

## Crystallographic calculations using an electronic digital computer

### I. Calculation of Patterson and electron density functions

By SVEN WESTMAN, GÖRAN BLOMQVIST and STIG ÅSRINK

With 3 figures in the text

#### ABSTRACT

A programme for the electronic computer BESK intended for the calculation of three-dimensional Fourier- and Patterson syntheses, sections, projections and bounded projections has been constructed. The possibility of calculating the Patterson or electron density function inside small "boxes" in space is stressed.

The scheme of the calculations could easily be used with any digital computer.

#### Introduction

The computation of electron density or Patterson functions at a sufficient number of points so as to render the result valuable for interpreting experimental structure factor data involves a considerable amount of work. Accordingly, it was thought desirable to carry out these calculations on the electronic computer BESK<sup>1</sup> (1) which is available to research workers in Sweden, especially those in Stockholm.

A programme, designated SI,<sup>2</sup> intended for such calculations was constructed at this institute by M. Edstrand in 1956. It can, however, only be used for electron density and Patterson projections. Moreover, a further limitation is that the presence of a centre of symmetry at the origin of the projections is required.

The present authors have constructed a programme which is not subject to these limitations and which, moreover, provides certain additional features. The programme is called SUPERMUS which is an abbreviation of: "Sekvens för Uträkning av Patterson- och Elektrotäthetsfunktioner i Rymden Med och Utan Symmetri-centrum".

The principles of programming explained in the following should be applicable to a variety of automatic digital computers.

<sup>1</sup> A description of the computer and instructions for programming are given in a manual published by the Swedish Board for Computing Machinery. See reference.

<sup>2</sup> The programme tape and instructions for the use of the programme are available at BESK.

### Scope of the programme

SUPERMUS was constructed to perform the following operations:

1. Automatic calculation of any chosen number of equidistant sections of desired area and point spacing through the electron density or Patterson functions making up a three-dimensional Fourier synthesis within a parallelepipedic portion of space, i.e. a sort of "box". The possibility of calculating small "boxes" with a cross sectional area of  $8 \times 8$  points should be taken advantage of whenever possible in order to save computing time. (See *warning* at end of section entitled organization.)

2. Computation of the contents of several "boxes", of variable size and variable spacing of points, placed at desired positions within the crystallographic unit cell. The structure factor data must be fed to the computer only once, viz. at the beginning of the calculations.

3. Automatic calculation of bounded projections of equidistant slabs of Patterson or "real" space. The slabs may be of any desired thickness and the points of projection may be spaced at will. Structure factor information must be furnished anew if the slab thickness is changed.

4. Calculation, as a special case of 1, of Patterson and electron density projections of desired area and point spacing. It must be pointed out, that the performance of S1 in these cases is generally faster.

5. Cyclic interchange of the unit cell axes so that the indices  $h, k, l$  of the structure factors can always be presented to the programme in the order mentioned. Sections parallel to and projections on to the  $xy$ ,  $yz$  and  $zx$  planes may thus be computed without rearrangement of the indices.

The  $z$  axis ( $w$  axis if the Patterson function is considered) is normally regarded as the "height" in a right-handed coordinate system where the order of the axes is taken as  $x, y, z$ . In the following, this praxis will be generally followed, but any statements made about the  $z$  axis in the setting  $x, y, z$  will, of course, apply equally well to the  $x$  axis in the setting  $y, z, x$  or to the  $y$  axis in the setting  $z, x, y$ . Further interchange of directions and planes in the three settings is obvious.

### Input data

The experimental information, just like the programme itself, is presented to the computer by means of a punched tape and is arranged in "blocks", each block except possibly the last one consisting of 32 items which is as much as can be stored on two lines of the magnetic drum memory of BESK. Each item consists of four or five entries which are, in correct order:

$$h, k, l, F(hkl)$$

or

$$h, k, l, |F(hkl)|, \alpha(hkl)$$

depending upon whether the structure treated is centrosymmetric or not. If the Patterson function is to be computed, the entries should of course be:

$$h, k, l, |F(hkl)|^2$$

The indices  $h, k, l$  must be integers, positive, negative, or zero. No other restrictions are placed on their magnitudes. The  $F$ 's should be scaled so that the absolute value

Table 1. Steering parameters of SUPERMUS III.

Punch:

$1 D$	Heading.	
$x_{\min}$	Origin of box	Positive integers $\leq 127$ or zero.
$y_{\min}$		
$z_{\min}$		
$x_{\max}$	"End" of box	
$y_{\max}$		
$z_{\max}$		
$\Delta x$	Spacing of points in $x, y$ and $z$ directions	
$\Delta y$		
$\Delta z$		
$N$	Number of blocks of $F$ values ( $N \leq 80$ ).	
$n$	Number of items in last block ( $n \leq 32$ ).	
$1000 V$	Scaling constant (see text).	
$S$	Number of digits in the result. $S = 2, 3, 4$ .	
$C$	$C = 0$ : Centrosymmetry. Data include $F$ ( $ F ^2$ ).	
	$C = 1$ : No centrosymmetry. Data include $ F , \alpha$ .	
$T$	$T = 0$ : Structure factors provided previously.	
	$T = 1$ : Structure factor input follows.	
$G$	$G = 0$ : Output $16 \times 16$ points at a time.	
	$G = 1$ : Output $8 \times 8$ points at a time.	
$P$	$P =$ Positive integer $\leq 64 =$ half the height of a slab of the bounded projection. If $P = 0$ , a section is obtained.	
$M$	$M = 1$ : Sections parallel to $yz$ . Height = $x$ .	
	$M = 2$ : Sections parallel to $zx$ . Height = $y$ .	
—	$M = 3$ : Sections parallel to $xy$ . Height = $z$ .	
$R$	Sum, with opposite sign, of entries from $x_{\min}$ to $M$ .	
$A$	The signal $A$ marks end of row.	
	First block follows if $T = 1$ .	

of the function computed does not anywhere exceed the number 4096. The scaling may be effected by means of the steering parameter  $V$  that is discussed in a following section.

The values of the phase angles,  $\alpha_i$ , are to be expressed as fractions of one cycle (i.e.  $2\pi\alpha$  is a value in radians). These are exactly the numbers computed by SNUSKMUS<sup>1</sup> (2).

Only observed structure factors in one half of the reciprocal lattice should be included in the data, i.e. structure factors unrelated by Friedel's law. In order to speed up the calculations, the data should, for reasons mentioned later, be assembled into groups, containing items punched in sequence, within which  $l$  is varied and  $h$  and  $k$  are kept constant (if the  $z$  axis is treated as the "height"). The groups should be made as large as possible, i.e. the division of the reciprocal lattice into halves should be, in this case, effected in such a way as to make both  $F(hkl)$  and  $F(hk\bar{l})$  remain in the data to be treated. This may be done most conveniently by including  $F(hkl)$  with  $h > 0$ ,  $F(0kl)$  with  $k > 0$ ,  $F(00l)$  with  $l > 0$  and  $\frac{1}{2}F(000)$ .

The programme can handle, at the one time, at most 80 blocks (2560 items) assembled into not more than 1536 groups of constant  $h$  and  $k$ .

Before the structure factors are fed into the machine memory, information must be provided concerning the exact procedure to be carried out by the computer. This is performed by punching, immediately before these data, a row of 18 steering parameters. The functions of the parameters are summarized in Table 1.

<sup>1</sup> SNUSKMUS is a programme for calculation of structure factors. See reference.

The coordinates and coordinate spacings given in the table are positive integers, representing 128ths of the unit cell edges. The differences  $x_{\max} - x_{\min}$  and  $y_{\max} - y_{\min}$  must be chosen to conform to the appearance of the output in "squares" comprising  $8 \times 8$  or  $16 \times 16$  points.

If a bounded projection is desired,  $z_{\min}$  should be put  $= \frac{1}{2}(z_1 + z_2)$  of the lowest "slab" and  $P = \frac{1}{2}(z_2 - z_1)$ . Several slabs can be obtained automatically if  $\Delta z > 0$ . Adjacent slabs are obtained if  $\Delta z = 2P$ .

The above instructions refer of course to the case  $M = 3$ .

Sections are obtained if  $P = 0$ . Projections of the whole height of the cell are calculated from the appropriate data with  $P = 0$ .

$V$  is a scaling constant for the  $F$ 's, the latter being stored by the machine as  $F/V$ . The value of  $V$  should be chosen sufficiently large so as to make the magnitude of the result always lie within the capacity of the programme. Too large a value must be avoided however, since it will have a bad effect on the accuracy of the result. This is because  $F/V$  is stored with a precision of  $1/8$  in the first decimal place.

If it is desired to obtain the electron density in absolute measure from absolute intensity data  $V$  should, if practicable, be put equal to half the unit cell volume.

### Punching of data

The data are punched in rows according to Tables 1 and 2. Each row is preceded by the heading  $1D$ , except for the row of  $\alpha$ 's where the heading is  $dD$  denoting the number of decimal places in the  $\alpha$  values. Thus  $4D$  means 4 decimal places.

At the end of each row, the sum, with the opposite sign, of the entries in the row is punched immediately followed by the signal  $A$ .

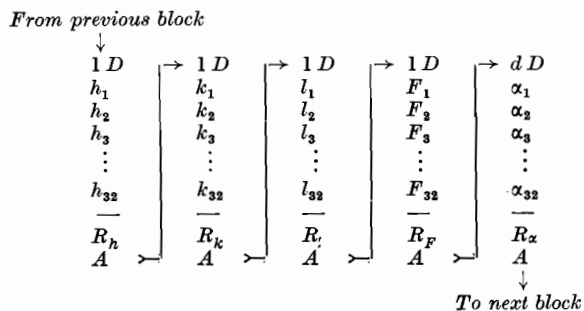
Each steering parameter contains one decimal, generally  $= 0$ . The indices  $h, k, l$  contain no decimals. Every  $F$ ,  $|F|$  or  $|F|^2$  contains one decimal, not necessarily  $= 0$ . The  $\alpha$ 's may contain a number  $d$  ( $1 \leq d \leq 11$ ) of decimals indicated as mentioned above.

Every positive number is followed by the signal  $D$  (denoting  $+$  sign) and every negative number by the signal  $B$  (denoting  $-$  sign).

Decimal points and zeros preceding the first significant digit may be omitted. The number zero is generally written:  $D$ .

An erroneous digit or signal may be deleted by punching over with an  $F$  signal.

Table 2. One block of structure factor data.



The rows of Table 2 are punched in the order indicated, the  $\alpha$  row being however excluded when the structure treated is centrosymmetric.

The last block contains in general less than 32 items. If the total number of items is less than 32 (i.e.  $N = 1, n < 32$ ), then the rows of the first block must be filled with zeros ( $D$ ) so that they comprise nevertheless 32 entries.

### Output

The output of the calculations is recorded decimally on a punched tape which can be automatically interpreted later on and the result typed out on an electric typewriter.

By means of the parameter  $G$  in Table 1, the output may be regulated so that portions, "squares", of a section or projection having "areas" of either  $16 \times 16$  or  $8 \times 8$  points can be computed and recorded as units. The latter possibility is used in order to obtain small boxes round maxima whose positions are already approximately known.

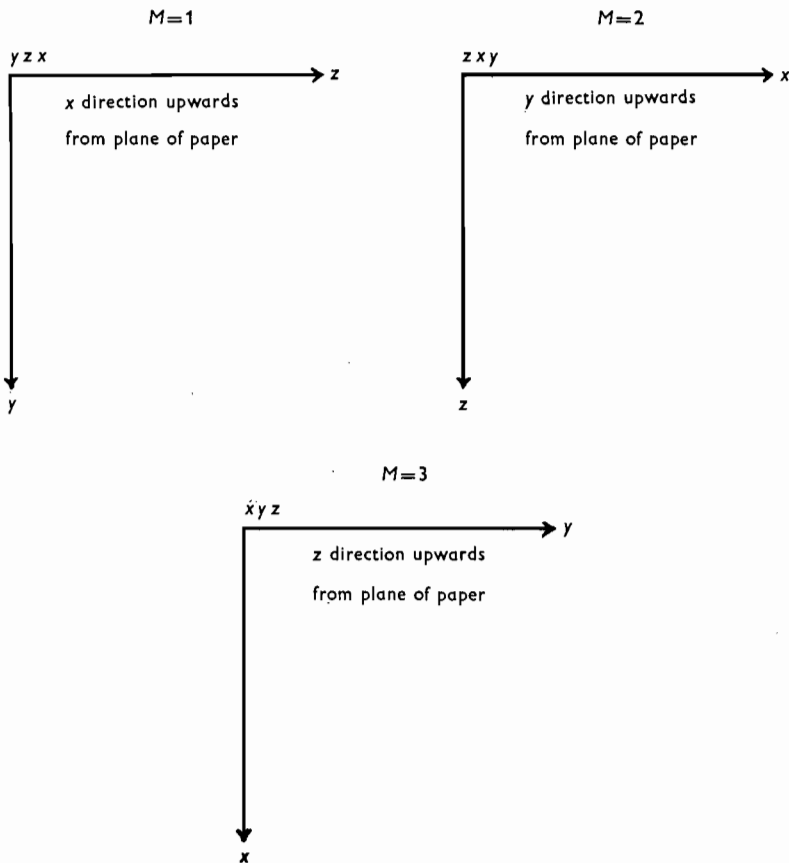


Fig. 1.

Adjacent portions are recorded in "strings" coherent in a direction downwards in the plane of paper. The strings extend from the minimum to the maximum value of the corresponding coordinate ( $x$  if  $M = 3$ ).

In the upper left corner of each string, the coordinates of that particular corner are written down. The order of these coordinates and the directions of the coordinate axes resulting from the choice of the parameter  $M$  in Table 1 are demonstrated in Fig. 1.

The coordinate marking serves as a guide when connecting the strings sidewise to make up whole sections or projections and when stacking the sections in the box.

It is possible of course to obtain one single section, for instance at the height  $z$ , by putting  $z_{\min} = z_{\max} = z$ .

The number of digits in the resulting  $\rho(xyz)$  or  $P(uvw)$  values may be reduced from the maximum 4 to 3 or 2 by means of the parameter  $S$ . The reduction is effected by dividing the result by 8 or 64 immediately before printing. The values are not rounded off.

It should be noted that it is the unreduced result (the one obtained if  $S = 4$ ) that must be kept less than 4096. If it is greater, say  $4096 + \Delta$ , it will be recorded as  $-4096 + \Delta$  or the reduced value thereof.

### Organization of the programme

SUPERMUS treats all structures as though they belonged to space group  $P1$ , i.e. no symmetry elements are taken into account in the calculations. Only the input mechanism is constructed so as to differentiate between centrosymmetric and non-centrosymmetric cases.

The construction of SUPERMUS is outlined in Figs. 2 and 3.

From the organization of the input mechanism, it is evident that once the programme has been used with one specified value of  $M$  (one axial setting) it cannot be used with any other value.

Likewise it is impossible to carry out bounded projections if any sort of run has been performed previously, with the one obvious exception that several slabs of equal thickness can be computed from the same structure factors read only once from the data tape.

The above restrictions are consequences of the lack of space for the programme. The sections of SUPERMUS regulating the input for the cases mentioned are destroyed by the calculations. It may thus be necessary to feed the programme tape repeatedly to the computer if the axial setting is changed or if several bounded projections are to be calculated.

It should also be noted that sections and bounded projections cannot be computed from the same structure factor input.

From the flow diagram in Fig. 3, it is moreover evident that the quantities  $A''$  and  $B''$  are sums of functions of subsequent  $A'$ 's and  $B'$ 's having  $h_1$  and  $h_2$  constant. If the "groups" in the input data are made as large as possible, the result will be the smallest possible number of  $A''$ ,  $B''$  pairs. This reduces the computing time, which may be estimated roughly to be 20 seconds per block of  $A''$ ,  $B''$  (i.e. 32 groups) and  $16 \times 16$  "square".

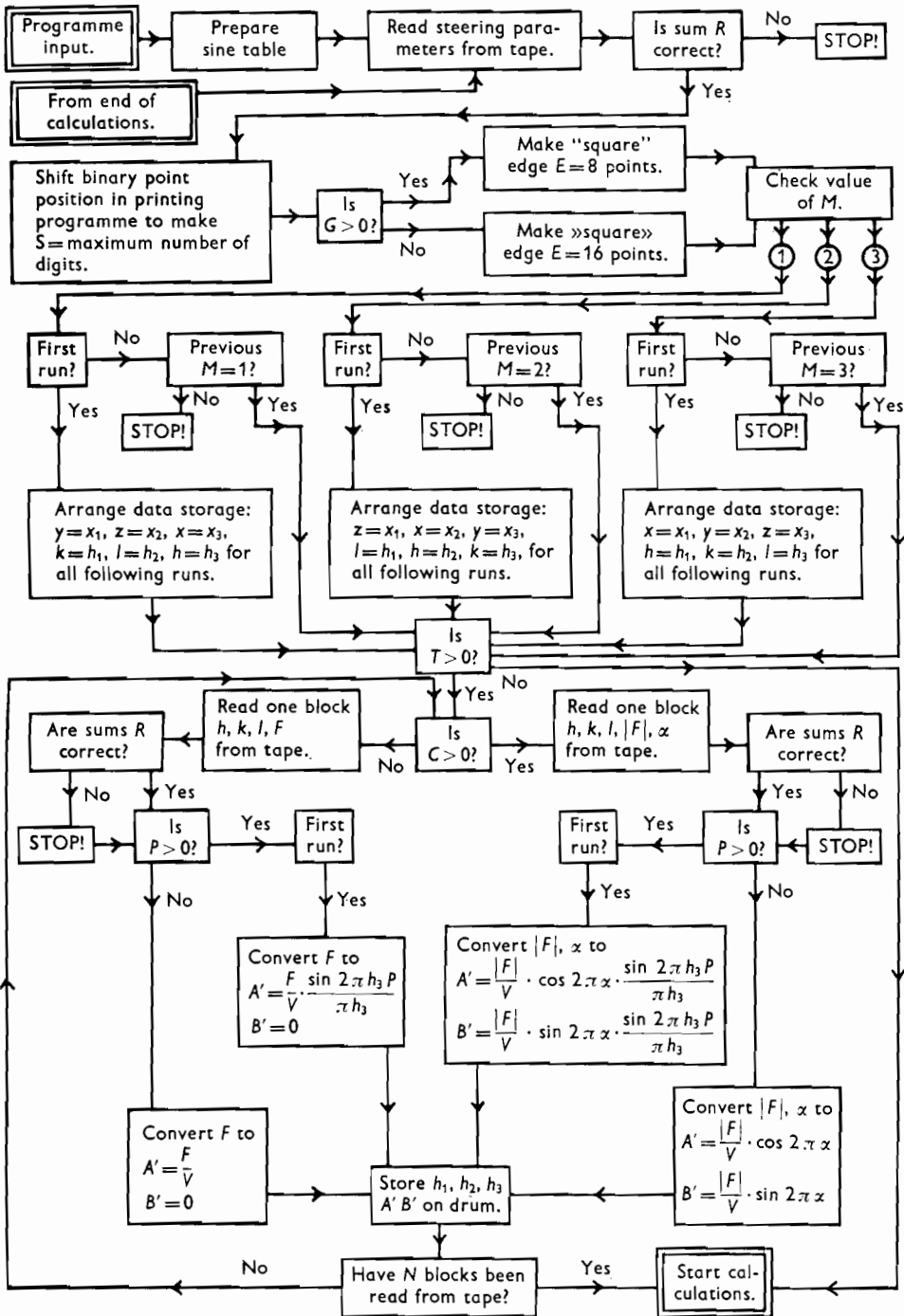


Fig. 2. Input mechanism.

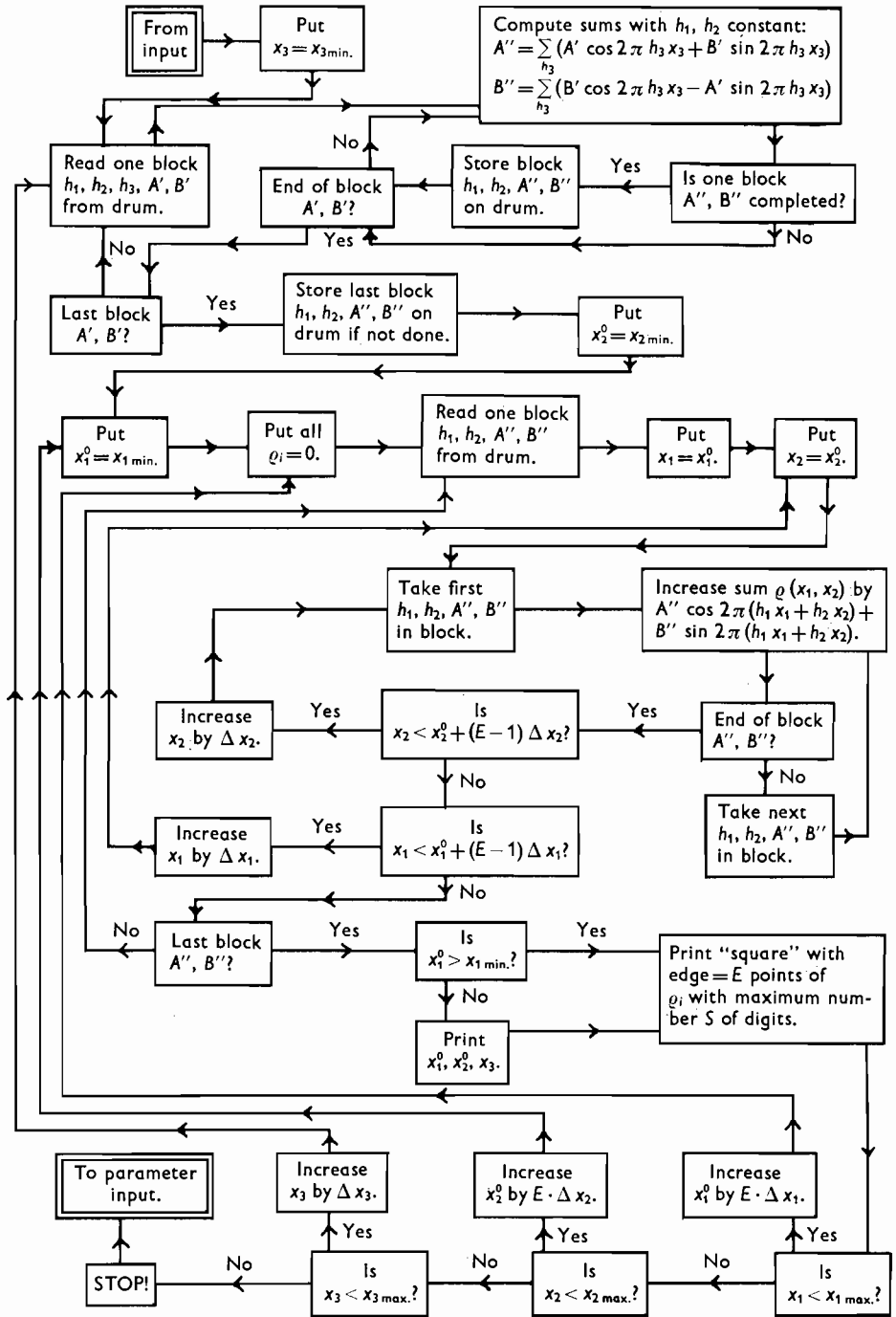


Fig. 3. Flow of calculations.

*Warning*

A complete three-dimensional synthesis with a maximum number of data (1536 groups) and minimum coordinate spacings (1/128) would require c:a 2200 hours = 3 months computing time, not counting the time for the output process.

**Checking**

A number of checks on the correctness of the programme and data are incorporated in SUPERMUS. Thus:

1. If the programme input is correct, the signal *AA* and nothing more will be printed by BESK. If the signal is changed, if it does not appear or if it is followed by one or more digits, then the input is in error.

2. If the sum,  $-R$ , of the steering parameters is not correct or if the parameters are not correctly read from the tape, then the input process will stop and a star ( $\ast$ ) will be printed by BESK. Further data input is barred until the programme tape is put in anew.

3. If any one of the sums,  $-R_i$ , of the rows of entries in the structure factor data is not correct or if a row is not correctly read from the tape, the input process will stop and BESK will print a signal consisting of (in the following order): a designation of the defect row, a star and the decimal number of the block.

The designations are: *A* - *h* row; *B* - *k* row; *C* - *l* row; *D* - *F* row; *E* -  $\alpha$  row. Thus the signal *B*  $\ast$  6 signifies that the *k* row of block 6 is in error.

The input process may be continued if the error is not considered serious. This may be judged from the difference *d* between *R* and the sum of the entries in the row. The number  $-|d|$  (all entries are considered integral) appears in the accumulator (*AR*) of BESK and can be read off in hexadecimal form from the control board of the computer.

4. If the value of *M* is changed from one run to the next, then the input process will stop. Further data input is barred until the programme tape has been put in anew.

*Warning*

An attempt to compute a bounded projection will lead to no definite result if any sort of run has been performed previously. BESK must then be stopped manually. Further data input is barred.

**Operation**

The construction of SUPERMUS is such that a minimum of manual operation is needed for a run. All instructions pertaining to the scope of the calculations are included in the steering parameters. The presence of a crystallographer during the actual run is not necessary.

The programme tape is available at BESK.

The instructions for the use of SUPERMUS III, the latest version of the programme, are given in Table 3.

It should be remarked, in addition, that it is generally advisable to use the lowest possible input speed for the data tape.

Table 3. Instructions for the use of SUPERMUS III with BESK.

1. Turn knob "UTMATNING" (output) to position "SKRIVMASKIN" (typewriter).
2. Put programme tape in tape reader and press button "O→MS". Then press button "START FRÅN REMSA" (start from tape).
3. When input has stopped, check signal typed by typewriter. If "AA" only has been typed, continue. Otherwise start again from 2.
4. Turn knob "UTMATNING" (output) to position "STANS" (punch).
5. Put data tape in tape reader. Press button "START".
6. Check amount of data input. If wrong, stop calculations. If input stops automatically, check signal punched and digits in "AR". If acceptable, press button "START". If not, or if no signal is printed, start again from 1.
7. When calculations have stopped, start again from 5, using next data tape.

#### 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. STIG ÅSBRINK, GÖRAN BLOMQVIST and SVEN WESTMAN: Crystallographic Calculations using an Electronic Digital Computer II. Calculation of Structure Factors. *Arkiv Kemi* 14, 545 (1959).

Tryckt den 20 oktober 1959

Uppsala 1959. Almqvist & Wiksells Boktryckeri AB