdot 2H\Lambda K^T) = \text{tr}(2K\Lambda K^T) = \text{tr}(\Lambda).$$

Hence, if the eigenvalues in Λ and $-\Lambda$ are arranged in decreasing

order of modulus, and if the determinant of M is negative, then exchanging the third and sixth columns of

$$\begin{pmatrix} H & H \\ K & -K \end{pmatrix}$$

produces a product $2KH^T$ with positive determinant (*i.e.* a proper rotation) at minimum cost in residual. Similarly, if M is singular and one or more eigenvalues in Λ vanishes it is necessary only to complete an orthonormal set of eigenvectors such that the determinants of H and K have the same sign.

(vi) MacKay's method. MacKay (1984) was the first to consider the rotational superposition problem in terms of the vector $\mathbf{r}$ of Section 3.3.1.2.1. Using quaternion algebra he showed that if a vector $\mathbf{x}$ is rotated to $\mathbf{X} = R\mathbf{x}$ then

$$(\mathbf{X} - \mathbf{x}) = \mathbf{r} \times (\mathbf{X} + \mathbf{x}),$$

where $|\mathbf{r}| = \tan(\theta/2)$ and the direction of $\mathbf{r}$ is the axis of rotation, as may also be shown from elementary considerations. MacKay then solves this for the vector $\mathbf{r}$ by least squares given the vector pairs $\mathbf{X}$ and $\mathbf{x}$. The individual errors are

$$e_{IA} = \varepsilon_{Ijk} r_j (X_{kA} + x_{kA}) - (X_{IA} - x_{IA})$$

and

$$E = W_a e_{ia} e_{ia}.$$

Setting $\partial E / \partial r_P = 0_P$ gives

$$W_a \varepsilon_{iPk} \varepsilon_{ilm} r_l (X_{ka} + x_{ka})(X_{ma} + x_{ma})$$
$$= W_a \varepsilon_{iPk} (X_{ka} + x_{ka})(X_{ia} - x_{ia})$$

which reduces to

$$2\mathbf{V} = -(Q + Q_0)\mathbf{r}$$

in which

$$Q = M + M^T - 2I \operatorname{tr} M$$
$$Q_0 = S + S' - I(\operatorname{tr} S + \operatorname{tr} S')$$
$$V_I = \varepsilon_{Ijk} M_{jk}$$
$$S_{IJ} = W_a x_{Ia} x_{Ja}$$
$$S'_{IJ} = W_a X_{Ia} X_{Ja}.$$

Thus a direct solution for $\mathbf{r}$ is obtained,

$$\mathbf{r} = -2(Q_0 + Q)^{-1}\mathbf{V},$$

the elements of which are u, v and w, and may be used to construct the orthogonal matrix as in Section 3.3.1.2.1. $Q + Q_0$ may be obtained directly from $X + x$.

If the requisite rotation is 180°, $(Q_0 + Q)$ is singular and cannot be inverted. In this case any row or column of the adjoint of $(Q_0 + Q)$ is a vector in the direction of the axis. Normalizing this vector to unity, giving l, gives the requisite orthogonal matrix as

$$R = 2ll^T - I.$$

Note that MacKay's residual E is quadratic in $\mathbf{r}$. E therefore has one minimum and no maximum, and the minimum is reached on the first cycle of least squares. By contrast, the objective function E that is minimized in methods (i), (ii), (v) and (vii) has one minimum, one maximum and two saddle points in the space of the vector $\mathbf{r}$, as shown in (vii).

It may be shown (Diamond, 1989) that if MacKay's solution vector $\mathbf{r}$ is denoted by $\mathbf{r}_M$ and that given by the other methods [except (iv)] by $\mathbf{r}_O$ then

$$\mathbf{r}_M = \mathbf{r}_O - A^{-1}B\mathbf{r}_O$$

in which A and B are real symmetric, positive semi-definite. A is positive definite unless all the individual vector sums $(\mathbf{X} + \mathbf{x})$ are parallel, as can happen when the best rotation is 180°. Thus the MacKay method only gives the same result as the other methods if:

(*a*) the initial orientation is optimal, for then $\mathbf{r}_O = \mathbf{0}$, or

(*b*) perfect fitting is possible, for then $B = \mathbf{0}$, or

(*c*) all the residual vectors (after fitting by $\mathbf{r}_O$) are parallel to $\mathbf{r}_O$, for then B is singular such that $B\mathbf{r}_O = \mathbf{0}$. In general, $|\mathbf{r}_M| \le |\mathbf{r}_O|$. $\mathbf{r}_O$ may be found by iterating $\mathbf{r}_M$, but $\mathbf{x}$ must be replaced by $R\mathbf{x}$ on each iteration.

(vii) Diamond's second method. This is closely related to MacKay's method, but uses a four-dimensional vector $\boldsymbol{\rho}$ with components λ, μ, ν and σ in which λ, μ and ν are the direction cosines of the rotation axis multiplied by $\sin(\theta/2)$ and σ is $\cos(\theta/2)$. In terms of such a vector Diamond (1988) showed that

$$E = E_0 - 2\boldsymbol{\rho}^T P \boldsymbol{\rho}$$

in which E is the weighted sum of squares of coordinate differences, as before, E_0 is its value before any rotation is applied and P is the matrix

$$P = \begin{pmatrix} Q & \mathbf{V} \\ \mathbf{V}^T & 0 \end{pmatrix}.$$

The rotation matrix R corresponding to the vector $\boldsymbol{\rho}$ is then the last of the forms for R given in Section 3.3.1.2.1.

The minimum E is therefore E_0 minus twice the largest eigenvalue of P since $\boldsymbol{\rho}^T \boldsymbol{\rho} = 1$, and a stationary value of E occurs when $\boldsymbol{\rho}$ is any of the four eigenvectors of P. E thus has a maximum, a minimum and two saddle points, in general, and its value may be determined before any coordinates are transformed. Diamond also showed that the orientations giving these stationary values are related by the operations of 222 symmetry. Equivalent results have also been obtained by Kearsley (1989).

As an alternative to solving a 4×4 eigenproblem, Diamond also showed that the vector $\mathbf{r}$, as in MacKay's solution, may be obtained by iterating

$$\alpha_0 = E_0/2$$
$$\mathbf{r}_n = (\alpha_n I - Q)^{-1}\mathbf{V}$$
$$\alpha_{n+1} = \frac{\mathbf{V} \cdot \mathbf{r}_n + \alpha_n r_n^2}{1 + r_n^2}$$

which has the property that if X and x are exactly superposable then α_0 is the exact solution and no iteration is necessary. If X and x are similar but not exactly superposable then a small number of iterations may be required to reach a stable $\mathbf{r}$ vector, though the matrix Q_0 is not required. As in MacKay's solution, $(\alpha I - Q)$ is singular at the end of the iteration if the required rotation is 180°, but the MacKay and Diamond methods both have the advantage that improper rotations are never generated by these means, and methods based on P and $\boldsymbol{\rho}$ rather than Q and $\mathbf{r}$ are trouble-free for 180° rotations. The iterative loop in this method does not require $R\mathbf{x}$ to be redetermined on each cycle.

Finally, it may be shown that if p_1, p_2, p_3, p_4 are the eigenvalues of P arranged in descending order and

$$p_1 - p_2 - p_3 + p_4$$

is negative, then a closer superposition may be obtained by reversing the chirality of one of the vector sets, and the R matrix constructed from $\boldsymbol{\rho}_4$ optimally superimposes $R\mathbf{x}$ onto $-X$, the enantiomer of X (Diamond, 1990*b*).