

The Use of a High Speed Automatic Calculator in the Refinement Stages of Crystal-Structure Determinations

BY T. R. THOMPSON, D. T. CAMINER AND L. FANTL

Lyons Electronic Office, Cadby Hall, London W. 14, England

AND W. B. WRIGHT AND G. S. D. KING*

The Lyons Laboratories, Hammersmith Road, London W. 14, England

(Received 5 October 1953 and in revised form 24 October 1953)

The differential Fourier method and the method of steepest descents have been adapted for iterative refinement of centrosymmetric crystal structures using the high speed automatic calculator built by J. Lyons and Company Limited. The methods have been applied to the refinement of the crystal structure of nicotinamide, which has 9 crystallographically independent atoms and 939 observed hkl reflexions. The differential refinement for a structure of this complexity takes 35 min. per iteration and the steepest-descents refinement 40 min., including the recording of the individual calculated structure-factor values on magnetic tape.

Introduction

The problem of refinement of crystal structures is mainly one of computation. Refinement can be said to begin when approximate atomic positions have been found and the phases of the most important structure factors are known. The best known methods available for refinement are:

1. The Fourier method.
2. The differential Fourier method.
3. The $(F_o - F_c)$ Fourier method.
4. The method of steepest descents and least squares.

The first two methods have the advantage that they require only the phases of the calculated structure factors. As the magnitudes of these quantities need not be determined, it is unnecessary to have the observed structure factors on an absolute scale or to know the temperature factor. It is therefore desirable to use one of these methods until the structure is sufficiently refined for a reliable temperature factor to be determined. These methods, however, do not necessarily give the 'correct' structure, which may be defined as the structure that gives the best agreement between observed and calculated structure factors. In order to obtain the 'correct' structure it is necessary, therefore, to take into account the magnitudes of the calculated structure factors either by using method (1) or (2) on the observed and calculated structure factors in parallel or by using method (3) or (4).

In any method to be used with an automatic calculator it is desirable that the refinement of atomic parameters should be a cyclic process so that the calculator will, without outside interference, carry

out as many iterations as may be necessary to give the final result, i.e. to proceed until the changes in parameters produced by one iteration are less than a predetermined value.

Bennett & Kendrew (1952) have prepared programmes for calculating Patterson syntheses with the EDSAC which would need only slight modification to adapt them for method (1). This method, which was the first considered, is by no means convenient for the primary refinement for the following reasons. At each iteration a complete set of significant ρ values is produced; but it is not until a new set of structure factors based on atomic coordinates obtained from these ρ values have been calculated, that a decision can be made as to whether a further iteration is required. Although the matter was not pursued, because the alternative method proved more attractive on other grounds, it was not immediately apparent how it would be possible to incorporate in a single iterative programme the calculation of ρ values, the recalculation of structure factors, and the final discrimination as to whether the required result had been obtained. Furthermore, though the formula was so broken down as to minimize the number of trigonometrical calculations, the time that the calculator would have taken to produce the ρ values for the substance under investigation was so large as to discourage the use of this method of refinement when the coordinates alone were required. Should the whole electron-density map be wanted, the large amount of calculator time involved in the Fourier method might then be justified. However, it is more appropriate to use method (2), the differential Fourier method, for the first stages of the refinement of coordinates, as it is very much faster in execution and can be made completely iterative. For similar reasons, method (4) was preferred for the final stage of refinement.

* Present address: Birkbeck College Research Laboratory, 21 Torrington Square, London W. C. 1, England.

The calculator

The automatic calculator employed was LEO.* Like other full-scale automatic calculators, LEO, having been supplied with a programme of orders for a given problem, is able to read-in a section of the problem data, carry out the calculations, print the results and then go back to deal with the next section of data, and so on, until it recognizes that the whole task is completed. Once set in motion LEO continues in automatic operation without any external control. When, as often occurs in a series of computations, different action is required according to the nature of the data or partial results, the calculator may be programmed so that it automatically examines the circumstances at any time and takes the correct course. A simple example of this fundamental feature occurs in the refinement calculations where the formula is varied according to whether $(h+k)$ is odd or even; the programme includes sequences to cope with both eventualities and the calculator uses one or the other according to the outcome of its examination of $(h+k)$.

LEO is a serial calculator using a single-address code and operating in the binary notation; it comprises a store, an arithmetic unit, a control unit and several input and output channels of various types and speeds. The store is composed of 64 mercury delay lines providing a capacity of 2048 'compartments', each of 17 binary positions; alternatively, pairs of compartments may be used to hold 'long numbers' of 35 binary positions, equivalent to signed numbers of the order 10^{10} . The LEO store is not divided into high- and medium-speed sections as are the stores of most other computers: all compartments are available in the average access time of 0.5 msec.

The automatic arithmetic facilities provided in LEO are addition, subtraction and multiplication. The addition or subtraction of two numbers and the replacement of the result in the store ready for later use can be accomplished at the rate of 300 per sec. and multiplications, including the setting of both factors and transfer back of the result, at 140 per sec. Both these speeds are irrespective of the size of the factors up to the normal maximum of 10^{10} . Division and calculation of trigonometrical functions are obtained by the use of special iterative sequences of programme orders. The speed of division depends on the size of divisor but 5 per sec. can generally be achieved. As an example of trigonometrical functions, cosines can be evaluated by means of power series at the rate of 14 a sec. In the programmes referred to in this paper, the division and trigonometrical sequences or 'sub-routines' as they are called, were those devised at the Cambridge University Mathematical Physics Laboratory for EDSAC, an earlier calculator on which LEO is largely based (Wilkes, Wheeler & Gill, 1951).

For the refinement calculations both orders and

programme data were fed to LEO in the form of perforations in paper tape, the tape being read by a photo-electric reader and converted from its binary-decimal form into the required full binary code inside the calculator. Results in the main were produced by teleprinter directly coupled to the calculator and LEO's special high speed output facilities were not employed until the later stages of the calculations, when the final F_c values were needed to assess the accuracy of the structure determinations. By using one of the high speed output channels to record on magnetic tape the 939 values as they were obtained during the calculations, the calculator time taken, over and above that which would have been taken had these values not been recorded, was of the order of 10 sec.

The scope of the programmes used

The programmes were worked out for nicotinamide, which has the space group $P2_1/a$. They could, however, be readily modified to make them applicable to other centrosymmetric space groups. Application of these methods to acentric space groups would be rather more complicated.

The time taken to carry out the programmes on the calculator depends on the number of atoms and h, k, l , values used. For nicotinamide, the number of atoms other than hydrogen is 9 and the ranges of h, k , and l values were:

$$\begin{aligned} -12 \leq h \leq 12, & \text{ total number of values } H = 25; \\ 0 \leq k \leq 20, & \text{ total number of values } K = 21; \\ 0 \leq l \leq 5, & \text{ total number of values } L = 6. \end{aligned}$$

All times, numbers of terms in tables etc. referred to in this paper are in relation to these numbers of values.

Notation

The following notation will be used:

a, b, c	unit cell dimensions in Ångström units;
β	the monoclinic interaxial angle;
V	the volume of the unit cell;
x_r, y_r, z_r	the assumed coordinates of the r th atom in radians;
ϵ_{xr} etc.	correction to be applied to the x coordinate of the r th atom;
$\rho(x, y, z)$	the electron density at the point x, y, z ;
Z	the atomic number;
f	the atomic scattering factor for any reflexion;
\hat{f}	the unitary atomic scattering factor for any reflexion = f/Z ;
F_o	the observed structure factor for any reflexion;
F_c	the calculated structure factor for any reflexion;
A_h etc.	$-(\partial \rho / \partial x)(V/4)$;
A_{hh} etc.	$-(\partial^2 \rho / \partial x^2)(V/4)$;

* A machine built at Cadby Hall for the Lyons Electronic Office.

$$\sum_3 \sum_{-h}^h \sum_0^k \sum_0^l \text{ over the relevant } h, k, l \text{ values;}$$

$$B_h \text{ etc. } \frac{1}{4} \sum_3 \frac{\partial F}{\partial x};$$

$$B_h^2 \text{ etc. } \frac{1}{16} \sum_3 \left(\frac{\partial F}{\partial x} \right)^2.$$

The differential Fourier method

Given the approximate positions of the maxima in the electron-density distribution, this method (Booth, 1946) arrives at their exact locations.

In the monoclinic system

$$\varepsilon_x = \frac{1}{\sin^2 \beta} \left[\frac{A_l c \cos \beta}{A_{ll} a} - \frac{A_h}{A_{hh}} \right],$$

$$\varepsilon_y = -\frac{A_k}{A_{kk}},$$

$$\varepsilon_z = \frac{1}{\sin^2 \beta} \left[\frac{A_h a \cos \beta}{A_{hh} c} - \frac{A_l}{A_{ll}} \right].$$

These equations assume spherical symmetry of the electron-density distribution so that off-diagonal terms (A_{hk} etc.) can be neglected. As the process is iterative this assumption does not affect the final result but only the rate of refinement.

For the space group $P2_1/a$

$$\rho(x, y, z) = \frac{4}{V} \left[\sum_3 \sum_{(h+k) \text{ even}} F_o \cos(hx+lz) \cos ky \right. \\ \left. - \sum_3 \sum_{(h+k) \text{ odd}} F_o \sin(hx+lz) \sin ky \right],$$

$$A_h = \sum_3 \sum_{(h+k) \text{ even}} hF_o \sin(hx+lz) \cos ky \\ + \sum_3 \sum_{(h+k) \text{ odd}} hF_o \cos(hx+lz) \sin ky,$$

$$A_k = \sum_3 \sum_{(h+k) \text{ even}} kF_o \cos(hx+lz) \sin ky \\ + \sum_3 \sum_{(h+k) \text{ odd}} kF_o \sin(hx+lz) \cos ky,$$

$$A_l = \sum_3 \sum_{(h+k) \text{ even}} lF_o \sin(hx+lz) \cos ky \\ + \sum_3 \sum_{(h+k) \text{ odd}} lF_o \cos(hx+lz) \sin ky,$$

$$A_{hh} = \sum_3 \sum_{(h+k) \text{ even}} h^2 F_o \cos(hx+lz) \cos ky \\ - \sum_3 \sum_{(h+k) \text{ odd}} h^2 F_o \sin(hx+lz) \sin ky.$$

The expressions for A_{kk} and A_{ll} are the same as that for A_{hh} with k and l respectively substituted for h .

In outline the procedure followed in refinement by this method is to take each reflexion in turn and, having calculated F_c , to attach its sign to the $|F_o|$. The contributions of the reflexions to the A_h and A_{hh} etc. for each atom are calculated and accumulated. From these functions ε_x etc. are calculated and added

to the old x_r to form the new values of the coordinates. The new x_r are used in a repetition of the cycle. The refinement is complete when the corrections for all atoms are less than a predetermined maximum.

External data

The data required for the whole refinement process are only the $|F_o|$ values, the initial x_r, y_r, z_r , the constants a, c, β and Z_r , and the minimum corrections required. The $|F_o|$ for reflexions of the type $h0l$ and $0k0$ are halved as in the normal Fourier method. In the case of nicotinamide the approximate x and y coordinates had been refined in the c -axis projection (Wright & King, 1954). The z coordinates were postulated assuming a planar configuration for the molecule, but could not be refined by two-dimensional methods owing to the high degree of overlapping in the a - and b -axis projections. As it was not considered advisable to assume that the z coordinates were within the limits required by the differential Fourier method, they were refined once by line syntheses parallel to the c axis before applying the differential method.

In addition it is necessary, as noted above, to arrive at the sign of F_o in the course of the calculations. This can be obtained by partly evaluating F_c . Assuming that the atomic scattering factor is proportional to the atomic number, which in the case of light atoms is approximately true, the sign of F_c is given by

$$\left(\frac{F_c}{4f} \right)_{(h+k) \text{ even}} = \sum_r Z_r \cos(hx_r+lz_r) \cos ky_r,$$

$$\left(\frac{F_c}{4f} \right)_{(h+k) \text{ odd}} = \sum_r Z_r [-\sin(hx_r+lz_r) \sin ky_r].$$

Construction of the programme

The organization of the programme is shown in Fig. 1. The outstanding feature of the expressions for A_h, A_{hh} etc. is the large number of sines and cosines that have to be evaluated. In fact the number of calculations of the form

$$\sin(hx+lz) \sin ky$$

amounts to approximately

$$\text{number of atoms} \times 3 \times H \times K \times L.$$

In the case of nicotinamide this is more than 80,000.

There are two methods available for making these calculations. Either a sub-routine may be used for evaluating individual cosines by power series or, since h, k and l increase by equal intervals, use may be made of addition formulae. The time taken by the two methods is 0.07 sec. and 0.024 sec. respectively, which when repeated many times represents a substantial difference. The latter method is obviously more acceptable and is used in the programme wherever it applies. Thus the basic trigonometric functions that

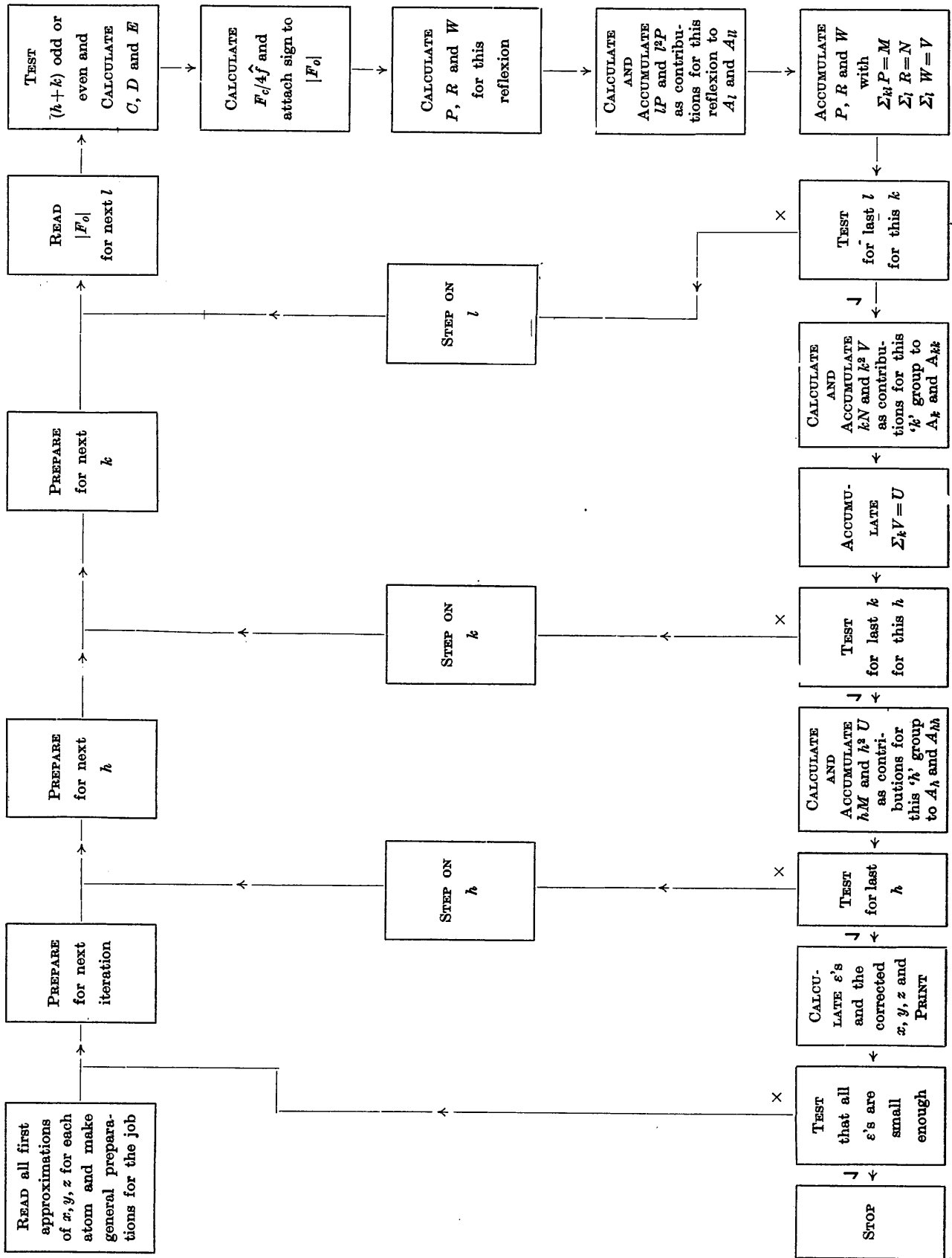


Fig. 1. Organization of the differential Fourier synthesis.

appear both in the formula for determining the sign of F_o and in the summations for A_h , A_{hh} etc. are built up as follows:

$$\begin{aligned} \sin (hx+lz) \cos ky \\ = (\sin hx \cos lz + \cos hx \sin lz) \cos ky, \quad (i) \end{aligned}$$

$$\begin{aligned} \cos (hx+lz) \sin ky \\ = (\cos hx \cos lz - \sin hx \sin lz) \sin ky, \quad (ii) \end{aligned}$$

$$\begin{aligned} \cos (hx+lz) \cos ky \\ = (\cos hx \cos lz - \sin hx \sin lz) \cos ky, \quad (iii) \end{aligned}$$

$$\begin{aligned} \sin (hx+lz) \sin ky \\ = (\sin hx \cos lz + \cos hx \sin lz) \sin ky. \quad (iv) \end{aligned}$$

Ideally it would be best to prepare the three tables of sines and cosines of hx , ky and lz , from which the expressions are constructed, at the start of each iteration and to hold them in the store throughout the iteration, selecting them as required. Such a table would, however, occupy a greater part of the store than could conveniently be spared for the purpose; in fact, it was decided to hold only the shortest of the three tables throughout each iteration, namely, the table of $\sin lz$ and $\cos lz$.

The order of using the structure factors

As the $\sin lz$ and $\cos lz$ values were to be made rapidly available, the greatest benefit may be obtained by carrying out the summations by first obtaining the contributions for all values of l for any hk , i.e. by varying l in the inside loop of the programme flow charts shown in Fig. 1. h was made the outer loop of the summation because the h values are symmetrical about zero and, if taken in the order $+1, -1, +2, -2$, make it particularly simple to arrive at the sine and cosine values of hx for each negative value of h merely by changing the signs of the sine values in the table for the preceding positive value of h . The reflexions were consequently dealt with in the order:

$$\begin{aligned} h = 0 \\ k = 0 \\ l = 0, 1, 2, 3, 4, 5 \\ k = 1 \\ l = 0, 1, 2, 3, \text{ etc until all } k \text{ values completed for } h = 0; \end{aligned}$$

then $h = +1$ with all l 's for the first k being dealt with and then k being stepped on to $1, 2, \dots, 20$;

then $h = -1$ with l and k being stepped on as before.

Making the summations

In dealing with each reflexion the first step taken in the programme is to test whether $(h+k)$ is even or odd and to evaluate the appropriate functions and store them. Three functions have to be evaluated for each reflexion. Using the references to the expressions set out above, they are:

	$(h+k)$ even	$(h+k)$ odd
C	(i)	(ii)
D	(ii)	(i)
E	(iii)	-(iv)

Thenceforward the programme required for the partial evaluation of F_c and for making the summations for each reflexion is independent of whether $(h+k)$ is even or odd. The calculation required for F_c is $\sum_r Z_r E$, giving the sign of F_c to be attached to $|F_o|$.

The values of A_h, A_{hh} for each atom expressed as summations in C, D and E are:

$$A_h = \sum_3 h F_o C = \sum_3 h P, \quad A_{hh} = \sum_3 h^2 F_o E = \sum_3 h^2 W,$$

$$A_k = \sum_3 k F_o D = \sum_3 k R, \quad A_{kk} = \sum_3 k^2 F_o E = \sum_3 k^2 W,$$

$$A_l = \sum_3 l F_o C = \sum_3 l P, \quad A_{ll} = \sum_3 l^2 F_o E = \sum_3 l^2 W,$$

where $F_o C = P, F_o D = R$, and $F_o E = W$.

Each reflexion having been read, C, D and E were calculated and the sign of F_o was determined. The next step was to evaluate P, R and W for each atom.

Since for any given combination of hk the values of l are taken in turn, it is natural to obtain the contributions to A_l and A_{ll} for each l by multiplying P and W respectively by l and l^2 . However, for the contributions to A_k and A_{kk} , it is more convenient to sum R and W for all l 's for each k and then to multiply the sub-totals by k and k^2 respectively; similarly for contributions to A_h and A_{hh} , it is convenient to sum P and W for all l 's and k 's for each h and then to multiply by h and h^2 respectively.

For each k , therefore, preparatory summations were made giving

$$\sum_l R = N \quad \text{and} \quad \sum_l W = V,$$

and for each h preparatory summations of

$$\sum_k \sum_l P = M \quad \text{and} \quad \sum_k V = U.$$

A contribution was therefore added to each of $\sum_3 l P, \sum_3 l^2 W, \sum_k \sum_l P, \sum_l R$ and $\sum_l W$ for each reflexion.

After the last l for any k , $\sum_l R$ was multiplied by k to give kN , the contribution to $\sum_3 k R$, and $\sum_l W$ was multiplied by k^2 to give $k^2 V$, the contribution to $\sum_3 k^2 W$.

A contribution was also added to $\sum_k V$.

After the last k for any h , $\sum_k \sum_l P$ was multiplied by h to give hM , the contribution to $\sum_3 h P$, and $\sum_k V$ was multiplied by h^2 to give $h^2 U$ the contribution to $\sum_3 h^2 W$.

After the last h , the ϵ 's were calculated for each atom using the values A_h, A_{hh} etc. obtained by the summations, and from these corrections revised approximations to the coordinates were obtained for each atom. Each ϵ was then compared by the calculator with the predetermined maximum error and if any one ϵ exceeded this maximum the whole process was repeated with the revised coordinates.

The information produced

At the end of each iteration the programme arranges for the printing in tabular form of

$\epsilon_{xr}, \epsilon_{yr}, \epsilon_{zr}$;
the new x_r, y_r, z_r ;
 A_{hh}, A_{kk} , and A_{ll} for each atom.

The A_{hh} etc. are required for the assessment of the accuracy of the final structure.

The storage space used and time taken in carrying out the programme

The total number of reflexions which could be recorded was 1301. In fact only 939 of these were strong enough to be observed. The programme was so arranged that no calculations were made for the 362 reflexions not observed. Although the $|F_o|$ values were required for each iteration of the programme it was inconvenient to hold a table of a thousand values throughout the calculations. The arrangement made, therefore, was to read and hold in the calculator at one time one $|F_o|$ value only and to carry out the summations for that reflexion. Then the next value was read and used. At the end of the iteration when the last $|F_o|$ value had been read the tape bearing the $|F_o|$ values was attached to the reader ready to start with the first value. Thus the values were again read for each iteration. The time taken in re-reading was not large; indeed had the new high speed input arrangements now available been in existence when the calculations were done, the reading of each $|F_o|$ value would have occurred simultaneously with the summations for the previous $|F_o|$ value, so that there would have been no loss of time at all.

The storage space used was made up as follows:

	Compartments
Programme	
569 orders + 74 orders for each atom	= 1235
Working locations	
98 compartments + 46 compartments for each atom =	512
Tables	
sin lz and cos lz for each atom for each $l=2 \times 9 \times 6 =$	<u>108</u>
	<u>1855</u>

The number of orders carried out in each iteration other than those concerned with reading data values and printing results was:

	Orders
For each combination of h, k, l , namely 939	
106 for each atom + 25 = $939 \times (9 \times 106 + 25) =$	919,281
For each combination of hk , namely 380	
67 for each atom + 22 = $380 \times (9 \times 67 + 22) =$	237,500
For each h , namely 25	
72 for each atom + 32 = $25 \times (9 \times 72 + 32) =$	<u>17,000</u>
	<u>1,173,781</u>

Estimated time at average of 1.25 msec. per order	24.5 min.
Add time for reading $ F_o $ values	3.7 min.
Add time for calculating results and printing	<u>2.0 min.</u>
	<u>30.2 min.</u>

It must be pointed out that the value of 1.25 msec. required to execute one order is only a rough average as the actual time required for an individual order depends on the nature of that order.

It is of interest to compare the times taken per iteration using the differential Fourier method for the nicotinamide refinements and the time that would have been taken using the Fourier method:

	Fourier method	Differential Fourier method
Data input	1½ hr.	3.7 min.
Calculations	11½ hr.	24.5 min.
Printing results	½ hr.	2.0 min.
Total	<u>13½ hr.</u>	<u>30.2 min.</u>

Discussion

The progress of the differential refinement, which for nicotinamide took 35 min. per iteration (as compared with the estimate of 30.2 min.), is shown in Table 1. The limits set to the corrections were 0.0006

Table 1. *Progress of the differential Fourier refinement*

Iteration number		(1)	(2)	(3)	(4)	(5)	(6)
Mean change in atomic positions (Å)	x	0.0142	0.0055	0.0011	0.0003	0.0005	0.00002
	y	0.0160	0.0051	0.0015	0.0005	0.0002	0.00007
	z	0.0405	0.0038	0.0012	0.0005	0.0001	0.00002
Maximum change in atomic positions (Å)	x	0.0299	0.0210	0.0039	0.0013	0.0011	0.00003
	y	0.0313	0.0096	0.0042	0.0011	0.0005	0.00020
	z	0.1483	0.0085	0.0036	0.0010	0.0002	0.00004

in ϵ_x , 0.0004 in ϵ_y and 0.0015 in ϵ_z , corresponding to 0.001 Å in all cases.

In the light of the steepest-descents refinement described below it will be seen that the apparent increase in accuracy brought about by the 4th, 5th and 6th differential refinements was artificial, and the limit to the corrections should have been set at 0.005 Å ($\epsilon_x = 0.003$, $\epsilon_y = 0.003$, $\epsilon_z = 0.008$). The limit had been set too low because nothing was known beforehand about the rate of refinement.

The final coordinates (set (6)) were used in conjunction with unitary scattering factors to calculate the hkl structure factors. The scattering factors employed were weighted means of the values for C, N and O given by Viervoll & Ögrim (1949). By comparison of these values with the observed structure factors, a temperature factor was determined and the scattering factors were modified accordingly in preparation for the method of steepest descents. This section of the work was done outside the calculator.

The method of modified steepest descents

By this method the atomic coordinates are refined until the best agreement between observed and calculated magnitudes of the structure factors $|F_o|$ and $|F_c|$ is obtained.

Qurashi's (1949) modification of this method has been used. The corrections ϵ_{xr} are given by

$$\epsilon_{xr} = \frac{\sum_3 (F_o - F_c) \partial F_c / \partial x_r}{\sum_3 (\partial F_c / \partial x_r)^2},$$

This expression is exactly equivalent to that obtained for the correction using the method of least squares with neglect of the off-diagonal terms.

In the space group $P2_1/a$

for $(h+k)$ even

$$F_c = 4 \sum_r \hat{f} Z_r \cos (hx_r + lz_r) \cos ky_r,$$

$$\partial F_c / \partial x_r = -4h \hat{f} Z_r \sin (hx_r + lz_r) \cos ky_r,$$

$$\partial F_c / \partial y_r = -4k \hat{f} Z_r \cos (hx_r + lz_r) \sin ky_r,$$

$$\partial F_c / \partial z_r = -4l \hat{f} Z_r \sin (hx_r + lz_r) \cos ky_r;$$

for $(h+k)$ odd

$$F_c = 4 \sum_r -\hat{f} Z_r \sin (hx_r + lz_r) \sin ky_r,$$

$$\partial F_c / \partial x_r = -4h \hat{f} Z_r \cos (hx_r + lz_r) \sin ky_r,$$

$$\partial F_c / \partial y_r = -4k \hat{f} Z_r \sin (hx_r + lz_r) \cos ky_r,$$

$$\partial F_c / \partial z_r = -4l \hat{f} Z_r \cos (hx_r + lz_r) \sin ky_r.$$

The above expressions are similar to the corresponding expressions in the differential Fourier method and

therefore the calculations proceed along similar lines. The two most significant differences are the introduction into the equations of \hat{f} and F_c .

External data

The data fed in for the beginning of the first iteration were the final x , y , z coordinates from the differential syntheses, the atomic number for each atom and the minimum correction. During each iteration the $|F_o|'$ values on an arbitrary scale and the \hat{f} values were fed in successively for each reflexion in turn. The $|F_o|'$ and \hat{f} values for the reflections of the type $h0l$ and $0k0$ were reduced outside the calculator by a factor of $1/\sqrt{2}$ in order to halve their contributions to the refinement. Each $|F_o|'$ value was multiplied by K to obtain the $|F_o|$ values, on an absolute scale, where K was 1 for the first iteration and was obtained from $K = \sum_3 |F_c| / \sum_3 |F_o|'$ for succeeding iterations.

F_c was calculated from

$$F_c = 4 \hat{f} \sum_{(h+k) \text{ even}} Z_r \cos (hx_r + lz_r) \cos ky_r,$$

$$F_c = -4 \hat{f} \sum_{(h+k) \text{ odd}} Z_r \sin (hx_r + lz_r) \sin ky_r.$$

The sign of F_c was attached to $|F_o|$ to obtain F_o .

Construction of the programme

The same method was used here as for the differential Fourier synthesis, the differences in expressions being given effect in the individual orders rather than in the organization of the computation (Fig. 2). It was possible to achieve some saving of space and time by holding a table of $Z_r \sin lz_r$ and $Z_r \cos lz_r$ as the parameter Z_r appears linearly in all the expressions. The basic components of the summations and of F_c are thus slightly modified:

$$Z \sin (hx + lz) \cos ky = [\sin hx(Z \cos lz) + \cos hx(Z \sin lz)] \cos ky, \quad (i)$$

$$Z \cos (hx + lz) \sin ky = [\cos hx(Z \cos lz) - \sin hx(Z \sin lz)] \sin ky, \quad (ii)$$

$$Z \cos (hx + lz) \cos ky = [\cos hx(Z \cos lz) - \sin hx(Z \sin lz)] \cos ky, \quad (iii)$$

$$Z \sin (hx + lz) \sin ky = [\sin hx(Z \cos lz) + \cos hx(Z \sin lz)] \sin ky. \quad (iv)$$

Making the summations

We may now again define the quantities

	$(h+k)$ even	$(h+k)$ odd
C	(i)	(ii)
D	(ii)	(i)
E	(iii)	-(iv)

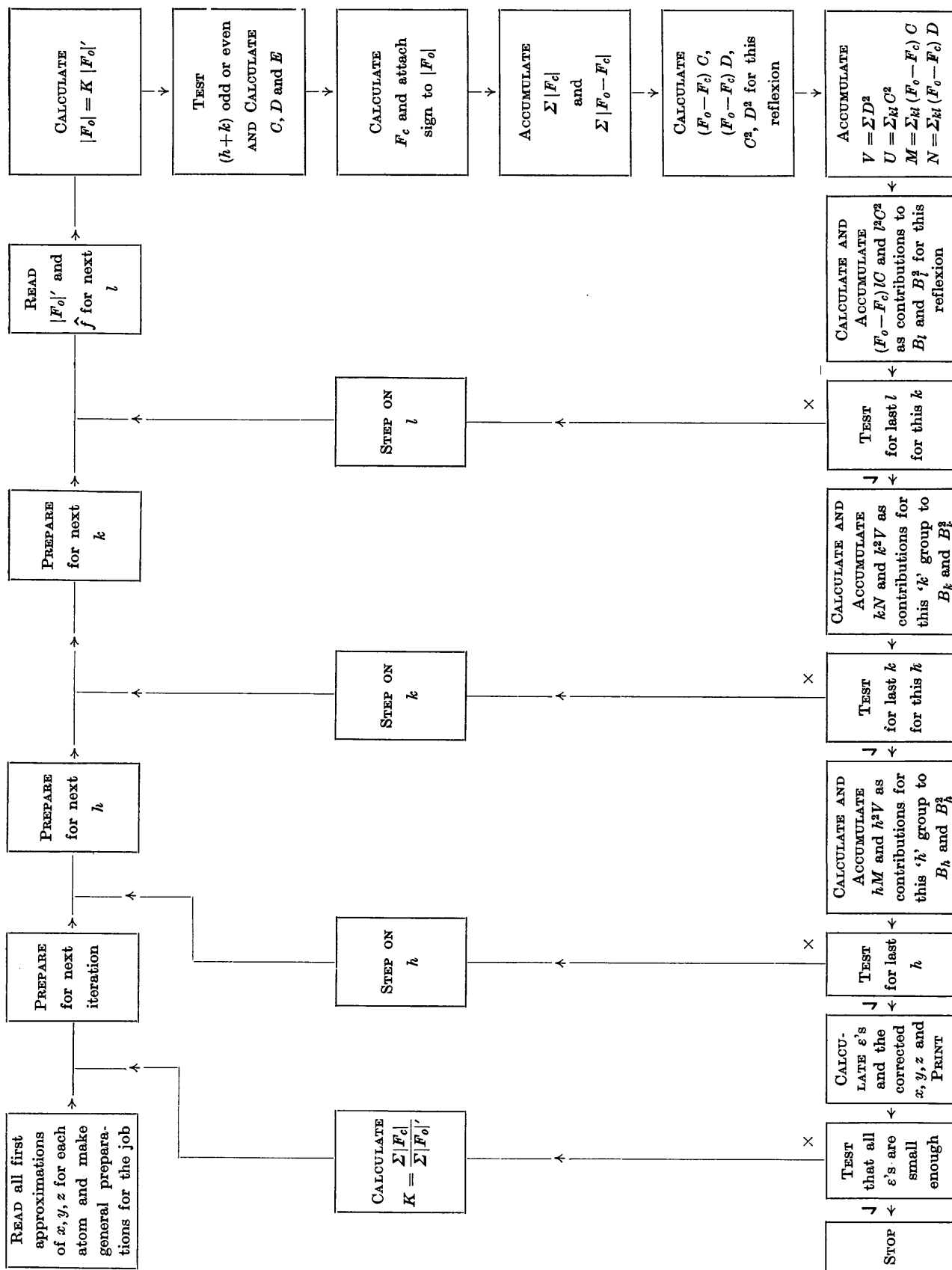


Fig. 2. Organization of the steepest-descent calculation.

Then the accumulations proceed as follows:

$$B_h = -\sum_h hM, \quad \text{where } M = \sum_{kl} (F_o - F_c)C,$$

$$B_k = -\sum_{hk} kN, \quad \text{where } N = \sum_l (F_o - F_c)D,$$

$$B_l = -\sum_{hkl} (F_o - F_c)lC,$$

$$B_h^2 = \sum_l h^2 U, \quad \text{where } U = \sum_{kl} C^2,$$

$$B_k^2 = \sum_{hk} k^2 V, \quad \text{where } V = \sum_l D^2.$$

$$\text{Finally } \epsilon_x = 4B_h/B_h^2.$$

Information produced

The information printed at the end of each iteration was:

$$\epsilon_{xr}, \epsilon_{yr}, \epsilon_{zr};$$

the new x_r, y_r, z_r ;

$$\sum_3 |F_c| \quad \text{and} \quad \sum_3 |F_o - F_c|.$$

In addition, during the final iteration the F_c values were recorded on magnetic tape for subsequent printing.

The storage space used and time of programme

The storage space used was made up as follows:

	Compartments
Programme	
669 orders + 66 orders for each atom	= 1263
Working locations	
141 compartments + 43 compartments for each atom	= 528
Tables	
sin lz and cos lz for each atom for each $l=2 \times 9 \times 6$	= 108
	<u>1899</u>

The number of orders carried out in each iteration, other than those concerned with reading data values and printing results, was:

	Orders
For each combination of h, k, l , namely 939	
134 for each atom + 104 = $939 \times (9 \times 134 + 104)$	= 1,230,090
For each combination of hk , namely 380	
87 for each atom + 24 = $380 \times (9 \times 87 + 24)$	= 306,660
For each h , namely 25	
206 for each atom + 136 = $25 \times (9 \times 206 + 136)$	= 49,750
	<u>1,586,500</u>

Estimated time at average of 1.25 msec. per order	33.1 min.
Add time for reading data values	7.5 min.
Add time for calculating results and printing	2.0 min.
	<u>42.6 min.</u>

Discussion

The progress of the steepest descents refinement which for nicotinamide took 40 min. per iteration (as compared with the estimated 42.6 min.) is shown in Table 2. The mathematical accuracy obtained is far greater than the actual accuracy of the results because of the errors in the estimation of intensities. In order to take full advantage of the method it would be necessary to have much more accurate values of the observed structure factors.

Checking arrangements

Precautions were, of course, taken to ensure the correctness of the results produced in both the differential Fourier and the steepest descents calculations. The checking arrangements for each of them fell into three phases:

- (1) To check that the programme as constructed gave results of the required accuracy.
- (2) To check, before starting an actual run of the programme, that the calculator was working with complete efficiency.
- (3) To check during the run, by self-checks imposed by the programme itself, that the general mechanism of the programme was working correctly.

The following are brief notes on each of these phases:

(i) The check on the construction of the programmes was effected by inventing an F_o table short enough for results to be computed by hand without undue labour but none the less sufficient to test all facets of the programme. The results produced by hand and by LEO were compared.

(ii) Before commencing each run of the programme (calculations were carried out not only for nicotinamide but also for a substance not discussed in this paper) the calculator was given its normal operational tests. These comprise sequences specially constructed to try each feature of the calculator under artificially induced physical conditions which are far more exacting than those that would be encountered in normal operation. Experience has shown that such tests give ample warning of valve or component changes that might otherwise give trouble at some time in the operational run that follows.

(iii) The special checks built into the programme included the following:

- (a) By means of a check-total of the $|F_o|$ and \hat{f}

Table 2. Progress of the steepest-descents refinement

Iteration number		(1)	(2)	(3)	(4)
Mean change in atomic positions (Å)	x	0.0103	0.0024	0.0005	0.0002
	y	0.0142	0.0019	0.0014	0.0003
	z	0.0075	0.0012	0.0005	0.0002
Maximum change in atomic positions (Å)	x	0.0190	0.0055	0.0010	0.0005
	y	0.0340	0.0066	0.0026	0.0005
	z	0.0265	0.0027	0.0014	0.0004

values for each h , it was possible to ensure that no $|F_o|$ value had been read wrongly.

(b) By means of check-totals for each h of the last $|F_o|$ value of each k , it was possible to ensure that k had been stepped on correctly.

(c) Each h value was checked before it was used by comparing the value read on the $|F_o|$ tape with that registered by a counter inside the calculator.

(d) Before calculating tables of sines and cosines of the new x_r , the routine used for this purpose was tested. It was programmed to calculate the cosine of a test angle (this happened to be $\frac{1}{2}\pi$) and the result was compared with the externally calculated cosine which was held in the calculator.

(e) All tables of sines and cosines held in the calculator for a complete iteration were checked periodically against a check-total derived in the first place by the calculator itself.

No special checking arrangements were made for the calculation of the values C, D, E etc. nor for the summations, since this is not a problem which lends itself easily to checks on separate stages of the calculations. However, the self-checks already referred to served not only individually to confirm the results of the processes but also collectively to show that the calculator was correctly reading and carrying out all the arithmetic and logical processes on the ground covered by the self-checks. Further, it was possible to check externally that ϵ_{xr} plus the old x_r equalled the new x_r and this, too, proved many more aspects of the operation of the calculator than its mere ability to add two numbers together. In addition, the convergence of the values of the coordinates for the atoms to definite values itself provided some measure of check.

Finally, the printing of the results was checked by using an output code in which the corruption of one element of the code would have caused the printing of a character other than a numeral, thus making the corruption immediately apparent. In fact, no such corruption occurred.

Conclusions

The great speed of operation of electronic calculators makes possible the rapid refinement of crystal structures as a routine method. Hitherto the practice

in structure determinations has been to use two-dimensional methods to their limit of accuracy before embarking on the laborious process of three-dimensional analysis. In cases where it is intended eventually to use three-dimensional methods, it would probably be quicker to use projections only as far as is necessary to obtain the atomic coordinates within the limits of accuracy (usually about 0.2 Å for light atoms) required to start the differential synthesis. Although the preliminary calculations (Fourier projections and structure-factor values) can easily be done on the calculator there is a great saving in time if the process can be made iterative.

There is no reason why the time taken should not be further reduced by making the whole process automatic from the beginning of the differential Fourier synthesis. The external data would then consist of the $|F_o|$'s, the initial parameters, the cell dimensions, the minimum corrections and a table of atomic scattering factors, perhaps a different set for each atomic species. At the end of the differential refinement a new programme sequence would be read in to replace part of the old one. The machine would then calculate the temperature factor and use it to modify the table of scattering factors before proceeding to the steepest descents refinement. Finally, the accuracy could be automatically assessed by some method such as that of Cruickshank (1949).

The authors wish to thank Dr L. H. Lampitt for his great interest in this work and the Directors of J. Lyons and Company for permission to publish.

References

- BENNETT, J. M. & KENDREW, J. C. (1952). *Acta Cryst.* **5**, 109.
 BOOTH, A. D. (1946). *Trans. Faraday Soc.* **42**, 444.
 CRUICKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 65.
 QURASHI, M. M. (1949). *Acta Cryst.* **2**, 404.
 VIERVOLL, H. & ÖGRIM, O. (1949). *Acta Cryst.* **2**, 277.
 WILKES, M. V., WHEELER, D. J. & GILL, S. (1951). *The Preparation of Programmes for an Electronic Digital Computer, with Special Reference to the EDSAC and the Use of a Library of Sub-Routines*. Cambridge, Mass.: Addison-Wesley Press.
 WRIGHT, W. B. & KING, G. S. D. (1954). *Acta Cryst.* **7**, 283.