

anomalous scatterer, $|F_{ij}^{\pm}|$ is uniquely determined by $|F_i^N|$, $|F_i^L|$, $\varphi_i^N - \varphi_i^L$, f_j' and f_j'' , but not so, however, by g_i , ε_i , f_j' and f_j'' . This suggests that the phase-determining information in the former procedure is more powerful than in the latter. (iii) Triplet invariants of which the contributing reflections have small g values cannot be estimated in Karle's procedure. However, in our procedure these invariants can also be estimated since for these reflections $F^N \approx F^L$ and $\nu \approx 0$, *i.e.* (1) can still be used.

Instead of using $|F_i^N|$, $|F_i^L|$ and $\varphi_i^N - \varphi_i^L$ it is possible to use $|F_i^N|$, $|F_i^H|$ and $\varphi_i^N - \varphi_i^H$ or $|F_i^L|$, $|F_i^H|$ and $\varphi_i^L - \varphi_i^H$ as conditional information (after normalization of the magnitudes). With each of these choices, $|F_{ij}^{\pm}|$ is determined if there is one type of anomalous scatterer. Hence with one type of anomalous scatterer the information contained in $|F_{ij}^{\pm}|$ is also present in each of the three choices mentioned. The choice for $|F_i^N|$, $|F_i^L|$ and $\varphi_i^N - \varphi_i^L$ was motivated by the fact that it resembles most closely the conditional information used in Hauptman's (1982*a*) distribution for single isomorphous replacement from which our equation (1) was derived.

In protein crystallography it is customary to try to solve the heavy-atom structure first from isomorphous replacement data, often supplemented by anomalous dispersion. With multi-wavelength data it is possible

to solve the heavy-atom structure by standard direct or Patterson methods using the calculated heavy-atom magnitudes g . Subsequently, with known heavy-atom positions, protein structure-factor phases can be determined except for reflections with small heavy-atom contributions. The procedure proposed in the present paper allows accurate estimation of triplet invariants and does not require the positions of the heavy atoms. The derivation of structure-factor phases from these triplet phase invariants will be the subject of a forthcoming paper.

References

- GIACOVAZZO, C. (1983). *Acta Cryst.* **A39**, 585-592.
 HAUPTMAN, H. (1982*a*). *Acta Cryst.* **A38**, 289-294.
 HAUPTMAN, H. (1982*b*). *Acta Cryst.* **A38**, 632-641.
 HAUPTMAN, H. (1983). Private communication.
 KARLE, J. (1980). *Int. J. Quantum Chem.* **7**, 357-367.
 KARLE, J. (1984). *Acta Cryst.* **A40**, 526-531.
 KLOP, E. A., KRABBENDAM, H. & KROON, J. (1987). *Acta Cryst.* **A43**, 810-820.
 MATTHEWS, B. W. & CZERWINSKI, E. W. (1975). *Acta Cryst.* **A31**, 480-487.
 PHILLIPS, J. C., WLODAWER, A., GOODFELLOW, J. M., WATENPAUGH, K. D., SIEKER, L. C., JENSEN, L. H. & HODGSON, K. O. (1977). *Acta Cryst.* **A33**, 445-455.
 SIEKER, L. C., ADMAN, E. T. & JENSEN, L. H. (1972). *Nature (London)*, **235**, 40-42.
 SINGH, A. K. & RAMASESHAN, S. (1968). *Acta Cryst.* **B24**, 35-40.

SHORT COMMUNICATIONS

Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 1000 words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible.

Acta Cryst. (1989). **A45**, 208-210

On the orthogonal transformation used for structural comparisons. By SIMON K. KEARSLEY, *Department of Chemistry, Yale University, 225 Prospect Street, New Haven, CT 06511, USA*

(Received 5 May 1988; accepted 11 August 1988)

Abstract

Rotation matrices that minimize or maximize the sum of the squared distances between corresponding atoms for two structures are found using a constrained least-squares procedure solved analytically as an eigenvalue problem in quaternion parameters.

Orthogonal transformations have been used widely to compare molecular structures (Cox, 1967; McLachlan, 1972, 1979; Rao & Rossmann, 1973; Nyburg, 1974; Hendrickson, 1979; Kenknight, 1984; Honzatko, 1986; Lesk, 1986). Such a transformation superimposes structures to facilitate visual comparisons and to give a quantitative measure of shape similarity as the root mean square (r.m.s.) deviation of

distances between corresponding atoms. A prerequisite for the comparison is that the atom-atom correspondence has been determined between molecules. Then a criterion for the best superposition of two molecular structures is that the sum of the squared distances between these atoms be a minimum. Finding the optimal orthogonal transformation requires determination of a rotation matrix R and a translation vector that will superimpose two sets of coordinates to meet this criterion. To obtain optimal overlap with respect to the translation vector the two coordinate sets should first have their centroids moved to the origin. The problem of finding the rotation matrix, the elements of which are not linearly independent, invariably involves some iterative optimization procedure, although Kabsch (1976, 1978) solves for R by a direct method using Lagrange multipliers to impose orthogonality constraints on the transformation

and Diamond (1988) with a method related to the present work.

Recently Mackay (1984) proposed that quaternion algebra would make it possible to determine R by a linear least-squares procedure. However, applying Mackay's algorithm to different relative orientations between structures gives discrepancies in the final r.m.s. deviation of the atoms. A simple remedy for this problem is to re-apply the algorithm to the rotated coordinates until convergence is achieved; this concatenation of rotations will give the best orthogonal transformation.

In this communication we show how the non-linear least-squares problem can be solved directly by means of an eigenvalue determination using quaternion algebra. Practical application of the method has several advantages. (i) It requires only the construction and subsequent diagonalization of a 4×4 symmetric matrix. (ii) Proper rotations are always produced. (iii) Unlike the methods of McLachlan (1972, 1979) and Kabsch (1976, 1978) there are no special cases which must be considered separately.

A quaternion can be considered as a row matrix of four numbers or the combination of a scalar with a 3D Cartesian vector; $Q = (q_1, q_2, q_3, q_4) \equiv [q_1, \mathbf{q}]$. The product of two quaternions P and Q is a quaternion and can be expressed by matrix multiplication or through a combination of vector products:

$$PQ = \begin{pmatrix} p_1 & -p_2 & -p_3 & -p_4 \\ p_2 & p_1 & -p_4 & p_3 \\ p_3 & p_4 & p_1 & -p_2 \\ p_4 & -p_3 & p_2 & p_1 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} \\ \equiv [p_1 q_1 - \mathbf{p} \cdot \mathbf{q}, p_1 \mathbf{q} + q_1 \mathbf{p} + \mathbf{p} \wedge \mathbf{q}].$$

The norm squared of Q , $|Q|^2$, is given by $(q_1^2 + \mathbf{q} \cdot \mathbf{q})$ and the inverse by $[q_1, -\mathbf{q}]/|Q|^2$. Also, the norm of the product of two quaternions is equal to the product of their individual norms, $|PQ| = |P||Q|$. A unit quaternion \hat{Q} (unit norm) can be used as a rotation operator to move \mathbf{x} to a new position \mathbf{x}^R as follows, where vectors are treated as quaternions with zero scalar component.

$$[0, \mathbf{x}^R] = \hat{Q}^{-1}[0, \mathbf{x}]\hat{Q} = [0, q_1^2 \mathbf{x} + (\mathbf{x} \cdot \mathbf{q})\mathbf{q} + 2q_1(\mathbf{x} \wedge \mathbf{q})]$$

$$\mathbf{x}^R = R\mathbf{x}$$

$$= \begin{pmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2(q_2 q_3 + q_1 q_4) & 2(q_2 q_4 - q_1 q_3) \\ 2(q_2 q_3 - q_1 q_4) & q_1^2 + q_3^2 - q_2^2 - q_4^2 & 2(q_3 q_4 + q_1 q_2) \\ 2(q_2 q_4 + q_1 q_3) & 2(q_3 q_4 - q_1 q_2) & q_1^2 + q_4^2 - q_2^2 - q_3^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Factorizing \mathbf{x} from the vector part of the $[0, \mathbf{x}^R]$ quaternion we arrive at the rotation matrix R in terms of quaternion components. Either \hat{Q} or R can be used for rotation. An advantage of using the quaternion formulation is that improper rotations are never generated since the determinant of R , which reduces to $(q_1^2 + q_2^2 + q_3^2 + q_4^2)^3$, is always greater than zero regardless of values taken by the quaternion components.

Mackay's method uses a unit quaternion to relate corresponding normalized position vectors of the atoms. Therefore the least-squares procedure does not explicitly minimize the sum of the squared distances between atoms but rather the sum of the angles between position vectors. These procedures can be shown to be equivalent if the appropriate

weighting scheme is used. Instead, we relate two vectors of arbitrary length \mathbf{x}' and \mathbf{x} using a general quaternion Q . Since the norm of Q is not restricted to unity, a quaternion can be found that will rotate and isotropically dilate \mathbf{x} to match \mathbf{x}' . When comparing structures Q will not exactly match all the relatable atomic position vectors so in general there will be a residual quaternion $[0, \mathbf{e}]$.

$$[0, \mathbf{e}] = [0, \mathbf{x}'] - Q^{-1}[0, \mathbf{x}]Q.$$

Q is determined by using all corresponding atomic position vectors and minimizing the sum of the squared magnitudes of these residual vectors $\sum |\mathbf{e}|^2 = \sum |[0, \mathbf{e}]|^2$. To make the problem tractable algebraically, first multiply through by Q ,

$$Q[0, \mathbf{e}] = Q[0, \mathbf{x}'] - QQ^{-1}[0, \mathbf{x}]Q = Q[0, \mathbf{x}'] - [0, \mathbf{x}]Q$$

then construct a slightly modified least-squares residual function ε .

$$\varepsilon = \sum |Q[0, \mathbf{e}]|^2 = \sum |Q|^2 |\mathbf{e}|^2 = |Q|^2 \sum |\mathbf{e}|^2 \\ = \sum |[-\mathbf{q} \cdot (\mathbf{x}' - \mathbf{x}), q_1(\mathbf{x}' - \mathbf{x}) + \mathbf{q} \wedge (\mathbf{x}' + \mathbf{x})]|^2.$$

To prevent the transformed structure from dilating, the norm of Q is now constrained to unity using the Lagrange multiplier procedure; this is reminiscent of the method of determining the best plane through a set of points (Scheringer, 1971). Since the norm of Q is constrained to unity, then the solutions for ε reduce the unmodified residual function, $\sum |\mathbf{e}|^2$. The residual ε in terms of the quaternion components is given by

$$\varepsilon = \sum \{ [q_2(x' - x) + q_3(y' - y) + q_4(z' - z)]^2 \\ + [q_1(x' - x) + q_3(z' + z) - q_4(y' + y)]^2 \\ + [q_1(y' - y) + q_4(x' + x) - q_2(z' + z)]^2 \\ + [q_1(z' - z) + q_2(y' + y) - q_3(x' + x)]^2 \\ + \lambda(1 - q_1^2 - q_2^2 - q_3^2 - q_4^2) \}$$

where λ is the Lagrange multiplier. The non-linear aspect of the least-squares refinement has been removed and concealed in the λ multiplier. Differentiating ε with respect to each Q component and setting to zero, we find that the resulting equations can be organized as an eigenvalue problem where $x_m = (x' - x)$, $x_p = (x' + x)$ with similar definitions for y_m, y_p, z_m and z_p .

$$\begin{pmatrix} \sum (x_m^2 + y_m^2 + z_m^2) & \sum (y_p z_m - y_m z_p) & \sum (x_m z_p - x_p z_m) \\ \sum (y_p z_m - y_m z_p) & \sum (y_p^2 + z_p^2 + x_m^2) & \sum (x_m y_m - x_p y_p) \\ \sum (x_m z_p - x_p z_m) & \sum (x_m y_m - x_p y_p) & \sum (x_p^2 + z_p^2 + y_m^2) \\ \sum (x_p y_m - x_m y_p) & \sum (x_m z_m - x_p z_p) & \sum (y_m z_m - y_p z_p) \end{pmatrix}$$

$$\begin{pmatrix} \sum (x_p y_m - x_m y_p) \\ \sum (x_m z_m - x_p z_p) \\ \sum (y_m z_m - y_p z_p) \\ \sum (x_p^2 + y_p^2 + z_m^2) \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} = \lambda \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix}.$$

Diagonalizing this symmetric matrix will give four orthogonal unit quaternions. The eigenvalues give the value of the residual for the rotation produced by application of the corresponding eigenvector. The r.m.s. deviation is given by $(\lambda/n)^{1/2}$ where n is the number of atoms compared. Smallest and largest eigenvalues give rotations that minimize and maximize the sum of the distances between all corresponding atoms.

References

- COX, J. M. (1967). *J. Mol. Biol.* **28**, 151-156.
 DIAMOND, R. (1988). *Acta Cryst.* **A44**, 211-216.
 HENDRICKSON, W. A. (1979). *Acta Cryst.* **A35**, 158-163.
 HONZATKO, R. B. (1986). *Acta Cryst.* **A42**, 172-178.
 KABSCH, W. (1976). *Acta Cryst.* **A32**, 922-923.
 KABSCH, W. (1978). *Acta Cryst.* **A34**, 827-828.
 KENKNIGHT, C. E. (1984). *Acta Cryst.* **A40**, 708-712.
 LESK, A. M. (1986). *Acta Cryst.* **A42**, 110-113.
 MACKAY, A. L. (1984). *Acta Cryst.* **A40**, 165-166.
 McLACHLAN, A. D. (1972). *Acta Cryst.* **A28**, 656-657.
 McLACHLAN, A. D. (1979). *J. Mol. Biol.* **128**, 49-79.
 NYBURG, S. C. (1974). *Acta Cryst.* **B30**, 251-253.
 RAO, S. T. & ROSSMANN, M. G. (1973). *J. Mol. Biol.* **76**, 241-256.
 SCHERINGER, C. (1971). *Acta Cryst.* **B27**, 1470-1472.

Acta Cryst. (1989). **A45**, 210

Space groups rare for organic structures. Erratum. By A. J. C. WILSON, *Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England.*

(Received 28 November 1988)

Abstract

There is an error in the Appendix of the paper by Wilson [*Acta Cryst.* (1988), **A44**, 715-724]. Equation (A3)

should read

$$\langle p \rangle = 1/(N+2) \approx N^{-1}. \quad (A3)$$

All relevant information is given in the *Abstract*.

International Union of Crystallography

Acta Cryst. (1989). **A45**, 210

Nominations for the Ewald Prize

The International Union of Crystallography is pleased to invite nominations for the Ewald Prize for outstanding contributions to the science of crystallography. The Prize is named after Professor Paul P. Ewald, in recognition of his significant contributions to the foundations of crystallography and to the founding of the International Union of Crystallography. Professor Ewald was the President of the Provisional International Crystallographic Committee from 1946 to 1948, the first Editor of the Union's publication *Acta Crystallographica* from 1948 to 1959 and the President of the Union from 1960 to 1963.

The Prize consists of a medal, a certificate and a financial award. It is presented once every three years during the triennial International Congresses of Crystallography. The first Prize was presented at the XIV Congress in Perth, Australia, in 1987. The second Prize, for which nominations are now being invited, will be presented at the XV Congress in Bordeaux, France, in July 1990.

Scientists who have made contributions of exceptional distinction to the science of crystallography are eligible for the Ewald Prize, irrespective of nationality, age or experience. The only exceptions are the current members of the Prize Selection Committee and the President of the Union, none of whom are eligible. No restrictions are placed on the time or the means of publication of the nominee's contributions. The Prize may be shared by more than one contributor to the same scientific achievement.

Nominations for the Ewald Prize should be submitted in writing, preferably using the Ewald Prize Nomination Form and accompanied by supporting documentation, to the Executive Secretary of the International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England, from whom copies of the Nomination Form, the names of the Selection Committee and advice on the submission of nominations may be obtained. **The closing date for nominations is 31 August 1989.**

M. NARDELLI
President

A. I. HORDVIK
General Secretary

Notes and News

Acta Cryst. (1989). **A45**, 210-211

European Microbeam Analysis Society

The European Microbeam Analysis Society (E-MAS) was founded in 1987 as a scientific society focusing on ultra-structural analysis methodology, with primary interests in education, communication and innovation. The Society summarizes its aims and scope as follows:

The Society has been founded to meet the growing need and demands of microbeam analysis users and scientists for further education, communication and counselling. The Society is independent but wants to cooperate in a synergic way with national groups and European societies with related interests. The Society applies to scientists and technicians active in the development and application of micro-