Acta Cryst. (1968). B24, 1165

The Refinement of Structures with Non-Crystallographic Molecular Symmetry

By E. N. Maslen

Department of Physics, University of Western Australia, Nedlands, Western Australia

(*Received* 29 *August* 1967 *and in revised form* 21 *November* 1967)

A method has been devised for improving the structure factor phase calculations for crystals where the unit of pattern contains symmetry elements which are not utilized in the packing of the structure. The phase corrections, which are derived by an iterative refinement process, take the form of small corrections to the phases obtained from the usual structure refinement. The changes in phase have no effect on the structural parameters, but give a substantial improvement in the difference density. The method has been verified by application to a model structure.

Introduction

The following notation is used, but not defined in the text:

α, α_n	Approximate structure factor phase
α_c, α_n^c	Calculated structure factor phase
α_o, α_n^o	Observed structure factor phase
$\Delta\varrho(\mathbf{r})$	Difference density at the position **r**
$\mathbf{F}$, $\mathbf{F}_n$	Approximate structure factor
$\mathbf{F}_c$, $\mathbf{F}_n^c$	Calculated structure factor
$\mathbf{F}_o$, $\mathbf{F}_m^o$, $\mathbf{F}_n^o$	Observed structure factor
$\mathbf{H}_m$, $\mathbf{H}_n$	Reciprocal space vector
J	Number of chemically equivalent units per asymmetric unit
N	Number of observed structure factors, including those for symmetrically equivalent planes
$\mathbf{r}$	Real space vector
$\mathbf{r}_j$	Vectors related by elements of molecular symmetry to **r**
$\sum\limits^{J}$	Sum over the chemically equivalent units
$\sum\limits^{N}$	Sum over the observed reflexions
$d\tau$	Element of volume
V	Unit-cell volume

Crystal structures where the asymmetric unit contains elements of symmetry which are not utilized in the packing have certain well known advantages for the purposes of accurate structure analysis. The assessments of accuracy may be verified by an examination of the agreement between chemically equivalent but crystallographically independent bonds and angles, and the effects of random error can be reduced by averaging over the chemically equivalent regions of the molecule. It is not generally realized, however, that other information of fundamental importance can be obtained from the symmetry properties of these structures. The structure factor phases may be further refined after the normal refinement has been terminated to give a set of phases which are largely independent of the theoretical models used to evaluate the calculated structure factors. The electron-density standard deviations may be estimated in a manner which has certain advantages over the conventional methods. These factors are of particular relevance in the study of electron-density distributions.

In this paper the theoretical basis of phase refinement and its application to a model structure are discussed. In an adjoining paper (Maslen, 1968*a*) the use of the method to examine the structure analysis of benzotrifuroxan by Cady, Larson & Cromer (1966) is described. The techniques proposed are more general than might at first be realized. They may be utilized, for example, in a planar structure where the molecular plane is not a space group mirror plane provided that the electron density may be regarded as symmetrical above and below this plane. The ideas are also readily extended to cases of partial symmetry. It is probably, however, that they will be more useful in structures of high molecular symmetry where they are of greater power, and where high accuracy in the mean structural parameters and electron densities is more readily achieved.

It is necessary to consider to what degree molecular symmetry may be distorted by packing forces in the crystal, which constitute a perturbation of lower symmetry. Kitajgorodskij (1965) has shown that the effects of intermolecular van der Waals and repulsive forces on the molecular geometry are generally small. Since these forces are greater than the interactions between the static dipoles and higher order multipoles, except perhaps for strong hydrogen bonds, molecular symmetry will usually be preserved quite closely in a crystal structure. This is confirmed by an examination of accurate structures which have been reported. Thus there is no instance in 1,3,5-triamino-2,4,6-trinitrobenzene (Cady & Larson, 1965) or biphenyl (Robertson, 1961) of a significant difference between a pair of chemically equivalent bonds or angles. The argument is equally valid when applied to electron density distributions. The overlap integrals between atomic orbitals, which determine the size of the changes in density due to an interaction, are an order of magnitude smaller for atoms making intermolecular contact than for covalently bonded atoms. The changes in electron density brought about by covalent bondings are small and have only recently been regarded as detectable. At the present level of accuracy attainable in structure analysis those due to intermolecular forces will be negligibly small.

Phase refinement

In the normal method for refining crystal structures the phases of the observed structure factors are set equal to the calculated structure-factor phases. Difference electron densities represent the combined effects of errors in the structure parameters, the measured amplitudes of the structure factors, and the phases, together with any inadequacies in the atomic scattering factors. For well refined centrosymmetric structures, where the phases with the centre of symmetry as origin are either 0 or π radians, the probability of phase errors is small, except for terms with very small magnitudes. Recently O'Connell, Rae & Maslen (1966) have shown that, where such structures have molecular symmetry within the asymmetric unit of the crystal, averaging the chemically equivalent regions of the difference density reduces the effect of random errors in the structure-factor amplitudes and structure parameters. The averaged electron difference density represents more closely the departure of the real density from the theoretical model used to evaluate the calculated structure factors.

For the non-centrosymmetric case, on the other hand, the observed phases are not exactly correct even in well refined structures, since the errors in the theoretical atomic form factors are transmitted to the phase calculations. The phase errors are not of great importance in evaluating the structural parameters, but they are a serious limitation in the study of difference densities. This may be illustrated by reference to a phase diagram for the observed and calculated structure factors [Fig. 1(*a*)]. In the structure refinement $\mathbf{F}_c$ may be calculated exactly, but only the locus of $\mathbf{F}_o$ is known. Ignoring experimental errors in the measurement of F_o[*] for the purposes of this discussion, the locus of $\mathbf{F}_o$ is the circumference of a circle on which the points nearest $\mathbf{F}_c$ have the highest probability. For a well refined structure the phase error $\Delta\alpha$ is small, but this is not true of α_D, the difference between the phase of $\Delta\mathbf{F}_o$, defined as $\mathbf{F}_o-\mathbf{F}_c$, and α_c. Ideally the difference electron density should be evaluated using the $\Delta\mathbf{F}_o$ as coefficients, but in practice only the $\Delta\mathbf{F}'$, which have magnitudes F_o-F_c and phases α_c are known. The $\Delta\mathbf{F}'$ are approximately equal to the components of the $\Delta\mathbf{F}_o$ parallel to the $\mathbf{F}_c$, but the components perpendicular to the $\mathbf{F}_c$ are usually indeterminate, and the difference map is strongly affected by their omission. The features indicating errors in the atomic form factors are reduced, and there is an increase in the random background of the synthesis. If this random background could be eliminated the true phases would be determined, since the Fourier transform of the remaining features would then give the direction of the $\Delta\mathbf{F}_o$ vectors. In general a complete elimination is not possible, but where the asymmetric unit has molecular symmetry the effects of random error may be reduced by averaging the electron difference densities in chemically equivalent regions. This suggests a method by which the structure-factor phases might be refined. In the ideal case of correct phases and no experimental errors in the measurement of the F_o, the $\Delta\mathbf{F}_o$ would be given by the Fourier transform of the difference density. In an actual case the $\Delta\mathbf{F}_o$ will be approximated more closely by the Fourier transform of the averaged, rather than the unaveraged, difference densities. The phases calculated from the Fourier transform of the averaged difference density will then give a better approximation to the true phases, which may be used to calculate a closer approximation to the correctly phased difference density. This phase refinement process may now be iterated, leading eventually to nearly correct phases. The procedure is mathematically equivalent to minimizing the sum of the squares of the differences between the observed structure factors and calculated values $\mathbf{F}'$ obtained from the sum of the original $\mathbf{F}_c$'s and the $\Delta\mathbf{F}$'s calculated from the Fourier transform of the averaged difference density. This may be shown as follows: Consider the intermediate stage of refinement represented by Fig. 1(*b*). Phase corrections $\delta\alpha_n$ may be calculated from the conditions for minimizing

$$R_F=\sum^{N}\|[\mathbf{F}_{n^o}-(\mathbf{F}_{n^c}+\Delta\mathbf{F}_n)]\|^2$$

$$=\sum^{N}\Big[\int\Big\{\Delta\varrho(\mathbf{r})-\frac{1}{J}\sum^{J}\Delta\varrho(\mathbf{r}_j)\Big\}\exp(2\pi i\mathbf{H}_n\cdot\mathbf{r}\,d\tau\Big]^2. \quad (1)$$

Now

$$\Delta\varrho(\mathbf{r}) = \sum_{m=1}^{N}\frac{1}{V}\Big[F_{m^o}\exp(i\alpha_{m^o})-F_{m^c}\exp(i\alpha_{m^c})\Big]$$
$$\times\exp(-2\pi i\mathbf{H}_m\,.\,\mathbf{r}) \quad (2)$$

and

$$\delta\alpha_m=\alpha_{m^o}-\alpha_m\,. \quad (3)$$

By a multivariate Taylor expansion, ignoring terms of higher than first order in $\delta\alpha$,

$$R_F=\sum^{N}\Big[\int\Big\{\Delta\varrho(\mathbf{r})-\frac{1}{J}\sum^{J}\Delta\varrho(\mathbf{r}_j)\Big\}\exp(2\pi i\mathbf{H}_n\,.\,\mathbf{r}d\tau)$$
$$+\frac{i}{V}\sum_{m-1}^{N}F_{m^o}\exp(i\alpha_m)$$
$$\times\Big\{\int\exp[2\pi i(\mathbf{H}_n\,.\,\mathbf{r}-\mathbf{H}_m\,.\,\mathbf{r})]\,d\tau$$
$$-\frac{1}{J}\int\sum^{J}\exp[2\pi i(\mathbf{H}_n\,.\,\mathbf{r}-\mathbf{H}_m\,.\,\mathbf{r})]d\tau\Big\}\delta\alpha_m]^2 \quad (4)$$

---

[*] F_o and F_c are the magnitudes of $\mathbf{F}_o$ and $\mathbf{F}_c$ respectively.

where the $\Delta\varrho$ are now evaluated using the approximate phases α. The integrals in the right-hand terms take on appreciable values only when $\mathbf{H}_m=\mathbf{H}_n$, and for the second integral when $\mathbf{r}_j=\mathbf{r}$. The normal equations matrix, obtained from the conditions $\partial R_F/\partial\alpha_n=0$ is diagonal and has solutions

$$\delta\alpha_n = -\frac{J}{J-1}\,\frac{i\exp(-i\alpha_n)}{F_{no}}\left\{\int \frac{1}{J}\sum^{J}\Delta\varrho(\mathbf{r}_j)\right.$$
$$\left.\times\exp(2\pi i\mathbf{H}_n\,.\,\mathbf{r})d\tau+\mathbf{F}_{nc}-\mathbf{F}_n\right\} \quad (5)$$

where $\mathbf{F}_n$ is the observed structure factor with the phase α_n. $\delta\alpha_n$ estimated in this way has two components. The first, which is real, is $J/(J-1)$ times the phase change calculated by adding $\Delta\mathbf{F}$ to $\mathbf{F}_c$. The effect of the factor $J/(J-1)$ will simply be to speed convergence. The second component is imaginary, and has no physical significance. Identical refinement equations are obtained for minimizing

$$\int\left\{\Delta\varrho(\mathbf{r})-\frac{1}{J}\sum^{J}\Delta\varrho(\mathbf{r}_j)\right\}^2 d\tau,$$

rather than $\sum^{N}|[\mathbf{F}_{no}-\mathbf{F}_{nc}-\Delta\mathbf{F}_n]|^2$. The refinement process is an iterative one because of the non-linearity in α of the equations for the residual. The technique is applicable both to centrosymmetric and non-centrosymmetric structures, the only difference being that in the former case the probability of a phase change is low.

It may be shown from equation (5) that the method is remarkably insensitive to errors in the scale factor for the F_o's. A fractional change of $\delta k/k$ in the scale alters $\mathbf{F}_n$ by an amount $(\delta k/k)\mathbf{F}_n$, and there will be a corresponding change of $(\delta k/k)\varrho(\mathbf{r}_j)$ in $\varrho(\mathbf{r}_j)$. Since $\varrho(\mathbf{r})$ will be approximately equal to $\varrho(\mathbf{r}_j)$ for all $\mathbf{r}$ and $\mathbf{F}_n$ is the Fourier transform of $\varrho(\mathbf{r})$ the changes in the terms within the brackets due to an error in scale will very nearly cancel. The magnitude of the F_{no} term in the denominator will also be altered, but since the $\delta\alpha_n$ are already small the change due to scale will be of second order, and can safely be ignored.

The validity of the method in general and the relative power of various possible refinement techniques were established by reference to a test problem, based on the model structure shown in Fig. 2. Six carbon atoms were set at the corners of a regular hexagon, having a side of 1·5 Å with three hydrogen atoms at the mid-points of alternate sides. The molecule was placed with an arbitrary orientation in a cell of dimensions $a=6\cdot923$, $b=19\cdot516$ and $c=6\cdot518$ Å and space group $Pna2_1$. The atoms were given isotropic temperature factor coefficients $B=3\cdot0$ Å^2. The 'observed' structure factors were calculated from all the atoms whereas the 'calculated' values were obtained from the carbon atoms alone. The data were terminated at an upper limit of $(\sin\theta/\lambda)^2=0\cdot30$. The effects of random error in measurement and refinement were simulated by rounding the structure factors to the nearest tenth of an electron, and the phases to the nearest degree. In the average process only those symmetry elements generated by the threefold axis and the mirror plane at right angles to it were utilized. The mirror planes parallel to the threefold axis were used to check the accuracy of the average density. The interval and range for the numerical evaluation of the Fourier transform of the difference densities were selected so as to reduce the computing time as much as possible without serious loss of accuracy. The optimum values were an interval of 0·5 Å, and a hexagonal-prism-shaped volume of integration with hexagonal side 3 Å and length 1·5 Å. This unit was centred on the mole-

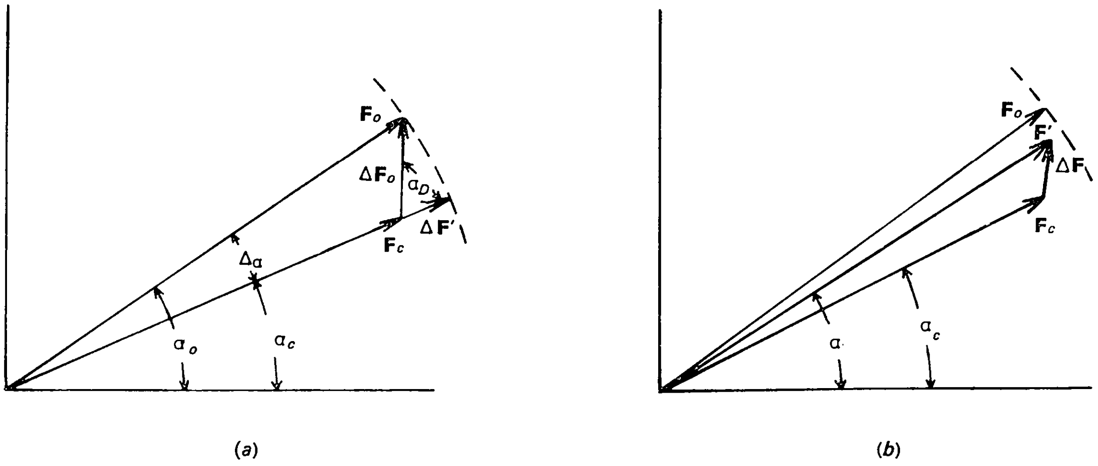

Fig. 1. (*a*) Phase diagram for the observed and calculated structure factors. (*b*) Phase diagram at an intermediate stage of phase refinement. $\Delta\mathbf{F}$ is calculated from the Fourier transform of the averaged electron density.

cular centre, with the unique axis perpendicular to the molecular plane. A number of acceleration devices were tried, but none gave more rapid convergence than the original method.

A factor which presented some difficulty at first was the detection of the end point of the refinement. The method was not expected to give exact phases because of the rounding errors in the structure factors, series termination and the approximations involved in the numerical integration. Convergence alone was not sufficient, since some plausible refinement techniques converged to what was clearly an intermediate stage in the process. Eventually two criteria were adopted. The first, which assumes that the α_D were randomly distributed for the non-centrosymmetric reflexions, was to check that $\langle|\sin \alpha_D|\rangle = \langle|\cos \alpha_D|\rangle$ for these terms. This assumption should be valid for the test case, although it may not hold for all cases of interest. The second criterion was obtained by applying the refinement procedure to the initially correct phases. Refinements from the calculated phases which gave phase agreements comparable to that of the 'refined' correct phases were accepted as correctly refined. There was close agreement between the assessments from the two criteria.

The data were refined through six cycles, at the conclusion of which the hydrogen atom peaks were within 1% of their correct height. After each round the following indices of refinement and standard deviations were evaluated:

$$\overset{N}{\Sigma}\,\big|F_{n^o}-|\mathbf{F}'|\big|/\overset{N}{\Sigma}\,F_{n^o},\ \overset{N}{\Sigma}\,|\mathbf{F}_{n^o}-\mathbf{F}_{n^c}|/\overset{N}{\Sigma}\,F_{n^o},$$

$$\langle|\sin \alpha_D|\rangle/\langle|\cos \alpha_D|\rangle,\ \langle|\alpha_o-\alpha|\rangle,$$

$$\sigma_1(\Delta\varrho)=\left[\frac{\dfrac{J}{J-1}\displaystyle\int\left\{\Delta\varrho(\mathbf{r})-\frac{1}{J}\overset{J}{\Sigma}\,\varrho(\mathbf{r}_j)\right\}^2 d\tau}{\displaystyle\int d\tau}\right]^{1/2},$$

$$\sigma_2(\Delta\varrho)=\frac{J\sqrt{2})}{V(J-1)}\left[\overset{N}{\Sigma}\left\{|F_{n^o}|-|F'|\right\}^2\right]^{1/2*}$$

The progress of the refinement is indicated in Table 1(*a*). The results of a one-round refinement from the correct phases is given in Table 1(*b*) for comparison. Except for the fact that it is defined vectorially the second refinement index is identical with the conventional R index, $\overset{N}{\Sigma}\,|F_{n^o}-F_{n^c}|/\overset{N}{\Sigma}\,F_{n^o}$, which takes no account of the phase errors. It should be noted that this index rises to a little less than $\sqrt{2}$ times its initial value as the refinement converges. The factor is almost exactly $\sqrt{2}$ if the terms from the centric projection are excluded from the summation. This is in accord with the assessment of the effect of phase errors on electron density standard deviations which is given in a subsequent paper (Maslen, 1968*b*). Sections through the plane of the structure in the correct difference synthesis and the corresponding mean synthesis are shown in Figs. 3(*e*) and (*f*) respectively. Similar sections at

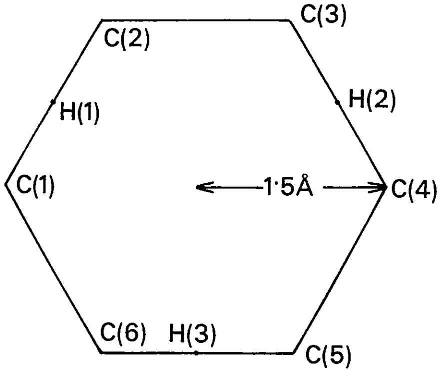

Fig. 2. Model structure used for testing the phase refinement method.

Table 1. *Refinement progress*

Cycle number	$\overset{N}{\Sigma}\,|F_{no}-F'| \,/\, \overset{N}{\Sigma}\,F_{no}$	$\overset{N}{\Sigma}\,|F_{no}-F_{n}c| \,/\, \overset{N}{\Sigma}\,F_{no}$	$\langle|\sin \alpha_D|\rangle^* / \langle|\cos \alpha_D|\rangle$	$\langle|\alpha_o-\alpha|\rangle^*$ (°)	$\sigma_1(\varrho)$ (e.Å^{-3})	$\sigma_2(\varrho)$ (e.Å^{-3})
(*a*) Approximate phases						
0	–	0·0465	0·00	2·83	0·020	0·072
1	0·0200	0·0510	0·70	1·68	0·011	0·035
2	0·0143	0·0560	0·83	1·16	0·008	0·030
3	0·0123	0·0586	0·90	1·07	0·007	0·025
4	0·0115	0·0604	0·96	1·03	0·006	0·023
5	0·0107	0·0615	0·98	0·95	0·006	0·021
6	0·0105	0·0622	1·00	0·90	–	0·019
(*b*) Correct phases						
0			1·00	0·0	0·006	
1	0·0091	0·0628	1·00	0·47	–	0·019

* The averages are taken over the non-centric terms only.

---

* The expressions for the standard deviations are taken from Maslen (1968*b*).

the beginning and end of the refinement are given in Figs. 3(*a*) to (*d*). One interesting feature of the refinement is the rapid improvement of the difference synthesis compared with that of the phases. The hydrogen peaks had attained more than 90% of their full height and the general lack of resolution had disappeared after three rounds of refinement. It is clear that the mean difference synthesis attains a nearly correct form as soon as the systematic nature of the phase errors is destroyed.

The argument given in the introduction of this paper regarding the preservation of symmetry in the molecular geometry cannot be applied to the thermal motion and there are small but significant departures from symmetry among the thermal parameters of accurate structures which have been reported. This in turn degrades the symmetry of the time averaged electron density which is observed in a Fourier synthesis. However it has been shown by Higgs (1953) that the density distribution is very sensitive to changes in thermal parameters near the atomic centres, but relatively insensitive in other regions of the structure. Thus symmetry will be conserved quite strongly in the difference density except in regions in the immediate vicinity of the atomic centres, which contribute very little to the integral in equation (5) for a well refined structure. Nevertheless there will be changes in the calculated phases due to the reduced symmetry. The effects of departures from symmetry among the thermal parameters on the phase refining method were therefore investigated and for the sake of completeness changes in the geometrical symmetry were also considered. The positional and thermal parameters of individual carbon atoms were varied by 0·02 Å and 1·0 $Å^2$ respectively and the phases were again refined. There was no deterioration in the rate of refinement or in the ultimate agreement factors.

The models used to assess the usefulness of the method are idealized in that the random errors introduced by rounding are small. In an analysis where the errors in the data are larger, or where there are small but genuine differences in the electron density due to

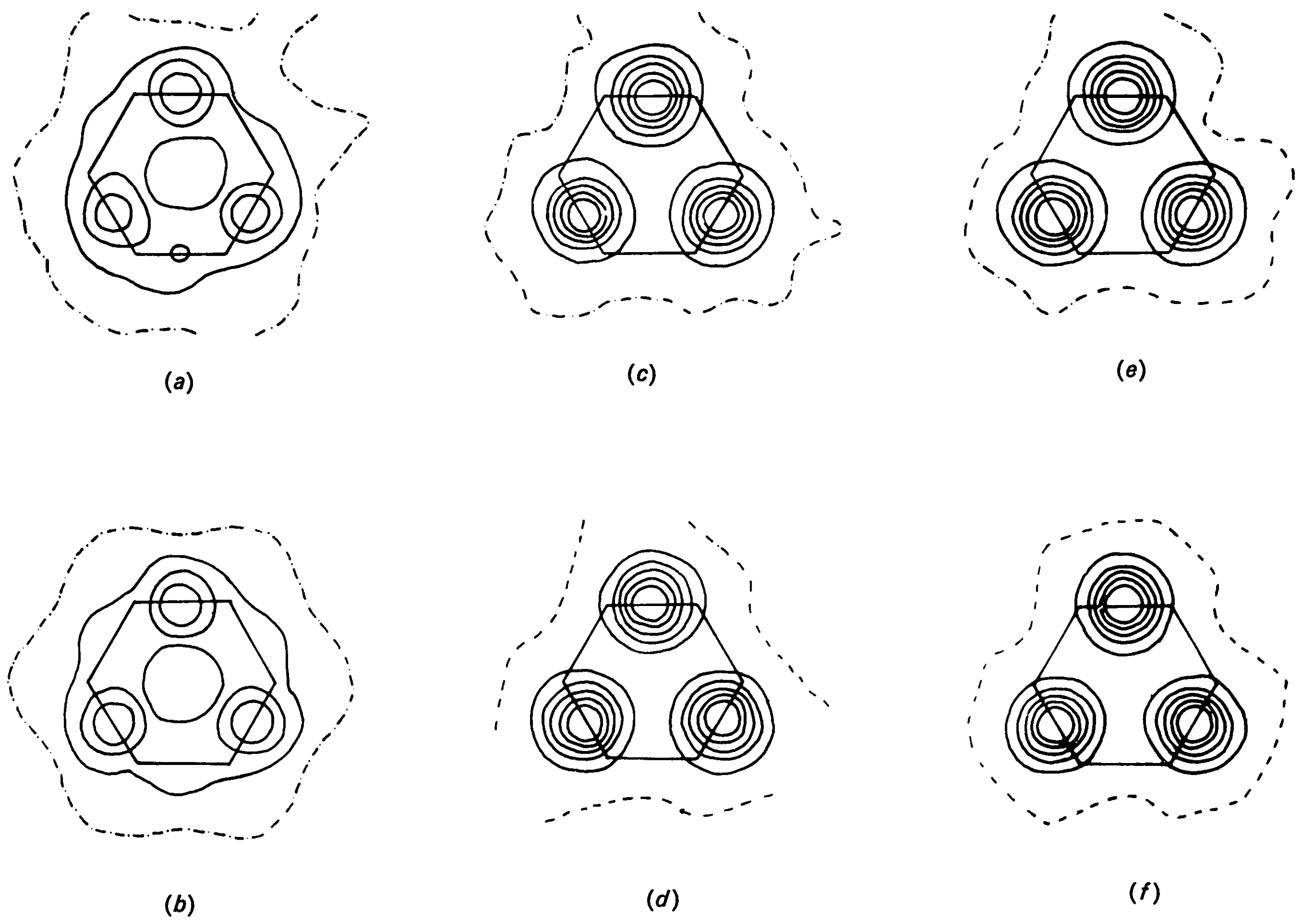

Fig. 3. (*a*) Section through the molecular plane of the normal difference synthesis. (*b*) Section of the averaged synthesis corresponding to (*a*). (*c*) Section through the molecular plane of the refined difference synthesis. (*d*) Section of the averaged synthesis corresponding to (*b*). (*e*) Section through the molecular plane of the ideal difference synthesis. (*f*) Section of the averaged synthesis corresponding to (*e*).

thermal motion or strong hydrogen bonding, a lesser degree of refinement is to be expected.

Thanks are due to the Atlas Computer Laboratory, Chilton, near Didcot, Berkshire, England for the use of the ATLAS computer and to the staff of the computing laboratory at the University of Western Australia for advice on computational problems. The receipt of a Nuffield Foundation Travelling Fellowship is very gratefully acknowledged.

References

CADY, H. H. & LARSON, A. C. (1965). *Acta Cryst.* **18**, 485.
CADY, H. H., LARSON, A. C. & CROMER, D. T. (1966). *Acta Cryst.* **20**, 336.
HIGGS, P. W. (1953). *Acta Cryst.* **6**, 232.
KITAJGORODSKIJ, A. I. (1965). *Acta Cryst.* **18**, 585.
MASLEN, E. N. (1968*a*). *Acta Cryst.* B**24**, 1170.
MASLEN, E. N. (1968*b*). *Acta Cryst.* B**24**, 1172.
O'CONNELL, A. M., RAE, A. I. M. & MASLEN, E. N. (1966). *Acta Cryst.* **21**, 208.
ROBERTSON, G. B. (1961). *Nature, Lond.* **191**, 593.

Acta Cryst. (1968). B**24**, 1170

A Phase Refinement of the Crystal Structure of Benzotrifuroxan

BY E. N. MASLEN

Department of Physics, University of Western Australia, Nedlands, Western Australia

(*Received* 29 *August* 1967 *and in revised form* 21 *November* 1967)

The structure-factor phases for benzotrifuroxan have been refined by a procedure which takes advantage of the molecular but non-crystallographic symmetry. The difference synthesis calculated with the refined phases has been interpreted in terms of the distribution of bonding electrons combined with the effects of disorder in the structure.

The crystal structure of benzotrifuroxan was determined by Cady, Larson & Cromer (1966). The analysis was accurate enough to justify an examination of the final difference synthesis for bonding-electron effects, were it not for the phase errors resulting from the use of isolated atom form factors and the non-centrosymmetric space group of the structure. The structure is closely related to that of 1,3,5-triamino-2,4,6-trinitrobenzene, for which the electron-density distribution has been studied by Cady & Larson (1965) and more closely by O'Connell, Rae & Maslen (1966).

The free benzotrifuroxan molecule contains one threefold axis and a mirror plane of symmetry which are not utilized by the space group of the crystal structure. For the reasons given in the preceding paper (Maslen, 1967*a*) it is expected that the $3/m$ symmetry of the free molecule will be preserved quite closely in the crystal, and if this is valid the structure factor phases may be further refined after the normal refinement has been terminated. The refined phases are then reasonably independent of the theoretical model used to evaluate the atomic form factors, and the systematic nature of the errors in a phase calculation based on an approximate model is destroyed. The examination of the difference synthesis for bonding-electron effects is then justified if the electron-density standard deviation is sufficiently low. In the structural data given by Cady, Larson & Cromer there are no significant differences between chemically equivalent bond lengths and angles. The molecule is also essentially planar although there are some significant deviations from the least-squares mean plane, the largest being 0·096 Å. The method has been shown to be insensitive to small departures from exact symmetry, and since the r.m.s. deviation from planarity (0·049 Å) is less than one tenth of the resolution the phase refinement method should give a satisfactory first order correction to the structure factor phases.

The data of Cady, Larson & Cromer were refined by using the procedure described in the preceding paper. The progress of the refinement is given in Table 1. Sections of the difference density and averaged difference density in the molecular plane at the start and the conclusion of the refinement are shown in Fig. 1(*a*), (*b*), (*c*) and (*d*). The numbering of the atoms in the molecule is given in Fig. 1(*e*).

Table 1. *Refinement progress**

Cycle number	$\frac{\sum^{N} \lvert F_{no}-F'\rvert}{\sum^{N} F_n^o}$	$\frac{\sum^{N} \lvert \mathbf{F}_n^o-\mathbf{F}_{nc}\rvert}{\sum^{N} F_n^o}$	$\sigma_1(\varrho)$ (e.Å^{-3})	$\sigma_2(\varrho)$ (e.Å^{-3})
0	—	0·0375	0·036	0·068
1	0·0279	0·0412	0·034	0·048
2	0·0271	0·0461	0·034	0·046
3	0·0278	0·0498	—	0·046

* Symbols are defined as in Maslen (1968*a*).

The standard deviation in the refined electron-difference density, estimated as described in the following paper (Maslen, 1968*b*), is 0·046 e.Å^{-3}. The standard deviations in the averaged difference density in and