A mathematical procedure for superimposing atomic coordinates of proteins. By A. D. McLachlan, *Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge, England*

A procedure is given which determines the best rigid-body rotation and translation that matches a given set of measured atomic coordinates to a fixed set of guide coordinates and minimizes the weighted sum of squared deviations.

It is often useful to compare two sets of coordinates for the same set of atoms in a protein by finding a rigid-body rotation and translation which makes them coincide as closely as possible. Examples are the comparison of chymotrypsinogen with chymotrypsin (Freer, Kraut, Robertus, Wright & Xuong, 1970), insect haemoglobin with myoglobin (Huber, Epp, Steigemann & Formanek, 1971), or the fitting of a rigid haem group to approximate measured coordinates. The problem can be stated as follows. Given two sets (*A* and *B*) of *N* vectors a_α, b_α, ($\alpha = 1, \ldots N$) find an orthogonal rotation matrix $\mathbf{R}$ with determinant $+1$, and a translation $\mathbf{t}$ which converts the coordinates $a_{i\alpha}(i=1,2,3)$ to

$$r_{i\alpha} = \sum_j R_{ij} a_{j\alpha} + t_i \qquad (1)$$

and minimizes the residual

$$E = \tfrac{1}{2} \sum_{i\alpha} w_\alpha (r_{i\alpha} - b_{i\alpha})^2 . \qquad (2)$$

Here w_α is a weight assigned to each atom. The translation can be removed from the problem by ensuring that the centroids of both sets *A* and *B* have already been shifted to the origin of coordinates. Diamond (1966) has described a procedure for finding a suitable rotation, but it does not necessarily lead to a minimum for *E*, and it breaks down if either set of atoms lies in a plane. The procedure given here overcomes these difficulties.

Consider a small additional rotation through an angle θ about the direction $\mathbf{l}$ with direction cosines l_1, l_2, l_3. To second order in θ, the vector $\mathbf{r}$ becomes

$$\mathbf{r} + \theta \wedge \mathbf{r} + \tfrac{1}{2}\theta \wedge (\theta \wedge \mathbf{r}) \qquad (3)$$

where $\theta = \mathbf{l}\theta$. The corresponding change in *E* is

$$\delta E = -\sum_i g_i \theta_i + \tfrac{1}{2} \sum_{i,j} \theta_i T_{ij} \theta_j . \qquad (4)$$

Here

$$\mathbf{g} = \sum_\alpha w_\alpha \mathbf{r}_\alpha \wedge \mathbf{b}_\alpha \qquad (5)$$

is the couple that would be produced by a force of strength $w_\alpha(\mathbf{b}_\alpha - \mathbf{r}_\alpha)$ acting at each point $\mathbf{r}_\alpha$.

Now introduce the matrices

$$U_{ij} = \sum_\alpha w_\alpha a_{i\alpha} b_{j\alpha} \quad V_{ij} = \sum_\alpha w_\alpha r_{i\alpha} b_{j\alpha} . \qquad (6)$$

The antisymmetric part of $\mathbf{V}$ gives the couple, since $g_1 = V_{23} - V_{32}$ and so on. Also the symmetric part yields the second derivative matrix

$$T_{ij} = v\delta_{ij} - \tfrac{1}{2}(V_{ij} + V_{ji}) \qquad (7)$$

with

$$v = V_{11} + V_{22} + V_{33} . \qquad (8)$$

The residual *E* can also be calculated directly from the diagonal sum of $\mathbf{V}$, since

$$E = \tfrac{1}{2} \sum_{i\alpha} w_\alpha (r_{i\alpha}^2 + b_{i\alpha}^2 - 2r_{i\alpha} b_{i\alpha})$$

$$= \tfrac{1}{2} \sum_\alpha w_\alpha (\mathbf{a}_\alpha^2 + \mathbf{b}_\alpha^2) - v . \qquad (9)$$

Another useful property of the matrix $\mathbf{V}$ is that it transforms simply under rotations, with

$$\mathbf{V} = \mathbf{RU} . \qquad (10)$$

The least-squares fitting problem thus reduces to one of finding a rotation $\mathbf{R}$ such that $\mathbf{V}$ is symmetric, $\mathbf{T}$ is positive definite, and the diagonal sum v is therefore a maximum. Once the initial matrix $\mathbf{U}$ has been formed the atomic coordinates need not be consulted again.

Iterative solution

A succession of rotations are applied. At positions far from the minimum one can use a steepest-descent method, suggested by Jacobi's method for diagonalizing a real symmetric matrix (Wilkinson, 1965; *Modern Computing Methods*, 1961). To derive it consider the situation in two dimensions

$$\mathbf{V} = \begin{pmatrix} \cos\theta - \sin\theta \\ \sin\theta \quad \cos\theta \end{pmatrix} \begin{pmatrix} U_{11} \quad U_{12} \\ U_{21} \quad U_{22} \end{pmatrix} . \qquad (11)$$

Here a rotation defined by

$$A \cos\theta = U_{11} + U_{22} , \quad A \sin\theta = U_{12} - U_{21} \qquad (12)$$

$$A^2 = (U_{11} + U_{22})^2 + (U_{12} - U_{21})^2 \qquad (13)$$

makes $\mathbf{V}$ diagonal and alters the diagonal sum from $A \cos\theta$ to A. In three dimensions the axis of rotation $\mathbf{l}$ is taken parallel to that of the couple $\mathbf{g}$, and the angle is given by

$$A \cos\theta = u - \sum_{ij} l_i U_{ij} l_j , \quad A \sin\theta = g \qquad (14)$$

where $u = U_{11} + U_{22} + U_{33}$. The rotation then reduces the component of $\mathbf{g}$ about the $\mathbf{l}$ axis to zero and decreases *E* by $\tfrac{1}{4}A \sin^2(\theta/2)$. A succession of such rotations always leads to a true minimum, but the process converges slowly.

Once the second derivative matrix $\mathbf{T}$ becomes positive definite, and the angle θ calculated from (14) becomes small, it is quicker to change over to a Newton–Raphson method, applying rotations defined by

$$\theta = \mathbf{gT}^{-1} . \qquad (15)$$

With these tactics 6 to 10 cycles are usually sufficient, and the whole calculation takes about 7 seconds on an IBM 360/44.

Analytic solution

A formal solution of (10) is $V = (U'U)^{1/2}$, and $R = VU^{-1}$, but this leaves the signs of the square roots undefined and fails when U is singular. (U' is the transpose of U)

More generally consider the matrices $U'U$ and UU'. They are both symmetric and positive definite. Also the diagonal sums of $(U'U)^n$ and $(UU')^n$ are the same, and it can be shown that both matrices have the same eigenvalues. Therefore orthogonal matrices H, K exist, with determinant $+1$, such that

$$UU' = H^{-1}D^2H, \quad U'U = K^{-1}D^2K \qquad (16)$$

where D^2 is a diagonal matrix with positive elements. The matrix

$$M = HUK^{-1} \qquad (17)$$

has the property that

$$MM' = M'M = D^2 \qquad (18)$$

and commutes with D^2. Hence, if D^2 is nondegenerate, M is already diagonal. If D^2 is degenerate, with eigenvalues D_λ^2, M and M' are block-diagonal, each block being of the form $M_\lambda = d_\lambda Q_\lambda$, where $d_\lambda^2 = D_\lambda^2$ and $Q_\lambda Q'_\lambda = I_\lambda$. Thus Q_λ is orthogonal, with determinant ± 1. The matrix $Q^{-1}M = d$ is therefore diagonal, and

$$V_1 = (K^{-1}Q^{-1}H)U = K^{-1}dK \qquad (19)$$

is symmetric. However, the transformation Q derived by this procedure may have determinant -1, and the signs of the elements d_λ may not maximize the diagonal sum of V_1. To correct this we now construct a matrix P with all its diagonal elements ± 1, such that Det $PQ = 1$ and form the diagonal matrix $D = Pd$, with $D_\lambda = \pm d_\lambda$. The correct solution is then

$$V = (K^{-1}PQ^{-1}H)U = K^{-1}DK \qquad (20)$$

$$R = K^{-1}PQ^{-1}H \qquad (21)$$

and

$$v = D_1 + D_2 + D_3 .$$

This is unique unless U is a singular matrix of rank 1, for the subspace belonging to the eigenvalue $D_\lambda^2 = 0$ is then of dimension 2, and the rotation is undefined. This happens when the atoms of either set lie on a line.

References

DIAMOND, R. (1966). Acta Cryst. 21, 253.
FREER, S. T., KRAUT, J., ROBERTUS, J. D., WRIGHT, H. T. & XUONG, NG. H. (1970). Biochem. Wash. 9, 1997.
HUBER, R., EPP, O., STEIGEMANN, W. & FORMANEK, H. (1971). Europ. J. Biochem. 19, 42.
Modern Computing Methods (1961). National Physical Laboratory, Notes on Applied Science No 16, 2nd edition. London: H. M. Stationery Office.
WILKINSON, J. H. (1965). The Algebraic Eigenvalue Problem. Oxford Univ. Press.

The *a priori* optimization of diffractometer data to achieve the minimum average variance in the electron density.

By R. C. G. KILLEAN, *Department of Physics, University of St. Andrews, St. Andrews, Scotland*

A discussion is given of the relationship between diffractometer data accuracy and the average variance of the electron density.

A consideration of constant-count-per-reflexion diffractometer experiments such as those performed, for example, by the Siemens AED System, has led Killean (1967) to show that to obtain data capable of yielding an R index of $0\cdot10$ the total number of counts per reflexion need not exceed twenty-five, for low background reflexions, or one-hundred and twenty-five counts for a peak to background ratio of three to two. Clearly in view of these small counts no problem exists in the high-speed collection of data required only for stereochemical determination or conformation. The collection of data becomes more time consuming when the data is to be used for deductions requiring highly accurate electron densities. This paper is concerned with the mode of data collection likely to make the most efficient use of the diffractometer time available.

The approach to the problem is entirely *a priori* and, as such, represents a basic planning approach to single-crystal diffractometry. It assumes low background counts and consequently it is likely that a monochromator would be used in the experiment. The separate, but related, attempts of Hamilton (1967) and Shoemaker (1968) to optimize the collection of data for least squares analysis are *a posteriori*

in that they require that the structure must have been solved before the data can be optimally recollected with respect to, for Hamilton's treatment, the variance of one positional or one thermal parameter and, for Shoemaker's analysis, a linear combination of the weights of the various parameters.

Cruickshank (1960) has summarized the requirements that must be satisfied for an accurate structure determination. It is assumed in the following analysis that these conditions are to be satisfied.

Cruickshank (1949) has derived an expression for the variance on the electron density for centrosymmetric space groups and Killean & Lawrence (1969b) have modified his approach to take account of random errors in the phase angles of the structure factors for non-centrosymmetric space groups. Essentially when the average variance in the electron density is to be considered it is sufficient to use Cruickshank's form for $P\bar{1}$

$$\langle \sigma^2(\varrho) \rangle = \tfrac{1}{2} \left(\frac{2}{V} \right)^2 \sum \sigma^2(\mathbf{h}) \qquad (1)$$

where $\sigma^2(\mathbf{h}) = $ variance of the structure factor.