CHAPTER II

EVALUATION OF ATOMIC POSITIONS

II.1 DETERMINATION OF PHASES

As mentioned earlier, the determination of atomic positions in the unit cell requires a knowledge of the structure amplitudes of as many reflections \((h,k,l)\) as possible. This is so because the structure amplitudes of the individual reflections depend upon the positions of the various atoms. In fact, if \(\rho(x,y,z)\) is the electron density at a point \((x,y,z)\) in the unit cell, then the structure amplitude \(F_{hkl}\) of \((1,2)\) is given by

\[
F_{hkl} = \frac{\nu}{a} \iint_{0}^{2\pi} \int_{0}^{\pi} \rho(x,y,z) \exp\left[2\pi i \left(\frac{h}{a}x + \frac{k}{b}y + \frac{l}{c}z\right)\right] \sin \alpha \sin \beta \cdot [2.1]
\]

In the atomic representation where the unit cell is filled with \(N\) atoms, the structure amplitude could be represented by

\[
F_{hkl} = \sum_{\nu=0}^{N} \alpha_{\nu} \exp\left[2\pi i \left(\frac{h}{a}x_{\nu} + \frac{k}{b}y_{\nu} + \frac{l}{c}z_{\nu}\right)\right] \cdot [2.2]
\]

where \(\alpha_{\nu}\) is the atomic scattering factor of atom \(\nu\) which has the co-ordinates \((x_{\nu}, y_{\nu}, z_{\nu})\). We see that \(F_{hkl}\), which is the ratio of the scattering from the plane \((h,k,l)\) to that from a single electron under identical condition, is a pure number and is in general complex. Hence we can write

\[
F_{hkl} = |F_{hkl}| \exp \left[i \phi_{hkl}\right] \cdot [2.3]
\]

where \(|F_{hkl}|\) is the amplitude and \(\phi_{hkl}\) the phase angle of \(F_{hkl}\).

Now, it is a fundamental drawback of X-ray diffraction experiments that we get only the intensities of the various reflections experimentally, so that one can obtain only the amplitudes \(|F_{hkl}|\) but
not the relative phases $\alpha_{kk'}$. To this extent, the building up of a picture of atomic arrangement from X-ray diffraction data alone is ambiguous. In the case of structures possessing a center of inversion the ambiguity in $\alpha$ is reduced to the two possibilities of either 0 or $\pi$. Even then the number of possible combinations with $n$ reflections (i.e., $2^n$) is so large that it is not possible to try them all out. Some other method of deducing the phase angles is necessary before we could evaluate the electron density distribution in the unit cell by Fourier methods. However, it would not be necessary to obtain the phases of all the reflections. It may be enough if we obtain the correct phases of the majority of the more important reflections by some method. Then it would be possible to make a preliminary Fourier synthesis using these phases and the experimental $|F_{kk'}|$ so that the main outline of the structure are obtained. Once the approximate atomic parameters are obtained it would be possible to calculate the phase angles of a larger number of reflections on the basis of these parameters and extend the synthesis using calculated phase angles and observed $|F|$ values and thus obtain the finer details of the electron density distribution in the unit cell by a method of successive approximations. Hence, it is essential in the application of Fourier methods, that to start with, we should by some method or other get the phases of a good number of the more important reflections. We shall now consider in sections 2-5 of this Chapter some of the methods used in obtaining the phase angles of reflections.
IL.2 DIRECT METHODS

There are few cases like the beautiful determination by Robertson (1936) of the structure of phthalocyanines, where the phases of almost all the reflections were obtained in a more or less direct way by comparing the intensities of the corresponding reflections from two isomorphous crystals with replaceable metallic atoms. The replacement of one atom by another causes the absolute values of the structure amplitude to be changed by a known amount but in a direction depending on the phase of the reflection so that by observing the direction of change from one crystal to the other it would be possible to determine the phases of almost every reflection and obtain a Fourier synthesis which at the first instance itself clearly gave the parameters of the atoms. As Robertson points out the analysis of this complex organic crystal is one of the most direct determinations based on X-ray methods alone where even the knowledge of discrete atoms have not been assumed.

Another favourable case when direct determination of phase angles is possible is when there is a single heavy atom at the center of inversion of a projected unit cell, the scattering factor of the heavy atom being far greater than that of any other atom in the unit cell. In such a case, we could reasonably assume that the phases of all the important reflections are determined by the contribution due to the heavy atom. For example, if the origin is chosen at the heavy atom then all the phases could be supposed to be positive and a preliminary Fourier synthesis with
all terms positive would usually indicate the positions of most other atoms. If the resulting co-ordinates change the signs of any of the reflections then a re-synthesis could be made using the new phase angles and this procedure repeated till re-calculation on the basis of co-ordinates obtained from Fourier synthesis do not change the phases of any more reflections. This method has been applied in the case of platinum phthalocyanine \( b \)-projection (Robertson & Woodward, 1940) as well as the \( b \)-projection of barium chlorate monohydrate (Part II).

II. 3 Vector Methods

By using an auto-correlation function of the electron density in the unit cell, Patterson (1935) has obtained valuable information regarding the interatomic vectors between the various atoms in the unit cell without any assumption regarding the phase angles. He defines the function

\[
P[u, \nu, \omega] = \frac{1}{V} \iiint_{0}^{c} \rho(x, y, z) \rho(x + u, y + \nu, z + \omega) \, dx \, dy \, dz
\]

which is a weighted average distribution of electron density about any point in the crystal and shows that this function, being equivalent to

\[
P[u, \nu, \omega] = \sum_{h, k, l}^{+\infty} |F_{h, k, l}|^2 \cos \left[ 2\pi \left( \frac{h}{a} + \frac{k}{b} + \frac{l}{c} \right) \right] (\nu, \omega)
\]

could be evaluated even in the absence of any knowledge regarding phase angles. This function \( P(u, \nu, \omega) \) which is called the 'Patterson function' has peaks at points corresponding to the interatomic vectors between atoms in the unit cell. But, when
there are a large number of atoms in the cell the number of peaks become very large and it is not directly possible to decide which pairs of atoms give rise to which vectors. However, this method is very useful when the unit cell contains a small number of heavy atoms which give really outstanding peaks that could be easily assigned. The fixing of the heavy atom positions by vector methods decides the phases of sufficient number of structure factors to give a recognizable picture of the atomic arrangement by a preliminary Fourier synthesis. It may be mentioned here that one difficulty in the determination of structures containing heavy atoms is that the intensities of reflections should be measured with a high degree of accuracy if the effect of the lighter atoms on the intensity is to be recognized over the swamping effect of the heavier atoms. Patterson has also given methods of removing unwanted peaks in the function as well as sharpening the vector peaks.

A very useful extension of Patterson's method was by Harker (1936) in the use of the symmetry elements of the crystal in obtaining some simplified sections of the three dimensional Patterson series given in (2,5). Here we get vectors between equivalent atoms and hence some of the co-ordinates of the atoms could be fixed directly since in these sections all other vectors are eliminated. Difficulties, however, may arise due to accidental coincidence of one or more co-ordinates of atoms not related by symmetry elements.
Since vector methods give an unprejudiced presentation of all the experimental diffraction results, the usual procedure in structure analysis is to first make the Patterson synthesis. This step is almost always found fruitful when there is a minority of heavy atoms in the structure. When this is not so the assigning of the vectors to the corresponding atoms is difficult, but comparison of vector diagrams of isomorphous substances, of photographs taken with two different X-ray wavelengths one on either side of the absorption edge of one type of atom in the crystal, etc. may give valuable clue. If these yield any information it may be sufficient in many cases to settle the phases of enough number of reflections to make a successful preliminary synthesis.

II.4 USE OF PHYSICAL AND CHEMICAL DATA

It is only in a very small number of cases that the above direct or semi-direct methods give all the necessary information. Hence, usually we have to fall back on trial and error methods based on physical and stereochemical data to obtain an approximate structure which can form the basis of Fourier refinement. Such an attempt is not as hopeless as might seem at first since in such methods we could make use of well established physical and chemical facts which none of the above methods take into account. With the knowledge gained by experience with related structures it becomes possible to tackle more and more complex substances.

Knowledge of space group and the symmetry of the molecule or ion might limit the number of possibilities. If the number of atoms
is smaller than the general position number, then they must be situated on the symmetry elements and should have the corresponding symmetry. If we know the molecular dimensions, the unit cell dimensions might restrict the possible arrangements of the pattern units. Impenetrability of atoms restrict the minimum distance at which they should lie from rotation axes or reflection planes to prevent overlapping with their equivalent atoms. If a few of the planes show outstanding intensities then it may be possible to live the approximate arrangement of molecules (whether in the form of sheets or rings) as well as their general orientation. Physical considerations like the packing of molecules, knowledge of bond lengths and bond angles, the condition of minimum potential energy of the arrangement, diffuse x-ray reflections, refractive indices and dispersion, rotation of plane or polarization, magnetic properties, pyro and piezo-electricity etc. might give valuable information regarding the general features of the structure; for, they may lead to knowledge of internal symmetry, to definite conclusions regarding the shape and orientation of the molecules or polyatomic ions or the distribution of forces between the various units. These considerations usually might be enough to enable us to postulate a trial structure which would give sufficient number of phases correctly to help us make a successful preliminary Fourier synthesis.

II. USE OF UNQUALIFIED

By use of more general methods than the heavy atom method
and the isomorphous series method it is sometimes possible to get the correct signs of some of the reflections. An important advance was made in this direction by Harker and Kasper (1948) when they showed that the use of certain mathematical inequalities may help to fix the signs of some at least of the reflections.

In their first paper they made use of Schwarz's inequality

$$\left| \int \mathbf{f} \cdot \mathbf{g} \, d \mathbf{r} \right|^2 \leq \left[ \int |\mathbf{f}|^2 \, d \mathbf{r} \right] \left[ \int |\mathbf{g}|^2 \, d \mathbf{r} \right]$$  \hspace{1cm} (2.6)

to obtain various useful relationships between the structure amplitudes for crystals which possess various symmetry elements.

For example, if for any reflection \((\mathbf{h}, \mathbf{k}, \mathbf{l})\) we postulate the quantity \(\hat{F}_{\mathbf{h}, \mathbf{k}, \mathbf{l}}\) given by

$$\hat{F}_{\mathbf{h}, \mathbf{k}, \mathbf{l}} = \frac{F_{\mathbf{h}, \mathbf{k}, \mathbf{l}}}{Z}$$  \hspace{1cm} (2.7)

then they obtained relationships of the form

$$\left| \hat{F}_{\mathbf{h}, \mathbf{k}, \mathbf{l}} \right|^2 \leq \frac{1}{2} + \frac{1}{2} \hat{F}_{2\mathbf{h}, 2\mathbf{k}, 2\mathbf{l}}^2$$

$$\left| \hat{F}_{\mathbf{h}, \mathbf{k}, \mathbf{l}} \right|^2 \leq \frac{1}{2} + \frac{1}{2} \hat{F}_{2\mathbf{h}, 0, 2\mathbf{l}}^2$$

$$\left| \hat{F}_{\mathbf{h}, \mathbf{k}, \mathbf{l}} \right|^2 \leq \frac{1}{2} + \frac{1}{2} \left[ -1 \right]^K \hat{F}_{2\mathbf{h}, 0, 2\mathbf{l}}^2$$  \hspace{1cm} (2.8)

For the lighter atoms that usually occur in organic crystals the assumption of similar atomic scattering factor curves helps us to obtain more stringent conditions. By defining the unitary structure factor \(U_{\mathbf{h}, \mathbf{k}, \mathbf{l}}\) for the plane \((\mathbf{h}, \mathbf{k}, \mathbf{l})\) by the equation

$$U_{\mathbf{h}, \mathbf{k}, \mathbf{l}} = \frac{F_{\mathbf{h}, \mathbf{k}, \mathbf{l}}}{Z} = \sum_{\mathbf{g}} \gamma_{\mathbf{g}} \exp \left[ 2\pi i \left( \frac{\mathbf{h} \cdot \mathbf{g}}{a} + \frac{\mathbf{k} \cdot \mathbf{g}}{b} + \frac{\mathbf{l} \cdot \mathbf{g}}{c} \right) \right]$$  \hspace{1cm} (2.9)

as a sum instead of as an integral as in (2.1) more powerful
inequalities could be obtained. In (2.9) \( \eta_j = \frac{3\gamma_j}{2} \) is the fraction of the electrons in the unit cell in the \( j \)-th atom and \( \hat{f} \) is the scattering factor of the atom divided by its atomic number. Now, Cauchy's inequality which holds for sums just as Schwarz's does for integrals could be written as

\[
\left| \sum_{j=1}^{N} a_j d_j \right|^2 \leq \left( \sum_{j=1}^{N} |a_j|^2 \right) \left( \sum_{j=1}^{N} |d_j|^2 \right)
\]

(2.10)

Using these, relations of the form

\[
\left| U_{h\kappa \ell} \right|^2 \leq \frac{1}{\kappa} + \frac{1}{\ell} \left| U_{2\kappa, 2\kappa, 2\ell} \right|
\]

\[
\left| U_{h} \pm U_{\kappa} \right|^2 \leq \left( 1 \pm U_{h} \right) \left( 1 \pm U_{h} \right)
\]

(2.11)

etc. could be obtained for crystals possessing a center of symmetry. When other symmetry elements are present more restricted conditions could be obtained.

Gillis (1948) has derived many types of structure factor inequalities using other mathematical inequalities.

For e.g:-

a) Hölder's inequality

If \( p, q \) are real and \( \frac{1}{p} + \frac{1}{q} = 1 \) then

\[
\left| \sum_{\gamma=1}^{n} a_{\gamma} d_{\gamma} \right| \leq \left\{ \sum_{\gamma=1}^{n} (a_{\gamma})^p \right\}^{\frac{1}{p}} \left\{ \sum_{\gamma=1}^{n} (d_{\gamma})^q \right\}^{\frac{1}{q}}
\]

(2.12)

b) If \( x_1, x_2, \ldots, x_n \) are positive, \( \xi_1, \xi_2, \ldots, \xi_n \) and \( \xi_1, \xi_2, \ldots, \xi_n \) are non-negative numbers such that \( \sum_{\gamma=1}^{n} \xi_\gamma = 1 \) then if

\[
M_\alpha = \left[ \sum_{\gamma=1}^{n} \xi_\gamma \right]^{\frac{1}{\alpha}}
\]
then \[ M_\alpha \leq M_\beta \quad \text{if} \quad 0 < \alpha < \beta \quad . \quad \text{(2.13)} \]

If \( 0 < x_i < 1 \), \( \xi > 0 \), \( \theta = 1, 2, 3 \)
then \[ \sum_{i=1}^{n} \xi \cdot x_i^\alpha \geq \sum_{i=1}^{n} \xi \cdot x_i^\beta \quad \text{(2.14)} \]
for \( \alpha < \beta \)

Using these mathematical inequalities Gills obtain structure factor inequalities of the form

\[ \left| U_{3H} + 3 U_H \right|^3 \leq \frac{1}{\alpha} (1 + U_{2H})(3 + 4 U_{2H} + U_{4H}) \]
\[ \left| U_{3H} + 3 U_H \right| \leq 2 (1 + U_{2H}) \quad \text{(2.15)} \]

 Recently Okaya and Nita (1962) have used the linear inequalities

\[ \sum_{i} \gamma_i \cdot a_i^1 + \sum_{i} \gamma_i \cdot b_i^2 \geq 2 \left| \sum_{i} \gamma_i \cdot a_i \cdot b_i \right| \quad \text{(2.16)} \]

in some structure determinations. From (2.16) they have derived a series of inequalities some of which are given below:

\[ 3 + U_{2H} \quad \geq \quad 4 |U_H| \]
\[ 2 + U_{2H} \quad \geq \quad 2 \sqrt{2} |U_H| \]
\[ 2 + U_{2H} + U_{2H'} \quad \geq \quad 2 |U_H + U_H'| \]
\[ 5 + 4 U_{2H} + U_{2H-H'} \quad \geq \quad 4 |U_H + U_H'| \quad \text{(2.17)} \]

They point out that these inequalities contain only linear terms of structure factors as against the higher order terms of Marker Kasper inequalities and hence these are easier to use.

Obviously one can multiply examples like this, but it is only by practice that one is able to choose the right inequality.
for the particular problem. Examples of the application of these methods are to be found in the case of the determination of the signs of the reflections in the case of oxalic acid dihydrate by Gillis (1948); orthorhombic decaborane by Kasper, Lucht and Harker (1950) etc. Gillis remarks that in many instances where both the signs satisfied the inequality, one by a larger margin than the other, the former was always found to be the correct one, suggesting that the method has ample 'reserves of power'. However, Hughes (1949) has shown that the power of inequality methods is reduced if the number of atoms in the unit cell is very large. He proved that in the case of a large number \( N \) of similar atoms in the unit cell, the intensities of the various reflections fall on a smooth curve, with a peak at \( \left( \frac{1}{|U_{4,4,4}|^2} \right)^{1/4} = \frac{1}{\sqrt{N}} \) and that the mean value of the unitary structure factor is given by

\[
\sigma = \left( \frac{1}{|U_{4,4,4}|^2} \right)^{1/4} = \frac{1}{\sqrt{N}} \tag{2.18}
\]

From this it is seen that the number of reflections whose intensity exceed 1.7 \( \sigma \) will be only about 10% of the total number of reflections. Now, since the inequalities are usually of the form \( U \leq A + Bu \) they would be of use only if the \( U_{4,4,4} \) are large enough, so that this method of finding the signs would work only for planes having large values of structure factors. To make a preliminary Fourier synthesis which could give the main outlines of the structure it is necessary that the signs of at least 30% of the more important reflections must be fixed. The complicated inequalities of Gillis are more powerful, but they require that at least some of the signs are known earlier even though it might be possible to arbitrarily assign the signs of one or two of the
reflections depending on the choice of the origin. Grisón (1951) has given a table describing the straightforward process by which the Harker-Kasper inequalities may be applied to any practical structure analysis.

In view of the apparent limitations of the inequality methods, the use of equalities or identities have been suggested by some authors. Very recently Zachariasen (1952) describes what he calls "a new analytical method of solving complex crystal structures" in which he uses the identity

\[ \sum a_3 a_k |^2 = \sum |a_3|^2 \sum |a_k|^2 - \frac{1}{2} \sum \sum |a_3 a_k|^2 \] (2.19)

for a crystal possessing a center of symmetry. Using (2.19) he obtains interesting equalities which he has used in the determination of the structure of \( \text{HgO}_3 \). He remarks that though it was not possible to obtain definitely the signs using the various other methods, the use of equalities solves the problem in a few days. In his paper he has also considered the relationship of this method to other methods like the Sayre's squaring method (1956). It may be mentioned that another method of using identities in the solution of atomic parameters was suggested by Banerjee as early as 1933 and has recently been extended by Hughes (1949) in calculating the signs of \( h\kappa\ell \) reflections correctly in the case of dicyanamide.

II.6 THE REFINEMENT OF ATOMIC CO-ORDINATES

By using some or all the methods described above, it will be possible in most cases to obtain the parameters of the various atoms in the unit cell to an accuracy which would be enough for
the assignment of chemical structure and in stereochemistry. But, the verification of the extension of quantum mechanical treatment to molecular systems by Heitler and London, Pauling and others necessitates more accurate data regarding bond lengths, angles etc. than were previously needed. By the extension of three dimensional Fourier methods in conjunction with rapid computational techniques it is now not very difficult to fix atomic positions with an accuracy of ±0.01 Å in a moderately complex crystal.

At the outset it is clear that the accuracy of determinations based on projection alone is limited and hence it is necessary to go to three dimensional syntheses using all observable reflections with the corresponding many fold increase in the labour of computation. In three dimensional series the density

\[
\rho(x, y, z) \text{ at the plane } z = z_1 \text{ is given by }
\]

\[
\rho(x, y, z_1) = \frac{1}{V} \sum_{k} \sum_{l} \sum_{\epsilon} |F_{k\ell\epsilon}| \cos \left[ 2\pi \left( \frac{kx}{a} + \frac{ly}{b} + \frac{z_1}{c} - \alpha_{k\ell\epsilon} \right) \right] (2.20)
\]

so that the summation has to be done for various points \((x, y)\) over the plane \(z = z_1\). For the various space-groups Londale's "Structure factor tables" (1936) help one to put the series in forms best suited for computation.

The labour of computation of complete three dimensional sectional synthesis could be considerably reduced without much reduction in accuracy by the methods of 'Section projections' and 'Projected sections' suggested by Booth (1946). In the method of section projection - or the method of 'Sounded projection' - all
the scattering material between the planes $z = z_1$ and $z = z_2$
are projected on to the c-face giving a quantity $B_{z_2}^{z_1}$ defined by

$$B_{z_2}^{z_1}(x, y) = \int_{z_1}^{z_2} f(x, y, z) \, dz$$

To calculate this projection $B(x, y)$ one requires all the observable reflections and hence this is more laborious to calculate than the Bragg projections where only zonal reflections are used. But, the bounded projection is very useful in getting a projection of a molecule free from overlap due to other molecules in the unit cell. This is especially so if the single molecule is planar and gives a projection free from overlapping. By judiciously choosing the orientation and the thickness of the two bounding planes it may be possible to obtain projections free from overlap or distortion, thus enabling high resolution to be achieved. In the method of projected sections the sections through atoms at, say, $z_1 - z_3$, are combined to give a sum

$$R(x, y; z_1, z_3, z_n) = \sum_{n=1}^{N} f(x, y, z_n)$$

This could be computed with an ease not much inferior to that of computing a single three dimensional section $f(x, y, z_n)$ but at the same time giving information regarding the co-ordinates of a number of atoms with about the same accuracy that is obtainable by 'n' different sections.

As mentioned earlier, the process of Fourier refinement proceeds in a number of stages of successive approximations using experimental $|F_{hkl}|$ values and phase constants $\alpha_{hkl}$ obtained from the co-ordinates of postulated approximate structures. In
practically, three-dimensional methods are worthwhile only after
the refinement has proceeded to the maximum possible extent
using the much simpler two-dimensional Bragg projections. This
is especially true if the structure does not possess a center
of inversion so that the ambiguity in the general \(\alpha_{k\ell m}\) is very
large, but in the case of projections there might be a center
of inversion making the computations much simpler.

When one needs only the position of maximum electron
density and not the actual density distribution throughout the
unit cell it is not necessary to perform the complete Fourier
syntheses. The labour of computation can be very much reduced
by using Booth's method of differential synthesis and refine-
ment (1946). The position of maximum electron density is fixed
by the condition that at the point \((x_v, y_v, z_v)\) at which it is
maximum the derivatives of \(\rho(x, y, z)\) with respect to the co-
ordinates must vanish, i.e.,

\[
\left( \frac{\partial \rho}{\partial x} \right)_{x=x_v} = \left( \frac{\partial \rho}{\partial y} \right)_{y=y_v} = \left( \frac{\partial \rho}{\partial z} \right)_{z=z_v} = 0 \quad (223)
\]

The conditions (2.23) could be put in the form of 'three linear
equations', the solution of which for each atom \(A\) gives the
necessary refinements \(\xi_x, \xi_y, \xi_z\) to the original co-
ordinates \(x, y, z\). The assumption of spherical symmetry
of atoms further reduces the labour of calculation of \(\xi_x, \xi_y, \xi_z\).
Booth has also given method of refining the phase angles \(\alpha_{k\ell m}\)
also at the same time as the atomic coordinates so that the
refinement by this method of differential synthesis could proceed
by a series of successive approximations. This method is particularly helpful in the analysis of structures possessing no center of inversion.

II.7 COMPUTATION TECHNIQUES

The extensive numerical work involved in the evaluation of multi-dimensional syntheses has been put into manageable form by the strip method of summation by Beavers and Lipson (1936). In the strip method, evaluation of a multi-dimensional series is reduced to that of a series of single summations. By the use of the recent strips (Beavers, 1952), which have three figure accuracy, it is possible to sum the series at points corresponding to intervals of 1/120-th the cell dimensions. The use of Robertson's sorting board, though it gives a higher accuracy, is more tedious than the strip method. The two dimensional phase factor tables of Beaucclair and Sincoowitz (1949) enable the summation to be done without having to split the two dimensional series into those of one dimension, but it has the disadvantage that the net covered by this method is too coarse for most practical purposes. The use of punched card methods (Shaffer, Schomaker & Pauling, 1946), which is a mechanization of the strip method, enable the syntheses to be made at much finer intervals (1/500-th unit cell) with much higher accuracy and speed and lesser chances of computational errors. The same method has also been used at Cavendish, but at intervals of only 1/120-th the unit cell. The details of extension of the punched card method to two dimensional cases have been worked out by the author (Chapter V) and the necessary
Electro-magnetic, but the method could not be completed because of financial difficulties involved in the production of the large number of master cards necessary. Various other ad hoc machines, mechanical, electro-mechanical and electronic, have also been used for Fourier syntheses in various laboratories. During the preliminary stage, the optical summation using the Bragg-Huggins masks, is very useful and has the advantage of extreme simplicity. The same principle of the combination of two dimensional harmonic distributions is done electronically in the analogue computer of Pepinsky with the additional advantage that the syntheses could be made within a few minutes which is the time necessary to feed in the data in the form of the amplitudes and phases of the reflections. The picture could be made to appear on the screen of the Cathode ray tube either intensity modulated or preferably with contour lines. The adjustment of unit cell shape and size could also be easily done. The ease and quickness with which various combinations of phases could be tried out and the result observed on the screen makes this method one of extreme elegance.

The calculation of the structure factors of the various reflections at each stage of the parameter refinement and comparison with the observed values also involves much labour. The use of strip method - with the help of the structure factor expressions given in Lonsdale's tables - considerably simplifies the labour; but since one can use only values of atomic parameters at discrete intervals the method is not capable of very high accuracy in the general case. However, the author has found in practice that the
use of linear combinations of strips could give the necessary accuracy in most cases when the atomic positions do not coincide exactly with the points of the net covered by the strip and this procedure has very recently been given by Beevers (1952). The use of Bragg - Lipson charts and Beevers - Lipson figure fields are of much use in the preliminary stages of analysis by trial and error because they show in a direct manner the effect of moving an atom in a known direction on the values of the various structure factors. This is particularly of much advantage when the structure has a high symmetry involving a large number of equivalent atoms. The fly's eye method suggested by Bragg is simple and rapid and could be used with advantage during the preliminary stages of analysis but is not capable of much accuracy. Recently, Bragg (1952) has described a compact and robust structure factor calculating machine where pivoted dials, one for each atom, cause corresponding moments and the sum of these is obtained by restoring the balance with an applied torque. The use of punched card method has the advantage that it is not adhoc and any accuracy is possible, but the equipment needed is elaborate and costly and not available to every worker. Various other adhoc machines have been used by Cox, Vand, Booth etc. as well as the crystallographic slide rule by Bunn for the calculation of structure factors. However, none of these seem to have yet acquired universal use, each author finding his own method best suited to him.

II.3 ERRORS IN FOURIER SYNTHESIS AND ACCURACY

The crystal structure determination is based on the
determination of intensities of as many planes as possible with a high degree of accuracy and the use of these intensity data in obtaining a picture of the distribution of electron density in the unit cell. The actual intensity data might be subjected to errors of two principal types viz.,

1) Systematic errors due to absorption and extinction from use of unsuitable crystal sizes and shapes.

ii) Random experimental errors.

Once the intensity data are obtained the usual procedure is to perform a Fourier synthesis after obtaining the phases in some way or other. In this process also certain errors may creep in, being not just computational errors, but errors due to the nature of the mathematical procedure. Apart from the errors arising due to rounding off during computation (Cochran, 1948) the most serious error which occur in the application of Fourier methods is due to the termination of the series. Obviously, to get an exact representation of the electron density distribution by Fourier synthesis we must take all the non-zero terms of the Fourier series. However, experimental difficulties set a limit to the \( \frac{\sin \theta}{\lambda} \) range in which we can observe the reflections, thus making it necessary for us to neglect the coefficients of the series beyond this range. In general it may be mentioned that the maximum error that can be introduced due to the 'termination error' is equal to the last term that has been included in the series, provided the terms
on an average become progressively less and less. The effects
of this error are of two kinds. Firstly, it may produce a distor-
tion of electron density distribution and secondly it may cause
a shift of the position of the maximum. A third one is in fact
related to the above two, namely that the resolving power is
limited by the termination of the series.

The problem of ultimate accuracy of the final structure
is intimately connected with the influence of these errors. From
optical analogy (Bragg & West 1930) it would seem that the resolution
would be limited to \((0.66 \lambda / 2 \sin \theta)\) which is usually of the order
of 0.5 Å. However, purely from a comparison of calculated and
observed structure factors, most crystallographers agree that the
atomic peaks could be fixed to a much higher degree of accuracy —
say, 0.05 Å or even 0.02 Å. The apparent contradiction can be
resolved as follows:— The resolving power of 0.5 Å only means that
if there are two atoms closer than that, then their peaks would
not be shown as separate. However, if the atoms are separated
by much greater distances so that no appreciable overlapping
occurs, then the positions of the peaks could be fixed to a very
high degree of accuracy from the intensities of high order re-
fections. Hence, it may be mentioned that when there is no over-
lap the finite series errors have little effect on the position
of the atomic peaks but is much more important on their effect
on the electronic charge distribution of the atoms. But, when
there are other atoms near enough, the peak positions of one
might be affected by the diffraction effects of another.
Using these qualitative ideas, Booth (1946, 1947, 1949), has given a mathematical theory of the influence of various errors in the accuracy of atomic co-ordinates derived from Fourier synthesis. To start with, he assumed that the spherically symmetrical atoms had a Gaussian distribution of charge density which could be represented by

\[ \varphi(r) = A e^{-\frac{Pr^2}{2}} \]

\[ Z \left( \frac{\varphi}{n} \right) e^{\alpha P} \left( -P \varphi^2 \right) \]

with \( P = 4.639 \)

If the probable error in \(|F_o|\) values is \( \Delta \varepsilon \), the standard deviation \( \sigma(|F|) \), and the error of a particular \(|F|\) value is \( \varepsilon \), we have the relations:

\[ \sigma = \sqrt{\frac{\sum \delta^2}{n-1}} \approx \sqrt{\frac{\delta^2}{n}} \]

and \( \Delta \varepsilon = 0.6745 \sigma(|F|) \)

Booth showed that if \( \varepsilon_x \) is the error in the co-ordinate \( x \) of the atom, then

\[ \varepsilon_x < \frac{90(\Delta \varepsilon)}{Z(\lambda P)^{1/2}} \]

provided the maximum values of \( \lambda, k, \lambda \) used in obtaining the Fourier synthesis are greater than 10. Later, Booth gave a more precise formula where

\[ \sigma(x) = 90 \sigma(|F|) / Z \left[ \sqrt{(\lambda P)^2} \right]^{1/2} \]

For a particular case the value of \( \sigma(x) \) comes out to be of the order of 0.004 \( \AA \). Other workers, notably Cruickshank (1949) gave slightly different formulae for the value of \( \sigma(x) \) which are in general larger by a factor 2. In any case the errors in parameters are all less than 0.01 \( \AA \).
The same theory also gives the standard deviation $\sigma(p)$
of the electron density. They are

$$\sigma(p) = \sigma(F) \left( \frac{2\pi^2}{3\sqrt{3}} \right)^{1/4} \propto A^{0.3}$$

(A 21)

and

$$\sigma(p) = \sigma(F) \frac{\pi}{\lambda} \sqrt{\frac{\pi}{A}} \propto A^{0.2}$$

for three dimensions and two dimensions respectively so that for
a cubic unit cell of side $10 \lambda$ the errors become

$$\sigma(p) = 0.10 \ \text{el.} \ \lambda^{-3}$$

and

$$\sigma(p) = 0.23 \ \text{el.} \ \lambda^{-2}$$

respectively for the two cases. Thus we find that the errors
introduced by experimental errors are practically negligible.

To test these theories, the case of dibenzyl, which has been
studied by two independent workers; one an early study by Robertson
(1935) and another a more accurate study by Jeffrey (1946) was used.
Comparison of the two data showed that even though $\sum |F_y| - |F_R|$ was only 0.05, for any particular author, the quantity $\sum |F_o| - |F_R|$ was about 0.15; i.e., the two independent experimental data agreed
remarkably with each other, but the calculated $|F_c|$ values do not
agree so well with either set of data. Hence it is clear that this
cause is not very important. However, one must remember that the
$|F_c|$ values are obtained on the basis of ideal spherically
asymmetrical atoms and this might explain a large part of the dis-
agreement between the observed and calculated values of the structure
amplitudes.

Booth has also considered the effects of termination error
in Fourier series and has obtained formulae for the mean error $\epsilon$
due to this cause. He finds that the value of \( \epsilon \) rapidly increases as \((\sin \theta / \lambda)\)range over which observation is made is decreased, being about 0.005 \( \AA \) for \( 2 \sin \theta = 2.0 \) and 0.02 for \( 2 \sin \theta = 1.5 \). These are only orders of magnitude, but it is clear that the error due to this cause is larger than those due to experimental errors.

An experimental investigation of these problems was made by Robertson and White (1947). They used a hypothetical structure of two rings and two close atoms and used the \(|F_c|\) values of this structure, after aplying the various possible types of errors, in re-synthesizing the structure. Comparison of the original hypothetical structure with the re-synthesized structure showed that in general the termination errors are of more importance than the experimental errors, say they be random or even systematic in nature.

II.9 CORRECTIONS FOR TERMINATION ERRORS.

Since the termination errors are the most important, many methods have been suggested to get over them or to obtain corrections for these effects. Since termination errors occur due to the slow convergence of the series and are of the order of the last term included in the series, one of the first attempts to get over these errors was to annihilate the high order terms by applying a convergency factor to the observed structure factors. Since the effect of the convergency factor is similar to that of thermal motion this method is known as the 'artificial temperature factor method'. Values of convergency factors equivalent
to about 5000 K was used by Brill, Grimm, Herman and Peters (1939) in the case of diamond. Even though the diffraction effects disappear we are sacrificing a lot of knowledge by thus decreasing the weight of the high order reflections by the use of this method and hence the electron density distribution becomes very diffuse and the error in determining the atomic peaks is consequently increased.

Another interesting method of eliminating diffraction effects has been suggested by von Reijen (1942). The termination error being due to regarding $F$ values over regions beyond $\sin \theta = 1$ to be zero, he used a suitable curve of extrapolation and assumed the $F$ values on an average to follow this curve for values beyond $\sin \theta = 1$. By using simple extrapolation function $F^*$ (such as $Ae^{-\alpha s}$, $Be^{-\beta s}$) whose Fourier transform $\hat{f}^*(x, y, z)$ from $-\infty$ to $+\infty$ could be easily evaluated analytically, he showed that $\hat{f}(x, y, z)$ corrected for termination errors is given by

$$\hat{f}(x, y, z) = \hat{F}^*(x, y, z)$$

$$+ \frac{1}{V} \sum_{-h}^{+h} \sum_{-k}^{+k} \sum_{-l}^{+l} (F^* F) \cos \left[ \frac{2\pi}{a} (h \frac{x}{a} + k \frac{y}{b} + l \frac{z}{c}) - \alpha_{hlk} \right]$$

Since in this summation only terms up to the maximum values of $h, k, l$ observed need be used, the labour of computation is no more than in an ordinary synthesis and at the same time diffraction effects are avoided. However, the method has the disadvantage that we are using an arbitrary function in extrapolation so that the finer details of the electron density distribution...
at the peak would depend on the nature of this function.

Booth has suggested a different method known as the 'back-correction method' for correcting the termination errors. In this method a re-synthesis is made using the $F_c$ values obtained for the computed structure but terminated at the same region as $F_0$. The difference $\epsilon_x^2$ between the $x$ co-ordinate of atom $\gamma$ in the first and second computed structures is applied in the opposite direction to obtain the corrected co-ordinate. This method is reasonably accurate for finding the atomic co-ordinates so long as the correction is not large. It is not good enough if the electron densities are to be corrected since the back correction in this case would depend on the particular electronic charge distribution of the atom.

In the case of crystals where there are a few heavy atoms, the error would be mainly due to the diffraction effects around the heavy atoms. The present author has used a simple method in such cases. It is possible to obtain the positions of the heavier atoms accurately from vector diagrams or even by the minimization methods. Hence we could extend the series by using $F_c$ values calculated from the positions of heavy atoms alone which are known accurately. Since in many cases the contributions of the lighter atoms in the range beyond $(\sin \theta/\lambda) = 0.8$ or $1.0$ would be negligible this assumption is allowable for extension of the series in this range. This method has been used to overcome the diffraction effects of the heavier atoms in the structures studied by the author (Part II). This method, which is similar
to von Reijen's method, has the disadvantage that more terms have to be introduced into the series thus making the computation more elaborate, but has the advantage that the extrapolation function used is not arbitrary but uses the theoretical scattering factor curves for the free atoms.

II. 10 MINIMIZATION METHODS

A fundamental disadvantage of the Fourier method is that no single non-zero $F$ factor must be omitted from the series used to evaluate $\rho(x, y, z)$. In trial and error methods it is not necessary to take every single reflection into consideration. If for a sufficiently large number of reflections the observed and calculated structure amplitudes show good agreement the validity of the proposed structure may be warranted. Another disadvantage of the Fourier method is that we could not use the experimental data in a weighted form. Even if we are sure that the $|F_o|$ values of some of the reflections are largely in error due to extinction etc., and we are unable to make detailed corrections, still we have to give the same weight for all the reflections in performing the Fourier synthesis. Further, as the structure analysed becomes more and more complex and the number of variable parameters increase correspondingly the labour needed in obtaining a trial structure to commence the process of Fourier refinement also increases.

In view of these defects of the Fourier method, the problem of crystal structure analysis could be posed in a slightly
different manner (Booth, 1949). We know the geometry of the unit cell and the number of different kinds of atoms in the unit cell as well as the atomic scattering factor $f_{\gamma}$ of atom $\gamma$ and the way in which they fall off with increasing scattering angle. From experimental data we obtain a set of numbers $F_{0 + k \ell}$ for each plane $(h, k, \ell)$. These $F$'s are the moduli of the number $F_c(h, k, \ell)$ defined by

$$F_c(h, k, \ell) = \sum_{\gamma=1}^{N} f_{\gamma} \exp \left[ 2\pi i \left( \frac{2h x_{\gamma}}{a} + \frac{2k y_{\gamma}}{b} + \frac{2\ell z_{\gamma}}{c} \right) \right] \quad (2.29)$$

Now, the problem of structure analysis is to determine the coordinates $(x_{\gamma}, y_{\gamma}, z_{\gamma})$ of atom $\gamma$ such that the function $R_\gamma$ defined by

$$R_\gamma = \sum_{k} \sum_{k} \sum_{\ell} G \left[ g(F_0) - g(F_c) \right] \quad (2.30)$$

is a minimum, the $G$ and $g$ being arbitrary weight functions. A solution which satisfies the minimization of (2.30) and which is probable from stereochemical and physical data could be taken as the correct structure. Put in this way the problem of structure analysis reduces to a minimization problem.

Several types of minimization functions may be chosen. The first one, was the so called 'least square method' (Hughes, 1941) where the function

$$R_1 = \sum V (F_0 - F_c)^2 \quad (2.31)$$

is minimized. Booth (1949) has suggested the functions

$$R_0 = \sum V |F_0| - |F_c| \quad (2.32)$$

and

$$R_2 = \sum V (F_0^2 - F_c^2)^2$$
and has shown how these could be minimized by the method of 'steepest descents'. In these cases the weighting factor \( W \) is introduced to allow for the fact that not all \(|F_0|\) values are observable with the same degree of accuracy. Booth has also shown that the three methods, viz., the Southwell's relaxation method, the Hughes's least square method and Booth's steepest descent method are all variations of the minimization method depending on the way in which the function to be minimized is expanded in terms of the atomic co-ordinates. He has given the actual practical procedure in the application of the steepest descent method and has pointed out its advantages over the other minimization methods as well as the Fourier methods.

He has also suggested the use of the steepest descent method to refine not only the atomic parameters but also the scale factor, the temperature factor, the atomic scattering factor etc. also simultaneously. (More details of these procedures will be given as application arises in Part II). Quarrth (1949) has investigated the optimum conditions for the convergence of steepest descents as applied to structure determinations. This method can be put in a form suitable for mechanical computation.

It must be emphasized that the minimization method could give only the best atomic co-ordinates, i.e., the points where the electron density becomes a maximum and around which the electron density is spherically symmetrical over a small region. Further, we have to make use of the theoretical structure factor expressions for free spherically symmetrical atoms calculated on the basis of wave mechanics in evaluating the minimization
functions. If it is desired to investigate the actual electron density distribution and not the position of the peak alone, then the Fourier method, and not the minimization method or the differential synthesis method, is the only one that could be used.

Further, Cochran (1948) has shown that if the Fourier method is carried out to the limit, i.e., with no termination errors, but with experimental errors, then the atomic positions that we have obtained are those corresponding to minimization of the function

\[ R_c = \sum \frac{1}{f_i} (F_o - F_c)^2 \]  \hspace{1cm} (2.33)

so that it is a least squares method where large angle reflections are given greater weight. Using this fact he has suggested a simple method by which weight could be given to the coefficients \( F_{hkl} \) of the Fourier series. One consequence of this is that when we have reason to doubt that the experimental values \( |F_o| \) of some reflections are in much error we could adopt the following procedure: First a Fourier synthesis using the observed data is obtained. The structure factors are then calculated and a second synthesis is made using \( WF_o + (1 - W)F_c \) for the doubtful structure amplitudes. In fact, one may even take the doubtful values to be \( F_c \), i.e., \( W = 0 \), without much error if these are very few in number as has been done by the present author for a very few reflections in his synthesis of the b-projection of barium chlorate monohydrate. In any case as the refinement proceeds the weight factor \( W \) of these reflections
could be reduced to zero without causing any appreciable error.

Taking these into account and remembering that in the Fourier method we do not even assume the theoretical atomic scattering factor, in the final analysis, Cochran comes to the conclusion that, provided termination errors are corrected for, the Fourier method is superior to the minimization methods.

II.11 THE FAITHFULNESS OF THE RESULT OF X-RAY ANALYSIS TO THE TRUE STRUCTURE

Since the experimental errors produce negligible errors in the positions of peaks and since the termination errors could be got over by any of the minimization methods, it is always possible to reduce the final error in atomic co-ordinates to less than 0.01 Å. This corresponds to an error of one or two in thousand for the atomic co-ordinates \((x, y, z)\) when the unit cell dimensions are of the order of 10 Å. However, the same cannot be claimed about the electron density distribution since only the Fourier series method gives this and the unavoidable finiteness of the series causes errors. It is obvious that no amount of computational jugglery can compensate for the fact that we are not observing the reflections beyond a finite \((\sin \theta/\lambda)\) range and to that extent our final picture of the electron density will be lacking in detail.

Still, the various methods that have been suggested to get over the termination errors have varying degrees of success. The use of artificial temperature factor, eventhough it gives
satisfactory electron counts under the peak and removes regions of negative electron density, results in the smearing out of all the finer details of the electron density distribution. The extrapolation method of van Heijen does eliminate the diffraction effects, but it is the arbitrary extrapolation function which determines the actual value of the electron density distribution at the peak.

However, the method of \((F_0-F_e)\) synthesis suggested by Cochran (1961) seems to be comparatively satisfactory and is useful after the atomic peak positions have been obtained to a very high degree of accuracy. This method depends on the fact that the \(\sum (F_0-F_e)\) is more due to the assumption of spherical atoms in evaluating \(F_e\) (an assumption which is not warranted because of the distorting influence of neighbouring atoms) than due to the experimental errors. Since the outer electrons are affected by the presence of neighbouring atoms in the crystal, the departure from the sphericity causes the difference in \(\sum (F_0-F_e)\). Hence, the synthesis of \(\Delta \rho\) using \((F_0-F_e)\) given by

\[
\Delta \rho(x,y,z) = \sum (F_0-F_e) \cos \left[ 2\pi \left( \frac{x}{a} + \frac{y}{b} + \frac{z}{c} \right) - \alpha_{\text{crystal}} \right]
\]

where the summation is over all observed \(F\)'s, gives the difference between the actual electron density distribution in the crystal and that given by the final structure. Cochran has applied this method and plotted contours of \(\Delta \rho\) and found that it has in general a maximum of 1 el. \(Å^{-2}\). Obviously the final distribution \(\rho(x,y,z)\) will be independent of diffraction errors also.

The \((F_0-F_e)\) synthesis of Cochran is identical in principle
to the extrapolation method of von Reijen, but has the advantage that the extrapolation function used being dependent on $P_0$ is based on the scattering factor of the idealized spherical atom. Since the scattering for high values of $(\sin \Theta / \lambda)$ would be mainly due to the central core of electrons in the atom this method gives the outer electron distribution accurately, but may lead to erroneous peak values.
REFERENCES

1. Albrecht, G., ...  ...  ... (1939) Rev. Sci. Instrum. 10, 221.
3. ...  ...  ... (1929) Proc. Roy. Soc. A 123, 575.
4. Banerjee, K., ...  ...  ... (1933), Proc. Roy. Soc. A 141, 192.
10. Booth, A.D., ...  ...  ... (1946) Trans. Farad. Soc. 42, 444.
11. ...  ...  ... (1946) Trans. Farad. Soc. 42, 617.
22. Buerger, M.J., ...  ...  ... (1942) X-ray Crystallography, New York. John Wiley.
24. ...  ...  ... (1951) Acta. Cryst. 4, 403.
28. Grison, E., ... ... (1951) Acta. Cryst. 4, 43.
29. Harker, D., ... ... (1936) J. Chem. Phy. 4, 381.
31. Hendershot, O.P., ... ... (1937) Rev. Sci. Instrum. 8, 324.
34. Hughes, A.W., ... ... (1941) Jour. Am. Chem. Soc. 63, 1737.
35. ... ... (1949) Acta. Cryst. 2, 34, 37.
37. Jay, A.H., ... ... (1941) J. Sci. Inst. 18, 128.
41. Lonsdale, K., ... ... (1936) 'Structure factor Tables', Bell and Sons.
42. Lonsdale, K., ... ... (1949) Acta. Cryst. 2, 42.
43. Okaya, Y., & Sitta, I., ... ... (1952) Acta. Cryst. 5, 564.
44. Patterson, A.W., ... ... (1935) Z. Kristallogr. 90, 517.
47. Renniger, J., ... ... (1937) Z. Kristallogr. 72, 95, 107.
49. ... ... (1936) J. Chem. Soc. 1195.
50. ... & Woodward, I., ... ... (1940) J. Chem. Soc. 36.

52. Robertson, J.M., ... ... (1943) J.Sci.Inst. 20, 175.


54. Sayre, D., ... ... (1952) Acta. Cryst. 5, 60.


56. van Heijen., ... ... (1942) Physica. 9, 461.

57. Wilson, A.J.C., ... ... (1942) Nat. 150, 151.

58. Wooster, W.A., ... ... (1936) Z.Kristallogr. 94, 376.