# The Use of Punched Cards in Molecular Structure Determinations.
## III. Structure-Factor Calculations of X-Ray Crystallography*

By Jerry Donohue and Verner Schomaker

Department of Chemistry, California Institute of Technology, Pasadena 4, California, U.S.A.

A method is described whereby the rapid calculation of the structure factors of X-ray crystallography is accomplished on punched cards and International Business Machines. The method is sufficiently flexible to allow its application to any space group. The principles set forth may also be applied to other formerly laborious calculations, such as the application of the method of least squares to crystal-structure determinations.

## Introduction

The use of punched cards with a minimal installation of International Business Machines for certain of the calculations which occur in the determination of molecular and crystal structures by X-ray and electron diffraction has been described in two communications from these Laboratories (Shaffer, Schomaker & Pauling, 1946 a, b), and several English groups have described similar methods adapted to somewhat different limited combinations of equipment (Cox, Gross & Jeffrey, 1947; see also L. J. Comrie in discussion on a paper by Beevers (1939)). Two additional types of machines having recently become available to us, we have now extended the scope of these punched-card methods to include the complete calculation of structure factors. We feel that this extension is important because the usual method of calculation with trigonometric tables, ordinary calculating machines, and possibly certain punched-card (Shaffer et al. 1946 a, b) or other aids, which have been adequate although tedious heretofore, is hardly sufficient for the determination of more complex structures, where, for example, the evaluation of the electron-density distribution throughout the unit cell is often required. It is then necessary to calculate the structure factors for the entire {hkl} set of reflections for each of the successive trial structures, whereas in the previously used projection methods only the structure factors of the several prism zones needed to be calculated repeatedly. In the application of the method of least squares (Hughes, 1941) this extension to the use of general reflections becomes equally essential. In any case, not only the number but also the complexity of the structure-factor calculations is increased, making more rapid methods desirable.

Although the methods described in this paper are very successful (but not yet as fast as could be wished),

it is not to be supposed that they will necessarily prove best for all space groups or all combinations of punched-card machines; nor is the range of possible similar applications, such as the evaluation of the derivatives $\partial F_{hkl}/\partial x_i$ required in the least-squares method, more than implied. In all these calculations punched-card methods have the considerable advantages of flexibility and adjustable, but essentially unlimited, precision as compared to specially constructed mechanical computers,* but quite possibly practical experience will show the latter to be faster. It may be hoped that modern super-computers will soon provide the ideal method, at least for the more complicated structures.

## Machines used in the calculations

The machines used in punched-card methods have been described in detail elsewhere (Shaffer et al. 1946 a, footnote 4). The present scheme makes use of the following standard equipment:

Electric key punch (type 12),
Electric sorting machine (type 80),
Automatic reproducing punch (type 513),
Automatic multiplying punch (type 601),
and
Alphabetic direct subtraction tabulating machine (type 405).

The use of a type 602 multiplying punch would speed the calculation considerably, and a collator (type 077) would be useful in some of the operations. The particular application of these two additional pieces of equipment will be obvious.

## The form of the calculation

In general the structure factor of the crystallographic plane (hkl) may be written as

$$F_{hkl} = \sum_j f_j \exp\{2\pi i(hx_j + ky_j + lz_j)\},$$

---

---

* Vand's (1949) new computer, for example, requires all the form factors to have the same dependence on $\sin\theta/\lambda$.

where $f_j$ is the form factor and $x_j$, $y_j$ and $z_j$ are the co-ordinates of the $j$th atom in the unit cell, the summation being carried over all the atoms in the cell. This expression is usually split into its real and imaginary parts:

$$A_{hkl} = \sum_j f_j \cos 2\pi(hx_j + ky_j + lz_j)$$

and

$$B_{hkl} = \sum_j f_j \sin 2\pi(hx_j + ky_j + lz_j),$$

such that $F_{hkl} = A_{hkl} + iB_{hkl}$.

Now in all but the triclinic space groups this form of the structure factor is usually not the one most suited to convenient calculation. More useful expressions for the structure factors in all 230 space groups have been given by Lonsdale (1936). For example, Lonsdale's simplified expression for the space group $D_2^1$-$P222$ is

$$A_{hkl} = 4\sum_j f_j \cos 2\pi hx_j \cos 2\pi ky_j \cos 2\pi lz_j$$

and

$$B_{hkl} = -4\sum_j f_j \sin 2\pi hx_j \sin 2\pi ky_j \sin 2\pi lz_j,$$

where the summation is carried over the non-equivalent atoms in the unit cell. The particular form to be chosen depends on a variety of factors, such as the space group, the number of equivalent atoms in the unit cell, whether or not any atoms lie on symmetry elements of the cell, and so on. The scheme is quite flexible, however, and may easily be adapted to suit best the structure at hand. Since the trigonometric parts of the structure factor may be expressed in a variety of ways, more than one detailed scheme for the calculation will in general be available. Careful study of these several possibilities should be made before embarking on a calculation in order to determine which one will be the most efficient.

## Preliminary work

Before proceeding with the calculation of a set of structure factors a certain amount of preliminary work is necessary. This includes the punching of the following packs of cards:

(1) a pack of 1000 cards on which is punched values of the argument $X$, running from $X = 0.000$ to $X = 0.999$ in intervals of $0.001$, together with the corresponding values of $\cos 2\pi X$ and $\sin 2\pi X$ and their signs. This master pack is then reproduced diagonally several times in order to bring the information into the proper columns indicated by the detailed scheme of the calculation;

(2) a pack of reflection cards (Shaffer et al. 1946a) on which is punched $h$, $k$, $l$, the corresponding values of the form factors of each of the kinds of atoms in the cell, and additional information such as observed intensity, observed $|F|$, and $\sin\theta$; and

(3) three packs of parameter cards.

The first parameter pack contains the following information: trial number, atom number, $h$, and $hx$, where $h$ runs from zero to the maximum value of $h$ observed in any $hkl$, and $x$ is the value of $x$ for the particular atom for the particular trial. The other two packs contain the same general information and, in different columns, the respective values of $k$ and $ky$ or $l$ and $lz$. (The parameter packs could perhaps be obtained more efficiently by the use of a summary punch and tabulator equipped for progressive totals.)

## Details of the calculation

We shall give in detail the procedure for the case, (1), where

$$A_{hkl} = \sum_j f_j(\cos 2\pi hx_j \text{ or } \sin 2\pi hx_j)$$
$$\times (\cos 2\pi ky_j \text{ or } \sin 2\pi ky_j)$$
$$\times (\cos 2\pi lz_j \text{ or } \sin 2\pi lz_j),$$

and $B_{hkl}$ is of similar form (or zero). This form of the structure factor occurs, for example, in all of the orthorhombic space groups and for atoms in certain of the special positions of some tetragonal and cubic space groups. We shall also indicate how the procedure is modified in the case, (2), of the important space group $C_{2h}^5$-$P2_1/c$.

The case (1) above requires that the master pack of $X$ v. $\cos 2\pi X$ and $\sin 2\pi X$ be diagonally reproduced three times to give packs which are designated as $hx$ — $\cos 2\pi hx$ — $\sin 2\pi hx$, $ky$ — $\cos 2\pi ky$ — $\sin 2\pi ky$, and $lz$ — $\cos 2\pi lz$ — $\sin 2\pi lz$. The columns in which $hx$, $ky$ and $lz$ are punched correspond respectively to the columns containing $hx$, $ky$ and $lz$ on the parameter packs.

The values of $\cos 2\pi hx$ and $\sin 2\pi hx$ are transferred to the $hx$ parameter pack by the operations of sorting, on the three columns containing $hx$, the

$$hx - \cos 2\pi hx - \sin 2\pi hx$$

pack with the $hx$ parameter pack and gang punching, the trigonometric cards serving as the interspersed masters. In the same way $\cos 2\pi ky - \sin 2\pi ky$ and $\cos 2\pi lz - \sin 2\pi lz$ are added to the $ky$ and $lz$ parameter packs. (In these operations a collator would prove more efficient.) The reflection pack is now partially reproduced as many times as there are non-equivalent atoms in the cell, only one of the several form factors on the master pack being punched, and always into the same columns on the new cards. For example, in the case of the calculations for the crystal threonine, $C_4O_3N(H_9)$, the reflection pack would be reproduced four times choosing the form factors of carbon, three times choosing the form factors of oxygen, and once choosing the form factors of nitrogen. An identifying atom number is gang punched during the reproduction. The cards obtained in this process are the atom reflection cards. A sample card showing all of the information eventually punched is presented in Fig. 1.

The atom reflection pack is then sorted on atom number, and $h$, $k$ or $l$ in turn with the three parameter packs, and the $hx$, $ky$ and $lz$, together with their cosines and sines, are gang punched, the parameter packs serving as masters.

The atom reflection pack is next sorted to separate it into the various classes, depending on relationship among the indices, into which the reflections fall.* The multiplications for one class may be made without changing the plugboard wiring. If three places of decimals are used, the capacity of the counters on the type 601 multiplying punch is such that two multiplications, e.g. $\cos 2\pi hx \cos 2\pi ky$ and $\sin 2\pi hx \sin 2\pi ky$ may be made simultaneously. If a 602 type multiplier

is used, the three multiplications are performed in one pass of the cards through the machine, effecting a considerable saving of time over the three separate multiplications required when the 601 type multiplier is used.

After the three successive multiplications have been made, the atom reflection pack is sorted for tabulation. We have found it convenient to tabulate the information in two different ways: first, listing the individual atomic

$A_{hkl}$ and $B_{hkl}$, and second, printing only the total values of the $A_{hkl}$ and $B_{hkl}$, as shown in Figs. 2 and 3. The number of atom reflection cards is checked for each $(hkl)$ by means of the tabulator card count.

The above scheme is easily modified for the case of the space group $C_{2h}^5$–$P2_1/c$, where

$$A_{hkl} = 4f_j \cos 2\pi(hx_j + lz_j) \cos 2\pi ky_j, \quad k+l = 2n$$

or $A = -4f_j \sin 2\pi(hx_j + lz_j) \sin 2\pi ky_j, \quad k+l = 2n+1.$

ATOM REFLECTION



Fig. 1. Sample atom reflection card showing the information punched during the calculation.



Fig. 2. Portion of a report showing, for each $(hkl)$, the values of the contributions of the individual atoms to $A_{hkl}$ and $B_{hkl}$, and the total values of $A_{hkl}$ and $B_{hkl}$ for these reflections. The CR (credit symbol) indicates a negative value.



Fig. 3. Portion of a report showing for each $(hkl)$ the totals $A_{hkl}$ and $B_{hkl}$. The values of $|F|^2 = A^2 + B^2$ may be computed manually from this report, or obtained by use of the multiplying punch if the information on this report is recorded with a summary punch also.

In this case a cross-footing operation to obtain $hx_j + lz_j$ on the atom reflection cards precedes the punching of the sines and cosines. It turns out, since $B$ is always zero, that there is sufficient space for two $F_{hkl}$ values, for example those of $F_{hkl}$ and $F_{h+1,k,l}$, to be calculated

* An illustration of this for the space group $D_2^4$–$P2_1 2_1 2_1$ is given by Lonsdale (1936, p. 31). A code punch is used to designate the class.

on the same card simultaneously, thereby reducing the number of operations required.

The general outline of the scheme is sufficiently flexible to enable its application to any space group. It should be mentioned that in some cases it is better to set up the calculating scheme in accordance with the original form of the structure factor, i.e.

$$A_{hkl} = \sum_j f_j \cos 2\pi(hx_j + ky_j + lz_j)$$

and

$$B_{hkl} = \sum_j f_j \sin 2\pi(hx_j + ky_j + lz_j),$$

obtaining the values of the argument by a cross-footing operation. In this instance it is also possible to calculate values for two reflections on one card simultaneously. In any case, the most efficient scheme for a particular job should be carefully worked out before embarking upon the calculation.

The wiring on all plugboards is of the conventional type. Wiring diagrams are available on request to these Laboratories.

This punched-card method reduces greatly the amount of time required to calculate a set of structure factors. It has thus far been applied in the calculation of the structure factors for $l$-threonine (space group $P2_12_12_1$), glycine (space group $P2_1/n$), and acetylglycine (space group $P2_1/c$). In the first case, 625 complex structure factors, depending on the parameters and form factors of eight non-equivalent atoms, were calculated in about 30 man-hours; in the second case, 250 real structure factors, involving five non-equivalent atoms, were

calculated in about 10 man-hours; and in the third case, over 1200 structure factors, involving eight non-equivalent atoms, were calculated in about 30 man-hours, excluding in all cases the time required in the preliminary operations. These times are about one-tenth of those required for the conventional method of calculation with the aid of trigonometric tables. and a motor-driven adding calculator. As pointed out above, these times are more than those which would be required if additional machines mentioned were available.

### References

BEEVERS, C. A. (1939). *Proc. Phys. Soc.* **51**, 660.

COX, E. G., GROSS, L. & JEFFREY, G. A. (1947). *Nature, Lond.*, **159**, 433.

HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.

LONSDALE, K. (1936). *Simplified Structure-Factor and Electron-Density Formulae for the* 230 *Space Groups of Mathematical Crystallography.* London: Bell.

SHAFFER, P. A. Jr., SCHOMAKER, V. & PAULING, L. (1946a). *J. Chem. Phys.* **14**, 648.

SHAFFER, P. A. Jr., SCHOMAKER, V. & PAULING, L. (1946b). *J. Chem. Phys.* **14**, 659.

VAND, V. (1949). *Nature, Lond.*, **163**, 169.

---

# An Improved Punched-Card Method for Crystal Structure-Factor Calculations

BY M. D. GREMS

*Application and Service Engineering Division, General Electric Co., Schenectady, N.Y., U.S.A.*

AND J. S. KASPER

*Research Laboratory, General Electric Co., Schenectady, N.Y., U.S.A.*

Using punched cards and International Business Machines, a method of crystal structure-factor calculations has been developed, which is not only rapid and accurate, but also general for all structures. The principles and application of the scheme are discussed.

## Introduction

The need for rapid and accurate means of performing laborious calculations attending a crystal-structure determination has been satisfied in great measure by Shaffer, Schomaker & Pauling (1946) with the successful application of punched-card methods, using International Business Machines. In particular, the scheme for the evaluation of Fourier series has been eminently successful. However, the method of structure-factor calculations outlined by these authors is not completely satisfactory, and a new scheme has been reported recently by Donohue & Schomaker (1948, 1949).

We wish to describe a significant improvement in the use of punched cards for this purpose, which results in a procedure that we believe to be more direct and more uniform for different crystals than any of those reported previously.

Considering the general expression for the structure factor,

$$F_{hkl} = \sum_{j=1}^{N} f_j \cos 2\pi(hx_j + ky_j + lz_j)$$
$$+ i \sum_{j=1}^{N} f_j \sin 2\pi(hx_j + ky_j + lz_j),$$