the sample, were taken from Hamilton & Abrahams (1972). The comparison of the molecular geometries [Fig. 1(*d*)] was with all the intramolecular distances ≤ 4.65 Å instead of bond distances and angles (De Camp, 1973).

The HNP plot for all the positional parameters [Fig. 1(*a*)] is reasonably linear, with nearly zero intercept, suggesting that no systematic error is present. The slope of the plot indicates that the standard deviations are correctly estimated. These results are confirmed by the HNP plot for the intramolecular distances [Fig. 1(*d*)].

The HNP plot for the U_{ij}'s of all the atoms (not reported here) was markedly non-linear. However, elimination of all the U_{ij}'s of the S atoms resulted in the plot of Fig. 1(*b*), which is linear with zero intercept. It shows that systematic errors are absent but the standard deviations are underestimated by a factor of 1·4.

A systematic error in one or both structure determinations is clearly shown by the HNP plot for the U_{ij}'s of the S atoms [Fig. 1(*c*)]. All the observed $\Delta p \geq 3.7$ are relative to the U_{ii} terms, whose correlation coefficient with the overall

scale factor is about 0·4. This suggests that the systematic error is caused by the procedure, used by LB, of changing the interlayer scaling factors during the isotropic refinement.

References

ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* A27, 157–165.
CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* 18, 104–109.
DE CAMP, W. H. (1973). *Acta Cryst.* A29, 148–150.
HAMILTON, W. C. & ABRAHAMS, S. C. (1972). *Acta Cryst.* A28, 215–218.
HANSON, H. P., HERMAN, F., LEA, J. D. & SKILLMAN, S. (1964). *Acta Cryst.* 17, 1040–1044.
LEE, J. D. & BRYANT, M. W. R. (1969). *Acta Cryst.* B25, 2094–2101.
STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* 42, 3175–3187.

An improved method for intensity data collection with a computer-controlled diffractometer. By I. J. TICKLE,

School of Biological Sciences, University of Sussex, Brighton BN1 9QG, England

A method for calculation of the integrated peak intensities of reflexions measured with a computer-controlled four-circle diffractometer is described. This is put forward as an alternative to the commonly used 'ordinate analysis' method for installations where the memory capacity of the computer is severely limited and external storage devices are not available. It is demonstrated by computer simulation that 'ordinate analysis' systematically overestimates weak reflexions whereas this effect is not significant for the method described.

This paper will be of interest mainly to users of a single-crystal X-ray diffractometer controlled by a computer with limited program storage. The method to be described has been programmed for a PDP-8I with neither 'extended memory' nor magnetic storage devices, and is currently in routine use on a Hilger and Watts Y290 four-circle diffractometer. The particular problem of concern here is that of deriving the integrated peak intensity from the measured step-scan reflexion profile. If the peak is consistently centred in the scan the usual 'background-peak-background' method will suffice. Protein crystals, however, when mounted in the usual way (*i.e.* in a capillary adhering by surface tension to the wall) are often subject to small movements which must be allowed for if frequent realignment is to be avoided.

One method in current popular use is 'ordinate analysis' (Watson, Shotton, Cox & Muirhead, 1970) in which a reflexion is scanned in $2n$ steps, starting n steps before the predicted peak position. The peak is taken as the consecutive n steps with the greatest sum. The background intensity is then taken as the sum of the counts for the remaining n steps. Net positive intensity will always be found by this procedure, and a statistical analysis will show that weak reflexions are systematically overestimated.

Alternative methods have been proposed: Diamond (1969) has suggested that the measured profile should be fitted to a stored peak profile which varies over reciprocal space. Vandlen & Tulinsky (1971) describe a scheme in which a realignment subroutine is automatically entered if the intensities of monitor reflexions fall below predesignated values, indicative of crystal motion. Lehmann & Larsen (1974) have shown that the peak can be defined as the set of consecutive steps for which $\sigma(I)/I$ is a minimum (I is the net integrated intensity for the assumed peak); the portion of the scan taken as peak is thus a characteristic of the scan profile, and not, as is common practice, a function only of the Bragg angle. Unfortunately these methods, although superior to 'ordinate analysis' in coping with crystal motion are either too sophisticated for a computer with limited memory capacity, or demand storage of the profiles for subsequent off-line processing, and many diffractometer users will not have the facilities to implement them.

The procedure proposed here can be envisaged in three stages: (i) the reflexion is scanned in $2n$ steps centred on the predicted position and the profile is stored in the computer; (ii) the actual peak position is computed, and (iii) the peak is taken as the n steps centred on the actual position, the

sum of the remaining steps giving the background intensity. In practice the program allows the possibility of unequal peak and background scan widths, which, however, must be preset at empirically determined optimum values. Stage (ii) requires further amplification; several procedures for deriving the peak position from the observed profile were considered; the most convenient seems to be to compute the peak position as the centroid $\bar{x} = \sum_{j=1}^{2n} (C_j - \bar{C}) \cdot x_j /$ $\sum_{j=1}^{2n} (C_j - \bar{C})$ (summations only for $C_j > \bar{C}$) where C_j is the count and x_j is the angular position for the jth step. $\bar{C}$ is a bias level which is taken as the mean of the C_j.

In order to test the ability of this method to locate weak reflexions with sufficient reliability, a simulation program in Fortran was written. The value of a Gaussian function superimposed on a constant background was calculated at each of 40 equally spaced intervals to represent the 'true' peak profile:

$$E(C_j) = E(B) + (2\pi)^{-1/2} \cdot w^{-1} \cdot E(I) \cdot \exp\left[-(j-j_0)^2/2w^2\right].$$

The Gaussian peak width parameter w was set at 3·33 in all calculations; thus the value of the Gaussian function at a position 10 steps from the maximum is 1·1 % of the value at the maximum. The expected background level $E(B)$ was set at 10 and the expected net integrated peak intensity $E(I)$ was varied to give values of $E(S)$ of 1, 2, 3, 5 and 10, where $S = I/\sigma(I)$ and $\sigma(I) = (I+400)^{1/2}$. The 'observed' counts C_j were selected by a random-number generator from the Poisson distribution with expectation $E(C_j)$. The peak position j_0 was also selected randomly from the uniform distribution in the range -10 to $+10$ to simulate crystal motion. In each experiment 500 profiles were generated and the net integrated intensity obtained both by 'ordinate analysis' and by the method described. In addition, the counts for 20 steps centred on j_0 were summed to give, after subtraction of the remaining background, a control value I_c. Typical distributions of S [for $E(S)=3$] are shown in Fig. 1.
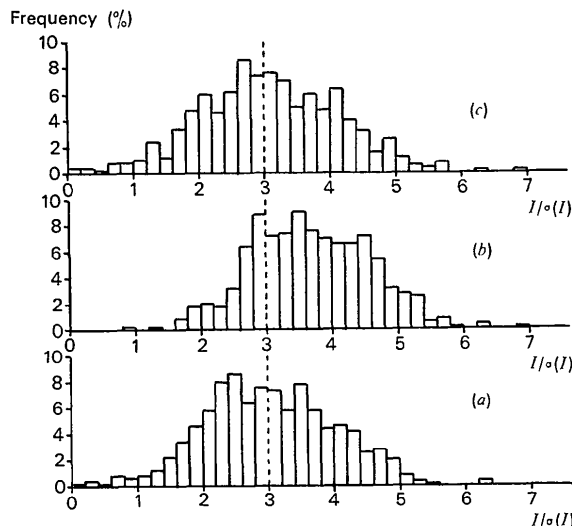

Fig. 1. Histograms of the distribution of $S[=I/\sigma(I)]$ for $E(S)=3$. (a) Control, (b) ordinate analysis, (c) new method.

In each run the following statistics were accumulated (Table 1): the mean net intensity/σ ratio $\bar{S}$, the r.m.s. deviation in S, the skewness in the distribution of S, and the conventional R values $R_1 = \sum|I - E(I)|/\sum E(I)$ and $R_2 = \sum|I - I_c|/\sum I_c$. R_1 is thus a measure of the absolute accuracy and R_2 of the accuracy of peak location. The frequency of the error in the calculated peak position for $E(S)=5$ is shown in Table 2. The measurements were divided into two groups: those with $S_c < E(S)$ and those with $S_c \geq E(S)$ to demonstrate that the largest errors in the peak position occur when I_c is smaller than the expected value.

Table 2. *Frequency of the error in the calculated peak position for $E(S)=5$*

Ordinate analysis

Position error:	0	1	2	3	4	5	6	7	8	9
$S_c < 5$ Frequency(%):	8	18	17	24	14	12	5	0	0	1
$S_c \geq 5$	17	26	21	21	11	2	1	0	0	0

New method

Position error:	0	1	2	3	4
$S_c < 5$ Frequency(%):	21	42	25	9	3
$S_c \geq 5$	36	47	15	2	0

It is clear that weak reflexions measured by the new method are not on average subject to the positive bias that is a serious defect of the ordinate analysis method. It is interesting in particular that the method allows the possibility of a net negative peak intensity. A further advantage of the method is that, since the peak position is calculated precisely for all but the very weak reflexions, any slight misalignment of the crystal will be immediately apparent. The method suffers, however, from the disadvantage of ordinate analysis in that there is clearly a limit to the amount of crystal movement that can be tolerated. It should be particularly noted that the method may give an incorrect result for a profile with a non-uniform background

Table 1. *Summary of statistics for the peak-scan simulation*

$E(S)$	Method*	$\bar{S}$	$\overline{(S-\bar{S})^2}^{1/2}$	Skewness	R_1	R_2
	Control	1·04	1·00	0·102	0·810	0
1	A	1·93	0·84	0·274	1·021	0·802
	B	1·07	1·10	−0·052	0·899	0·587
	Control	2·04	1·01	0·108	0·405	0
2	A	2·78	0·89	0·117	0·478	0·347
	B	2·02	1·14	−0·054	0·463	0·267
	Control	3·03	1·01	0·094	0·273	0
3	A	3·68	0·91	0·134	0·305	0·203
	B	3·08	1·09	−0·004	0·292	0·145
	Control	5·02	1·00	0·109	0·163	0
5	A	5·56	0·92	0·114	0·177	0·100
	B	5·03	1·03	0·175	0·172	0·054
	Control	9·99	0·96	0·122	0·082	0
10	A	10·33	0·91	0·117	0·084	0·034
	B	10·00	0·98	0·152	0·085	0·009

* Method A is ordinate analysis, method B is that described in this paper.

and/or an asymmetric peak, as is often encountered in a θ/ω scan. Data collection by the ordinate analysis method is, however, normally carried out by means of an ω scan, which usually gives a symmetric peak profile, since experience indicates that intensities measured in this way are less sensitive to crystal motion.

References

DIAMOND, R. (1969). Acta Cryst. A25, 43–55.
LEHMANN, M. S. & LARSEN, F. K. (1974). Acta Cryst. A30, 580–584.
VANDLEN, R. L. & TULINSKY, A. (1971). Acta Cryst. B27, 437–442.
WATSON, H. C., SHOTTON, D. M., COX, J. M. & MUIRHEAD, H. (1970). Nature, Lond. 225, 806–811.

On the method of assessing the number of atoms in the unit cell of an organic crystal.* By R. SRINIVASAN and S. S. RAJAN, Centre of Advanced Study in Physics, University of Madras, Guindy Campus, Madras-600025, India

It is shown that the rough method of arriving at the number of non-hydrogen atoms from the volume of the unit cell of an organic crystal [Kempster & Lipson (1972). Acta Cryst. B28, 3674] can be improved by including hydrogen atoms. The empirical relation $V = 8·9N$, where V = volume of the unit cell and N = number of atoms including hydrogen, is arrived at by a least-squares fit of data from 114 compounds.

Kempster & Lipson (1972), hereinafter referred to as KL, obtained empirically a relationship between the volume (V) of the unit cell of an organic crystal and the number (N) of atoms it contains, and gave the relation $N \simeq V/18$. They, however, excluded hydrogen atoms. Although the above is a rough relation and gives the number of non-hydrogen atoms with reasonable accuracy we felt that the hydrogen atoms cannot be excluded. The van der Waals radius for hydrogen is about 1·2 Å whereas carbon, nitrogen and oxygen have a value around 1·6 Å. We therefore carried out the following calculations on a total of 114 organic compounds (aliphatic). In the first case the calculations of KL were repeated and a least-squares fit was obtained for the parameter A' in the relation, $V = N_1 A'$ where N_1 now is the total number of atoms excluding hydrogen. In the second case the value of A in the relation $V = NA$ was obtained by a least-squares calculation where N denotes the total number of atoms including hydrogen. In the third case a two-parameter fit was attempted by least-squares calculation for the relation $V = B_1 N_1 + B_2 N_2$ where B_1 and B_2 are the parameters to be determined and N_1 and N_2 are the number of non-hydrogen atoms (C,N,O) and hydrogen atoms respectively. The values obtained are given in Table 1. For comparison the r.m.s. value of the fractional error $\langle[(V - V_c)/V]^2\rangle^{1/2}$ where V_c is the calculated value of the volume using the appropriate parameters was also obtained for each of the cases and is given in Table 1.

It may be seen that the value of A' (18·5) is close to that reported by KL. The use of a single parameter (including hydrogen atoms) may be seen to yield an appreciable improvement in the fit since the r.m.s. error is much lower (0·098) than for the earlier case (0·158). The two-parameter case seems to give a slightly better result with a standard error of 0·093 but the improvement is perhaps not as

pronounced as might be expected from the use of two parameters.

Fig. 1. gives $\langle(V - V_c)/V\rangle$ as a function of $\langle N_1/N_2\rangle$ where the average is over selected groups of compounds with approximately equal number in the different ranges of N_1/N_2. The agreement may be seen to be poor for the case
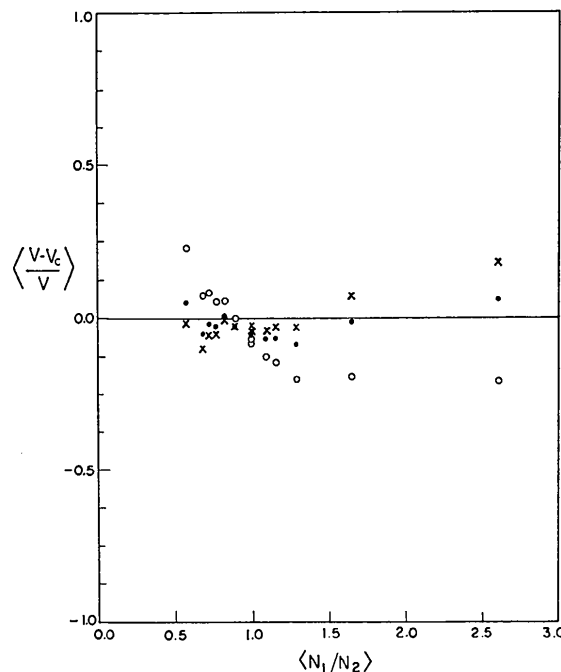

Fig. 1. Average fractional error in the calculated volume as a function of $\langle N_1/N_2\rangle$ for the cases with A' (○), A (×) and B_1 and B_2 (●).