Least-squares refinement at high speed.

By P. H. Friedlander,* W. Love and D. Sayre,† *Johnson Foundation for Biophysics, University of Pennsylvania, Philadelphia 4, Pennsylvania, U.S.A.*

This note describes a program for the IBM 701 computer by which least-squares refinement of a crystal structure may be performed at high speed. The program was written to facilitate the two crystal structure determinations we were then engaged in doing, namely 1,2-benzanthracene (P.H.F.) and 9,10-dimethyl-1,2-benzanthracene (D.S.), and its principal characteristics were determined accordingly. These are: (1) that the program is designed for three-dimensional refinement, and accordingly has no provisions for dealing with unresolved atoms; (2) that any space group, with or without a center of symmetry, except for the two triclinic space groups, can be handled; and (3) that all atoms considered are to be identical. This last characteristic, however, can easily be modified; in any case, a program which will be free of this restriction, and which will embody the experience gained with the present program, is contemplated for the IBM 704, an even more powerful machine.

The input to the program is two-fold: (1) a magnetic tape, with a list of $h, k, l, |F_o(h, k, l)|$, and f's, written from standard IBM punched cards by a small preliminary program; and (2) a list of starting coordinates, and a starting scale-factor, on binary punched cards.

At the end of a refinement the machine prints out the coordinate shifts Δx_i, the new coordinates x_i, the discrepancy factor R, and the new scale factor K, and then stops. Pressing the START button causes the next refinement to begin. During each refinement the calculated quantities $A_c(h, k, l)$, $B_c(h, k, l)$, and $|F_c(h, k, l)|$ are written on a second magnetic tape, and they, too, may be printed out at the end.

The expressions which the machine evaluates during a refinement are of the form

Δx_i, etc.

$$= \frac{\sum_H [K|F_o(H)| - |F_c(H, x_i\text{'s})|][\partial|F_c(H, x_i\text{'s})|/\partial x_i, \text{ etc.}]}{\sum_H [\partial|F_c(H, x_i\text{'s})|/\partial x_i, \text{ etc.}]^2}.$$

This expression is derived under three assumptions: (1) that the quantity to be minimised is

$$\sum_H [K|F_o(H)| - |F_c(H, x_i\text{'s})|]^2 ;$$

(2) that all second and higher derivatives of $|F_c(H, x_i\text{'s})|$ may be neglected; and (3) that in the set of linear algebraic equations in the shifts which then result only the terms on the principal diagonal need be retained. Cruickshank (1952) has shown that assumption (3) is justified when the atoms are resolved and the axes are orthogonal or nearly so. In any case, assumptions (2) and (3) cannot affect the final structure, although they may slow down the approach to it.

The calculated structure factors are obtained by the method of Grems & Kasper (1949), and their derivatives by an almost identical method.

As an illustration of the method we present in Table 1 a summary of a succession of refinements on 9,10-dimethyl-1,2-benzanthracene. This molecule crystallises in the non-centrosymmetrical orthorhombic space-group $P2_1nb$, with $a = 7.62$, $b = 8.62$, $c = 21.11$ Å, 4 molecules of $C_{20}H_{16}$ per cell. The starting coordinates for the 20 carbon atoms had been derived mainly by trial and error and, through the courtesy of Dr Pepinsky, by one three-dimensional Fourier summed on X-RAC. The reflections entered into the calculation were the 476 independent reflections with $2 \sin \theta < 1.3$; of these, 7 were too weak to have been measured and they were given one-half their minimum measurable value.

Table 1

Refine-ment	R.m.s. atomic shift (Å)	Maximum atomic shift (Å)	Discrepancy R before refinement (%)
1st	0.148	0.251	29.5
2nd	0.098	0.206	22.9
3rd	0.056	0.087	19.0
4th	0.028	0.052	17.4
5th	0.015	0.029	17.1
6th	0.012	0.019	16.7
All 6	0.275	0.489	—

From Table 1 may be drawn two interesting conclusions about the dynamics of convergence of the least-squares method when applied to a non-centrosymmetrical structure: (1) that the method can 'pull in' even a quite inaccurate structure, and therefore may be applied quite early in a structure determination, thus by-passing a very large amount of hand calculation; and (2) that when a very accurate final structure is required, repeated refinements are necessary.

Each refinement took just over 10 min. In general, the time for one refinement of a structure without a center of symmetry is approximately $RN/1000$ min., where R is the number of independent reflections considered and N is the number of atoms in the asymmetric unit. The rental cost of the 701 computer at the IBM Data Processing Center in New York is $300 per hour. The method is therefore inexpensive as well as fast.

Further information can be obtained from one of us (D.S.).

---

* Present address: Industrial Cellulose Research Limited, Hawkesbury, Ontario, Canada.

† Present address: Electronic Data Processing Center, International Business Machines Corporation, 590 Madison Avenue, New York 22, New York, U.S.A.

References

Cruickshank, D. W. J. (1952). *Acta Cryst.* **5**, 511.
Grems, M. D. & Kasper, J. S. (1949). *Acta Cryst.* **2**, 347.