similar features. Both are triclinic with two substituted pyrimidine rings per unit cell. The significance tests have shown that their mean molecular dimensions are identical. Both are in the triketo configuration with an equal number of single and double bonds. Perhaps the most striking feature is the similarity of the puckering of the ring; the displacements of the atoms being in the same sense. Even in their physiological actions both are diabetogenic.

The hydrogen-bonded network, which provides the attractive forces holding the molecules together in a compact arrangement, gives rise to comparatively high-density crystals and shorter than normal van der Waals approaches. Alloxan has two distinct types of hydrogen bonds whereas alloxantin has five. The hydrogen-bonded layers in alloxan, separated by $\frac{1}{2}c$, are interlinked by hydrogen bonds, whereas in alloxantin such layers, separated by $\frac{1}{2}b$, are linked only through the C(5)–C′(5) bond.

There is, however, a fundamental difference in the substituents at position 5. In alloxan there is a *gem*-dihydroxy group, whereas in alloxantin there is only one hydroxyl group. Alloxantin has two molecules of water in its structure, while alloxan has none.

The author is grateful to Professor Sir Nevill Mott, F.R.S. and Dr W. H. Taylor for the provision of research facilities and for interest and encouragement; to Dr W. Cochran, F.R.S., and Dr W. H. Taylor, my supervisors, for interest, advice and helpful criticism. Computations were carried out on EDSAC II by kind permission of Dr M. V. Wilkes, F.R.S., and Mr E. N. Mutch. My thanks are due to Dr M. Wells for kindly making available his crystallographic programs and to the staff of the Crystallographic Laboratory, Cavendish Laboratory for assistance.

The author wishes to acknowledge with gratitude the award of the 1957 Queen's Scholarship by the Federation of Malaya Government during the tenure of which this work was performed.

References

Chatar Singh (1965*a*). *Acta Cryst.* **19**, 759.
Chatar Singh (1965*b*). *Acta Cryst.* In the press.
Cochran, W., Srinivasan, R. & Tollin, P. (1963). *Crystallography and Crystal Perfection.* London: Academic Press.
Fuller, W. (1959). *J. Phys. Chem.* **63**, 1705.
Howells, E. R., Phillips, D. C. & Rogers, D. (1950). *Acta Cryst.* **3**, 210.
Maubasher, R. & Othman, A. M. (1950). *J. Amer. Chem. Soc.* **72**, 2667.
McDonald, T. R. R. (1956). *Acta Cryst.* **9**, 162.
McDonald, T. R. R. & Beevers, C. A. (1952). *Acta Cryst.* **5**, 654.
Merritt, L. L. & Schroeder, E. D. (1956). *Acta Cryst.* **9**, 801.
Pauling, L. (1960). *The Nature of the Chemical Bond.* 3rd Ed. Ithaca: Cornell Univ. Press.
Piloty, O. & Finckh, K. (1904). *Liebigs Ann.* **333**, 22.
Retinger, J. M. (1917). *J. Amer. Chem. Soc.* **39**, 1059.
Slimmer, M. & Stieglitz, J. (1904). *Amer. Chem. J.* **31**, 677.
Wöhler, F. & Liebig, J. (1838). *Liebigs Ann.* **26**, 241, 292.

Acta Cryst. (1965). **19**, 774

A Theoretical Study of the Least-Squares Refinement of Flexible Long-Chain Molecules, with Special Reference to α-Helical Structures

By Robert Diamond

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

(*Received* 24 *February* 1964 *and in revised form* 13 *January* 1965)

We present the theory of a method of refinement which, like Scheringer's (1963), uses a minimum number of parameters, and also gives reason to believe that by partial refinement, *i.e.* by the use of selected eigenvectors, it may be possible to refine the main features of a chain at a much earlier stage in the analysis than is normally possible.

The method starts direct from the atomic coordinates, rather than from a matrix representation of these; it guarantees the integrity of the chains and preserves all the chemical information (bond angles *etc.*) which is implanted in the trial structure.

1. Introduction

Scheringer (1963) has shown clearly the advantages to be gained by refining certain group parameters describing the structure in preference to the x, y, z parameters of each atom. This approach is justified whenever one's foreknowledge of the configuration of various atomic groupings is more accurate than the results to be expected from a conventional refinement, or when the initial coordinates are far enough from the truth to raise doubts about convergence. With large molecules it becomes increasingly difficult to determine x, y and z as independent variables with satisfactory accuracy and the choice of some other parameters becomes imperative; this is partly because the various covalent bond lengths and angles are likely to be known in advance to greater accuracy than the values implied by the independently measured x, y and z, and partly because the reduction in the number of parameters to be handled is highly desirable.

It is our first purpose in this paper to lay the groundwork of a new system of refinement. This system starts from a set of trial coordinates, which must be constructed to be exactly compatible with chemical knowledge, and it preserves the implanted chemical information during the refinement. This is achieved by regarding the angles of rotation about single bonds as the independent variables so that all bond lengths and inter-bond angles are conserved during refinement, without the need to impose constraints, although the molecule as a whole is flexible. It will also be our purpose to exhibit the characteristics of such a scheme and to indicate means of dealing with these. Most of what follows will be related to α-helical peptides for which it is possible to make an analytical study of these characteristics, but the underlying principles are equally relevant to non-helical peptides, or to any other chain.

As is well known, the normal equations of the least-squares problem may be formulated as

$$\mathbf{M}\,\Delta\mathbf{p} \equiv \left[\frac{\partial\overline{\mathbf{F}}^T}{\partial\mathbf{p}}\,\frac{\partial\mathbf{F}}{\partial\mathbf{p}} + \frac{\partial\mathbf{F}^T}{\partial\mathbf{p}}\,\frac{\partial\overline{\mathbf{F}}}{\partial\mathbf{p}}\right]\Delta\mathbf{p} = \frac{\partial\overline{\mathbf{F}}^T}{\partial\mathbf{p}}\,\Delta\mathbf{F} + \frac{\partial\mathbf{F}^T}{\partial\mathbf{p}}\,\Delta\mathbf{F} \qquad (1)$$

in which $\Delta\mathbf{F}$ is a column matrix containing the (complex) quantities $(|F_o|/|F_c|-1)F_c$, $\Delta\mathbf{p}$ is a column matrix containing the (real) shifts in the parameters p which determine the structure and $\partial\mathbf{F}/\partial\mathbf{p}$ is a rectangular matrix for which the number of rows equals the number of observations, and the number of columns equals the number of parameters. The superscript T denotes a transpose and the bar denotes a complex conjugate. The normal matrix $\mathbf{M}$ is real symmetric, being the sum of a Hermitian matrix and its transpose, and it must necessarily be inverted implicitly or explicitly if the matrix $\Delta\mathbf{p}$ is to be found. As will soon become apparent, the matrix $\mathbf{M}$ is expected to be ill-conditioned, and to contain high correlation among the parameters, *i.e.* the matrix $\mathbf{M}$ is nearly singular. Scheringer (1964) has shown that high correlation need not necessarily defeat a least-squares analysis, and Diamond (1956, 1958) satisfactorily solved a highly correlated problem by the methods which are outlined below.

If we wish to solve the normal equations

$$\mathbf{M}\Delta\mathbf{p}=\mathbf{r}$$

for the parametric shifts $\Delta\mathbf{p}$ [here $\mathbf{r}$ represents the right hand side of equation (1)] then ordinarily one writes

$$\Delta\mathbf{p}=\mathbf{M}^{-1}\mathbf{r}$$

and difficulties will be encountered in evaluating $\mathbf{M}^{-1}$ if the determinant of $\mathbf{M}$ is small. To combat this situation we transform our variables to a new set, each of which is a linear combination of the original parameters. These new variables may be constructed to be uncorrelated, and although they may not all be measurable, the elimination of correlation permits the important ones to be measured reliably, and the unimportant and inaccurate ones to be suppressed. This suppression, or filtering, has the twin advantages of effectively reducing the number of parameters to be handled, and of excluding spurious detail from the results. Diamond (1958) describes and illustrates filtering more fully than can be done here, and in Appendix 1 we discuss the connection between correlation coefficients and the rotation of axes to which the transformation corresponds. The algebra of the process, however, is as follows.

If $\mathbf{A}$ is a real orthogonal matrix having the property

$$\mathbf{A}\mathbf{A}^T=\mathbf{I}$$

the identity matrix, then premultiplying the normal equations by $\mathbf{A}^T$ and interposing $\mathbf{A}\mathbf{A}^T$ gives

$$\mathbf{A}^T\mathbf{M}\mathbf{A}\mathbf{A}^T\Delta\mathbf{p}=\mathbf{A}^T\mathbf{r}$$

and

$$\mathbf{A}^T\Delta\mathbf{p}=(\mathbf{A}^T\mathbf{M}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{r}$$

which is easy to calculate if $\mathbf{A}$ is such as to make $\mathbf{A}^T\mathbf{M}\mathbf{A}$ diagonal. In the case of an α-helical structure an exact orthogonal transformation can be found analytically which will yield an $\mathbf{A}^T\mathbf{M}\mathbf{A}$ which is approximately diagonalized and has some large and some very small elements on the diagonal. The elements of $\mathbf{A}^T\Delta\mathbf{p}$ corresponding to large elements in $\mathbf{A}^T\mathbf{M}\mathbf{A}$ may then be determined, whilst we choose not to refine those for which the diagonal elements of $\mathbf{A}^T\mathbf{M}\mathbf{A}$ are small because the agreement between F_o and F_c is little affected by them and their accuracy is very low. The elements of $\mathbf{A}^T\Delta\mathbf{p}$ which have been determined are then premultiplied by $\mathbf{A}$ to express the shifts in terms of the original parameters.

Our second purpose, therefore, is to seek a transformation $\mathbf{A}$ which exactly diagonalizes blocks on the diagonal of an idealized matrix $\mathbf{M}$ and will therefore approximately diagonalize corresponding blocks of an actual $\mathbf{M}$. It will be found that the geometry of an α-helix permits such an $\mathbf{A}$ to be written down. Once $\mathbf{A}$ is obtained, filtering, which is essential with an ill-conditioned problem, becomes possible and the inversion of $\mathbf{M}$ is greatly facilitated. The transformation which is found is almost certainly good enough as it stands to validate the use of the procedure outlined in § 7 (where off-diagonal blocks of $\mathbf{M}$ are considered) when working at low resolution with a helical molecular fragment. At high resolution or in the non-helical case, numerical diagonalization is called for if the best results are to be obtained. In any case, it is considered that the transformation given here illustrates the nature of the problem sufficiently well to be worthy of inclusion, as it shows just what kinds of disturbance the structure is sensitive to, and how these may be found. The various simplifying assumptions made in the derivation of **A** do not in themselves prejudice the values obtained for the shifts $\mathbf{\Delta p}$; they merely enable the crystallographer to work with those linear combinations of parameters to which a helical structure is particularly sensitive, and to ignore those combinations to which it is insensitive and for which no reliable measurements can be made.

The remainder of the paper is made up as follows. In § 2 we define our chosen parameters and evaluate the derivatives $\partial F/\partial p$; in § 3 we estimate a somewhat idealized matrix **M** from these derivatives and a simple representation of an α-helix; in § 4 we find a transformation **A** from this **M** which will also nearly diagonalize a real **M**; in § 5 we give the physical interpretation of the transformation so found, from which, it is hoped, the reader will gain a 'feel' for the nature of the problem; and in the remaining sections we discuss a number of points arising.

2. The selection of parameters and the evaluation of derivatives

Suppose four atoms are covalently linked in a chain and that free rotation is chemically possible about the bond from atom 2 to atom 3. Then atoms 1, 2 and 3 define one plane, 2, 3 and 4 define another, and the angle between these planes will be called p. In this work angles of the type p will be the only variables apart from scale factor, temperature factor and the coordinates and orientation of the end of the chain, although other parameters, such as inter-bond angles, may be included if desired (see § 5). A polypeptide chain contains two such parameters per peptide; one is between C_α and the carbonyl carbon and the other between the nitrogen and the next C_α atom. The bond between N and the carbonyl carbon is not regarded as having a parameter associated with it since the peptide grouping is regarded as planar. A molecule such as myoglobin also has, on the average, two such parameters per side chain, not counting the haem group.

A number of writers, notably Eyring (1932), Mizushima (1954), Nagai & Kobayashi (1962), Hughes & Lauer (1959), and Scheringer (1963) have considered chain molecules, especially helices, in terms of matrices which define the rotational relationship between one link of the chain and the next, or between one link and a link in a standard orientation, and for such a treatment a knowledge of the angles p is necessary since they occur implicitly or explicitly in the matrices. In the present treatment, however, we are not interested in the actual values of these angles, but only in the quantities Δp, the amounts by which these angles alter during refinement. Accordingly the matrices are not found to be necessary and the angles p need never be determined (except in so far as they may be considered interesting in themselves).

Fig. 1 represents schematically a portion of a polypeptide chain in which it is supposed that all the bond lengths and angles have already been made consistent with chemical knowledge, and it is our purpose, by modifying only the angles p, to improve the agreement with X-ray observations. We suppose that the left-hand end of the molecule is anchored, (though the anchor itself will be assigned refinable parameters) and consider what happens when a rotation Δp_k takes place about the kth bond, the vector length of this bond being $\mathbf{r}_k$ as shown. The sense of this rotation is to be considered positive if a right handed screw, when given a positive rotation Δp_k translates in the positive direction of $\mathbf{r}_k$ as indicated by the arrow. We divide the molecule into two parts, a head H_k and a tail T_k, such that everything in T_k rotates about the axis $\mathbf{r}_k$ when Δp_k is applied, the atoms in the head being unaffected. Then for the purposes of evaluating the *partial* derivatives $\partial F/\partial p_k$ we consider the tail to wag and the head to stay still. Vectors such as $\mathbf{r}_k$ will be referred to as spindle vectors.

We identify an atom t in the tail by its position vector $\mathbf{R}_t$ and the difference vector $\mathbf{r}_{kt}$ as shown, and for analytical purposes these vectors will be referred to Cartesian axes with components measured in Å rather than in crystallographic fractional coordinates. The practical use of crystallographic coordinates and the effects of symmetry are considered in Appendix II.

Now

$$F_\alpha = \sum_t^{H+T} f_{t\alpha} \exp 2\pi i \mathbf{R}_t \cdot \mathbf{s}_\alpha ,$$

in which α denotes a reflexion and t an atom, $\mathbf{s}_\alpha$ being the Cartesian reciprocal space vector. If the tail moves, then

$$\delta F_\alpha = \sum_t^{T} \frac{\partial F_\alpha}{\partial \mathbf{R}_t} \cdot \mathbf{\delta R}_t = \sum_t^{T} (\mathbf{grad}_t F_\alpha) \cdot \mathbf{\delta R}_t ,$$

in which $\mathbf{grad}_t F_\alpha$ is a vector in real space whose components are $\partial F_\alpha/\partial x_t$, $\partial F_\alpha/\partial y_t$ and $\partial F_\alpha/\partial z_t$, x, y and z being Cartesian. Now $\mathbf{R}_t$ and $\mathbf{r}_{kt}$ differ by a vector which is constant with respect to p_k, so that

$$\mathbf{\delta R}_t = \delta p_k \mathbf{n}_k \times \mathbf{r}_{kt} , \tag{2}$$

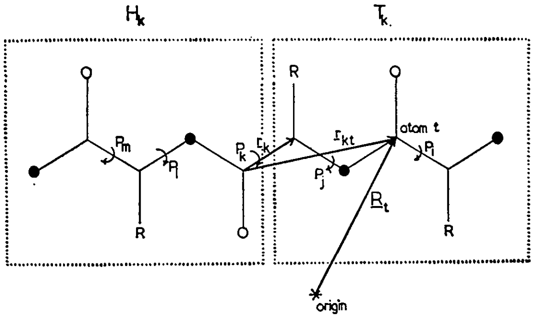

Fig. 1. Indicating the vectors used to describe the movement resulting from a rotation p_k about the bond $\mathbf{r}_k$. The chain represents a polypeptide chain with the solid circles representing nitrogen and the open circles oxygen.

$|\delta p_k|$ being assumed small enough to permit this linear approximation, $\mathbf{n}_k$ being a dimensionless unit vector parallel to $\mathbf{r}_k$, and

$$\frac{\partial F_\alpha}{\partial p_k} = \sum_t^{T_k} (\mathbf{grad}_t\, F_\alpha)\,.\,(\mathbf{n}_k \times \mathbf{r}_{kt})\,.$$

Equation (2), when summed over all parameters in the head, gives the first-order shift of the atom t. Large rotations (such as may be required in a model-building programme) may be expressed by writing

$$\boldsymbol{\delta}\mathbf{R}_t = \sum_1^{\infty} \mathbf{r}_m$$

where

$$\mathbf{r}_m = \frac{1}{m}\,\delta p_k\, \mathbf{n}_k \times \mathbf{r}_{m-1}$$

with

$$\mathbf{r}_0 = \mathbf{r}_{kt}\,.$$

The terms of this vector sum are then analogous to the terms in the series for sine and cosine. The use of this series is rapid for small rotations and accurate for any rotation up to 2π.

Now

$$\mathbf{grad}_t\, F_\alpha = 2\,\pi\, i\, \mathbf{s}_\alpha f_{t\alpha} \exp 2\,\pi\, i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha$$

$$\therefore\ \frac{\partial F_\alpha}{\partial p_k} = 2\,\pi\, i\, \mathbf{s}_\alpha\,.\sum_t^{T_k} (\mathbf{n}_k \times \mathbf{r}_{kt})\, f_{t\alpha} \exp 2\,\pi\, i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha$$

or

$$\frac{\partial F_\alpha}{\partial p_k} = 2\,\pi\, i\, (\mathbf{s}_\alpha \times \mathbf{n}_k)\,.\sum_t^{T_k} \mathbf{r}_{kt}\, f_{t\alpha} \exp 2\,\pi\, i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha$$

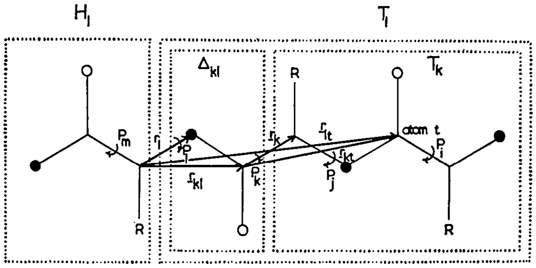

Fig. 2. Showing the vectors used to express the movement resulting from a rotation p_l about the bond $\mathbf{r}_l$ in terms of a rotation p_k about $\mathbf{r}_k$. For a full explanation see text.

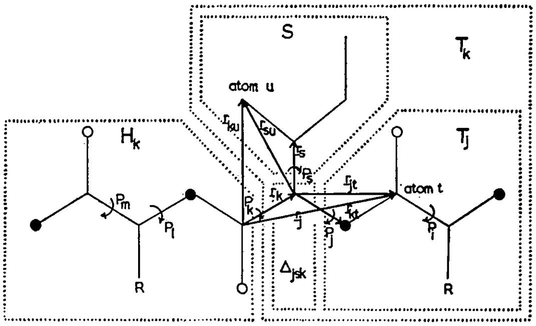

Fig. 3. Showing the vectors used to include side chains in the calculations. The vectors $\mathbf{r}_{sk}$, $\mathbf{r}_{su}$ and $\mathbf{r}_{ku}$ play roles equivalent to $\mathbf{r}_{kl}$, $\mathbf{r}_{kt}$ and $\mathbf{r}_{lt}$ respectively in Fig. 2. In this example $\mathbf{r}_{sk} = \mathbf{r}_{jk} = \mathbf{r}_k$.

i.e.

$$\left.\begin{aligned} \frac{\partial F_\alpha}{\partial p_k} &= 2\,\pi\, i\, (\mathbf{s}_\alpha \times \mathbf{n}_k)\,.\,\mathbf{Q}_{k\alpha} \\ \text{in which}\qquad & \\ \mathbf{Q}_{k\alpha} &= \sum_t^{T_k} \mathbf{r}_{kt}\, f_{t\alpha} \exp 2\,\pi\, i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha\,. \end{aligned}\right\} \quad (3)$$

i.e. $\mathbf{Q}_{k\alpha}$ is a vector in real space, measurable in e.Å, whose magnitude is complex so that it has six components.

Now if p_l is the next parameter, moving towards the head (Fig. 2), and $\mathbf{r}_{kl}$ is the indicated difference vector, then

$$\mathbf{Q}_{l\alpha} = \sum_t^{T_l} \mathbf{r}_{lt}\, f_{t\alpha} \exp 2\,\pi\, i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha$$

$$= \sum_t^{T_k} (\mathbf{r}_{kt} + \mathbf{r}_{kl}) f_{t\alpha} \exp 2\pi i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha + \sum_t^{\Delta_{kl}} \mathbf{r}_{lt} f_{t\alpha} \exp 2\pi i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha$$

in which Δ_{kl} is that part of the molecule which is in T_l but not T_k, *i.e.*

$$\mathbf{Q}_{l\alpha} = \mathbf{Q}_{k\alpha} + \mathbf{r}_{kl}\, F_\alpha(T_k) + \sum_t^{\Delta_{kl}} \mathbf{r}_{lt}\, f_{t\alpha} \exp 2\,\pi\, i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha\,, \qquad (4)$$

where $F_\alpha(T_k)$ is the contribution of T_k to F_α. Thus one $\mathbf{Q}$ may be built upon to provide the next, an essential prerequisite if the $\partial F_\alpha/\partial p_k$ are to be computed rapidly. Note also that all the vectors $\mathbf{r}$ appearing in these expressions are directly obtainable from the coordinates from which we start, and that the quantity $F_\alpha(T_k)$, which must also be carried, becomes F_c when the end of the molecule is reached.

Side chains may also be incorporated into the scheme. Whenever a side chain consisting of more than C_β is encountered, the current value of $\mathbf{Q}$, $\mathbf{Q}_{j\alpha}$ say, is shelved and a fresh start is made at the free end of the side chain. When all the side chain atoms have been included the corresponding $\mathbf{Q}$ for the side chain is $\mathbf{q}_{s\alpha}$ (Fig. 3). The two chains are then united to continue on the main chain by writing

$$\mathbf{Q}_{k\alpha} = \mathbf{Q}_{j\alpha} + \mathbf{q}_{s\alpha} + \mathbf{r}_{jk}\, F_\alpha(T_j) + \mathbf{r}_{sk}\, F_\alpha(s) + \sum_t^{\Delta_{jsk}} \mathbf{r}_{kt}\, f_{t\alpha} \exp 2\,\pi\, i\, \mathbf{R}_t\,.\,\mathbf{s}_\alpha$$

and

$$\frac{\partial F_\alpha}{\partial p_k} = 2\,\pi\, i\, (\mathbf{s}_\alpha \times \mathbf{n}_k)\,.\,\mathbf{Q}_{k\alpha}$$

as before.

The vector $\mathbf{Q}_{k\alpha}$ is, relative to the rotating bond, a weighted mean position vector of all the atoms in T_k, the weights given to the atoms being $f_{t\alpha} \exp 2\pi i \mathbf{R}_t\,.\,\mathbf{s}_\alpha$, *i.e.* their contributions to F_α. To a first approximation, (valid especially at low resolution) we may regard $\mathbf{Q}_{k\alpha}$ as running towards the centroid of the tail. If the tail is extensive, however, it is easy to see that this approximation breaks down, since the weights of some atoms will be negative, depending on $\mathbf{s}_\alpha$. Nevertheless, useful results can be obtained for the α-helix by supposing that the tail is a straight piece of α-helix, so that all the $\mathbf{r}_{kt}$ are approximately parallel to the helix axis whence, for almost any weighting system, both the real and imaginary parts of $\mathbf{Q}$ will be predominantly parallel (or anti-parallel) to the helix axis. This would be exactly the case if the atoms in the tail were strung out along a line, and the assumption that $\mathbf{Q}$ is parallel to the helix axis is equivalent to replacing the tail by a line structure. We shall make this approximation, however, and return in § 5 to consider its consequences, which are surprisingly simple and interesting.

3. Estimation of the normal matrix M

The normal matrix $\mathbf{M}$ [equation (1)] is twice the real part of $(\partial\mathbf{F}^T/\partial\mathbf{p}) \,.\, (\partial\overline{\mathbf{F}}/\partial\mathbf{p})$, likewise the right hand side is twice the real part of $\partial\mathbf{F}^T/\partial\mathbf{p}\overline{\Delta\mathbf{F}}$, and henceforth the factor 2 will be dropped and we consider only real parts denoted by the symbol $\mathcal{R}$. The elements of $\mathbf{M}$ are then given by

$$M_{kl} = \mathcal{R} \sum_\alpha \frac{\partial F_\alpha}{\partial p_k} \frac{\partial \overline{F}_\alpha}{\partial p_l}$$

$$= \mathcal{R} \sum_\alpha 4\pi^2 [(\mathbf{s}_\alpha \times \mathbf{n}_k) \,.\, \mathbf{Q}_{k\alpha}] [(\mathbf{s}_\alpha \times \mathbf{n}_l) \,.\, \overline{\mathbf{Q}}_{l\alpha}] \,. \quad (5)$$

From here until equation (8) we refer only to the k, l element of $\mathbf{M}$, drop the subscripts on the vectors and use subscripts to denote components, thus

s_i is the ith component of $\mathbf{s}_\alpha$
n_i is the ith component of $\mathbf{n}_k$
n'_i is the ith component of $\mathbf{n}_l$
q_i is the real part of the ith component of $\mathbf{Q}_{k\alpha}$
q'_i is the real part of the ith component of $\mathbf{Q}_{l\alpha}$
p_i is the imaginary part of the ith component of $\mathbf{Q}_{k\alpha}$
p'_i is the imaginary part of the ith component of $\mathbf{Q}_{l\alpha}$.

We also introduce the moduli:

$$s = |\mathbf{s}_\alpha| \,, \quad \mathcal{Q}^2 = \sum_i q_i^2 \,, \quad \mathcal{Q}'^2 = \sum_i q_i'^2 \,,$$

$$\mathcal{P}^2 = \sum_i p_i^2 \,, \quad \mathcal{P}'^2 = \sum_i p_i'^2$$

and the direction cosines:

$$f_i = q_i/\mathcal{Q} \,, \quad f'_i = q'_i/\mathcal{Q}' \,, \quad g_i = p_i/\mathcal{P} \,, \quad g'_i = p'_i/\mathcal{P}' \,.$$

In these terms

$$\left.\begin{aligned} M &= 4\pi^2 \sum_\alpha \varepsilon_{ijk} s_i n_j f_k \varepsilon_{lmn} s_l n'_m f'_n \,.\, \mathcal{Q}\mathcal{Q}' \\ &+ 4\pi^2 \sum_\alpha \varepsilon_{ijk} s_i n_j g_k \varepsilon_{lmn} s_l n'_m g'_n \,.\, \mathcal{P}\mathcal{P}' \end{aligned}\right\} \quad (6)$$

in which ε_{ijk} is the alternating tensor and the subscripted variables except s_i are now direction cosines, of which the n are constant during the summation over α, s ranges isotropically over all directions of a sphere, and the f and g are predominantly the direction cosines of the direction of the tail.

If we ignore the minor variations in the direction of $\mathbf{Q}$ as expressed by f and g from one reflexion to another, the first term in equation (6) may then be written as the product of two tensors

$$M = T_{il} U_{il}$$

where

$$T_{il} = 4\pi^2 \varepsilon_{ijk} \varepsilon_{lmn} n_j f_k n'_m f'_n$$

$$U_{il} = \sum_\alpha s_i s_l \mathcal{Q}\mathcal{Q}'$$

plus a similar term with g replacing f and $\mathcal{P}$ replacing $\mathcal{Q}$.

If the distribution of the scalar $\mathcal{Q}\mathcal{Q}'$ in reciprocal space is isotropic, then, for any given pair i and l with $i \neq l$, the distribution of $s_i s_l\, \mathcal{Q}\mathcal{Q}'$ has four quadrants which are similar except for alternation of sign. Summation over α (*i.e.* all quadrants) then gives zero for the expected value of U_{il}, $i \neq l$. For $i = l$ we obtain three similar terms on the diagonal, for which the trace is $\sum_\alpha s^2 \mathcal{Q}\mathcal{Q}'$; hence U_{il} may be represented by

$$U_{il} = \delta_{il} \,.\, \tfrac{1}{3} \sum_\alpha s^2 \mathcal{Q}\mathcal{Q}' \,. \quad (7)$$

Closer inspection of the form of $\mathcal{Q}\mathcal{Q}'$ shows that in fact its distribution is oblate with the unique axis parallel to the axis of the tail. This is because the exponential which occurs in the definition of $\mathbf{Q}$ [equation (3)] alternates most slowly when $\mathbf{s}$ is perpendicular to the tail. If this is taken into account, then it may be shown that the factor $\frac{1}{3}$ in equation (7) should be increased towards a limiting value of $\frac{1}{2}$, without other alteration. Numerical multipliers are inconsequential in what follows, so we shall continue to write $\frac{1}{3}$.

Thus

$$M = T_{il} U_{il}$$

$$= \frac{4\pi^2}{3} \,.\, \varepsilon_{ijk} \varepsilon_{lmn} \delta_{il} n_j f_k n'_m f'_n \,.\, \sum_\alpha s^2 \mathcal{Q}\mathcal{Q}' \,.$$

But

$$\varepsilon_{ijk} \varepsilon_{lmn} \delta_{il} = \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km} \,.$$

$$\therefore\; M = \frac{4\pi^2}{3} (n_j n'_j f_k f'_k - n_j f'_j f_k n'_k) \sum_\alpha s^2 \mathcal{Q}\mathcal{Q}'$$

$$= \frac{4\pi^2}{3} (\cos\mu_1 \cos\mu_3 - \cos\mu_2 \cos\mu_4) \sum_\alpha s^2 (\mathcal{Q}\mathcal{Q}' + \mathcal{P}\mathcal{P}') \quad (8)$$

when both parts of equation (6) are included, where μ_1 is the angle between $\mathbf{n}_k$ and $\mathbf{n}_l$ (reverting to the earlier notation), μ_3 is the angle between the predominant direction of the tail, as seen from parameter k and as seen from parameter l, *i.e.* zero in the helical case.

μ_2 is the angle between $\mathbf{n}_k$ and the tail ,
and μ_4 is the angle between $\mathbf{n}_l$ and the tail.

The derivation of equation (8) from equation (5) is shorter than one found originally by the author, and is based on some notes of Mr G. C. Fox.

If it is supposed that the vectors $\mathcal{Q}$ and $\mathcal{P}$ (*i.e.* q_1, q_2, q_3 and p_1, p_2, p_3) are scattered randomly around the central direction of the tail (with a scatter which increases as s increases) then it may be shown that equation (8) continues to represent a systematic contribution to the element M_{kl} if it is multiplied by $\overline{\cos^2 \varrho}$ where ϱ is the angle between any one $\mathbf{Q}_{k\alpha}$ and the predominant direction of the tail. Consequently, any difference in direction of the individual $\mathcal{Q}$ and $\mathcal{P}$ may be ignored [we have implicitly set $g_k = f_k$ in equation (8)] and $\mathbf{Q}_{k\alpha}$ may be represented as the product of a real unit vector and a complex scalar $Q_{k\alpha}$, and

$$\mathcal{Q}_{k\alpha}\mathcal{Q}_{l\alpha} + \mathcal{P}_{k\alpha}\mathcal{P}_{l\alpha} = \mathcal{R} Q_{k\alpha}\bar{Q}_{l\alpha} .$$

Furthermore, it follows that as resolution increases, the systematic effects of the geometry of the molecule, as expressed by equation (8), become gradually less important and increasingly overlaid by more random effects due to the finer details of the structure.

We now apply equation (8) to an α-helical chain, taking the parameters p_k in the order in which they occur in the molecule, with the object of illustrating the properties of such a system and of finding transformations to aid in its solution.

Fig. 4 shows on a stereogram the directions of 8 consecutive parameter axes in the chain, these being grouped in pairs, one pair per peptide, at angles ω apart, ω being the screw angle which is closely equal to 100° in the α-helix. The centre of the figure represents the helix axis. We also show two angles τ_1 and τ_2 giving the inclination of these bonds to the helix axis. We note that $\mu_3 = 0$, μ_2 and μ_4 take the values of τ_1 and/or τ_2, and μ_1 is the angle between the two rotation parameter axes concerned; thus

$$\cos \mu_1 = \cos \tau_i \cos \tau_j + \sin \tau_i \sin \tau_j \cos n\omega ,$$

where i and j take the values 1 or 2 as appropriate and n is an integer. Thus the first few expected values of M_{ij} are

$$M_{11} = \frac{4\pi^2}{3} (\cos^2 \tau_1 + \sin^2 \tau_1 - \cos^2 \tau_1) \sum_\alpha s_\alpha^2 |Q_{1\alpha}|^2$$

$$M_{12} = \frac{4\pi^2}{3} (\cos \tau_1 \cos \tau_2 + \sin \tau_1 \sin \tau_2 - \cos \tau_1 \cos \tau_2) \; \mathcal{R} \sum_\alpha s_\alpha^2 Q_{1\alpha} \bar{Q}_{2\alpha}$$

$$M_{13} = \frac{4\pi^2}{3} (\cos^2 \tau_1 + \sin^2 \tau_1 \cos \omega - \cos^2 \tau_1) \; \mathcal{R} \sum_\alpha s_\alpha^2 Q_{1\alpha} \bar{Q}_{3\alpha}$$

etc. Hence, defining

$$P_{ij} = \frac{4\pi^2}{3} \sum_\alpha s_\alpha^2 \mathcal{R} Q_{i\alpha} \bar{Q}_{j\alpha} , \tag{9}$$

M takes the form

$$\begin{array}{ll} & J=1 \\ I=1 & \left(\begin{array}{ll} \sin^2 \tau_1 P_{11} & \sin \tau_1 \sin \tau_2 P_{12} \\ \sin \tau_1 \sin \tau_2 P_{21} & \sin^2 \tau_2 P_{22} \\ \sin^2 \tau_1 \cos \omega P_{31} & \sin \tau_1 \sin \tau_2 \cos \omega P_{32} \\ \sin \tau_1 \sin \tau_2 \cos \omega P_{41} & \sin^2 \tau_2 \cos \omega P_{42} \\ \sin^2 \tau_1 \cos 2\omega P_{51} & \sin \tau_1 \sin \tau_2 \cos 2\omega P_{52} \\ \sin \tau_1 \sin \tau_2 \cos 2\omega P_{61} & \sin^2 \tau_2 \cos 2\omega P_{62} \end{array} \right. \\ I=2 & \\ I=3 & \end{array}$$

$$\begin{array}{ll} & J=2 \\ I=1 & \left. \begin{array}{ll} \sin^2 \tau_1 \cos \omega P_{13} & \sin \tau_1 \sin \tau_2 \cos \omega P_{14} \cdots \\ \sin \tau_1 \sin \tau_2 \cos \omega P_{23} & \sin^2 \tau_2 \cos \omega P_{24} \\ \sin^2 \tau_1 P_{33} & \sin \tau_1 \sin \tau_2 P_{34} \\ \sin \tau_1 \sin \tau_2 P_{43} & \sin^2 \tau_2 P_{44} \\ \sin^2 \tau_1 \cos \omega P_{53} & \sin \tau_1 \sin \tau_2 \cos \omega P_{54} \\ \sin \tau_1 \sin \tau_2 \cos \omega P_{63} & \sin^2 \tau_2 \cos \omega P_{64} \end{array} \right) \\ I=2 & \\ I=3 & \end{array}$$

etc. (10)

i.e. **M** consists of partitions each of them 2 × 2. If these partitions are given indices I and J then each partition is the same as the first *vis à vis* τ_1 and τ_2 and contains also a factor $\cos (I-J)\, \omega$.

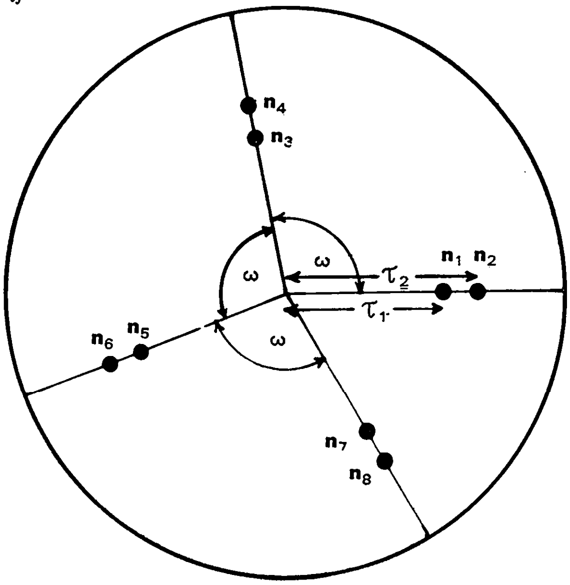

Fig. 4. Stereogram showing the directions of eight consecutive bonds of free rotation in an α-helix.

4. Estimation of eigenvectors

Consider first two parameters p_1 and p_2 which are associated with the two rotatable bonds in a single peptide. These two bonds are very nearly parallel, and are physically close together. Now suppose that Δp_1 and Δp_2 are equal in magnitude and sign, then the effect of applying these rotations is, almost exactly, to rotate the tail through an angle $2\Delta p$ about the mean line of the two bonds. This is a gross disturbance of the structure and the F's will be sensitive to it. If, however, Δp_1 and Δp_2 are made equal and opposite, then the structure is practically undisturbed, there being only second order displacements due to the non-collinearity of the two bonds and displacements of those few atoms which occur between the bonds. Evidently, therefore, we may expect to be able to measure $(\Delta p_1 + \Delta p_2)$ but not $(\Delta p_1 - \Delta p_2)$, and we should therefore seek those linear combinations (such as $\Delta p_1 + \Delta p_2$) which can be measured, and to sort out, and suppress if necessary, those linear combinations of parameters (such as $\Delta p_1 - \Delta p_2$) which cannot be measured. Δp_1 and Δp_2 are examples of correlated variables; a value cannot be assigned to Δp_1 alone unless Δp_2 is reliably known (because only their sum is measurable), but the quantities $(\Delta p_1 + \Delta p_2)$ and $(\Delta p_1 - \Delta p_2)$ are uncorrelated, the former being measurable despite the uncertainty in the latter.

$(\Delta p_1 + \Delta p_2)$ and $(\Delta p_1 - \Delta p_2)$ are the eigenvectors of this two-parameter system and their eigenvalues are respectively large and vanishingly small, this being the mathematical expression of the above considerations.

In general, there are n eigenvectors for a problem having n parameters. Myoglobin, for example, has nearly 600 such parameters, and, ideally, one would determine all 600 eigenvectors – a job which can only be done numerically, and it is clearly out of the question to attempt this on grounds of time, costs and machine size.

It is therefore necessary to seek a compromise technique and for this purpose we consider the normal matrix for a helical fragment and partition it into blocks such that each block on the diagonal is a square accounting for N peptides (*i.e.* order $2N$). We then find a perfect real orthogonal transformation matrix **A** which will approximately diagonalize each of the diagonal partitions in **M**. The off-diagonal partitions are also altered by the transformation, after which they provide the interaction terms between the diagonal elements in one diagonal block with those in another. As will be seen later, the significant interaction terms are few in number, and the net effect is to produce, for the whole helical fragment, a band matrix of very much reduced order.

For a chain which is not entirely helical the same approach is valid, but each diagonal block requires its own transformation to be found numerically.

The properties of the matrix (10) are readily studied if we approximate P_{ij} by $\sqrt{P_{ii}P_{jj}}$ within any one diagonal block. The extent to which this holds governs the maximum value of N for which the resulting transformation is valuable. Within this approximation each diagonal block of M becomes a diadic, being the sum of two diads

$$\mathbf{M} = \mathbf{u}\mathbf{u}^T + \mathbf{v}\mathbf{v}^T \tag{11}$$

in which the row vectors $\mathbf{u}^T$ and $\mathbf{v}^T$ are of the form

$$\begin{aligned} \mathbf{u}^T = \cdots, &\sqrt{P_{2n+1,\,2n+1}}\,\sin\tau_1\cos(n\omega+\varphi),\\ &\sqrt{P_{2n+2,\,2n+2}}\,\sin\tau_2\cos(n\omega+\varphi),\cdots\\ \mathbf{v}^T = \cdots, &\sqrt{P_{2n+1,\,2n+1}}\,\sin\tau_1\sin(n\omega+\varphi),\\ &\sqrt{P_{2n+2,\,2n+2}}\,\sin\tau_2\sin(n\omega+\varphi),\cdots \end{aligned} \tag{12}$$

in which φ is an arbitrary constant and the integer n increases by 1 in alternate elements, *i.e.* with each peptide. The vectors **u** and **v** have a simple physical significance: the trigonometrical factors in each component are proportional to the displacement of the tail per radian of displacement in the corresponding parameter, measured in each of two directions at right angles on a plane normal to the length of the tail (see below), and the factors $\sqrt{P}$ represent the sensitivity of the X-ray observations to such displacements. The angle φ represents a reference direction for azimuthal measurements in this plane and may be set to zero without loss of generality.

Now, the eigenvectors, of which there are only two with non-vanishing eigenvalue, must each be of the form $a\mathbf{u}+b\mathbf{v}$, being coplanar with **u** and **v**.

i.e.
$$(\mathbf{u}\mathbf{u}^T + \mathbf{v}\mathbf{v}^T)(a\mathbf{u}+b\mathbf{v}) = \lambda(a\mathbf{u}+b\mathbf{v})$$

where λ is the eigenvalue. Rearranging gives

$$\mathbf{u}(a\mathbf{u}^T\mathbf{u} + b\mathbf{u}^T\mathbf{v} - \lambda a) + \mathbf{v}(a\mathbf{v}^T\mathbf{u} + b\mathbf{v}^T\mathbf{v} - \lambda b) = 0\,,$$

in which the round brackets are both scalars. A linear combination of two non-parallel vectors can only vanish if their coefficients vanish; *i.e.*

$$a\mathbf{u}^T\mathbf{u} + b\mathbf{u}^T\mathbf{v} - \lambda a = a\mathbf{v}^T\mathbf{u} + b\mathbf{v}^T\mathbf{v} - \lambda b = 0\,,$$

whence

$$\frac{a}{b} - \frac{b}{a} = 2A\,, \qquad \frac{a}{b} = A \pm \sqrt{1+A^2} \tag{13}$$

where

$$A = \frac{\mathbf{u}^T\mathbf{u} - \mathbf{v}^T\mathbf{v}}{\mathbf{u}^T\mathbf{v} + \mathbf{v}^T\mathbf{u}}$$

$$= \frac{\sum\limits_{n=0,1\ldots} (P_{2n+1,\,2n+1}\sin^2\tau_1 + P_{2n+2,\,2n+2}\sin^2\tau_2)\cos 2n\omega}{\sum\limits_{n=0,1\ldots} (P_{2n+1,\,2n+1}\sin^2\tau_1 + P_{2n+2,\,2n+2}\sin_2\tau_2)\sin 2n\omega}\,.$$

Defining θ by

$$\theta = \arg \sum_{n=0,1\ldots} (P_{2n+1,\,2n+1}\sin^2\tau_1 + P_{2n+2,\,2n+2}\sin^2\tau_2)\,e^{2in\omega}, \tag{14}$$

then
$$A = \cot\theta \quad a = \cos\theta \pm 1 \quad b = \sin\theta$$

are unnormalized solutions to equations (13). The two choices of sign lead to the two eigenvectors.

If the dependence of the round bracket in (14) on n is ignored (*i.e.* if the 'weight' of the tail is regarded as essentially constant over N peptides) then

$$\theta = \arg \sum_{n=0}^{N-1} e^{2in\omega} = (N-1)\omega;$$

if, on the other hand, we attempt to represent the variation in the weight of the tail within the group of N peptides by writing

$$(P_{2n+1,\,2n+1}\sin^2\tau_1 + P_{2n+2,\,2n+2}\sin^2\tau_2) = c + nd,\ c,\ d\ \text{const}\,,$$

then it may be shown that

$$\theta = (N-1)\omega - \alpha \tag{15}$$

where

$$\alpha = \tan^{-1}\left[\frac{N\cot N\omega - \cot\omega}{2c/d + (N-1)}\right],$$

but for the sake of clarity of illustration we shall ignore this refinement, except to point out that N-peptide groupings near the head end or near the tail end of a helix differ principally in their values of c.

The unnormalized eigenvectors are then

$$
\begin{aligned}
&a\mathbf{u}^T + b\mathbf{v}^T \\
&\quad = [\cos (N-1)\omega \pm 1] \\
&\quad \times (\cdots, \sqrt{P} \sin \tau_1 \cos n\omega, \sqrt{P} \sin \tau_2 \cos n\omega, \cdots) \\
&\quad + \sin (N-1)\omega \\
&\quad \times (\cdots, \sqrt{P} \sin \tau_1 \sin n\omega, \sqrt{P} \sin \tau_2 \sin n\omega, \cdots) \\
&\quad n = 0, 1 \cdots N-1
\end{aligned}
$$

from which the normalized eigenvectors may be shown to be

$$
\cdots, \frac{\sin \tau_1 \cos\left(\frac{N-1}{2} - n\right)\omega}{\sqrt{(\sin^2\tau_1 + \sin^2\tau_2) \sum\limits_{n=0}^{N-1} \cos^2 \left(\frac{N-1}{2} - n\right)\omega}},
$$

$$
\frac{\sin \tau_2 \cos\left(\frac{N-1}{2} - n\right)\omega}{\sqrt{(\sin^2\tau_1 + \sin^2\tau_2) \sum\limits_{n=0}^{N-1} \cos^2 \left(\frac{N-1}{2} - n\right)\omega}}, \cdots
$$

for one vector with sin replacing cos in the other one, the dependence on τ_1 and τ_2 being unaltered.

The eigenvalues may then be obtained from

$$\lambda = \mathbf{e}^T(\mathbf{u}\mathbf{u}^T + \mathbf{v}\mathbf{v}^T)\mathbf{e} = (\mathbf{u}^T\mathbf{e})^2 + (\mathbf{v}^T\mathbf{e})^2$$

where $\mathbf{e}$ is a normalized eigenvector. This leads to

$$\lambda = 2P (\sin^2\tau_1 + \sin^2\tau_2) \sum_{n=0}^{N/2-1} \cos^2 \left(\frac{N-1}{2} - n\right)\omega$$

for N even for the first eigenvalue, with sin replacing cos for the second one. If N is odd we find

$$\lambda = P (\sin^2\tau_1 + \sin^2\tau_2) \left[1 + 2 \sum_{n=0}^{(N-3)/2} \cos^2 \left(\frac{N-1}{2} - n\right)\omega \right]$$

for one eigenvalue, and

$$2P (\sin^2\tau_1 + \sin^2\tau_2) \sum_{n=0}^{(N-3)/2} \sin^2 \left(\frac{N-1}{2} - n\right)\omega$$

for the other.

Now, we take as an example the case $N=4$, giving diagonal blocks in $\mathbf{M}$ of order 8. The foregoing analysis has shown that such a block approximates a matrix of rank 2; there must also be six other eigenvectors which, with the first two, form an orthogonal set, but which have eigenvalues negligible in comparison with the first two. These are not unique, but a suitable set may be found by rearranging components in the two already found, giving the transformation matrix $\mathbf{A}$ below. Here each row relates to a parameter, numbered as in Fig. 4, and the columns are the eigenvectors. The ordering of the columns is arbitrary, and the choice here is such that the ε vector is the first vector found above and the α vector is the second. These vectors will be referred to by the Greek letters, as indicated.

$$
\mathbf{A} = \begin{array}{c|cccc}
 & \alpha & \beta & \gamma & \delta \\
\hline
p_1 & +a_1a_2 & -b_1 & -a_1b_2 & 0 \\
p_2 & +b_1a_2 & +a_1 & -b_1b_2 & 0 \\
p_3 & +a_1b_2 & 0 & +a_1a_2 & -b_1 \\
p_4 & +b_1b_2 & 0 & +b_1a_2 & +a_1 \\
p_5 & -a_1b_2 & 0 & -a_1a_2 & 0 \\
p_6 & -b_1b_2 & 0 & -b_1a_2 & 0 \\
p_7 & -a_1a_2 & 0 & +a_1b_2 & 0 \\
p_8 & -b_1a_2 & 0 & +b_1b_2 & 0 \\
\hline
 & \varepsilon & \zeta & \eta & \theta \\
\hline
p_1 & +a_1b_3 & 0 & +a_1a_3 & 0 \\
p_2 & +b_1b_3 & 0 & +b_1a_3 & 0 \\
p_3 & +a_1a_3 & 0 & -a_1b_3 & 0 \\
p_4 & +b_1a_3 & 0 & -b_1b_3 & 0 \\
p_5 & +a_1a_3 & -b_1 & -a_1b_3 & 0 \\
p_6 & +b_1a_3 & +a_1 & -b_1b_3 & 0 \\
p_7 & +a_1b_3 & 0 & +a_1a_3 & -b_1 \\
p_8 & +b_1b_3 & 0 & +b_1a_3 & +a_1
\end{array} \tag{16}
$$

in which

$$
a_1 = \frac{\sin \tau_1}{\sqrt{\sin^2\tau_1 + \sin^2\tau_2}} \quad a_2 = \frac{\sin 3\omega/2}{\sqrt{2\left(\sin^2 \frac{\omega}{2} + \sin^2 \frac{3\omega}{2}\right)}}
$$

$$
b_1 = \frac{\sin \tau_2}{\sqrt{\sin^2\tau_1 + \sin^2\tau_2}} \quad b_2 = \frac{\sin \omega/2}{\sqrt{2\left(\sin^2 \frac{\omega}{2} + \sin^2 \frac{3\omega}{2}\right)}}
$$

$$
a_3 = \frac{\cos \omega/2}{\sqrt{2\left(\cos^2 \frac{\omega}{2} + \cos^2 \frac{3\omega}{2}\right)}}
$$

$$
b_3 = \frac{\cos 3\omega/2}{\sqrt{2\left(\cos^2 \frac{\omega}{2} + \cos^2 \frac{3\omega}{2}\right)}}.
$$

The transformed matrix

$$\mathbf{\Lambda} = \mathbf{A}^T\mathbf{M}\mathbf{A}$$

within an 8×8 diagonal block is then given approximately by

$$
\begin{aligned}
\mathbf{\Lambda} = \operatorname{diag} \Big(& 2(\sin^2\tau_1 + \sin^2\tau_2) \\
& \times P \left(\sin^2 \frac{\omega}{2} + \sin^2 \frac{3\omega}{2}\right), 0,0,0, 2 (\sin^2\tau_1 + \sin^2\tau_2) \\
& \times P \left(\cos^2 \frac{\omega}{2} + \cos^2 \frac{3\omega}{2}\right), 0,0,0 \Big).
\end{aligned}
$$

5. The interpretation of the eigenvectors

If we insert the approximate values of $\omega = 100°$ and $\tau_1 = \tau_2^*$, $\mathbf{A}$ may be evaluated as

---

* In any real computation τ_1, τ_2 and ω would be given their real values, but for the purpose of discussing the characteristics of the problem the difference of a few degrees between τ_1 and τ_2 may be ignored.

$$
\begin{array}{c|cccccccc|}
 & \alpha & \beta & \gamma & \delta & \varepsilon & \zeta & \eta & \theta \\
p_1 & 0{\cdot}2732 & -0{\cdot}7080 & -0{\cdot}4185 & 0 & -0{\cdot}4022 & 0 & 0{\cdot}2985 & 0 \\
p_2 & 0{\cdot}2732 & 0{\cdot}7080 & -0{\cdot}4185 & 0 & -0{\cdot}4022 & 0 & 0{\cdot}2985 & 0 \\
p_3 & 0{\cdot}4185 & 0 & 0{\cdot}2732 & -0{\cdot}7080 & 0{\cdot}2985 & 0 & 0{\cdot}4022 & 0 \\
p_4 & 0{\cdot}4185 & 0 & 0{\cdot}2732 & 0{\cdot}7080 & 0{\cdot}2985 & 0 & 0{\cdot}4022 & 0 \\
p_5 & -0{\cdot}4185 & 0 & -0{\cdot}2732 & 0 & 0{\cdot}2985 & -0{\cdot}7080 & 0{\cdot}4022 & 0 \\
p_6 & -0{\cdot}4185 & 0 & -0{\cdot}2732 & 0 & 0{\cdot}2985 & 0{\cdot}7080 & 0{\cdot}4022 & 0 \\
p_7 & -0{\cdot}2732 & 0 & 0{\cdot}4185 & 0 & -0{\cdot}4022 & 0 & 0{\cdot}2985 & -0{\cdot}7080 \\
p_8 & -0{\cdot}2732 & 0 & 0{\cdot}4185 & 0 & -0{\cdot}4022 & 0 & 0{\cdot}2985 & 0{\cdot}7080
\end{array}
$$

The combination of rotations $+0{\cdot}2732°$ about $\mathbf{n}_1$, $+0{\cdot}2732°$ about $\mathbf{n}_2$, *etc.*, as given in the α column, gives a combination of displacements which may be measured with an error proportional to $1/\sqrt{\lambda_\alpha}$ where λ_α is the α eigenvalue, which appears in the 11 position in $\mathbf{\Lambda}$, *i.e.* it is measurable. $\mathbf{\Lambda}$ also indicates that the ε combination is also measurable, slightly more accurately, but that a virtually infinite error is associated with the measurement of the remaining six combinations.

It was indicated at the beginning of § 4 that we could not expect to measure $(\Delta p_1 - \Delta p_2)$ *etc.*, and this fact is expressed in the above formulation by the vanishing eigenvalues associated with the β, δ, ζ and θ combinations. Thus any experimentally determined combination of rotations will contain an indeterminate contribution from these four eigenvectors unless we filter them out.

We next consider why it is that the γ and η eigenvalues are vanishing whilst the α and ε ones are not, and from this it will be clear why there are only two non-vanishing ones, and the consequences of our linear tail approximation (*i.e.* all Q's parallel to the axis) and of our neglecting of differences among the P_{ij} will become clearer.

In Fig. 5 we construct figures illustrating the movement of a point on the helix axis somewhere in the tail. Starting from O in the α diagram the first arrow represents the linear displacement arising from a rotation of $+0{\cdot}273°$ about $\mathbf{n}_1$ and about $\mathbf{n}_2$, the next arrow represents the displacement due to a rotation of $+0{\cdot}418°$ about $\mathbf{n}_3$ and $\mathbf{n}_4$, these axes being 100° round from the first, the figure representing the trace of the helix axis on a plane normal to it. We continue in this way to find the resulting displacements, from which it is clear that the α and ε combinations represent tilts of the tail about the 4-peptide grouping, in two planes which are at right angles, the magnitudes of these tilts being proportional to the square roots of the corresponding eigenvalues. When we have thus accounted for two tilts in planes at right angles to one another there remains no other possible tilt of the tail which could be mathematically orthogonal (and therefore also geometrically orthogonal) to the first two. It is for this reason that the γ and η combinations are self-cancelling with respect to tilts of the tail and have correspondingly vanishing eigenvalues. These considerations indicate that whatever number of peptides be included in the group, we shall never expect to get more than two large eigenvalues. This has its value in that the configuration of an α-helix may be refined *at an early stage*, with the use of only these two largest eigenvectors per group, which is sufficient to allow the helix as a whole to wriggle a little without allowing the situation to become confused by insignificant movements within each group, and for this purpose groups larger than four peptides may be worth while.

In the above derivations we have supposed, for the purpose of estimating $\mathbf{A}$, that the tail is a line structure, so that the direction of each $\mathbf{Q}$ is along the helix axis. This being so, the above tilts are the only meaningful rotations of the tail about axes in the groups of peptides. As soon as the tail is regarded as having thickness, however, rotations about the helix axis become meaningful. Since all the movements are compounded of rotations, and since the α and ε tilts can be regarded as rotations about mutually perpendicular axes normal to the helix axis, it follows that there can be *only one* other rotation which is mathematically (and geometrically) orthogonal to these two, this being about the helix axis. This means that if we had made allowance for the finite thickness of the tail, then we should have found one and only one combination of rotations equivalent to a rotation of the tail part of the helix about its own axis, but we have not done this, and the resulting combinations are therefore not orthogonalized with respect to this type of rotation. It follows that we may

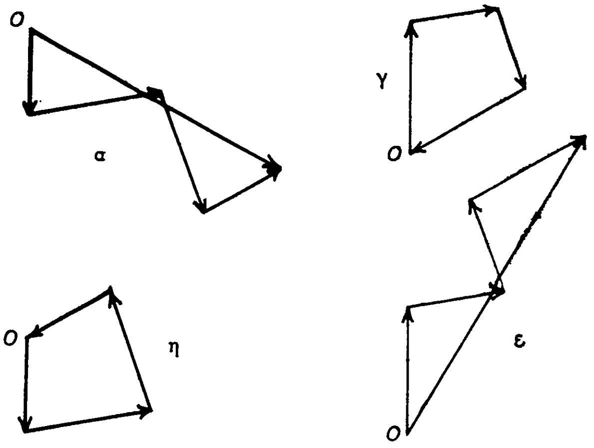

Fig. 5. These figures represent the trace of the helix axis on a plane normal to it at a point somewhere in the tail when certain combinations of rotations occur in a group of four peptides. The α and ε combinations of rotations result in tilts of the helix in two planes at right angles, whereas the γ and η combinations of rotations are self cancelling and leave the helix straight (if it was originally straight).

expect any of the combinations we have found to include some twisting, and it turns out that the ε and a combinations do, to different extents, and that the others are all free of it. Thus by using the idea of a linear tail, we have not been prevented from finding eigenvectors which involve rotations about the helix axis; we have simply been prevented from estimating their eigenvalues correctly. The zero value which we have found for the η eigenvalue corresponds to the fact that a rotation of a line about its length is undetectable.

If an actual numerical matrix **M** is computed from the coordinates of a structure, however, it will contain information dependent on the tail thickness. Therefore if such a genuine matrix is subjected to the transformation **A** it must yield a small but positive element on the diagonal in the η position.

For the purpose of estimating the movement of the tail some distance from the peptide grouping, it is sufficient to regard the rotation axes in the group as being grouped together in a region of negligible volume, *i.e.* as being concurrent. However, no amount of rotation about concurrent axes can produce an elongation or contraction of the helix along its length. In order to estimate this effect it is necessary to recognize the finite spatial extent of the group of peptides. When this is done, it is found that the elongation produced by any combination of rotations is proportional to the radius of the helix, to sin τ, and to the sum of the components of the combination concerned. Thus the ratios of the elongations produced by the α, γ, ε and η eigenvectors are 0:0: −0·1037: +0·7007. Thus the η vector, in addition to rotating the tail about the helix axis, also alters the helix length. Since the ε and η vectors both affect length it is clear that they are not mathematically orthogonal (*i.e.* they remain slightly correlated). The matrix algebra given above has failed to orthogonalize these vectors with respect to elongation, because we have ignored the radius of the helix and supposed all the **Q** vectors to be parallel, whereas peptides on opposite sides of the helix must have **Q** vectors whose directions differ by an amount depending on the diameter of the helix. If this effect were estimated it could be included in the analysis by modification of the values for the angles μ substituted in equation (8). Here we have, in effect, replaced the group of peptides by a structureless pivot, rotations about which cannot produce elongation.

We summarize these results in Table 1, where the entries are proportional to the diplacements concerned.

Table 1 reveals one adverse characteristic of the system, however. It shows that the ratio of twist to dilation is the same for all the combinations, so that it is not possible to produce a dilation *independently* of a twist, and *vice versa*, though both these movements may be made independent of tilting. The reason for this is that in this scheme inter-bond angles have been held constant at their initial values, only swivelling movements being permitted because these are the chemically permissible movements, and the treatment rests on the assumption that the starting coordinates have the correct local stereochemistry. However, the deficiency may be made up, if necessary, by introducing spindle vectors **n** which, instead of being aligned along the freely rotating bonds, are situated at C_α atoms and normal to the C–C_α–N plane. Parameters associated with these rotation axes would then decouple twisting from dilation by allowing variations in the inter-bond angles at C_α to take place. In the present scheme, however, it is apparent that with groupings of four peptides it is only likely to be worthwhile refining the α, ε and η vectors, *i.e.* three quarters of a parameter per peptide.

Table 1. *Summary of results*

Combination	Tilt (*i.e.* bending of the helix about the 4-peptide group)	Twist (*i.e.* rotation of tail part of helix about its axis)	Dilation (*i.e.* linear expansion parallel to helix length) (ϱ = radius of helix)
α	In plane containing helix axis and bisector of acute angle between $\mathbf{n}_1$ and $\mathbf{n}_8$ $\sqrt{2P(\sin^2\tau_1+\sin^2\tau_2)\left(\sin^2\frac{\omega}{2}+\sin^2\frac{3\omega}{2}\right)}$	0	0
γ	0	0	0
ε	in plane $\perp$ to α tilt $\sqrt{2P(\sin^2\tau_1+\sin^2\tau_2)\left(\cos^2\frac{\omega}{2}+\cos^2\frac{3\omega}{2}\right)}$	$\dfrac{\cos\tau\left(\cos\frac{\omega}{2}+\cos\frac{3\omega}{2}\right)}{\sqrt{\cos^2\frac{\omega}{2}+\cos^2\frac{3\omega}{2}}}$	$\dfrac{\varrho\sin\tau\left(\cos\frac{\omega}{2}+\cos\frac{3\omega}{2}\right)}{\sqrt{\cos^2\frac{\omega}{2}+\cos^2\frac{3\omega}{2}}}$
η	0	$\dfrac{\cos\tau\left(\cos\frac{\omega}{2}-\cos\frac{3\omega}{2}\right)}{\sqrt{\cos^2\frac{\omega}{2}+\cos^2\frac{3\omega}{2}}}$	$\dfrac{\varrho\sin\tau\left(\cos\frac{\omega}{2}-\cos\frac{3\omega}{2}\right)}{\sqrt{\cos^2\frac{\omega}{2}+\cos^2\frac{3\omega}{2}}}$

6. Anchor refinement

So far we have considered only parameters within the chain, but it is clearly also necessary to allow the anchor (or nose) of the molecule freedom of translation and rotation.

The three translational degrees of freedom are simple to handle, the derivatives $\partial F_\alpha/\partial x$, $\partial F_\alpha/\partial y$, $\partial F_\alpha/\partial z$ being components of the vector $2\pi i\mathbf{s}_\alpha F_\alpha$ where F_α denotes the scattering factor of the *molecule*, which may or may not be the crystallographic F, depending on the number of molecules in the cell. This will lead to a 3×3 block on the diagonal of $\mathbf{M}$, which is expected to be well conditioned, because, generally, there is no direction of translation (for the entire molecular fragment attached to the anchor) to which the F's are insensitive, though exceptions to this can occur.

For the orientation freedom we permit rotations of the whole molecule about axes parallel to the adopted Cartesian axes passing through some chosen anchor point. The centroid of the tail is a good choice as this minimizes correlation between translation of the whole molecule and rotation of it. If the molecule or molecular fragment being refined is rod-like with direction cosines $\cos\sigma_1$, $\cos\sigma_2$, $\cos\sigma_3$, then it may be shown that the transformation matrix

$$\mathbf{A}=\begin{bmatrix} \cos\sigma_1 & -\sin\sigma_1 & 0 \\ \cos\sigma_2 & \cot\sigma_1\cos\sigma_2 & \dfrac{-\cos\sigma_3}{\sin\sigma_1} \\ \cos\sigma_3 & \cot\sigma_1\cos\sigma_3 & \dfrac{\cos\sigma_2}{\sin\sigma_1} \end{bmatrix}$$

represents, in the first column, a rotation of the rod about its length, for which the corresponding diagonal element in $\mathbf{A}^T\mathbf{M}\mathbf{A}$ is small, and the other two columns represent tilts of the rod about axes perpendicular to its length, to which the X-ray observations are sensitive even at low resolution.

7. Application

The foregoing discussion has been primarily concerned with blocks on the diagonal of $\mathbf{M}$ and we now wish to propose an approach to the solution of $\mathbf{M}$ as a whole.

Suppose that the parameters in the main chain of the molecule are arranged in $\mathbf{M}$ in the order in which they occur in the molecule, and that $\mathbf{M}$ is divided into partitions $\mathbf{m}_{IJ}$ with the $\mathbf{m}_{II}$ square though not necessarily all of the same size. Then for the transformation $\mathbf{A}^T\mathbf{M}\mathbf{A}$ write

$$\Lambda=\begin{pmatrix} \Lambda_{11} & \Lambda_{12}\cdots \\ \Lambda_{21} & \Lambda_{22} \\ \vdots & \end{pmatrix}$$

$$=\begin{pmatrix} \mathbf{A}^T_{11}\mathbf{m}_{11}\mathbf{A}_{11} & \mathbf{A}^T_{11}\mathbf{m}_{12}\mathbf{A}_{22}\cdots \\ \mathbf{A}^T_{22}\mathbf{m}_{21}\mathbf{A}_{11} & \mathbf{A}^T_{22}\mathbf{m}_{22}\mathbf{A}_{22} \\ \vdots & \end{pmatrix}$$

$$=\begin{pmatrix} \mathbf{A}^T_{11} & \mathbf{O}\cdots \\ \mathbf{O} & \mathbf{A}^T_{22} \\ \vdots & \end{pmatrix}\begin{pmatrix} \mathbf{m}_{11} & \mathbf{m}_{12}\cdots \\ \mathbf{m}_{21} & \mathbf{m}_{22} \\ \vdots & \end{pmatrix}\begin{pmatrix} \mathbf{A}_{11} & \mathbf{O}\cdots \\ \mathbf{O} & \mathbf{A}_{22} \\ \vdots & \end{pmatrix}$$

so that each $\mathbf{A}_{II}$ diagonalizes the corresponding $\mathbf{m}_{II}$, being found either numerically or by the foregoing approach, and solve for those elements of $\mathbf{A}^T\Delta\mathbf{p}$ for which the corresponding diagonal elements of Λ are large. Elements of $\mathbf{A}^T\,\Delta\mathbf{p}$ for which the diagonal elements of Λ are very small may be ignored (set to zero) provided that they are not severely correlated with any in another partition for which the diagonal element is large and for which a determination is being made. Estimates of the partitions $\mathbf{A}^T_{II}\mathbf{m}_{IJ}\mathbf{A}_{JJ}$ become less reliable away from the diagonal, but with this reservation it may be shown that two diagonal blocks of order 8, ($N=4$), in an α-helical part of the molecule are related by an off-diagonal block $\mathbf{A}^T_{II}\mathbf{m}_{IJ}\mathbf{A}_{JJ}$ of the form

$$\begin{bmatrix} \sqrt{\lambda_\alpha\lambda'_\alpha}\cos 4(I-J)\omega & \vdots & \sqrt{\lambda_\alpha\lambda'_\varepsilon}\sin 4(I-J)\omega & \vdots \\ \cdots\cdots & & \cdots\cdots & \\ \sqrt{\lambda_\varepsilon\lambda'_\alpha}\sin 4(I-J)\omega & \vdots & \sqrt{\lambda_\varepsilon\lambda'_\varepsilon}\cos 4(I-J)\omega & \vdots \\ \cdots\cdots & & \cdots\cdots & \end{bmatrix}$$

expanded to order 8 by insertion of three rows (columns) of zeros at each row (column) of dots, where λ_α is the α eigenvalue in one diagonal block, λ'_α its value in the other and similarly for λ_ε and λ'_ε.

This indicates that the two eigenvectors in any diagonal partition which represent a bending of the helix as a whole are strongly correlated with the corresponding vectors in a neighbouring partition, but that those vectors which are self cancelling in the displacements they produce are not appreciably correlated with the important eigenvectors, α and ε of other blocks. This means that we may, for example, refine a structure at low resolution by deleting those rows and columns in Λ which are represented above by zeros and the resulting solution is then one in which the helix as a whole is permitted freedom to adjust itself to fit the X-ray observations by bending in any direction, with one pair of bend measurements for each turn of the helix. The deletions mentioned above are equivalent to ignoring minor alterations of structure of a local nature, such as the γ-vector, which represents a combination of rotations around bonds of a kind which transmits no disturbance to the tail. This concentration of attention on those combinations of displacements which affect the course of the chain as a whole is just what is required in the earlier stages of refinement.

We are then left with a band matrix of much reduced order and one must consider what width of band is required. The above estimate of the blocks $\mathbf{A}^T_{II}\mathbf{m}_{IJ}\mathbf{A}_{JJ}$ is based on a calculation which supposes that neighbouring groups of N peptides carry essentially the same tail, *i.e.* the geometrical factors are taken into account, but variations in the quantities P are not. Consequently the above estimate indicates that the α and ε vectors of one partition are together totally correlated with those in another, for any $(I-J)$, *i.e.*, all such off-diagonal blocks are important. However, the situation is not as bad as this. To say that the correlation of these vectors in neighbouring blocks is total is equivalent to saying that a bend of a few degrees in the helix as a whole located at one turn of the helix is indistinguishable, X-ray-wise, from the configuration produced by the same bend applied to a neighbouring turn of the helix. Clearly, for neighbouring turns, these two displacements *are* difficult to distinguish, and therefore correlated, but taken over several turns the distinction must improve on account of the scattering material located between the bending sites (correlation must fall because of variations in P). It must therefore be a matter of experience to determine the number of off-diagonal blocks required, but it is expected that a maximum value of $(I-J)$ about 4 may be sufficient for 4-peptide groupings ($N=4$).

8. Summary

It has been the object of this paper to show that by choosing appropriate parameters, p, and by choosing appropriate linear combinations of these, we may not only reduce the number of unknowns to be determined, but may also make 'refinement' a worthwhile operation at an early stage of analysis in a structure rich in α-helix, and in other structures if numerical diagonalization is used.

Certainly an electron density map at 6 Å resolution does not in itself determine atomic coordinates, but it does delineate the course of an α-helix. It follows that any combination of parameters (*e.g.* the α and ε vectors) which affects the course of the helix as a whole is refinable at this resolution. Similarly, if the resolution is high enough for the thread on the screw to be discernible, then parameters (such as the η vector) which relate to dilation and twisting are refinable. The generalization is that we may say that if we properly diagonalize the matrix $\mathbf{M}$ to yield Λ, then just as one proceeds from low to high resolution in solving the structure, so may we simultaneously proceed from refinement of a few eigenvectors with the largest eigenvalues towards those with smaller eigenvalues, the process being limited at every stage by the resolution with which the structure is determined.

Clearly this presupposes that it is possible to write a set of chemically reasonable starting coordinates which are nearly enough correct for refinement to converge. This may be achieved in the α-helical regions by inserting coordinates corresponding to some standard α-helix in the positions indicated roughly by a low resolution electron density map. The coordinates resulting from refinement of these will correspond well to a helix with the right course, but, of course, they will only correspond to the true atomic coordinates when resolution is high enough to determine also the azimuthal angle of the helix about its own axis. If this condition is satisfied, then the resulting atomic coordinates must already be very close to the true values unless the true structure is a good deal less regular than is to be expected. For example, if the planes in which the peptide groups lie *genuinely* make a variety of angles with the helix axis, making a somewhat ragged helix, then such a structure could not be represented by a standard α-helix except by modifying the latter with displacements such as the β combination, which we expect to be unmeasurable – only in such unlikely circumstances would the refined atomic coordinates fail to come close to their true values when the β, γ, δ, ζ and θ vectors are suppressed.

I should like to make acknowledgement to the referee whose comments on the first draft have led to significant improvements.

APPENDIX I

On the relationship between correlation coefficients and the eigenvalues and eigenvectors of two-dimensional sections of the representational ellipsoid of M

If the parameters shifts, Δp, which are found by a least-squares analysis are modified slightly by amounts δp, then the sum of the squares of the errors, σ, (which is minimized by $\mathbf{\Delta p}$) is increased by an amount $\delta\sigma$ given by

$$\delta\sigma = \boldsymbol{\delta}\mathbf{p}^T\,\mathbf{M}\,\boldsymbol{\delta}\mathbf{p}$$

(*cf.* Diamond, 1958). Thus the surface in p-space $\boldsymbol{\delta}\mathbf{p}^T\mathbf{M}\boldsymbol{\delta}\mathbf{p}=\text{const.}$ is a contour of constant error in the vicinity of the least-squares solution. If we suppose that the least-squares analysis is done many times with many sets of data, then the solutions obtained (points in p-space) will be normally distributed about the ideal solution, forming a cloud of points with density

$$A\exp-\pi\,\frac{\boldsymbol{\delta}\mathbf{p}^T\,\mathbf{M}\,\boldsymbol{\delta}\mathbf{p}}{B} \tag{17}$$

where A is a normalization constant and B is a variance given by

$$B = 2\pi\,\Sigma\,\frac{w\,(\Delta F)^2}{(n-f)}\,,$$

w being a weighting factor and $(n-f)$ the number of observations minus the number of independent parameters determined.

Now the correlation coefficient between two variates x and y in statistics is defined by

$$C_{xy} = \mu_{11}/\sqrt{\mu_{02}\,\mu_{20}} \tag{18}$$

where

$$\mu_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\bar{x})^i\,(y-\bar{y})^j f(x,y)\,dx\,dy\,, \tag{19}$$

where $f(x, y)$ is the distribution of x and y, $\bar{x}$ and $\bar{y}$ are the means of x and y and μ_{ij} is the i, jth moment of the distribution (Cruickshank, 1959). Thus to determine the correlation coefficient between two parameters p_k and p_l, say, we consider these parameters to be normally distributed variates according to (17) and evaluate (18) using (19). Now (17) describes the over-all distribution of all the parameters of the problem, whereas the correlation coefficient relates to these in pairs, so that we must consider just two among the many involved in (17), or figuratively, we must consider two-dimensional sections of the poly-dimensional ellipsoid $\boldsymbol{\delta}\mathbf{p}^T\mathbf{M}\boldsymbol{\delta}\mathbf{p}=\text{const}$*. Such a two-dimensional section is necessarily an ellipse, the principal axes of which do not necessarily coincide either in magnitude or direction with the principal axes of the ellipsoid of which it is a section. This is easy to see in three dimensions. Then

$$\mu_{11} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta p_k\, \delta p_l^2\, A_{kl} \exp\Big(-\frac{\pi}{B}[\delta p_k^2\, M_{kk} + 2\,\delta p_k\, \delta p_l\, M_{kl} + \delta p_l^2\, M_{ll}]\Big)\, d\delta p_k\, d\delta p_l\,, \qquad (20)$$

in which the contents of the square brackets now represent the elliptical section, and the normalization constant A_{kl} now depends only on M_{kk}, M_{ll}, M_{kl} and B. We have taken a central section of the ellipsoid (*i.e.* set $\delta p_m=0$ for $m\neq k$ or l) as this simplifies the algebra. Non-central sections need only be considered if the total correlation coefficient is sought.

Now suppose the semi-principal axes of the elliptical section make an angle θ with the axes p_k and p_l and are of lengths a and b, and then transform variables to α and β according to

$$\delta p_k = \alpha \cos\theta + \beta \sin\theta$$
$$\delta p_l = -\alpha \sin\theta + \beta \cos\theta$$

with unit Jacobian, then (20) becomes

$$\mu_{11} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\frac{\beta^2-\alpha^2}{2}\sin 2\theta + \alpha\beta\cos 2\theta\right] \times A_{kl} \exp -\frac{\pi}{B}\left(\frac{\alpha^2}{a^2}+\frac{\beta^2}{b^2}\right) d\alpha\, d\beta\,,$$

in which the term in $\alpha\beta$ contributes nothing to the integral in the range $-\infty\rightarrow\infty$ and integration gives

$$\mu_{11} = \frac{A_{kl}\, B^2 a\, b \sin 2\theta}{4\pi}(b^2-a^2)\,.$$

Similarly

$$\mu_{20} = \frac{A_{kl}\, B^2 ab}{2\pi}(a^2\cos^2\theta + b^2\sin^2\theta) \qquad (21)$$

and

$$\mu_{02} = \frac{A_{kl}\, B^2 ab}{2\pi}(b^2\cos^2\theta + a^2\sin^2\theta)\,,$$

whence

$$C = \frac{\mu_{11}}{\sqrt{\mu_{02}\,\mu_{20}}} = \frac{b/a - a/b}{\sqrt{\cot^2\theta + b^2/a^2 + a^2/b^2 + \tan^2\theta}} \qquad (22)$$

from which it is clear that the correlation coefficient vanishes if the ratio a/b is unity (ellipse circular; *sectional* eigenvalues equal) or if θ is a multiple of $\pi/2$, *i.e.* if the ellipse is aligned with its axes parallel to the axes p_k and p_l, which occurs if the element M_{kl} is zero. In the foregoing paper we have attempted to minimize correlation by working with combinations of parameters which in p-space are nearly parallel to the principal axes of the ellipsoid.

By the substitutions $y=\log(b/a)$ and $z=\log\tan\theta$, C may be written

$$C = \frac{\sinh y}{\sqrt{\cosh(y+z)\cosh(y-z)}}$$

which, for $\theta=45°$ gives $C=\tanh\log(b/a)$. In Fig. 6 we give a graph of the partial correlation coefficient as a function of θ for values of $\log(b/a)$ corresponding to ratios of *sectional* eigenvalues up to 1:180. These curves show in particular that the greater the ratio of sectional eigenvalues, the more precisely must θ be determined in order to yield a small correlation coefficient, which could mean that the angle α of equation (15), § 4, should not be ignored and the resulting transformation (16) should be modified accordingly. However, the transformation (16) contains implicitly the transformation

$$\begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix}$$

in four positions. This is equivalent to a rotation through an angle θ (in the notation of this appendix) given by

$$\cos\theta = a_1 = \frac{\sin\tau_1}{\sqrt{\sin^2\tau_1 + \sin^2\tau_2}}$$

and it may be shown that, if variations in the quantities P are taken into account to find an ideal rotation angle θ_i then the angle θ given above differs from the ideal value by a residual angle θ_r given by

$$\theta_r \simeq \tfrac{1}{4}\,.\,(P_{11}-P_{22})/P_{12}\,.$$

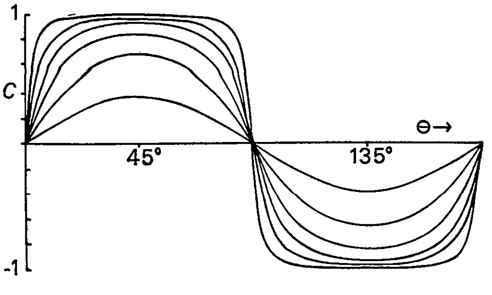

Fig. 6. Plot of the partial correlation coefficient between two variables as a function of θ for values of $\log(b/a)=\frac{1}{2}\log(\lambda_1/\lambda_2)$ by intervals of 0·4 up to 2·0 and at 2·6. The limiting forms are $(b/a-a/b)\sin 2\theta$, when $b\simeq a$ and the square wave when b/a approaches 0 or ∞.

---

* Strictly speaking such a two-dimensional *section* of the ellipsoid determines the *partial* correlation coefficient between p_k and p_l. The total correlation coefficient between p_k and p_l would be obtained by considering a *projection* of the distribution (17) on the p_k-p_l plane. See, for example, Weatherburn (1949), Ch. XII.

i.e. $\theta_r \ll 15°$ if the quantities P are closely similar, as is to be expected, especially near the head end of the molecule.

Now, if $\sqrt{\mu_{20}}$ is the r.m.s. error in the estimation of the combination $(p_1 \cos\theta + p_2 \sin\theta)$ then it may be shown that

$$\sqrt{\frac{(\mu_{20})_{\theta_r}}{(\mu_{20})_{\theta_r=0}}} = \cos\theta_r \sqrt{1 + \frac{\lambda_1}{\lambda_2}\tan^2\theta_r}$$

where λ_1/λ_2 is the ratio of sectional eigenvalues, *i.e.* eigenvalues of

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

This function is plotted in Fig. 7, and though this treatment is limited to partial correlation coefficients, it does indicate for a large ratio λ_1/λ_2, *i.e.* for an ill-conditioned problem such as this, that provided the transformation angles θ are correct to within a few degrees then the determination of an important combination of variables is achieved with a standard deviation greater than its ideal value by a factor which appears as ordinate in Fig. 7. This is clearly not an exhaustive study of the characteristics of similarity transformations which only approximately diagonalize a matrix, but it does give some indication of the accuracy required of the rotation angles of which any similarity transformation may be made up.

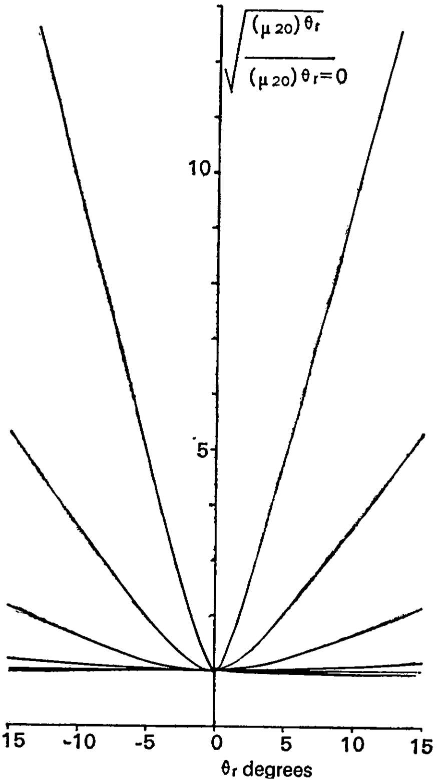

Fig. 7. Plot of the standard deviation of the determination of a parameter as a function of the residual angle θ_r between the parametric axes and the axes of the ellipse, expressed relative to its value at $\theta_r = 0$, for values of log (b/a) equal to 0 (horizontal straight line), 1, 2, 3 and 4. The curve for log $(b/a) = -\infty$ occurs just below the horizontal line.

APPENDIX II

Space-group effects

Space-group effects may conveniently be divided into two parts as follows:

(1) *The effect of non-Cartesian coordinates.*

The analysis of this paper has been given entirely in terms of Cartesian coordinates. The results may be carried over into the crystallographic system provided we can evaluate the quantities

$$\frac{\partial F_\alpha}{\partial p_k} = 2\pi i\,(\mathbf{s}_\alpha \times \mathbf{n}_k) \,.\, \mathbf{Q}_{k\alpha}$$

and

$$\mathbf{Q}_{k\alpha} = \sum_t^T \mathbf{r}_{kt} f_{t\alpha} \exp 2\pi i\, \mathbf{R}_t \,.\, \mathbf{s}_\alpha ;$$

i.e. a formulation of the vector triple product must be found. Here

s is a Cartesian vector measurable in Å^{-1}.
r is a Cartesian difference vector measurable in Å.
R is a Cartesian position vector measurable in Å.
Q is a complex Cartesian vector measurable in e.Å, and
n is a Cartesian vector whose magnitude is the pure number unity.

Let s_i, r_i, R_i, Q_i, n_i be the components of these vectors, and let h_i, x_i, X_i, q_i, u_i be their crystallographic counterparts; *i.e.* the first three are pure numbers and

$$q_i = \Sigma x_i f \exp 2\pi i\, \mathbf{h} \,.\, \mathbf{X} \,.$$

Then **r**, **R**, and **Q** transform according to

$$Q_i = A_{il} q_l$$

in which the elements of **A** are in Å. Then if

$$n_i = A_{il} u_l$$

is to be dimensionless, u_l must be in Å^{-1}, and if we define a vector **v** whose components are the (pure number) differences of the crystallographic coordinates of points which define the ends of the spindle vector **n** then

$$u_l = v_l / v$$

where $v = |A_{il} v_l|$ is the spindle length in Å.

Finally $s_i = B_{il} h_l$, where **A** and **B** are related by $\mathbf{AB}^T = \mathbf{I}$, the identity matrix.

We give the results for the general triclinic case using the usual angles α, β and γ and also angles A, B, Γ which are the angles in the corners of a spherical triangle whose sides are α, β, γ. The matrices **A** and **B** may then be written

$$A = \begin{pmatrix} a \sin\beta \sin\Gamma & 0 & 0 \\ a \sin\beta \cos\Gamma & b \sin\alpha & 0 \\ a\cos\beta & b\cos\alpha & c \end{pmatrix}$$

$$B = \begin{pmatrix} (a \sin\beta \sin\Gamma)^{-1} & (-b \sin\alpha \tan\Gamma)^{-1} & (-c \sin\alpha \tan B)^{-1} \\ 0 & (b \sin\alpha)^{-1} & (-c \tan\alpha)^{-1} \\ 0 & 0 & c^{-1} \end{pmatrix} .$$

The required triple product is then

$$\varepsilon_{ijk}\, Q_i\, s_j\, n_k = \varepsilon_{ijk}\, A_{il}\, B_{jm}\, A_{kn}\, q_l\, h_m\, u_n = \zeta_{lmn}\, q_l\, h_m\, u_n\,,$$

where

$$\zeta_{lmn} = \varepsilon_{ijk}\, A_{il}\, B_{jm}\, A_{kn}$$

from which we obtain

$$\zeta_{123} = -\zeta_{321} = \frac{ac \sin B}{b \sin A \sin\Gamma}$$

$$\zeta_{231} = -\zeta_{132} = \frac{ab \sin\Gamma}{c \sin A \sin B}$$

$$\zeta_{312} = -\zeta_{213} = \frac{bc \sin A}{a \sin B \sin\Gamma}$$

$$\zeta_{122} = -\zeta_{221} = \zeta_{331} = -\zeta_{133} = a/\tan A$$

$$\zeta_{233} = -\zeta_{332} = \zeta_{112} = -\zeta_{211} = b/\tan B$$

$$\zeta_{311} = -\zeta_{113} = \zeta_{223} = -\zeta_{322} = c/\tan\Gamma$$

and the remaining nine elements of ζ (those with first and last suffix equal) vanish.

In computation of $\partial F_\alpha/\partial p_k$ it is convenient to evaluate all the derivatives of a particular reflexion at a time, *i.e.* k varies in the inner loop and α in the outer loop. Hence the derivatives may be written

$$2\pi i (q_x, q_y, q_z) \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix} \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}$$

where the quantities

$$\xi_1 = h\,.\,\frac{bc \sin A}{a \sin B \sin\Gamma} - k\,.\,\frac{c}{\tan\Gamma} - l\,.\,\frac{b}{\tan B}$$

$$\xi_2 = -h\,.\,\frac{c}{\tan\Gamma} + \frac{k\,.\,ac \sin B}{b \sin A \sin\Gamma} - l\,.\,\frac{a}{\tan A}$$

$$\xi_3 = -h\,.\,\frac{b}{\tan B} - k\,.\,\frac{a}{\tan A} + \frac{l\,.\,ab \sin\Gamma}{c \sin A \sin B}$$

need only be calculated in the outer loop, and contain only constant multiples of h, k and l.

(2) *The effect of the plurality of asymmetric units*

In structure factor calculations one must take account of symmetry-related positions; in this work one must also take account of symmetry-related directions. Unfortunately, it is not feasible to give a generalized treatment here, but we give the results for space group $P2_1$, and others may be dealt with similarly.

In this paragraph we use unprimed variables to denote one asymmetric unit, and primed variables to denote the corresponding quantities in the second asymmetric unit. Thus

$$\frac{1}{2\pi i}\,.\,\frac{\partial F}{\partial p}$$

$$= (q_x, q_y, q_z) \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix} \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}$$

$$\oplus (q'_x, q'_y, q'_z) \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix} \begin{pmatrix} u'_x \\ u'_y \\ u'_z \end{pmatrix}$$

in which the encircled sign would be reversed if the relationship between asymmetric units were enantiomorphous.

Now in $P2_1$

$$u'_x = -u_x \quad u'_y = u_y \quad u'_z = -u_z$$

whence

$$\begin{aligned} \frac{1}{2\pi i}\,\frac{\partial F}{\partial p} = {} & u_x [\xi_3 (q_y - q'_y) - \xi_2 (q_z - q'_z)] \\ & + u_y [\xi_1 (q_z + q'_z) - \xi_3 (q_x + q'_x)] \\ & + u_z [\xi_2 (q_x - q'_x) - \xi_1 (q_y - q'_y)]\,, \end{aligned}$$

so that we must calculate the five complex quantities

$$(q_x + q'_x),\ (q_x - q'_x),\ (q_y - q'_y),\ (q_z + q'_z) \text{ and } (q_z - q'_z).$$

Now

$$q_x = \Sigma\, x f \exp 2\pi i\, (hX + kY + lZ)\,,$$

$$q'_x = -\Sigma\, x f \exp 2\pi i\, (-hX + kY - lZ + k/2)$$

$$q_y = \Sigma y f \exp 2\pi i (hX + kY + lZ)\,,$$

$$q'_y = \Sigma y f \exp 2\pi i (-hX + kY - lZ + k/2)$$

$$q_z = \Sigma z f \exp 2\pi i (hX + kY + lZ)\,,$$

$$q'_z = -\Sigma z f \exp 2\pi i (-hX + kY - lZ - k/2)$$

from which we obtain,

for k even for k odd

$$\begin{aligned} (q_x + q'_x) &= 2\Sigma x f \sin 2\pi (hX + lZ) \\ &\quad \times [-\sin 2\pi k Y + i \cos 2\pi k Y] = (q_x - q'_x) \\ (q_x - q'_x) &= 2\Sigma x f \cos 2\pi (hX + lZ) \\ &\quad \times [\cos 2\pi k Y + i \sin 2\pi k Y] = (q_x + q'_x) \\ & 2\Sigma y f \cos 2\pi (hX + lZ) \\ &\quad \times [\cos 2\pi k Y + i \sin 2\pi k Y] = (q_y - q'_y) \\ (q_y - q'_y) &= 2\Sigma y f \sin 2\pi (hX + lZ) \\ &\quad \times [-\sin 2\pi k Y + i \cos 2\pi k Y] \\ (q_z + q'_z) &= 2\Sigma z f \sin 2\pi (hX + lZ) \\ &\quad \times [-\sin 2\pi k Y + i \cos 2\pi k Y] = (q_z - q'_z) \\ (q_z - q'_z) &= 2\Sigma z f \cos 2\pi (hX + lZ) \\ &\quad \times [\cos 2\pi k Y + i \sin 2\pi k Y] = (q_z + q'_z) \end{aligned}$$

$$F=2\Sigma f\cos 2\pi(hX+lZ)\times[\cos 2\pi kY+i\sin 2\pi kY]$$

$$2\Sigma f\sin 2\pi(hX+lZ)\times[-\sin 2\pi kY+i\cos 2\pi kY]=F\,,$$

which may be summarized by saying that if $\mathcal{F}$ is the geometrical structure factor for any k, and $\mathcal{G}$ is the geometrical structure factor calculated as if the parity of k were opposite to its true value, then for odd or even k

$$F=\Sigma f\mathcal{F},\quad (q_x+q'_x)=\Sigma xf\mathcal{G},\quad (q_x-q'_x)=\Sigma xf\mathcal{F},$$

$$(q_y-q'_y)=\Sigma yf\mathcal{G}\,,\ (q_z+q'_z)=\Sigma zf\mathcal{G}\,,\ (q_z-q'_z)=\Sigma zf\mathcal{F}\,.$$

References

CRUICKSHANK, D. W. J. (1959). In *International Tables for X-ray Crystallography*, Vol. II, 88. Birmingham: Kynoch Press.
DIAMOND, R. (1956). Ph. D. Thesis, Univ. of Cambridge.
DIAMOND, R. (1958). *Acta Cryst.* **11**, 129.
EYRING, H. (1932). *Phys. Rev.* **39**, 746.
HUGHES, R. E. & LAUER, J. L. (1959). *J. Chem. Phys.* **30**, 1165.
MIZUSHIMA, S. I. (1955). *Advanc. Protein Chem.* **9**, 299.
NAGAI, K. & KOBAYASHI, M. (1962). *J. Chem. Phys.* **36**, 1268.
SCHERINGER, C. (1963). *Acta Cryst.* **16**, 546.
SCHERINGER, C. (1964). Private communication.
WEATHERBURN, C. E. (1949). *A First Course in Mathematical Statistics*. Cambridge Univ. Press.

Acta Cryst. (1965). **19**, 789

The Modes of Hydrocarbon Chain Packing

BY E. SEGERMAN*

Lever Brothers Company, Research and Development Division, 45 *River Road, Edgewater, New Jersey, U.S.A.*

(*Received* 27 *July* 1964 *and in revised form* 19 *February* 1965)

The symmetry of a saturated hydrocarbon chain in extended *trans* configuration and indeterminate length is reviewed. A uniform row of such chains is defined so that the chain axes are parallel, coplanar and equidistant, and that the planes of the carbon atom zigzags of all the chains are parallel. The symmetry within such a row is stated, and all the possible symmetry relations between such rows when they are parallel, adjacent and identical is investigated. Making reasonable assumptions concerning dimensions, packing efficiency and stability, four different relations are found when the chain axes of neighboring rows are parallel, and seven different relations are found when the axes are not parallel. All the possible chain-packing subcells that could be made up of such rows are constructed, subject to the limitations that the subcells extend over no more than two rows, and that the same symmetry relation exists between all adjacent rows. Ten different subcells are thus generated when all the chain axes are parallel, and thirty-one when they are not. It is found that the eight reported chain-packing modes are all included among these. It is shown that the symmetry of three of the six reported subcells in which the chain axes are all parallel is higher than was previously assumed. An attempt to relate the frequency of occurrence of the various subcells to the van der Waals energy as calculated by the Salem method failed.

Introduction

Eight hydrocarbon-chain packing arrangements have been described from single-crystal structure determinations, six of which have all the chain axes parallel and the remaining two have the chain axes of alternate layers crossed. Another packing arrangement, the hexagonal, has been postulated from powder data but its structure is unknown. Still another, the 'ideal', has had its structure described in detail (Kitaigorodskii, 1961) but it has never been observed.

The purpose of this communication is to relate the described modes by generating them from symmetry operations between uniform rows of chains. This approach not only leads to simple classification of all the known packing modes, but it may indicate what new ones might be found.

General considerations

The lateral van der Waals forces between hydrocarbon chains within a crystal structure are rarely so strong, compared with end-packing forces, as to make the effect of the end packing completely negligible. This effect becomes evident when the total crystal symmetry is lower than the chain packing symmetry. Conversely, these forces are rarely so weak in comparison with the forces in other parts of the structure that the chains are not able to assume a uniform, compact, stable and

---

* Present address: Unilever Research Laboratories, Chemical Physics Division, Port Sunlight, Cheshire, England.