

Crystallographic calculations using an electronic digital computer

II. Calculation of structure factors

By STIG ÅSRINK, GÖRAN BLOMQVIST and SVEN WESTMAN

With 1 figure in the text

ABSTRACT

A programme has been constructed for the automatic calculation of structure factors of centrosymmetric ($F(hkl)$) and non-centrosymmetric ($|F(hkl)|$ and $\alpha(hkl)$) structures. The number of atoms in the unit cell must not exceed 127 (in the centrosymmetric case 254) and the number of atomic species must not be larger than 12. The calculations can be made on any reflection (hkl) with the following limitation: $-127 \leq h, k, l \leq 127$. The structure factors are obtained grouped in one of three alternate ways corresponding to the three ways for grouping the structure factor data in the programme, which has been developed by our programming team, for the calculation of the electron density function. The possibility of calculation with individual isotropic temperature factors is emphasized. A further enlargement of the programme in order to make possible the use of anisotropic temperature factors is planned.

The principles employed in this work should be applicable also to other digital computers than the present one (BESK).

Introduction

In order to facilitate the computational work associated with crystallographic structure determinations, a few programmes have been constructed to be used with the digital computer BESK (1). The one to be described in this article calculates structure factors and is called SNUSKMUS ("Sekvens för Numerisk Uträkning av Strukturfaktorer i Kristallstrukturer Med och Utan Symmetricentrum"). The principles on which it is based could also be used in programmes for many other computers.

The functions calculated

The programme distinguishes between centrosymmetric and non-centrosymmetric structures and the centre of symmetry is also the only symmetry element that is taken into consideration. The different functions that are calculated in the both cases are given below. The following symbols are used:

T : the temperature factor = $\exp(-\beta/\lambda^2 \cdot \sin^2\theta)$ where β is a constant, λ the wavelength of the radiation used and θ the glancing angle for the reflection (hkl).

i : an index for the numbering of the atomic species.

I : the number of atomic species.

f_i : the atomic scattering factor of the i :th element.

N_i : the number of atoms of the i :th element.

M_{i+1} : the number of atoms of the i first elements; $M_1 = 0$, $M_{i+1} = M_i + N_i$.

x_n, y_n, z_n : the coordinates of the n :th atom.

In structures with a centre of symmetry, the structure factor, F , is calculated according to:

$$F(hkl) = 2 \cdot T \cdot \sum_{i=1}^I \left[f_i \cdot \sum_{M_{i+1}}^{M_i+N_i} \cos 2\pi(hx_n + ky_n + lz_n) \right]$$

The summation is made over half the unit cell.

In structures without a centre of symmetry, the absolute value of the structure factor, $|F|$, and the phase angle, α , are calculated as follows:

$$|F(hkl)| = T \cdot \sqrt{A'^2 + B'^2}$$

where

$$A' = \sum_{i=1}^I \left[f_i \cdot \sum_{M_{i+1}}^{M_i+N_i} \cos 2\pi(hx_n + ky_n + lz_n) \right]$$

$$B' = \sum_{i=1}^I \left[f_i \cdot \sum_{M_{i+1}}^{M_i+N_i} \sin 2\pi(hx_n + ky_n + lz_n) \right]$$

$$\alpha(hkl) = \frac{1}{2\pi} \cdot \arcsin \frac{B'}{\sqrt{A'^2 + B'^2}} \quad (\text{cycles})$$

Here the summation is made over the whole unit cell.

The atomic scattering factor, f_i , is calculated according to a three-term expression given by Appel (2):

$$f_i = A_i \cdot \exp(-a_i/\lambda^2 \cdot \sin^2\theta) + B_i \cdot \exp(-b_i/\lambda^2 \cdot \sin^2\theta) + C_i \cdot \exp(-c_i/\lambda^2 \cdot \sin^2\theta)$$

where $A_i, B_i, C_i, a_i, b_i, c_i$ is a set of constants characteristic for the element concerned. It is also possible to add a fourth term, D_i , corresponding to the real term of a dispersion correction (3).

$\sin^2\theta$ is calculated according to the formula:

$$\sin^2\theta = A \cdot h^2 + B \cdot k^2 + C \cdot l^2 - 2Dkl - 2Ehl - 2Fhk.$$

The programme is made up of a number of subroutines, viz. a self-loading routine, a routine that reads the decimal data from the tape and translates them to binary form, routines for calculation of cosine and sine, exponential functions, arc sine and square roots, a routine for decimal printing and the main routine.

Input data

The data that are needed for the calculations are grouped in three sets. The constants of the first set include the administrative constants and the constants which

Table 1. Input data for SNUSKMUS IV.

1.

Punch:

$d_1 D$	Heading.				
Q	$Q = 0$ if $\sin^2 \theta$ wanted, $Q = 1$ if $\sin^2 \theta$ not wanted.				
P	$P = 1$: h is varied in each group of the result.				
	$P = 2$: k is varied in each group of the result.				
	$P = 3$: l is varied in each group of the result.				
S	$S = 0$: Centrosymmetric structure.				
	$S = 1$: Non-centrosymmetric structure.				
G	Number of decimals wanted in F .				
I	Number of atomic species in the structure.				
	$I \leq 12$.				
N	Number of atoms in (half) the unit cell.				
	$N \leq 127$.				
A B C $-2 D$ $-2 E$ $-2 F$	Constants for the calculation of $\sin^2 \theta$.				
		h_{\min} k_{\min} l_{\min} h_{\max} k_{\max} l_{\max}	- 127 $\leq h, k, l \leq 127$. If for example the calculations are to be made on the $h k 0$ -reflections, one must put $l_{\min} = l_{\max} = 0$.		
				β/λ^2	For the calculation of the temperature factor. $\beta/\lambda^2 \geq 0.001$.

2.	$d_2^1 D$	$d_2^2 D$	$d_2^3 D$	$d_2^4 D$	$d_2^5 D$	$d_2^6 D$	$d_2^7 D$	$d_2^8 D$
	N_1	A_1	B_1	C_1	D_1	a_1/λ^2	b_1/λ^2	c_1/λ^2
	N_2	A_2	B_2	C_2	D_2	a_2/λ^2	b_2/λ^2	c_2/λ^2
	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
	N_I	A_I	B_I	C_I	D_I	a_I/λ^2	b_I/λ^2	c_I/λ^2
	R_2^1	R_2^2	R_2^3	R_2^4	R_2^5	R_2^6	R_2^7	R_2^8
	A	A	A	A	A	A	A	A

N_i = the number of atoms of the i -th element in (half) the unit cell.

3.	$d_3^1 D$	$d_3^2 D$	$d_3^3 D$
	x_1	y_1	z_1
	x_2	y_2	z_2
	\vdots	\vdots	\vdots
	x_N	y_N	z_N
	R_3^1	R_3^2	R_3^3
	A	A	A

The order in which the atomic coordinates are punched follows that of the N_i 's.

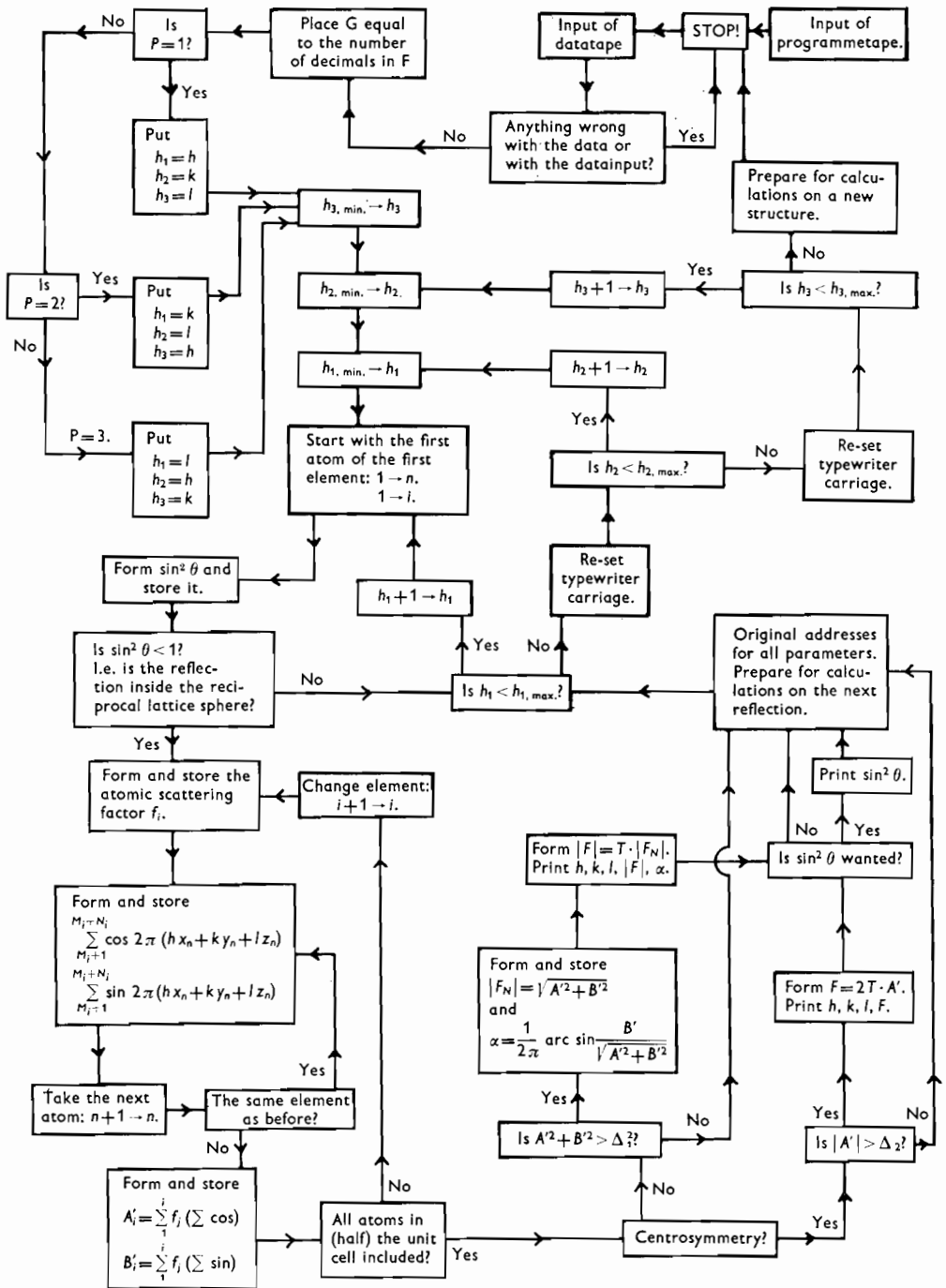


Fig. 1. Flow diagram.

depend on the structure in question except for the constants for the calculation of the f_i 's (which belong in the second set) and the atomic coordinates (which form the third set). The disposition of the data is shown in Table 1.

Punching of data

The data as well as the programme are presented to BESK as punched hole-combinations on a tape. The data tape is punched according to Table 1. In front of each group of data, dD is punched, where d denotes the number of decimals with which the values in the group must be punched, $d \geq 1$. The number of digits ≤ 11 . Each number is, if necessary, completed with zeros so as to contain d decimals. Decimal points and zeros before the first significant digit may be omitted. Every positive number is followed by the signal D (denoting + -sign) and every negative number by the signal B (denoting - -sign).

At the end of each group of data, the signal R is punched, viz. the sum with the opposite sign of all the numbers in the group (not including the heading). The signal A then immediately follows. An erroneous digit or letter may be deleted by punching it over with the signal F .

Organization of the programme

How the programme works is shown in the flow-diagram (*fig. 1*). The sum of each group of data is checked by the machine against the corresponding R -value and if they do not agree the computer stops. The first choice depends on the value of P . This constant determines the order in which the structure factors will be printed. I.e. the F -values are always gathered in groups with varying h_1 -values and constant h_2 - and h_3 -values. When changing from one such group to another, the index h_2 is first varied, then h_3 . The value of P now determines which one of the three indices h, k, l is identical with h_1 . Which index is h_2 or h_3 is determined by cyclic permutation, so there are in all three possibilities: $P = 1, 2$ or 3 . In this way SNUSKMUS connects with SUPERMUS (4), the programme for electron density calculations, the structure factor data of which can be arranged in three different ways.

The calculations begin with the reflection ($h_{\min} k_{\min} l_{\min}$) and $\sin^2\theta$ is first of all calculated. If $\sin^2\theta$ is larger than 1, the reflection cannot be registered with the radiation used. The calculation on this reflection is therefore broken off and one or more of the indices are changed, as shown in the flow-diagram, in preparation for the calculation on the next reflection. If $\sin^2\theta$ is less than 1, the reflection is inside the reciprocal lattice sphere and the next step in the calculations is to form the atomic scattering factor, f_i , which is then stored. Thereafter the expressions $\frac{\cos}{\sin} 2\pi(hx_n + ky_n + lz_n)$ are formed and stored. If the next atom is of the same atomic species as the former, the contribution of the former to the sums $\sum \cos$ and $\sum \sin$ is calculated and stored. This procedure is continued until "the next atom" is of a different atomic species. The contribution of the i :th element to the sums A'_i and B'_i is then calculated and the latter are formed and stored. If not all of the atoms in (half) the unit cell have been used, the i :th element is replaced by the $(i + 1)$:st and f_{i+1} is calculated and so on. If, on the other hand, no more atoms are left, the value of the constant S is examined: If it is zero the structure has a centre of symmetry, if it is not zero ($S = 1$) the structure lacks a centre of symmetry. In both cases, it is tested

whether the reflection is extinguished or not by comparing the structure factor or its absolute value with the very small numbers δ_2 or δ_1 respectively. If it is extinguished, preparations are made for calculations on a new one. If it is not, then, in the centrosymmetric case A' is multiplied by a temperature factor, T , and doubled to form F (the doubling depends on the fact that in the centrosymmetric case the calculations are made on only one half of the unit cell content). Then h, k, l, F and, possibly, $\sin^2\theta$ (if $Q = 0$) are printed. In the non-centrosymmetric case, the absolute value of the structure factor, $|F| = T \cdot \sqrt{A'^2 + B'^2}$, and the phase angle, α , are formed, the last mentioned in parts of a cycle, and $h, k, l, |F|, \alpha$ and, if required, $\sin^2\theta$ are printed. As regards the printing, one gets h, k, l without decimals, α with five decimals and $\sin^2\theta$ with four. The number of decimals in F is determined by the value of the constant G .

In both cases preparations are then made for computations on the next reflection. If, however, the just calculated reflection had the indices $h_{\max}, k_{\max}, l_{\max}$, preparations are made for a possible computation on a new structure and then the machine stops.

The programme can handle any structure for which the number of atomic species does not exceed 12 and for which the number of atoms in the unit cell does not exceed 127 for the non-centrosymmetric case and 254 for the centrosymmetric one. The calculations can be made on every reflection (hkl) whose indices lie in the interval $[-127, 127]$. In the case of centrosymmetry, one puts in only one half of the unit cell content in such a way that if the atom (x, y, z) is included the atom $(1-x, 1-y, 1-z)$ is not. *N.B.* If an atom is situated in a centre of symmetry, it must be treated separate from atoms of the same kind occupying acentric positions. The corresponding atomic form factor constants A_i, B_i, C_i and D_i (but not $a_i, b_i, c_i!$) should thus be given half their normal values.

The constant T is an over-all temperature factor. It is also possible, however, to use individual isotropic temperature factors. The k -th individual temperature constant of the i -th element may be written:

$$\beta_{i,k} = \beta_0 + \Delta\beta_{i,k} \quad (\Delta\beta \geq 0)$$

(Note: β is usually denoted by B but we have chosen this notation β since we have already used B for one of the constants in the expression for $\sin^2\theta$.) One must now first replace the last of the constants of the first set (*cf.* Table 1), β/λ^2 , by β_0/λ^2 and then add the constant $\Delta\beta_{i,k}$ to each of the constants a_i, b_i and c_i . Formally, atoms of the same kind but of different temperature factors are thus treated as different atomic species.

N.B. The fact that $\sin^2\theta/\lambda^2$, which enters into the exponents in the three term expression for f , is a constant depending only on the indices h, k, l of the reflection in question and on the unit cell dimensions of the structure makes possible a shortening of the computing time in certain cases. If, namely, the X-ray exposures have been made with radiation of a relatively short wavelength, e.g. Mo-radiation, the damping effect of the temperature factor will lead to the result that only the reflections from the interior of the reciprocal lattice sphere are registered. If, furthermore, one or more of the unit cell angles of the structure differs considerably from 90° , it is not possible to limit the range for the calculations by putting $|h_{\min}|, |k_{\min}|, |l_{\min}|, |h_{\max}|, |k_{\max}|$

and/or $|l_{\max}|$ sufficiently small. One would then of course lose some of the reflections of the type (hkl) with a large value for one or more of the indices. One may now choose a new suitable artificial wavelength, λ' , which is larger than the one used and then calculate a new set of cell constants A, B, C, \dots , using λ' . In this way, one can "shrink" the reciprocal lattice sphere and thus avoid calculations on a lot of the non-observed reflections.

Computing time

As regards computing time, the following information may be given: The input of the programme tape takes about 20 seconds. The time for the input of the data tape varies with the number of data but is of no importance in comparison with the real computing time. The latter, excluding the time for the output, can be estimated to be $(N + 2I) \cdot 0,01$ seconds/reflection. The time for the output cannot be easily estimated. An approximate value is $0,40 + (S - Q) \cdot 0,11$ seconds/reflection. A rough estimate of the total computing time, t , is thus:

$$t = (N + 2I)(U + V) \cdot 0,01 + [(S - Q) \cdot 0,11 + 0,40] \cdot U \text{ (seconds)}$$

where U is the number of non-extinguished reflections and V the number of extinguished reflections. (The remaining symbols have already been specified in Table 1.)

The programme tape is available at BESK. The instructions for the use of SNUSK-MUS are given in Table 2.

Table 2.

1. Turn knob "UTMATNING" (output) to position "SKRIVMASKIN" (typewriter). Set highest input-speed.
2. Put programme tape in tape reader. Press button " $O \rightarrow MS$ ". Then press button "START FRÅN REMSA" (start from tape).
3. When input has stopped, check signal typed by typewriter. If only "AA" has been typed, continue. Otherwise start again from 2.
4. Turn knob "UTMATNING" (output) to position "STANS" (punch). Set lowest input-speed.
5. Put data tape in tape reader. Press button "START". If anything is wrong with the data or with the data input, the machine stops immediately. Try once again from 5 and if the same thing happens, break off the calculations. If the data and the data input are correct, the calculations are performed and thereafter the machine stops.
6. Begin again from 5 with the next data tape.

Final remarks

The applicability of the principles employed in the programme described above is by no means restricted to the computer BESK for which it has been constructed. They should be applicable to various types of electronic digital computers.

The present programme should be applicable, without any modification, to the computer model FACIT EDB which has just been installed at the Swedish Board for Computing Machinery. The larger core-memory of the latter computer as compared with BESK may, however, make possible an extension of the programme to include e.g. anisotropic temperature factors.

ACKNOWLEDGEMENTS

The support given to this work by the Swedish Natural Science Research Council is gratefully acknowledged.

The authors are also indebted to the Swedish Board for Computing Machinery for having put computer time at their disposal. Thanks are especially due to Dr. G. Dahlquist for his support of the work and to the staff of BESK for their able assistance.

We would also like to express our gratitude to Professor A. Ölander for his kind interest and to Dr. A. Magnéli for his continuing encouragement and many helpful discussions.

Institute of Inorganic and Physical Chemistry, University of Stockholm, Stockholm, Sweden.

REFERENCES

1. Matematikmaskinnämndens arbetsgrupp: Kodning för BESK. 2nd ed. 1958.
2. Available as "Technical Note from the Quantum Chemistry Group, Uppsala University, Rundelsgränd 2 A, Uppsala".
3. C. H. DAUBEN and D. H. TEMPLETON, *Acta Cryst.* 8, 841 (1955).
4. SVEN WESTMAN, GÖRAN BLOMQVIST and STIG ÅSBRINK: "Crystallographic Calculations using an Electronic Digital Computer I. Calculation of Patterson and Electron Density Functions." *Arkiv Kemi* 14, 535 (1959).

Tryckt den 20 oktober 1959

Uppsala 1959. Almqvist & Wiksells Boktryckeri AB