Acta Cryst. (1963). **16**, 981

The Refinement of Atomic Parameters by direct Calculation of the Minimum Residual

By A. K. Bhuiya and E. Stanley

Department of Physics, College of Science and Technology, Manchester 1, *England*

(*Received* 15 *November* 1962 *and in revised form* 8 *February* 1963)

A new method of refining projections is proposed in which the parameters, both positional and thermal, are varied one at a time by specified increments and limited to a predetermined range. The best value of each parameter is taken as that which gives the lowest value of the residual. The method has been tried out on a variety of structures of various degrees of complexity and severity of overlap, and starting at various states of refinement. In all cases further refinement was achieved.

Introduction

Approximate structures obtained by one of the many available methods are tested for correctness by adjusting the atomic parameters so as to reduce the differences between the observed and calculated structure factors. The accuracy of the final structure can be related to the degree of agreement reached between the observed and calculated data. Fourier methods (F_o synthesis) are capable of giving good results, in projection, only when all the atoms are well resolved. When this is not so, refinement by F_o synthesis soon reaches its limit since the peak positions no longer correspond to atomic sites. Further refinement must proceed by difference-Fourier (Cochran, 1951) or by least-squares (Hughes, 1941; Booth, 1947) methods. The former method runs into difficulties when there is overlap and the interpretation of the difference maps may become ambiguous. The latter method, even if the full matrix is used, refines to the correct structure only if it is already in the right parabolic region to start with. If this is not so, the method will make the best of the situation in the wrong 'hole'.

Some successes have been reported (Curtis, 1959; Cruickshank, Pilling *et al.*, 1961) where the maximum corrections to coordinates have been of the order 0·5–0·6 Å, but these should be regarded as exceptions. In the latter case success was attained because the rest of the atomic coordinates were almost correct. If all the parameters are in error by such large amounts the very strong interactions at the early stages of refinement may very well lead to the wrong structure (Geller, 1961). If thermal parameters are included in a least-squares refinement it sometimes happens that the thermal parameters assume control, and although the final residual is low it may correspond to the wrong structure. Whenever a structure, because of overlap in projection, fails to refine although there are good reasons to suppose that it is correct, it is common to resort to structure-factor graphs (Bragg & Lipson, 1936) in an effort to improve the agreement between the observed and the calculated data.

Principles of the new method

In an effort to overcome these difficulties the present methods have been devised. Parameters are allowed to explore, or scan, the immediate neighbourhood of their supposed values, in search of values giving better agreement between the observed and calculated data. Although it would be desirable to explore, or scan systematically, the p-dimensional volume, where p is the number of independent parameters, such a process would be too time-consuming. One way of reducing the amount of work would be to restrict the exploration to the parameters of each atom in turn. Even this restriction fails to make the calculation a reasonable project, and the only way left to make the process more economical is to treat every parameter independently.

Booth (1949) investigated the theoretical possibility of a relaxation method of refinement in which the corrections to the parameters were calculated from the incremental relationship,

$$\Delta\varphi = (\partial\varphi/\partial u_j)\Delta u_j$$

where u_j is any structural parameter and φ is some measure of the agreement between the observed and calculated structure factors. Equations such as this are correct *only* if Δu_j is very small. However, the expression

$$\Delta\varphi = \varphi(u_1, u_2, \ldots u_j, \ldots) - \varphi(u_1, u_2, \ldots u_j + \Delta u_j, \ldots)$$

is always correct whatever the magnitude of Δu_j. This is an important consideration if large corrections are involved.

Although there are other criteria which might, on theoretical grounds, be preferred to the residual

$$R = \Sigma||F_o| - |F_c||/\Sigma|F_o| ,$$

as a test of the correctness of a structure, it still remains the criterion most familiar to crystallographers. It was decided to minimize R in this work.

If it can be assumed that all the parameters are independent we can vary each parameter systematically from $u_j - n\Delta u_j$ to $u_j + n\Delta u_j$ in $2n$ steps of Δu_j. The value of R is calculated for each value. That value of the parameter, say u'_j, which corresponds to the lowest of the $2n+1$ values of R is taken as a better value for this parameter and is used in all subsequent calculations. The same process is repeated for the next parameter u_{j+1} and so on. In this way the value of R can be systematically reduced. The values of the increment, Δu_j, can be adjusted to be appropriate to the circumstances—larger increments at the early stages of refinement, small increments at the later stages. The limits of variation of the parameter $\pm n\Delta u_j$ can also be varied by adjusting n so that $n\Delta u_j$ corresponds to the maximum change to be tolerated in any parameter at any stage of refinement. Some idea of the useful range of exploration can be obtained from the figures given by Luzzati (1952) for the value of the residual as a function of the average error in coordinate.

Refinement of atomic coordinates

Any process such as has been described is very time-consuming and a program has been written for the calculation to be performed on the Manchester University Mercury Computer. The program initially computes and sums the contributions to all the structure factors of all the atoms except the one being refined, and stores these quantities. Suppose x_1 is one coordinate of the first atom to be refined. The coordinate is initially set at $x_1 - n\Delta x_1$ and the contribution to all the structure factors calculated and added to the stored contributions for the remainder of the atoms to give the calculated structure factors for all the reflexions. The corresponding value of R is calculated and stored. The coordinate is then increased to $x_1 - (n-1)\Delta x_1$ and the new R-value calculated in the same way. The latest R-value is compared with the previously stored R-value and the lower of the two values is preserved, together with the corresponding value of the coordinate. The coordinate is then increased to $x_1 - (n-2)\Delta x_1$ and the process continues. In this way the smallest value of R within the range of variation of the coordinate is preserved together with the corresponding value of the coordinate.

When the scan in the x direction is complete the y coordinate is similarly treated. When both the x and y coordinates of the first atom have been dealt with the new values of x, y and R are printed. The next atom is then dealt with in the same way, and so on until all the atoms have been included. The corrected parameters of the atoms are always used in subsequent calculations. After the last atom has been dealt with a second cycle can be started. The analytical atomic scattering factors of Forsyth & Wells (1959) are used in calculating the structure factors, and individual isotropic temperature coefficients can be included if required. The scaling factor and the overall temperature factor can be adjusted at the end of each cycle using existing programmes. At present the program will accommodate up to 180 reflexions and up to 30 atoms of 5 different types. If the number of increments, n, is 5, *i.e.* the value of R is calculated at 11 points for each positional parameter, the calculation takes 32 minutes for one cycle including 7 minutes for the initial input of the data and for preliminary calculation and storage of the structure factors. Two cycles can be completed in one hour. A comparable problem using the full matrix least-squares method on the same computer would take about 25 minutes per cycle.

Refinement of the isotropic temperature factors

After several cycles of coordinate refinement (and normally not before) the individual temperature factors, B_j, can be refined by the same type of process. The value of B_1, the temperature factor of the first atom, is varied from $B_1 - n\Delta B_1$ to $B_1 + n\Delta B_1$ in $2n$ steps of ΔB_1. The residuals are calculated and compared as before and the lowest value of R is taken to define the best value of the temperature factor. As each atom is dealt with in turn the new values of B_j and R are printed. The present program will accommodate problems of the same size as the coordinate refinement program, and takes about 13 minutes per cycle.

Applications of the method

The methods have been tried on a wide variety of problems, always with success, the most complex problem having 25 atoms in the asymmetric unit projected along a 12·4 Å axis, and the simplest having only 8 atoms in the asymmetric unit and projected along a 6 Å axis. The largest initial R-value was 0·35, which was refined to 0·12, and the smallest initial R-value was 0·11, which refined to 0·078.

1. *Naphthocinnoline*

The methods were first applied to this structure, which is at present being studied by the authors. The space group is *Fdd*2 with 16 molecules in the unit cell and 18 atoms in each molecule. The trial structure of the centrosymmetric [001] projection, plane group *pgg*, was obtained by the method described by Bhuiya & Stanley (to be published) and refined by Fourier methods to an R-value of 0·37. The projection axis was 10·75 Å and there was considerable overlap. Further refinement by full matrix least-squares methods only reduced the residual to 0·33 in two cycles, after which it would not further refine. Starting from the same initial residual of 0·37 the new method of coordinate refinement described here, with a range of scan 0·1 Å, reduced the residual to 0·29 in the first cycle, an initial rate of refinement much greater than that of least-squares. Three further cycles reduced the residual to 0·24 but after adjusting the scale and the overall temperature factors, the next cycle reduced the residual to 0·18. Subsequent refinement of the individual temperature factors by the new method and further cycles of coordinate refinement reduced to residual to 0·16. During the refinement from a residual of 0·37 to the final value of 0·16 one of the atoms was moved by 0·9 Å, three other atoms were moved by more than 0·5 Å and the remaining fourteen atoms were all moved by the order of 0·2–0·3 Å. During this first trial of the new method difference-Fourier maps were calculated at successive stages of refinement in order to confirm that the corrections made were sensible in that they led to maps which were obviously improving. Individual changes in the values of R were quite small. Even for a shift of 0·1 Å, ΔR was of the order 0·005.

The [010] projection has the non-centrosymmetric plane group *cm*1 and the projection axis is again 10·75 Å. Using the model obtained from the [001] projection, the initial residual value was 0·35. No conventional refinement was undertaken at all. After refinement by the new methods the final residual was 0·12. The changes in the common coordinate were all very small.

2. *Triphenyl phosphate*

This structure has been refined by Davies & Stanley (1962) using Fourier-difference methods. The space group is $P2_1/a$ and there are 23 atoms of three different types in each of the four asymmetric units. The [100] projection, along a 17 Å axis, shows a great deal of overlap and the Fourier-difference methods of refining had reduced the residual to 0·19. Four stages of coordinate and two stages of temperature coefficient refinement reduced the residual to 0·10. During this refinement the maximum change in a coordinate was 0·17 Å and the average change was 0·05 Å.

3. *Methylammonium alum*

The structure of a low-temperature form of this salt has been determined by Fletcher & Steeple (to be published). The space group is $Pca2_1$. There are 25 atoms in the asymmetric unit, of five different types. The [001] projection has plane-group symmetry *pmg* and the length of axis of projection is 12·4 Å. There is a great deal of overlap in projection. The structure—an orthorhombic modification of the cubic, room temperature, form (Fletcher & Steeple, 1961, 1962)—was known to be substantially correct although refinement by Fourier-difference and by the full matrix least-squares methods had failed to reduce the residual below 0·24. The present methods of coordinate refinement reduced the residual to 0·13 in 3 cycles. During refinement the maximum change in any coordinate was 0·24 Å and the mean change was 0·06 Å.

4. *Dipotassium ethylenetetracarboxylate*

This structure has been determined in detail by Kumra & Darlow (to be published). The space group is *Pbca* with 8 atoms of 3 different types in the asymmetric unit. At the stage when Fourier-difference methods had reduced the residuals of the [100] and the [010] projections to 0·13 and 0·11 respectively the new methods were tried. Both projections were further refined by two cycles of coordinate and temperature factor refinement using the new methods. The residual of the [100] projection, axis of projection 9·75 Å, fell to 0·083 and that of the [010] projection, axis of projection 6·42 Å, fell to 0·078. During the course of refinement the maximum coordinate change was 0·044 Å and the mean change was 0·011 Å. At the stage where the new methods were applied there was still evidence in the difference-Fourier maps of changes in temperature coefficients and of small coordinate changes, but it is fairly certain that it would have taken a good deal longer to extract and test these corrections.

Discussion

The new methods have provided a means of refining structures at all stages from the relatively rough model of naphthocinnoline at a residual of >0·35 to the highly refined dipotassium ethylenetetracarboxylate at a residual of 0·11. In all cases further refinement was achieved. The disadvantage that Booth (1949) associated with his relaxation method, of always descending parallel to a unit-cell edge instead of along an axis of the countour ellipse as in the case of steepest-descents or least-squares methods, appears to be much less serious than he suspected.

Although the method was originally devised to deal with overlapped atoms in projection it is, in fact, a quite general method of refining any structure. The methods have been shown to work even when very large coordinate changes (>0·5 Å) are necessary, as in the case of naphthocinnoline, and have been shown to work when other methods have failed, as in the case of methylammonium alum. It is considered that these methods will be of most use in getting the rough structure into the right 'hole'. The final refinement will be most efficiently completed using the full matrix least-squares method.

Intuitively it might be supposed that the order in which the atoms are dealt with is important. In practice it has been found that the method is insensitive to changes in the order even in a single cycle. The assumption that the parameters are all independent is not, of course, true in that more than one cycle is required even when the changes in coordinates are within their permitted range of exploration. The method does, however, converge rapidly with both centrosymmetric and non-centrosymmetric projections. In non-orthogonal projections there will be some interaction between the two positional parameters. This should not be serious provided the interaxial angle is not too different from 90°. The method has not yet been applied to such a projection so that the actual effect cannot be assessed.

The structures so far refined have not involved large variations in the individual temperature factors but in all the cases where the new method has been used the greater part of the refinement of the temperature factor refinement has been completed in a single cycle, provided the overall temperature factor and the scaling factor were correct. It might be expected that the new method would deal more easily with large variations than the existing least-squares methods, which tend to overestimate the reductions and underestimate the increases in temperature factors and necessarily need more than one cycle. Provided the coordinate refinement is well advanced there should be no problem of missing the 'hole'.

There is no reason why the methods could not be extended to deal with three-dimensional data and anisotropic temperature coefficients except that the existing methods may well be adequate.

It is considered that the methods will be most useful in two-dimensional work where there is a great deal of overlap which makes it very difficult to get the structure into the right 'hole' to start with.

We are grateful to Prof. H. Lipson for his interest in the development of the methods, to Dr S. F. Darlow and Mr S. K. Kumra and to Drs H. Steeple and R. O. W. Fletcher for the use of their data before publication, to Mr S. S. Hasan and Mr P. Main for many helpful discussions on the programming, and to Prof. F. C. Williams for the use of the facilities of the Manchester University Computing Laboratory.

One of us (A. K. B.) is indebted to the Pakistan Council of Scientific and Industrial Research for the grant of study leave and to the Government of Pakistan for financial assistance while the present work was carried out.

References

Booth, A. D. (1947). *Nature, Lond.* **160**, 196.

Booth, A. D. (1949). *Proc. Roy. Soc.* A, **197**, 336.

Bragg, W. L. & Lipson, H. (1936). *Z. Kristallogr.* **95**, 323.

Cochran, W. (1951). *Acta Cryst.* **4**, 408.

Cruickshank, D. W. J., Pilling, D. E. *et al.* (1961). *Computing methods and the phase problem in X-ray crystal analysis*, p. 48. London: Pergamon Press.

Curtis, A. R. (1959). *A Mercury computer programme for two-dimensional least-squares crystallographic refinement.* London: H.M.S.O.

Davies, W. O. & Stanley, E. (1962). *Acta Cryst.* **15**, 1092.

Fletcher, R. O. W. & Steeple, H. (1961). *Acta Cryst.* **14**, 891.

Fletcher, R. O. W. & Steeple, H. (1962). *Acta Cryst.* **15**, 960.

Forsyth, J. B. & Wells, M. (1959). *Acta Cryst.* **12**, 412.

Geller, S. (1961). *Acta Cryst.* **14**, 1026.

Hughes, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.

Luzzati, V. (1952). *Acta Cryst.* **5**, 802.

Acta Cryst. (1963). **16**, 984

The Crystal and Molecular Structure of Lauric Acid (Form A_1)

By T. R. Lomer

Physics Department, The University, Birmingham 15, *England*

(*Received* 12 *November* 1962)

Form A_1 of lauric acid, $C_{12}H_{24}O_2$, is triclinic, with $a = 7{\cdot}45$, $b = 5{\cdot}40$, $c = 17{\cdot}47$ Å; $\alpha = 96° 53'$, $\beta = 113° 8'$, and $\gamma = 81° 7'$. The space group is $P\bar{1}$. The crystal is built of lauric acid dimers, which are packed together with the planes of all hydrocarbon chains parallel to each other (triclinic subcell), and with the carboxyl groups of one dimer adjacent to methyl groups of neighbouring dimers. The angle of tilt is 67° 20′.

Introduction

The polymorphism of the normal fatty acids is well known and the structures of most of the crystal forms have been reviewed by von Sydow (1956*a*). Form C (or α) which is obtained when lauric acid is crystallized from the melt or from ethyl alcohol solution at room temperature was described by Vand, Morley & Lomer (1951). Lomer (1955) described the slow, spontaneous transformation, in the solid state, of form C to the form A-super (or γ) and attempted to determine the unit cell dimensions of this latter form from powder photographs. Von Sydow (1956*b*) showed that Lomer's cell dimensions for the form