The authors thank Professor Kupcik from the Mineralogisch-Kristallographisches Institut der Universität Göttingen for allowing them to carry out measurements on the five-circle diffractiometer installed at HASYLAB/DESY Hamburg.

References

BACHMANN, R., KOHLER, H., SCHULZ, H. & WEBER, H.-P. (1985). *Acta Cryst.* A**41**, 35-40.
BACHMANN, R., KOHLER, H., SCHULZ, H., WEBER, H.-P., KUPCIK, W., WENDSCHUH-JOSTIES, M., WOLF, A. & WULF, R. (1983). *Angew. Chem.* **95**, 1011-1012.
BECKER, P. J. & COPPENS, P. (1974). *Acta Cryst.* A**30**, 129-153.
BECKER, P. J. & COPPENS, P. (1975). *Acta Cryst.* A**31**, 417-425.
BEIMGRABEN, O., GRAEFF, W., HAHN, U., KNABE, J., KOCH, E. E., KUNZ, C., MATERLIK, G., SAILE, V., SCHMIDT, W., SONNTAG, B. F., SPRÜSSEL, G., WEINER, E. W. & ZIETZ, R. (1981). *Phys. Bull.* **37**, 1-10.
COOPER, M. J. & ROUSE, K. D. (1970). *Acta Cryst.* A**26**, 214-223.
COPPENS, P. & HAMILTON, W. C. (1970). *Acta Cryst.* A**26**, 71-83.
DARWIN, C. G. (1922). *Philos. Mag.* **43**, 800-829.
KUPCIK, V., WULF, R., WENDSCHUH, M., WOLF, A. & PAEHLER, A. (1983). *Nucl. Instrum. Methods*, **208**, 519-522.
MATHIESON, A. MCL. (1977*a*). *Acta Cryst.* A**33**, 610-617.
MATHIESON, A. MCL. (1977*b*). *Acta Cryst.* A**33**, 133-136.
SCHNEIDER, J. R. (1983). *Fortschr. Mineral.* **61**, 85-151.
ZACHARIASEN, W. H. (1945). *Theory of X-ray Diffraction in Crystals.* New York: Wiley.
ZACHARIASEN, W. H. (1967). *Acta Cryst.* **23**, 558-564.
ZACHARIASEN, W. H. (1968). *Acta Cryst.* A**24**, 421-427.

Acta Cryst. (1986). A**42**, 110-113

A Toolkit for Computational Molecular Biology. II. On the Optimal Superposition of Two Sets of Coordinates

BY ARTHUR M. LESK*

Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

(*Received* 25 *March* 1985; *accepted* 4 *September* 1985)

Abstract

The determination of the optimal transformation to superpose two sets of points has many applications to the analysis of structures of proteins and nucleic acids. A new formulation of this problem is presented, which reduces it to the unconstrained maximization of a function of a single variable. This method is currently being applied in investigations of common substructures of proteins.

1. Introduction

The superposition of two coordinate sets is the basis of a number of techniques for the analysis and comparison of molecular structures. Considerable effort has been made to develop fast algorithms (McLachlan, 1972; Diamond, 1976; Kabsch, 1976, 1978; McLachlan, 1979, 1982; Mackay, 1984; Ken-Knight, 1984).

Several efficient algorithms are known, and programs based on them are already fast enough for calculations in which only a few superpositions are required. However, in searching two or more protein structures for common substructures, a very large number of superpositions are required to test all combinations of segments from each pair of structures. Such analyses can sometimes be organized to superpose a succession of related substructures, so that the optimal transformation determined for one superposition may be nearly optimal for the next. In this case, optimization methods make it convenient to apply the results of one calculation to speed up the next. The method described here, a development of those used by McLachlan (1972, 1982), has this feature.

2. Statement of the problem

This section follows McLachlan's analysis very closely (McLachlan, 1972; *cf.* Golub & Van Loan, 1983). Let $\mathbf{x}_i$, $i = 1, \ldots, N$, and $\mathbf{y}_i$, $i = 1, \ldots, N$, be two sets of points in 3-space. We wish to superpose them by means of a rigid-body motion of the $\mathbf{y}_i$ into the set $\mathbf{y}'_i$ such that the sum of the squares of deviations

$$D = \sum_{i=1}^{N} (\mathbf{x}_i - \mathbf{y}'_i)^2$$

is minimized. Any rigid-body motion in 3-space may be decomposed into a rotation and a translation. The optimal translation is that which brings the mean positions (colloquially, 'centers of gravity') of the two sets into coincidence. We therefore may assume without loss of generality that the mean positions of the two sets of points coincide at the origin. The problem is to determine the proper rotation matrix

* Permanent address: Fairleigh Dickinson University, Teaneck-Hackensack Campus, Teaneck, NJ 07666, USA.

0108-7673/86/020110-04$01.50

R (R_{ij} real, $R^T = R^{-1}$, det $R = +1{\cdot}0$) such that

$$D = \sum_i (\mathbf{x}_i - R\mathbf{y}_i)^2$$

is a minimum.

By rewriting D in the form

$$D = \sum_i (\mathbf{x}_i - R\mathbf{y}_i)^T . (\mathbf{x}_i - R\mathbf{y}_i)$$
$$= \sum_i (|\mathbf{x}_i|^2 + |\mathbf{y}_i|^2) - \sum_i 2\mathbf{x}_i^T . R\mathbf{y}_i,$$

in which we use the orthogonality property of rotation matrices: $(R\mathbf{y}_i)^T . R\mathbf{y}_i = \mathbf{y}_i^T . \mathbf{y}_i$; and recognizing that the first and second terms depend only on the *original* coordinate sets, independent of R, we see that the minimization of D is equivalent to the maximization of

$$E = \sum_i \mathbf{x}_i^T . R\mathbf{y}_i.$$

Explicitly, with $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3})$ and $\mathbf{y}_i = (y_{i1}, y_{i2}, y_{i3})$:

$$E = \sum_i \left[\sum_{k=1}^{3} x_{ik} . (Ry_i)_k \right]$$
$$= \sum_i \sum_{k=1}^{3} \sum_{j=1}^{3} [x_{ik} R_{kj} y_{ij}]$$
$$= \sum_{k=1}^{3} \sum_{j=1}^{3} R_{kj} \sum_i x_{ik} y_{ij} = \text{trace}\,(RA),$$

where $A_{jk} = \sum_i x_{ij} y_{ik}$. A is called the correlation matrix.

3. Computational methods

Several computational approaches have been pursued:

(1) Optimization techniques, in some cases specialized to make use of features of this particular problem (McLachlan, 1982).

(2) Use of the polar decomposition or singular value decomposition of the correlation matrix.

If A is a real nonsingular matrix, then $A = UH$, where U is a unique orthogonal matrix and H is Hermitian and positive definite (Diamond, 1976; Golub & Van Loan, 1983). Higham (1984) has proposed a fast algorithm to determine U and H. Suppose $H = Q^{-1}DQ$, where $D = \text{diag}\, d_i$ and $d_i \geq 0$. Then $\text{tr}(RA) = \text{tr}(RUH) = \text{tr}(RUQ^{-1}DQ) = \text{tr}(QRUQ^{-1}D) = \text{tr}(VD)$, where $V = QRUQ^{-1}$. But because $d_i \geq 0$, max $\text{tr}(RA) = \text{tr}\, D$ and the maximum is obtained if $QRUQ^{-1} = I$, which is equivalent to $R = U^T$.

If an orthogonal starting approximation V to the orthogonal matrix U is available, when we may apply the decomposition to V^TA to produce WH, where $W \sim 1$, and then $U = VW$.

(3) Explicit solution of the least-squares problem using Lagrange multipliers (Kabsch, 1976, 1978). This requires calculating the eigenvalues of A^TA.

Johnston (1982) has compared the numerical characteristics of some of these methods.

4. Introduction of an explicit representation of the rotation matrix

In this section we present an approach, based on an explicit representation of the rotation matrix, which reduces the problem to a minimization of a function of a single variable.

Expressing the general three-dimensional rotation matrix $R(l, m, n, \theta)$ as a function of the direction cosines l, m, n of the axis and the angle of rotation θ, with l, m, n obeying the constraint $l^2 + m^2 + n^2 = 1$ (Lomont, 1959), we can express the quantity E (see § 2) in the following form:

$$E = \text{tr}\,(Aa) - [\text{tr}(Aa) - \text{tr}\, A] \cos\theta + \text{tr}\,(Ab) \sin\theta,$$

where a and b are the 3×3 matrices

$$a = \begin{pmatrix} l^2 & ml & nl \\ ml & m^2 & mn \\ nl & mn & n^2 \end{pmatrix}; \quad b = \begin{pmatrix} 0 & -n & m \\ +n & 0 & -l \\ -m & l & 0 \end{pmatrix}.$$

For any given direction cosines l, m and n, the maximum of this quantity (with respect to θ) is

$$E = \text{tr}(Aa) + [(\text{tr}\, Aa - \text{tr}\, A)^2 + (\text{tr}\, Ab)^2]^{1/2}.$$

Thus, given an axis of rotation, the determination of the optimal *angle* of rotation is immediate. We must maximize E with respect to the *direction* of the axis of rotation. Let $\hat{\mathbf{n}} = (lmn)$ denote the unit vector along some trial axis of rotation. Then $\text{tr}\, Aa = \hat{\mathbf{n}} A^s \hat{\mathbf{n}}$, and $\text{tr}\, Ab = \hat{\mathbf{n}} . \mathbf{p}$, where $A^s = \frac{1}{2}(A^T + A)$ and $\mathbf{p} = (A_{23} - A_{32}, A_{31} - A_{13}, A_{12} - A_{21})$. We can express E as

$$E = \hat{\mathbf{n}}^T A^s \hat{\mathbf{n}} + [(\hat{\mathbf{n}}^T A^s \hat{\mathbf{n}} - \text{tr}\, A)^2 + (\hat{\mathbf{n}} . \mathbf{p})^2]^{1/2}$$

or

$$E = F + [F^2 + t^2 |\mathbf{p}|^2]^{1/2} + \text{tr}\, A,$$

where $F = \hat{\mathbf{n}}^T A^s \hat{\mathbf{n}} - \text{tr}\, A$, $\hat{\mathbf{p}} = \mathbf{p}/|\mathbf{p}|$, and $t = \hat{\mathbf{n}} . \hat{\mathbf{p}}$. Here t represents the component parallel to the vector $\mathbf{p}$ of the unit vector along the axis of rotation.

If $|\mathbf{p}| = 0$, A is symmetric, and $\theta = 0$ or 180°. In this case $E(\hat{\mathbf{n}}, \theta) = \hat{\mathbf{n}} A^s \hat{\mathbf{n}} - [\hat{\mathbf{n}} A^s \hat{\mathbf{n}} - \text{tr}\, A] \cos\theta$. For any $\hat{\mathbf{n}}$, max $E(\hat{\mathbf{n}}, \theta) = \hat{\mathbf{n}} A^s \hat{\mathbf{n}} + |\hat{\mathbf{n}} A^s \hat{\mathbf{n}} - \text{tr}\, A|$.

$$\max_{\theta, \hat{\mathbf{n}}} E(\hat{\mathbf{n}}, \theta) = \max \{\text{tr}\, A, 2 \max_{\hat{\mathbf{n}}} \hat{\mathbf{n}} A^s \hat{\mathbf{n}} - \text{tr}\, A\}$$
$$= \max \{\text{tr}\, A, 2\lambda_1 - \text{tr}\, A\},$$

where λ_1 is the largest eigenvalue of A^s. If $\text{tr}\, A > 2\lambda_1 - \text{tr}\, A$, $\cos\theta = 1$, $\theta = 0$ and $R =$ the identity matrix. If $\text{tr}\, A < 2\lambda_1 - \text{tr}\, A$, $\cos\theta = -1$, $\theta = 180°$, and the axis of rotation is the eigenvector of A^s corresponding to λ_1.

If $|\mathbf{p}| \neq 0$, we want to solve the following:

$$\underset{-1 < t < 1, \hat{\mathbf{n}} . \hat{\mathbf{p}} = t}{\text{maximize}} \max \{F + (F^2 + t^2 |\mathbf{p}|^2)^{1/2} + \text{tr}\, A\}.$$

Consider the maximization with respect to $\hat{\mathbf{n}}$ for any fixed t. The function $f(x) = x + (x^2 + c)^{1/2}$, for any positive constant c, is monotonically increasing, since $f'(x) = 1 + x(x^2 + c)^{-1/2} > 0$. Therefore, given any

value of t, we can maximize E with respect to $\hat{\mathbf{n}}$ by maximizing $F = \hat{\mathbf{n}}^T A^s \hat{\mathbf{n}}$. Then if $F_{\max}(t) = \max \hat{\mathbf{n}}^T A^s \hat{\mathbf{n}}$, subject to the constraints $|\hat{\mathbf{n}}| = 1$, $\hat{\mathbf{n}} \cdot \hat{\mathbf{p}} = t$, the problem can be stated:

$$\underset{-1 \le t \le 1}{\text{maximize}} \{F_{\max}(t) + [F_{\max}(t)^2 + t^2 |\mathbf{p}|^2]^{1/2} + \text{tr } A\}.$$

To remove the constraint on t, let $t = (1-v^2)/(1+v^2)$. Then $(1-t^2)^{1/2} = 2v/(1+v^2)$. Now define a right-handed coordinate system with its positive x axis lying along $\mathbf{p}$. We may use any such system, of which there is a family, the members of which differ by the rotation of the y and z axes around the vector $\mathbf{p}$. A numerically safe way to construct one is as follows: Assuming that $|\mathbf{p}| \neq 0$ we define the axes $\hat{e}_i^{\text{new}}$, $i = 1, 2, 3$, of the new coordinate system in terms of $\mathbf{p}$ and the axes $\hat{\mathbf{e}}_i^{\text{old}}$, $i = 1, 2, 3$, of the original coordinate system. Let $\hat{\mathbf{e}}_1^{\text{new}} = \hat{\mathbf{p}} = \mathbf{p}/|\mathbf{p}|$. Find j such that $|\mathbf{p} \cdot \hat{\mathbf{e}}_j^{\text{old}}|$ is the minimum, and set $\hat{\mathbf{e}}_2^{\text{new}} = (\mathbf{p} \times \hat{\mathbf{e}}_j^{\text{old}})/|\mathbf{p} \times \hat{\mathbf{e}}_j^{\text{old}}|$. Then set $\hat{\mathbf{e}}_3^{\text{new}} = \hat{\mathbf{p}} \times \hat{\mathbf{e}}_2^{\text{new}}$.

Let T be the matrix that multiplies the components of a vector expressed in the new coordinate system to give the components of the vector expressed in the original coordinate system. Observe that T is independent of t. The most general unit vector for which the component along $\mathbf{p}$ is t is - expressed in the *new* coordinate system - of the form:

$$\hat{\mathbf{n}} = \begin{pmatrix} (1-v^2)/(1+v^2) \\ [2v/(1+v^2)] \cos \varphi \\ [2v/(1+v^2)] \sin \varphi \end{pmatrix}.$$

Let S be the scaling matrix:

$$S = \begin{pmatrix} (1-v^2)/(1+v^2) & 0 & 0 \\ 0 & 2v/(1+v^2) & 0 \\ 0 & 0 & 2v/(1+v^2) \end{pmatrix}$$

and let $B = S^T T^T A^s T S$. Advantage can be taken, in the implementation of this method, of the fact that B is symmetric.

Then

$$F = \hat{\mathbf{n}} A^s \hat{\mathbf{n}} = (1 \ \cos \varphi \ \sin \varphi) B \begin{pmatrix} 1 \\ \cos \varphi \\ \sin \varphi \end{pmatrix}.$$

To maximize this quantity, introduce another variable u such that

$$\cos \varphi = (1-u^2)/(1+u^2), \qquad \sin \varphi = 2u/(1+u^2).$$

Then, in terms of u,

$$\begin{aligned} (1+u^2)^2 F'(u) = \ & (B_{23} + B_{32} - B_{12} - B_{21}) u^4 \\ & -2(2B_{33} - 2B_{22} + B_{12} + B_{21}) u^3 \\ & -6(B_{23} + B_{32}) u^2 \\ & +2(2B_{33} - 2B_{22} - B_{12} - B_{21}) u \\ & +(B_{13} + B_{31} + B_{23} + B_{32}). \end{aligned}$$

Observe that $(1+u^2)^2 F'$ is a quartic polynomial in u.

We therefore are led to suggest the following procedure for maximizing $E(\hat{\mathbf{n}}, t)$ subject to $|\hat{\mathbf{n}}| = 1$, $\hat{\mathbf{n}} \cdot \hat{\mathbf{p}} = t$:

Utilize a program for finding the maximum of a function of *one* variable, v. When this program calls for a function evaluation, use v to determine the matrix S and then the matrix B. From the elements of B, determine the coefficients of the quartic equation $(1+u^2)^2 F'(u) = 0$. Solve this. For each real root compute F and choose the largest value. From this maximum value of F compute the corresponding value of E and return this as the value of the function of v that we seek to maximize.

The equation $F' = 0$ might be solved in a closed form (Abramowitz & Stegun, 1965), possibly leading to further simplification of the problem. In particular, this might be useful to facilitate the analytic calculation of dE/dv. Alternatively, advantage can be taken of the fact that the roots of the polynomial $(1+u^2)F'(u)$ are continuous functions of v (Householder, 1970).

In two special cases there are explicit solutions for the maximum of F: (1) for $t = 0$ ($v = 1$; the axis of rotation is perpendicular to $\mathbf{p}$):

$$E_{\max} = F_0 + |F_0 - \text{tr } A|,$$

where $F_0 = \frac{1}{2}(B_{22} + B_{33}) + [(B_{22} - B_{33})^2/4 + B_{23}^2]^{1/2}$. The corresponding value of u is $\tan\{0{\cdot}25 \times \arctan[2B_{23}/(B_{22} - B_{23})]\}$.

(2) For $t = 1$ ($v = 0$, axis of rotation parallel to $\mathbf{p}$):

$$E_{\max} = B_{11} + [(B_{11} - \text{tr } A)^2 + |\mathbf{p}|^2]^{1/2}.$$

The corresponding value of $\mathbf{u}$ is undefined.

These are useful in estimating a starting value of t with which to enter the optimization routine.

When the maximum with respect to v is located, use the corresponding optimal value of u to calculate

$$\hat{\mathbf{n}} = TS \begin{pmatrix} 1 \\ \cos \varphi \\ \sin \varphi \end{pmatrix}.$$

This determines the axis of rotation. The angle of rotation around this axis is given by $\arctan[\hat{\mathbf{n}} \cdot \mathbf{p}/(\text{tr } A - \hat{\mathbf{n}}^T A^s \hat{\mathbf{n}})] = -t|\mathbf{p}|/F_{\max}$.

Specification of the axis and angle of rotation completely determines the desired rotation matrix R. Because the formulation of the problem defines an explicit rotation around an explicit axis, there is no danger of inverting the hand of the structure.

If all that we wish to compute is the minimum r.m.s. deviation of the superposed sets of points, it is unnecessary to compute the axis or angle of rotation.

This algorithm has been implemented in a Fortran program, making use of programs from the Harwell subroutine library. In practice it is unnecessary to solve a quartic equation 'from scratch' for each value

of t, because the roots – the candidate values of u – are continuous functions of v, and indeed we have found that the optimum value of u varies very slowly with v.

The work was supported in part by US National Science Foundation research grant PCM83-20171.

References

ABRAMOWITZ, M. & STEGUN, I. A. (1965). (Editors.) *Handbook of Mathematical Functions*, pp. 18-19. New York: Dover.
DIAMOND, R. (1976). *Acta Cryst.* A**32**, 1-10.
GOLUB, G. H. & VAN LOAN, C. F. (1983). *Matrix Computations*, pp. 425-426. Baltimore, MD: The Johns Hopkins Univ. Press.
HIGHAM, N. J. (1984). *Computing the Polar Decomposition – With Applications.* Numerical Analysis Report No. 94, Department of Mathematics, Univ. of Manchester, November 1984.
HOUSEHOLDER, A. S. (1970). *The Numerical Treatment of a Single Nonlinear Equation*, section 2.3. New York: McGraw-Hill.
JOHNSTON, R. L. (1982). *Numerical Methods: a Software Approach*, section 2.3.3. New York: John Wiley.
KABSCH, W. (1976). *Acta Cryst.* A**32**, 922-923.
KABSCH, W. (1978). *Acta Cryst.* A**34**, 827-828.
KENKNIGHT, C. E. (1984). *Acta Cryst.* A**40**, 708-712.
LOMONT, J. S. (1959). *Applications of Finite Groups*, p. 14. New York: Academic Press.
MACKAY, A. L. (1984). *Acta Cryst.* A**40**, 165-166.
MCLACHLAN, A. D. (1972). *Acta Cryst.* A**28**, 656-657.
MCLACHLAN, A. D. (1979). *J. Mol. Biol.* **128**, 49-79.
MCLACHLAN, A. D. (1982). *Acta Cryst.* A**38**, 871-873.

Acta Cryst. (1986). A**42**, 113-116

Temperature Effect of X-ray Diffraction Intensities for a Perfect Crystal across the Absorption Edge: Laue Case

BY M. YOSHIZAWA

Saitama Institute of Technology, Ohsato, Saitama 369-02, *Japan*

T. KAWAMURA

Department of Physics, Yamanashi University, Kofu, Yamanashi 400, *Japan*

T. FUKAMACHI

Saitama Institute of Technology, Ohsato, Saitama 369-02, *Japan*

AND K. HAYAKAWA

Advanced Research Laboratory, Hitachi Ltd, Kokubunji, Tokyo 185, *Japan*

(*Received* 8 *August* 1985; *accepted* 7 *October* 1985)

Abstract

By using a Ge perfect crystal, the 422 X-ray reflection intensities have been measured across the Ge K absorption edge in the Laue case. A conspicuous temperature dependence has been observed in the ratio of the integrated diffraction intensities below and above the absorption edge, which is peculiar to an absorbing perfect crystal. This temperature dependence is much larger in the Laue case than in the Bragg case.

Introduction

When the integrated reflection intensities are measured as a function of X-ray energy, the ratio of the intensities below and above the absorption edge changes when the temperature of a perfect crystal is varied. The temperature dependence (we abbreviate it as 'the temperature effect' hereafter), which is not expected for a mosaic crystal, has been studied theoretically by Kawamura & Fukamachi (1979) and experimentally by Fukamachi, Kawamura, Hayakawa, Nakano & Koh (1982). Kawamura & Fukamachi (1979) have studied the temperature effect theoretically by using the $\bar{5}\bar{5}\bar{5}$ reflection from a GaAs(111) perfect crystal in the Bragg case as a model across the K absorption edge of Ga. Fukamachi *et al.* (1982) have measured the 777 reflection intensities from InSb(111) across the In K absorption edge for both a perfect and a mosaic crystal, and confirmed the temperature effect. In these two cases, the reflection intensities in the Bragg case have been studied, although a similar or larger temperature effect may be expected in the Laue case.

In the present paper, we report on the temperature effect in the Laue case by measuring 422 reflection intensities from a Ge(111) perfect-crystal face. The integrated reflection intensities are measured by energy-dispersive X-ray diffractometry with a solid-state detector (SSD) (Fukamachi, Hosoya &

0108-7673/86/020113-04$01.50