Acta Cryst. (1958). **11**, 669

Generalized Programmes for Crystallographic Computations

By F. R. Ahmed and W. H. Barnes

Division of Pure Physics, National Research Council, Ottawa, Canada

(*Received* 17 *March* 1958)

The general principles of some machine computation programmes which are applicable to all space groups without modification are outlined. The procedure takes advantage of the space group symmetry elements, and eliminates the time and labour involved in the preparation and testing of separate routines for the different space groups. The appropriate equations of the space group are represented by a simple pseudo code and the necessary adjustments to the main programme are directed by a short interpretive routine.

Most programmes for crystallographic computations on automatic digital computers have been written for specific space groups, although a generalized one for two-dimensional calculations in the triclinic, monoclinic, and orthorhombic systems has been described by Fowweather (1955). A programme for three-dimensional calculations in all space groups has been reported by Sparks, Prosen, Kruse & Trueblood (1956) which, however, basically treats all structures as $P1$ or $P\bar{1}$.

Over the past three years a number of generalized programmes applicable to all space groups have been written in this laboratory for the Ferranti computer, FERUT. They have been used successfully in routine calculations of structure factors, Fourier and differential syntheses, and accuracy estimations. They are designed to take advantage of the symmetry characteristics of the individual space groups, and to avoid wasteful repetition of calculations for crystallographically equivalent atoms and reflections. The principles on which they are based can be applied with advantage to programmes for digital computers generally. Their most important feature is the inclusion of simple pseudo codes to represent the appropriate equations given in the *International Tables* (1952). The experimental data are arranged in groups according to the reflections to which a particular mathematical expression applies. Whenever a new group of reflections is sensed by the machine the pseudo code causes specific instructions to be planted in certain vacant lines of the programme depending on the number of terms in the expression, their signs and trigonometric functions. The routine is thereby converted into a specialized one for the evaluation of a specific equation, and hence only those terms pertinent to a specific group of reflections are computed. The interpretive parts of the programmes have no equivalent in those designed for individual space groups. It is important to bear in mind, however, that the length of a written programme is not necessarily indicative of the speed or efficiency with which the machine will carry out a computation because the longest time is consumed by those sections which must be obeyed most frequently. In the present routines the instructions required by the use of the pseudo codes are obeyed only at the start of each group of reflections in contrast to those which are followed for each reflection within a group. The proportionate increase in machine times due to this interpretive feature, therefore, is very small compared with the unproductive time consumed by the use of the general equations of $P1$ and $P\bar{1}$ for all space groups regardless of higher symmetry.

Storage space in the machine, and computation time, are conserved by the omission of data for the non-observed reflections. This requires identification of the observed data by the appropriate indices which is not necessary in a sequential arrangement of the data with no omissions. No additional storage space for the indices, however, is required because they are accomodated on the same lines as the corresponding structure factors. Data for those reflections which are observed but are equivalent because of changed signs of indices, or because one or more indices are zero, are included once only and allowance is made for this fact in the multiplicity factors employed. Data for all those reflections which are equivalent because of permutation of indices in the cubic, tetragonal, and hexagonal systems, however, are included to avoid the evaluation of the more complicated expressions which generally would otherwise be required.

The programmes contain intermediate checks at conveniently short intervals. They may be employed singly, or suitable ones may be combined for continuous machine operation. The computations may be halted at any stage without loss of results already reached or of the sequence of operations.

Depending on the programme, or sequence of programmes, to be followed, some or all of the following information is stored on the magnetic drum before starting the calculations: the directories, the routines, the equations for the space group under consideration,

the structure factors and the corresponding indices (arranged in groups according to the applicable equations), the co-ordinates of the atoms in the asymmetric unit, the appropriate atomic scattering factor curves for zero temperature factor, the isotropic temperature factors, the cell constants, the multiplicity factors, and the details of any Fourier meshes to be computed. Each f-curve is represented by 49 points spaced at small intervals for low values of $\sin\theta$ and at larger intervals for higher values of $\sin\theta$ to permit linear interpolation between points with sufficiently high accuracy. The present programmes accomodate isotropic temperature factors only.

The real parts, and the imaginary parts if not zero, of the structure factors are calculated according to expressions of the form

$$F(hkl) = C\left\{\sum_r f_r(hkl) \times \left[\sum_{s=1}^{n_r} G_s(hkl)\right]_r\right\} \qquad (1)$$

where C is an integer, and

$$f_r(hkl) = [f_0(hkl)]_J \times \exp\left[(-B_r/\lambda^2)\sin^2\theta(hkl)\right]. \quad (2)$$

If the space group is centrosymmetric the signs of the calculated F's are transferred automatically to the observed structure amplitudes; if it is non-centrosymmetric the machine computes $|F_c|$, A_o, and B_o.

Even within the same space group, G and C may not be the same for all groups of reflections.

When expanded, G takes one of the following forms:

$$\left.\begin{aligned}
&\text{(i)}\quad \sum \begin{matrix}\cos\\ \sin\end{matrix} 2\pi h(x)\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi k(y)\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi l(z) \\
&\qquad\qquad + \text{similar terms} \\
&\text{(ii)}\quad \sum \begin{matrix}\cos\\ \sin\end{matrix} 2\pi h\begin{pmatrix}x\\ y\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi k\begin{pmatrix}y\\ x\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi l(z) \\
&\qquad\qquad + \text{similar terms} \\
&\text{(iii)}\quad \sum \begin{matrix}\cos\\ \sin\end{matrix} 2\pi h\begin{pmatrix}x\\ y\\ z\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi k\begin{pmatrix}y\\ z\\ x\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi l\begin{pmatrix}z\\ x\\ y\end{pmatrix} \\
&\qquad\qquad + \text{similar terms} \\
&\text{(iv)}\quad \sum \begin{matrix}\cos\\ \sin\end{matrix} 2\pi i\begin{pmatrix}x\\ y\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi k\begin{pmatrix}y\\ x\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix}\begin{pmatrix}0^\circ\\ 2\pi lz+120^\circ\\ 240^\circ\end{pmatrix} \\
&\qquad\qquad + \text{similar terms} \\
&\text{(v)}\quad \sum \begin{matrix}\cos\\ \sin\end{matrix} 2\pi h\begin{pmatrix}x\\ y\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix} 2\pi i\begin{pmatrix}y\\ x\end{pmatrix}\cdot \begin{matrix}\cos\\ \sin\end{matrix}\begin{pmatrix}0^\circ\\ 2\pi lz+120^\circ\\ 240^\circ\end{pmatrix} \\
&\qquad\qquad + \text{similar terms}
\end{aligned}\right\}. \qquad (3)$$

For all space groups in the triclinic, monoclinic, and orthorhombic systems, G has the form (3, i) with a maximum of four terms; in the tetragonal system it has the form (3, ii) with a maximum of eight terms; in the cubic system it has the form (3, iii) with a maximum of twelve terms, and in the hexagonal and rhombohedral systems it has the forms (3, iii), (3, iv), (3, v) with a maximum of 24 terms. Furthermore, for general indices h, k, l, all the reflection types (hkl), $(\bar{h}kl)$, $(h\bar{k}l)$, $(hk\bar{l})$ are non-equivalent in the triclinic system, two are non-equivalent in the monoclinic system, and there is only one non-equivalent type in the orthorhombic system; either one or two are non-equivalent in the tetragonal system, and either two or four are non-equivalent in the hexagonal and rhombohedral systems, depending upon the space group.

The structure factor programme, therefore, consists of a major routine common to all space groups together with one of the following subsidiary routines, the function of which is to interpret the pseudo code and direct the computation of the trigonometric expressions, G, for

(1) the triclinic, monoclinic, and orthorhombic systems (form (3, i); max. 4 terms; 4 reflection types),
(2) the tetragonal system (form (3, ii); max. 8 terms; 2 reflection types),
(3) the cubic system (form (3, iii); max. 12 terms; 1 reflection type),
(4) the hexagonal and rhombohedral systems (forms (3, iii), (3, iv), (3, v); max. 24 terms; 4 reflection types).

The fourth subsidiary routine could be employed for each of the 230 space groups but would require longer machine times in all cases for which the first three were devised.

The pseudo codes represent the information which is variable in equation (1), namely, (*a*) the number of terms, n, (*b*) the multiplication factor, C, (*c*) whether the Bravais–Miller index, i, is present or not, (*d*) whether the function $\begin{matrix}\cos\\ \sin\end{matrix}\left(2\pi lz+\begin{matrix}120^\circ\\ 240^\circ\end{matrix}\right)$ is present or not, (*e*) the three trigonometric functions, and the sign, of each term in equations (3). Thus (*b*) and (*c*) ensure that the functions $\begin{matrix}\cos\\ \sin\end{matrix}\left(2\pi i\begin{matrix}x\\ y\end{matrix}\right)$ and $\begin{matrix}\cos\\ \sin\end{matrix}\left(2\pi lz+\begin{matrix}120^\circ\\ 240^\circ\end{matrix}\right)$ will be computed only if they are required. The codes for (*e*) enable the machine to distinguish among the various possible forms $\begin{matrix}\cos\\ \sin\end{matrix}\left[2\pi\begin{pmatrix}hi\\ kl\end{pmatrix}\begin{pmatrix}x & z & z+\frac{1}{3}\\ y & z & z+\frac{2}{3}\end{pmatrix}\right]$. The computation comes to an end as soon as the exact number of terms in the expression under consideration has been calculated even when this is less than the maximum number for which allowance has been made in the subsidiary routine.

Before entering the Fourier programme the structure factors (or the A and B parts if the space group is non-centrosymmetric) are modified by the appropriate multiplicity factors by means of a short routine. The programme for the calculation of electron densities is capable of computing the most general expression

$$\varrho(X, Y, Z) = C'/V_c\{F(000)$$
$$+\sum_{0}^{\infty}\sum\sum [R_1(hkl) \begin{matrix}\cos\\ \sin\end{matrix} 2\pi hX \begin{matrix}\cos\\ \sin\end{matrix} 2\pi kY \begin{matrix}\cos\\ \sin\end{matrix} 2\pi lZ$$
$$+\text{similar terms for } R_2, R_3, R_4 \text{ for the first group of planes]}$$
$$+\text{similar series for the imaginary parts } I_1, I_2, I_3, I_4 \text{ for the first group of planes}$$
$$+\text{similar series for the other groups of planes}\} \quad (4)$$

where C' is an integer (2, 4, or 8), and

$$\left.\begin{array}{ll} R(hkl) = \pm A(hkl) \pm A(\bar{h}kl) \pm A(h\bar{k}l) \pm A(hk\bar{l}) & \text{(i)} \\ \text{or } R(hkl) = \pm A(hkl) \pm A(\bar{h}kl) & \text{(ii)} \\ \text{or } R(hkl) = \pm A(hkl) & \text{(iii)} \end{array}\right\} \quad (5)$$

and similarly for the amplitudes $I(hkl)$

but, again through the use of pseudo codes, only the terms required for the particular space group under consideration are evaluated. Provision is also made for the synthesis to be carried out at any desired intervals of $N/120$, where N is an integer, either throughout the asymmetric part of the unit cell or only in selected regions. Furthermore, neither the number of points nor the intervals need be the same in all three directions.

The pseudo codes for the Fourier programme represent the information which is variable in equation (4), namely, (*a*) the constant C'/V_c, (*b*) the zero term, $F(000)$, (*c*) the number of terms in each series (i.e., the number of non-equivalent types of reflections: 4 for (5, i), 2 for (5, ii), 1 for (5, iii)), (*d*) for each term, the signs preceding the amplitudes in equations (5), and the three appropriate trigonometric functions.

Details of the Fourier meshes to be computed (which must be included with the input data) are the total number of separate meshes and, for each mesh, the co-ordinates (X_0, Y_0, Z_0) of the starting point, the total number of points, and the intervals $(\Delta X, \Delta Y, \Delta Z)$ in the directions of the crystallographic axes, with the co-ordinates and intervals expressed as the integers N.

The machine first calculates and stores a table of $\cos 2\pi(X/120)$ for $X = 0, 1, 2, \ldots, 60$ from which it calculates sines, and other cosines, as required. Summations are performed in three stages taking full advantage of the order in which the data are arranged.

For the differential synthesis programme the machine is supplied with the assumed atomic co-ordinates, and the electron density equation (coded in the same way as for the Fourier programme). The machine interprets these codes, modifies certain instructions in the routine and evaluates the required summations. It then determines the observed shifts, or the observed, calculated, and corrected shifts for each atom according to the instructions. Again, full advantage is taken of the order in which the data are arranged (Ahmed & Cruickshank, 1953).

A short programme, also applicable to all space groups, has been written for calculating the discrepancy factor $R = \Sigma|KF_o - F_c|/\Sigma|KF_o|$ (where K is an assumed scale factor), a new scale factor, K', and the sums $\Sigma(\Delta F)^2$, $\Sigma h^2(\Delta F)^2$, $\Sigma k^2(\Delta F)^2$, $\Sigma l^2(\Delta F)^2$ which are required for the estimation of $\sigma(x)$ and $\sigma(\varrho)$. If the machine is instructed to obey this routine a second time it automatically computes the sums with the new scale factor, K', thus showing immediately the effect of this factor on the value of R.

Grateful acknowledgement is made to Prof. E. G. Cox and to Dr D. W. J. Cruickshank for details of the programmes prepared at the University of Leeds for the Manchester MARK II, to Prof. W. H. Watson and to Dr C. C. Gotlieb for making time available on FERUT at the Computation Centre of the University of Toronto, and to Mrs M. E. Pippy of this laboratory for desk-machine and hand calculations in connection with the development of the present programmes.

References

AHMED, F. R. & CRUICKSHANK, D. W. J. (1953). *Acta Cryst.* **6**, 765.

FOWWEATHER, F. (1955). *Acta Cryst.* **8**, 633.

International Tables for X-ray Crystallography (1952), vol. 1. Birmingham: Kynoch Press.

SPARKS, R. A., PROSEN, R. J., KRUSE, F. H. & TRUEBLOOD, K. N. (1956). *Acta Cryst.* **9**, 350.