Acta Cryst. (1967). **23**, 888

Statistical Variance of Line-Profile Parameters. Measures of Intensity, Location and Dispersion

By A. J. C. Wilson

Physics Department, The University of Birmingham, P.O. Box 363, Birmingham 15, England

(*Received* 13 *March* 1967)

Diffractometric line profiles are ordinarily of little use in themselves, but are measured in order to derive from them some parameter of physical interest, such as the centroid position, the peak position, the variance, the integral breadth, the Fourier coefficients, *etc.* Since the counting rate I_j at the diffractometer setting $2\varphi_j$ is a random variable, all parameters derived from the I_j are also random variables, and in order to assess their reliability it is necessary to know their variances. Two methods of diffractometer operation, fixed-time counting and fixed-count timing, are in common use, and a third, minimum-variance counting, has been suggested. Expressions for the variances of measures of intensity, location and dispersion are derived or quoted for all three methods of operation. Since for fixed-time counting the variances depend linearly on the I_j, whereas those for fixed-count timing depend on I_j^2 and those for minimum-variance counting depend on $I_j^{\frac{3}{2}}$, the expressions for fixed-count timing are generally the simplest. Background interpolation is always a complicating factor.

1. Introduction

1.1. Accurate measurements of X-ray intensities are now usually obtained by the use of quantum counters, Geiger, proportional, or scintillation. The powder diffractometer, for example, is set at a particular angle, $2\varphi_j$, and the intensity is measured by counting the quanta detected. This process is repeated at a sequence of angles, distinguished by the suffix j, and the various parameters of the reflexion are calculated from the counting rate I_j observed at $2\varphi_j$, a background correction being made if necessary. Among the parameters in general or prospective use are the following.

I. Measures of intensity:
  (i) the peak intensity,
  (ii) the integrated intensity.
II. Measures of location:
  (iii) the peak position,
  (iv) the median position,
  (v) the centroid position.
III. Measures of dispersion:
  (vi) the half width,
  (vii) the integral breadth,
  (viii) the mean absolute deviation from the median,
  (ix) the variance,
  (x) the Fourier coefficients.
IV. Measures of coherence:
  (xi) the particle-size distribution,
  (xii) the strain distribution,
  (xiii) the distribution of mistakes and stacking faults.

The parameters measuring intensity, location and dispersion can all be obtained in a more or less routine fashion from the actual observations, and are so quite objective, except perhaps for a little human intervention in the estimation of background level. Unfortunately it has not yet been possible to reduce the estimation of the measures of coherence to a routine, and one may view with some caution any detailed interpretation of such distribution functions. Since I_j is a random variable all quantities obtained by manipulating it are also random variables, and for the proper design of experiments it is necessary to have at least an order-of-magnitude estimate of the statistical variance of the derived quantities. General statistical theory leads to the conclusion that the variance of a function of several variables, $F(x_1,x_2,x_3,\ldots)$, is given by

$$\sigma^2(F) = \sum_{i,j} \frac{\partial F}{\partial x_i}\frac{\partial F}{\partial x_j}\operatorname{cov}(x_i,x_j)\,, \tag{1}$$

where $\operatorname{cov}(x_i,x_j)$ is the covariance of x_i and x_j if $i\neq j$ and is the variance of x_i, $\sigma^2(x_i)$, if $i=j$ (Cramér, 1946, p. 295). The measured intensities, I_j, are statistically independent, so that $\operatorname{cov}(I_i,I_j)=0$, and where F is expressed directly as a function of the I_j equation (1) reduces to

$$\sigma^2(F) = \sum_j \left(\frac{\partial F}{\partial I_j}\right)^2 \sigma^2(I_j)\,. \tag{2}$$

These equations are strictly correct only if F is a linear in the x's or if the variance of each x is very small. Except for the integrated intensity, the functions of interest are non-linear, but the contribution of each I_j is in general sufficiently small to make the application of equations (1) and (2) plausible. Sands (1966) has recently used equation (1) in a discussion of the covariance of bond lengths and bond angles.

1.2. In this paper the estimation of the variance of parameters of intensity, location and dispersion is discussed, and expressions are derived for some not previously investigated. The following assumptions are made:

(i) The observations have been made at an odd number $R=2r+1$ of equally spaced positions, numbered from $-r$ to $+r$,

(ii) The background is determined by linear interpolation between observations made just outside the above range, so that the background is statistically independent of the line profile,

(iii) The range is adjusted, whenever it is convenient, so that the peak or centroid, whichever is convenient, is less than half a step from $2\varphi_0$, and

(iv) The effect of variations in the step length is negligible.

The first assumption is convenient, and it is clear that the use of an odd instead of an even number of points does not alter the problem in any essential. The interpolated background,

$$G_j=g+kj\,, \tag{3}$$

is ordinarily determined by averaging observations at several steps, or counting for a longer than usual time, so that the variance of G_j is a small fraction (in principle a negligible fraction) of G_j itself (Appendix A). The assumption that the linearly interpolated background is the true background, on the other hand, is probably a large source of experimental error, as it makes no allowance for possible non-linear variation of general scattering, unsuppressed white radiation, the peaking-up of thermal scattering under the Bragg peak, and similar troublesome effects. Though a source of experimental error, this assumption is not likely to affect seriously the limited objective of estimating statistical fluctuations in the derived parameters. Choosing the range so that it is approximately symmetrical is frequently convenient, and symmetry about the centroid is a necessary (though not sufficient) condition for the Fourier components to be purely real. Centring within half a step can always be achieved by dropping a few observations at one end or other of the profile (Pike & Wilson, 1959). The effect of random variations in the step length on the peak position has been discussed by Wilson (1965).

1.3. The commonest method of determining the counting rate at the diffractometer setting $2\varphi_j$, fixed-time counting, is to accumulate counts during a predetermined time, τ, and divide the number accumulated, N_j, by τ. A variation of this method, fixed-dose counting, is to accumulate counts during the time required by a monitor counter fed by the same X-ray source to accumulate a predetermined number of counts. This method gives a somewhat higher variance than fixed-time counting, but is useful when it is impracticable to stabilize the X-ray output (see, *e.g.* Eastabrook & Hughes, 1952). Now that highly stabilized X-ray sources are available it is of less importance. The second basic procedure for determining the counting rate is to measure the time, t_j, required to accumulate a predetermined number of counts c. The two expressions for the counting rate are thus

$$I_j=N_j/\tau \tag{4}$$

for fixed-time counting, and

$$I_j=c/t_j \tag{5}$$

for fixed-count timing. It is convenient to begin by recalling the statistical properties of I_j as determined by the two methods.

2. Counting-rate distribution functions

2.1. The statistical distribution functions for the two methods are closely related. The probability of obtaining c counts in a time τ is given by the discrete Poisson distribution

$$p_\tau(c) = \frac{(\lambda\tau)^c \exp(-\lambda\tau)}{c!} \tag{6}$$

where λ is an adjustable parameter. It is easily shown that the mean counting rate is

$$\langle I\rangle=\langle c\rangle/\tau=\lambda\,, \tag{7}$$

with a variance of

$$\sigma^2(I)=\langle(c-\langle c\rangle)^2\rangle/\tau^2=\lambda/\tau\,. \tag{8}$$

Fixed-count timing, on the other hand, gives rise to a continuous distribution. The probability of requiring a time between t and $t+dt$ to accumulate c counts is

$$p_c(t)dt = \frac{(\lambda t)^{c-1} \exp(-\lambda t)}{(c-1)!}\, d(\lambda t)\,. \tag{9}$$

This looks very like the equation for the Poisson distribution, but the roles of variable and parameter are reversed; in equation (6) τ is fixed and c is determined experimentally, whereas in equation (9) c is fixed and t is determined experimentally. The exact expressions for the mean counting rate and its variance are

$$\langle I\rangle = \frac{c\lambda}{c-1} \tag{10}$$

and

$$\sigma^2(I) = \frac{c^2\lambda^2}{(c-1)^2(c-2)}\,, \tag{11}$$

but in most practical applications c is a large number and the expression for the mean intensity and its variance can be written as

$$\langle I\rangle=\lambda \tag{12}$$

and

$$\sigma^2(I)=\lambda^2/c\,. \tag{13}$$

If c is not large equations (10) and (11) have to be used; the difference is 1% in counting rate and 4% in variance for $c=100$.

2.2. In diffractometric applications the value of the parameter λ is ordinarily unknown. It could be obtained with considerable accuracy by averaging repeated observations at each setting of the diffractometer, but this would make the time required for measuring a line profile so long as to be impracticable. It is thus necessary to take a single measurement I_j of the counting rate at the setting $2\varphi_j$ as the only available estimate of the corresponding λ. The estimate of variance is then

$$\sigma^2_{\rm ft}(I_j) = I_j/\tau \tag{14}$$

for fixed-time counting, and

$$\sigma^2_{\rm fc}(I_j) = I_j^2/(c-2) \tag{15}$$

$$\sim I_j^2/c \tag{16}$$

for fixed-count timing. The variances of a derived parameter F are thus, from equation (2),

$$\text{(fixed time)} \ \sigma^2_{\rm ft}(F) = \frac{1}{\tau}\sum_j \left(\frac{\partial F}{\partial I_j}\right)^2 I_j\,, \tag{17}$$

and

$$\text{(fixed count)} \ \sigma^2_{\rm fc}(F) = \frac{1}{c}\sum_j \left(\frac{\partial F}{\partial I_j}\right)^2 I_j^2\,. \tag{18}$$

Often, the available equipment will function in only one of the two modes of intensity measurement, and the question of choosing between them will not arise. If both are available, the first problem in experimental design will be to decide which will give the smaller value of $\sigma^2(F)$, c and τ being determined so that each method would require about the same total time for collecting the observations. In principle it is possible to adjust τ_j or c_j as a function of I_j so that the variance of any desired parameter is minimized. This was shown by Zevin, Umanskij, Khejker & Pančenko (1961) for the centroid, and generalized by Wilson, Thomsen & Yap (1965). They found that the general expression for the minimum variance of the parameter F is

$$\sigma^2_{\rm min}(F) = \frac{1}{T}\left\{\sum_j \left|\frac{\partial F}{\partial I_j}\right| I_j^{\frac{1}{2}}\right\}^2\,, \tag{19}$$

where T is the total time allowed for collecting the observations. It is remarkable that the minimum variance is a perfect square (aside from the factor T), so that the standard deviation takes a simple form. Unfortunately no instrument capable of producing the necessary continuous adjustment of the counting time is yet in existence. In terms of the total counting time equations (17) and (18) take the form

$$\sigma^2_{\rm ft}(F) = \frac{R}{T}\sum_j \left(\frac{\partial F}{\partial I_j}\right)^2 I_j \tag{20}$$

and

$$\sigma^2_{\rm fc}(F) = \frac{R\langle I_j^{-1}\rangle}{T}\sum_j \left(\frac{\partial F}{\partial I_j}\right)^2 I_j^2 \tag{21}$$

for fixed time and fixed count respectively. It is clear that parameters that change rapidly with I where I is small and slowly with I where I is large will tend to have variances smaller than those of parameters with the opposite relation to I. Parameters dependent only on the line profile in the neighbourhood of the peak will have much the same variance whatever method is used, since I_j itself does not vary greatly in this region.

2.3. Since fixed-time counting leads to variances depending directly on I_j, these variances tend to be more simply related to the properties of the line profile than do the variances resulting from fixed-count timing or minimization techniques. The latter involve similar relations to the square or square root of the line profile, and are thus less easy to visualize. There is perhaps some analogy with the simplicity of an electron-density synthesis compared with a Patterson synthesis.

3. Measures of intensity

3.1. *The peak intensity*

The calculation of the variance of the peak intensity is trivial, whatever method of measurement is chosen, and gives effectively I_0/T. Mack & Spielberg (1958) have discussed in some detail the best use of the available counting time when the quantities of interest are nett peak heights (actual height minus background), or differences or ratios of nett peak heights, as in quantitative analysis. The variance of the nett peak height is, of course, the sum of the variances of the peak height and of the background at the peak position (Appendix A).

3.2. *The integrated intensity*

The integrated intensity is given by

$$L_0 = \sum_j (I_j - G_j)\,, \tag{22}$$

so that $\partial F/\partial I_j$ is unity. The fixed-time variance is thus the total intensity of line plus background divided by τ, and for lines of ordinary profile will be less than the fixed-count variance, which is the total of (line plus background)^2 divided by c. The explicit equations are

$$\sigma^2_{\rm ft}(L_0) = (L_0 + Rg)/\tau + \tfrac{1}{4}R^2[\sigma^2_{\rm ft}(G_R) + \sigma^2_{\rm ft}(G_L)]\,, \tag{23}$$

and

$$\sigma^2_{\rm fc}(L_0) = \sum_j I_j^2/c + \tfrac{1}{4}R^2[\sigma^2_{\rm fc}(G_R) + \sigma^2_{\rm fc}(G_L)]\,, \tag{24}$$

where the background variances are given in Appendix A. The minimum variance [equation (19)] is

$$\sigma^2_{\rm min}(L_0) = \frac{1}{T}\{\sum_j I_j^{\frac{1}{2}}\}^2 + \tfrac{1}{4}R^2[\sigma^2_{\rm min}(G_R) + \sigma^2_{\rm min}(G_L)]\,. \tag{25}$$

The integrated intensity is not affected by the background slope, but it is sensitive to the mean background. The variance arising from this source is indicated explicitly in equations (23) to (25), but in all reasonable cases it is nearly enough R^2g/T_g, where T_g is the total time devoted to the determination of the background level.

4. Measures of location

4.1. *Variance of the peak position*

The variance of the peak position has been discussed by Wilson (1965), who considered three methods of peak location: fitting a least-squares parabola, bisecting horizontal chords, and using a type of scanning photometer. Only the first of these lends itself to a strict estimate of variance of the peak position. If the least-squares parabola has the equation

$$I_j = A + B\delta j + C\delta^2 j^2\,, \tag{26}$$

Wilson showed that the variance of the peak position is approximately

$$\sigma^2_{\text{ft}}(-B/2C\delta) = \frac{(P_2^2 - P_0P_4)^2\delta^5 M_2}{4P_2^2(P_2\delta^2 M_0 - P_0 M_2)^2}\,, \tag{27}$$

where

$$P_k = \sum_{j=-p}^{p} j^k\,, \tag{28}$$

δ is the step length in 2φ, $P = P_0 = 2p+1$ is the number of observations to which the parabola is fitted, and M_0 and M_2 are the zero and the second moment about the peak of the portion of the line profile used in fitting the parabola. This expression was derived for fixed-time counting, but the range of observations P will be much smaller than R and the intensities I_j will vary by only a small amount, perhaps 5% or 10%, so that other methods of observation would not give appreciably different variances, provided that the count is adjusted to give the same total time for the portion of the profile actually used.

The effect of background slope on the peak position is discussed in Appendix B.

4.2. *Variance of the median*

The median position divides the line in such a way that half the intensity lies to the left of it and half to the right. It is then at the step m that most nearly satisfies the equation

$$\sum_{j=-r}^{m-1} (I_j - G_j) = \sum_{j=m+1}^{r} (I_j - G_j)\,. \tag{29}$$

Though m is determined by the I_j, it is not a differentiable function of them, and so equation (2) cannot be applied. The following argument should, however, give a rough estimate of the variance of the median position. Suppose that the true median is at m and the corresponding true intensity is I_m. Any particular set of measurements is unlikely to satisfy equation (29). Suppose that the actual value of the left-hand side of (29) is L_1 and of the right-hand side is L_2. The step actually determined for the median by these measurements will not be m, but some other position n given approximately by

$$n = m + \tfrac{1}{2}\,\frac{L_2 - L_1}{I_m}\,, \tag{30}$$

where it is assumed that I_j does not vary rapidly near $j = m$. The variance of the median step is thus

$$\sigma^2(n-m) = \frac{1}{4I_m^2}\,\sigma^2(L_2 - L_1) \tag{31}$$

$$= \frac{1}{4I_m^2}\,\{\sigma^2(L_1) + \sigma^2(L_2) - 2\,\text{cov}(L_1, L_2)\} \tag{32}$$

$$= \frac{1}{4I_m^2}\,\{\sigma^2(L_0) - 4\text{cov}(L_1, L_2)\} \tag{33}$$

where L_0 is the total line intensity. The covariance of L_1 and L_2 arises only through the interpolated background, and is thus small. To a first approximation, therefore, the variance of the median angle is

$$\sigma^2(2\varphi_m) = \delta^2\sigma^2(n-m) = \frac{\delta^2}{4I_m^2}\,\sigma^2(L_0) \tag{34}$$

where $\sigma^2(L_0)$ is given by the appropriate one of equations (23) to (25).

4.3. *Variance of the centroid*

4.3.1. The centroid of the line profile is given by

$$\langle 2\varphi\rangle = \frac{\sum_j (I_j - g - kj)(2\varphi_j)}{\sum_j (I_j - g - kj)} \tag{35}$$

$$= \frac{\sum_j (I_j - kj)(2\varphi_j)}{\sum_j (I_j - g)}\,, \tag{36}$$

the terms in g vanishing from the numerator and those in k from the denominator because of the symmetry of the range chosen. Since g and k are not independent variables (Appendix A) it is simpler to write them in terms of the backgrounds at the ends of the range, G_L and G_R, and thus avoid covariance terms. Equation (36) becomes

$$\langle 2\varphi\rangle = \frac{\sum_j [I_j - (G_R - G_L)j/R](2\varphi_j)}{\sum_j I_j - \frac{1}{2}R(G_R + G_L)}\,, \tag{37}$$

so that

$$\frac{\partial\langle 2\varphi\rangle}{\partial I_j} = \frac{2\varphi_j}{L_0} - \frac{\langle 2\varphi\rangle}{L_0}\,, \tag{38}$$

$$\frac{\partial\langle 2\varphi\rangle}{\partial G_R} = \frac{-\sum j(2\varphi_j)}{RL_0} + \frac{R\langle 2\varphi\rangle}{2L_0}\,, \tag{39}$$

$$\frac{\partial\langle 2\varphi\rangle}{\partial G_L} = \frac{\sum j(2\varphi_j)}{RL_0} + \frac{R\langle 2\varphi\rangle}{2L_0}\,, \tag{40}$$

where L_0, as in equation (22), is the integrated intensity of the line alone.

These expressions could now be substituted in equation (2), giving a long and clumsy expression. The range has been chosen, however, so that $\langle 2\varphi\rangle$ is less than $\frac{1}{2}\delta$, and the terms in $\langle 2\varphi\rangle$ can thus be ignored in the estimation of the variance of the centroid. Since $2\varphi_j = j\delta$ the sum in equations (39) and (40) is

$$\delta \sum_{j=-r}^{r} j^2 = \tfrac{1}{3} r(r+1)(2r+1)\delta \quad (41)$$

$$= \frac{1}{12}(R-1)R(R+1)\delta$$

$$\sim \frac{1}{12} R^3 \delta \,. \quad (42)$$

The variance of the centroid is then

$$\sigma^2(\langle 2\varphi \rangle) = \frac{1}{L_0^2}\left\{\sum_j (2\varphi_j)^2 \sigma^2(I_j) + \frac{R^4\delta^2}{144}[\sigma^2(G_R)+\sigma^2(G_L)]\right\}. \quad (43)$$

The first term is closely related to the variance of the line, being the variance of (line plus background) divided by τ when fixed-time counting is used. In this case

$$\sigma^2_{\mathrm{ft}}(\langle 2\varphi \rangle) = \frac{1}{L_0^2 \tau}\left\{\sum_j (2\varphi_j)^2 (I_j - G_j) + \sum_j (2\varphi_j)^2 G_j\right\} + \frac{R^4\delta^2}{144 L_0^2}[\sigma^2_{\mathrm{ft}}(G_R)+\sigma^2_{\mathrm{ft}}(G_L)]$$

$$= \frac{W}{L_0\tau} + \frac{G_R+G_L}{2L_0^2\tau}\delta^2 \sum_j j^2 + \frac{R^4\delta^2}{144L_0^2}[\sigma^2_{\mathrm{ft}}(G_R)+\sigma^2_{\mathrm{ft}}(G_L)]$$

$$= \frac{W}{L_0\tau} + \frac{(G_R+G_L)R^3\delta^2}{24L_0^2\tau} + \frac{R^4\delta^2}{144L_0^2}[\sigma^2_{\mathrm{ft}}(G_R)+\sigma^2_{\mathrm{ft}}(G_L)]\,, \quad (44)$$

where W is the variance of the line alone. This differs from equation (20) of Pike & Wilson (1959) in notation, especially in the use here of the variance of the line instead of the variance of (line plus background), but not otherwise.

4.3.2. Since W is ultimately proportional to the range of observation used, the variance of the centroid is an increasing function of R. For the accurate determination of lattice parameters, therefore, it is advantageous to use as short a range as is consistent with including all essential features of the line (such as $K\alpha_3$), to reduce the mean background as far as possible (proportional or scintillation counters with pulse-height discrimination instead of Geiger counters; crystal-reflected instead of filtered radiation), and to use long counting times in determining the background so that the variance of the background is small.

4.3.3. The corresponding fixed-count variance is, from equation (43),

$$\sigma^2_{\mathrm{fc}}(\langle 2\varphi \rangle) = \frac{1}{L_0^2}\left\{\frac{1}{c}\sum_j (2\varphi_j)^2 I_j^2 + \frac{R^4\delta^2}{144}[\sigma^2_{\mathrm{fc}}(G_R)+\sigma^2_{\mathrm{fc}}(G_L)]\right\}. \quad (45)$$

The sum in the first term of this is the variance of the profile (line plus background)^2 multiplied by the integrated intensity of (line plus background)^2, and is not easy to represent in terms of any simpler concepts. Representing these factors by W_2 and L_2 gives

$$\sigma^2_{\mathrm{fc}}(\langle 2\varphi \rangle) = \frac{L_2 W_2}{L_0^2 c} + \frac{R^4\delta^2}{144L_0^2}[\sigma^2_{\mathrm{fc}}(G_R)+\sigma^2_{\mathrm{fc}}(G_L)]\,. \quad (46)$$

The minimum variance, from equations (19) and (38), is

$$\sigma^2_{\min}(\langle 2\varphi \rangle) = \frac{1}{TL_0^2}\left\{\sum_j |2\varphi_j| I_j^{\frac{1}{2}}\right\}^2 + \frac{R^4\delta^2}{144L_0^2}[\sigma^2_{\min}(G_R)+\sigma^2_{\min}(G_L)]\,. \quad (47)$$

The first term is related to the square of the mean absolute deviation from the centroid, and is perhaps easier to evaluate numerically than to visualize.

5. Measures of dispersion

5.1. *The half-width*

The variance of the half-width (full width at half height) may be estimated by a variation of the method used by Wilson (1965) for the variance of the mid-chord. The experimenter's eye may be assumed to average out statistical fluctuations over a few points, say s, in the neighbourhood of the half height, in a manner approximating to fitting a least-squares straight line to the points, and the abscissa x_l of the left-hand end of the chord at half height will be given by

$$a_l + b_l x_l = \tfrac{1}{2}(I - G_l) + G_l\,, \quad (48)$$

where a and b are the coefficients of the least-squares line, I is the peak height and G is the background. The abscissa of the right-hand end will be given by a similar equation with subscripts r instead of l, so that the half width w is given by

$$w = x_r - x_l$$

$$= \tfrac{1}{2}I(b_r^{-1} - b_l^{-1}) + \tfrac{1}{2}(G_r b_r^{-1} - G_l b_l^{-1}) - (a_r b_r^{-1} - a_l b_l^{-1})\,. \quad (49)$$

Of the quantities appearing on the right, only I is statistically independent of all the others. The variance of w arising from I is thus

$$\sigma_I^2(w) = \tfrac{1}{4}(b_r^{-1} - b_l^{-1})^2 \sigma^2(I)\,, \quad (50)$$

where $\sigma^2(I)$ is as discussed in § 3.1. In normal cases b_r and b_l are approximately equal and opposite, so that

$$\sigma_I^2(w) \sim \sigma^2(I)/b^2\,. \quad (51)$$

Since G_r and G_l both depend on G_L and G_R (Appendix A) the latter are the independent variables, and the variance arising from the background is

$$\sigma_G^2(w) = \left(\frac{1}{2b_r}\frac{\partial G_r}{\partial G_L} - \frac{1}{2b_l}\frac{\partial G_l}{\partial G_L}\right)^2 \sigma^2(G_L)$$

+ similar term in G_R

$$= \tfrac{1}{4}\left\{\left[\frac{R-2x_r}{2Rb_r} - \frac{R-2x_l}{2Rb_l}\right]^2 \sigma^2(G_L) + \left[\frac{R+2x_r}{2Rb_r} - \frac{R+2x_l}{2Rb_l}\right]^2 \sigma^2(G_R)\right\} \quad (52)$$

$$= \left\{\left[\tfrac{1}{4}\left(\frac{1}{b_r} - \frac{1}{b_l}\right)^2 + \frac{1}{R^2}\left(\frac{x_r}{b_r} - \frac{x_l}{b_l}\right)^2\right]\sigma^2(g) - \frac{1}{4R}\left(\frac{1}{b_r} - \frac{1}{b_l}\right)\left(\frac{x_r}{b_r} - \frac{x_l}{b_l}\right)[\sigma^2(G_L) - \sigma^2(G_R)]\right\}. \quad (53)$$

If b_r, b_l; x_r, x_l; are approximately equal and opposite this becomes

$$\sigma_G^2(w) \sim \sigma^2(g)/b^2 . \quad (54)$$

The variance of x_l arising from a_l and b_l was evaluated by Wilson (1965) as

$$\sigma_{a,b}^2(x_l) \sim \frac{s^4\delta^5 M_0}{144 M_1^2} \quad (55)$$

in the present notation, where M_0 and M_1 are the zero and first moment, about the central point, of the portion of the line profile used in determining equation (48). The appropriate value of M_0 is $s\delta a_l$, and of M_1 is $s^3\delta^3 b_l/12$, so that equation (55) becomes

$$\sigma_{a,b}^2(x_l) \sim a_l/sb_l^2 , \quad (56)$$

and the total variance of w arising from the a's and b's is

$$\sigma_{a,b}^2(w) \sim (a_l/b_l^2 + a_r/b_r^2)/s . \quad (57)$$

The full variance of w is the sum of equations (51), (54) and (57), or in approximate form

$$\sigma_{\text{ft}}^2(w) \sim \frac{1}{b^2\tau}\left[\frac{I}{q} + \frac{g}{p} + \frac{2a}{s}\right] \quad (58)$$

for fixed-time counting, where $q\tau$ is the counting time used in fixing the peak height and $p\tau$ that used in fixing the background. Fixed-count timing would give expressions of different form, but as the range of intensities within each of the separate determinations (I,g,a) is not great the final result would not be greatly different if c and τ were adjusted so that the times required were approximately equal.

5.2. *The integral breadth*

The integral breadth is given by the total intensity divided by the nett peak intensity,

$$\beta = L_0/(I-g) , \quad (59)$$

in units of δ. A complication arises in that the I_j used in determining I are the same as those that contribute most to L_0, and hence the variance is less than if I and L_0 were statistically independent. The effective statistical equation is

$$\beta = \frac{\sum_{j=-r}^{r} I_j - Rg}{\frac{1}{q}\sum_{j=-\frac{1}{2}(q-1)}^{\frac{1}{2}(q-1)} I_j - g} , \quad (60)$$

so that

$$\sigma^2(\beta) = \sum_{j=-r}^{-\frac{1}{2}(q+1)} \frac{\sigma^2(I_j)}{(I-g)^2} + \sum_{j=\frac{1}{2}(q+1)}^{r} \frac{\sigma^2(I_j)}{(I-g)^2} + \sum_{j=-\frac{1}{2}(q-1)}^{\frac{1}{2}(q-1)} \left[\frac{1}{I-g} - \frac{L_0}{(I-g)^2 q}\right]^2 \sigma^2(I_j) + \left[\frac{L_0}{(I-g)^2} - \frac{R}{I-g}\right]^2 \sigma^2(g) \quad (61)$$

$$= \frac{1}{(I-g)^2} \sum_{j=-r}^{r} \sigma^2(I_j) - \frac{2L_0}{q(I-g)^3} \sum_{j=-\frac{1}{2}(q-1)}^{\frac{1}{2}(q-1)} \sigma^2(I_j) + \frac{L_0^2}{q^2(I-g)^4} \sum_{j=-\frac{1}{2}(q-1)}^{\frac{1}{2}(q-1)} \sigma^2(I_j) + \frac{(R-\beta)^2}{(I-g)^2}\sigma^2(g) . \quad (62)$$

This rather complex expression may be simplified in appropriate ways. For fixed-time counting it becomes

$$\sigma_{\text{ft}}^2(\beta) = \frac{L_0 + Rg}{(I-g)^2\tau} - \frac{2L_0 I}{(I-g)^3\tau} + \frac{L_0^2 I}{q(I-g)^4\tau} + \frac{g(R-\beta)^2}{p\tau(I-g)^2} , \quad (63)$$

$$\sim \frac{L_0^2}{qI^3\tau}\left(1 - \frac{qI}{L_0}\right) + \frac{g[Rp + (R-\beta)^2]}{pI^2\tau} \quad (64)$$

if the background is small enough. For fixed-count timing equation (62) becomes

$$\sigma_{\text{fc}}^2(\beta) = \frac{K_0}{(I-g)^2 c} - \frac{2L_0 I^2}{(I-g)^3 c} + \frac{L_0^2 I^2}{q(I-g)^4 c} + \frac{g^2(R-\beta)^2}{p(I-g)^2 c} , \quad (65)$$

where K_0 is the total of (line + background)^2. For small g it reduces to

$$\sigma_{\text{fc}}^2(\beta) \sim \frac{L_0^2}{qI^2 c}\left(1 - \frac{2qI}{L_0} + \frac{qK_0}{L_0^2}\right) + \frac{g^2(R-\beta)^2}{pI^2 c} . \quad (66)$$

5.4. *Mean absolute deviation from the median*

The mean absolute deviation from the median is

$$M \equiv \langle|2\varphi - 2\varphi_m|\rangle = \frac{\sum_j |2\varphi_j - 2\varphi_m|(I_j - G_j)}{\sum_j (I_j - G_j)} , \quad (67)$$

so that

$$\frac{\partial M}{\partial I_j} = \frac{|2\varphi_j - 2\varphi_m|}{L_0} - \frac{M}{L_0} , \quad (68)$$

$$\frac{\partial M}{\partial G_R} = \frac{1}{L_0} \sum_j (M - |2\varphi_j - 2\varphi_m)\left[\tfrac{1}{2} + \frac{j}{R}\right] , \quad (69)$$

$$\frac{\partial M}{\partial G_L} = \frac{1}{L_0} \sum_j (M - |2\varphi_j - 2\varphi_m|)\left[\tfrac{1}{2} - \frac{j}{R}\right] . \quad (70)$$

If the median is centred the term $2\varphi_m$ will be less than half a step length, and for any reasonable line the median will not be far from the centre of the range even if it is the centroid or the peak that is centred. Neglecting $2\varphi_m$ in comparison with M is therefore legitimate in estimating the variance of M. Equations (68) to (70) become

$$\frac{\partial M}{\partial I_j} = (|2\varphi_j| - M)/L_0 \tag{71}$$

and

$$\frac{\partial M}{\partial G_R} = \frac{\partial M}{\partial G_L} = \frac{RM}{2L_0} - \frac{\delta}{2L_0}\sum_j |j| \tag{72}$$

$$= \frac{RM}{2L_0} - \frac{(R-1)(R+1)}{8L_0} \tag{73}$$

$$\sim \frac{RM}{2L_0} - \frac{R^2\delta}{8L_0}\,. \tag{74}$$

These bear a close resemblance to equations (38) to (40), with M replacing $\langle 2\varphi \rangle$. The contribution of the intensities to the variance of M is

$$\sigma_I^2(M) = \frac{1}{L_0^2}\sum_j (|2\varphi_j| - M)^2 \sigma^2(I_j) \tag{75}$$

$$= \frac{1}{L_0^2}\sum_j [(2\varphi_j)^2 - 2M|2\varphi_j| + M^2]\sigma^2(I_j)\,. \tag{76}$$

For fixed-time counting this is

$$\sigma_{I,ft}^2(M) = \frac{1}{\tau L_0^2}\sum_j [(2\varphi_j)^2 - 2M|2\varphi_j| + M^2] I_j \tag{77}$$

$$= \frac{1}{\tau L_0}\left\{W + \frac{g}{L_0}\sum_j (2\varphi_j)^2 - 2M^2 - \frac{2Mg}{L_0}\sum_j |2\varphi_j| + M^2 + \frac{RgM^2}{L_0}\right\}$$

$$\sim \frac{1}{\tau L_0}\left\{W - M^2 + \frac{R^3 g\delta^2}{12L_0} - \frac{MgR^2\delta}{2L_0} + \frac{RgM^2}{L_0}\right\}, \tag{78}$$

where W is the variance of the line profile. In addition there is the variance arising from the background:

$$\sigma_G^2(M) = \frac{R^2}{4L_0^2}(M - \tfrac{1}{4}R\delta)^2[\sigma^2(G_R) + \sigma^2(G_L)]\,. \tag{79}$$

Fixed-count timing and minimum-variance techniques give expressions based on the square and the square root of the line profile respectively.

5.5. *The variance*

The variance is defined by

$$W = \frac{\sum_j (I_j - G_j) j^2}{\sum_j (I_j - G_j)} - \left[\frac{\sum_j (I_j - G_j) j}{\sum_j (I_j - G_j)}\right]^2, \tag{80}$$

in units of (step length)². The second term is just the square of the centroid position, $\langle 2\varphi \rangle$, but it has been written in full to emphasize its dependence on the counting rates I_j. With the range symmetrical, to the nearest half step, about the centroid, the second term cannot exceed one-quarter. Differentiation gives

$$\frac{\partial W}{\partial I_j} = \frac{j^2}{L_0} - \frac{W}{L_0} - \frac{\langle 2\varphi \rangle^2}{L_0} - \frac{2\langle 2\varphi \rangle j}{L_0} + \frac{2\langle 2\varphi \rangle^2}{L_0}\,. \tag{81}$$

The terms involving the centroid are clearly small compared with the others, and may be ignored. The variance thus becomes

$$\sigma_{ft}^2(W) = \frac{R}{TL_0^2}\sum_j (j^4 - 2Wj^2 + W^2) I_j$$

$$= \frac{R}{TL_0}[M_4 - W^2] + \frac{RI_g}{TL_0^2}[M_{4g} - 2WW_g + W^2]\,, \tag{82}$$

where M_4 is the fourth moment of the line profile, M_{4g} is the fourth moment of the background, W_g is the variance of the background, and I_g is the total background intensity.

The variances for fixed-count timing and minimum variance are easily written down as

$$\sigma_{fc}^2(W) = \frac{R\langle I_j^{-1}\rangle}{TL_0^2}\sum_j (j^4 - 2Wj^2 + W^2) I_j^2 \tag{83}$$

and

$$\sigma_{\min}^2(W) = \frac{1}{TL_0^2}\left\{\sum_j |j^2 - W| I_j^{\frac{1}{2}}\right\}^2, \tag{84}$$

where $\langle I_j^{-1}\rangle = c/T$ is the average reciprocal intensity. As usual, these are more easily calculated than visualized.

The variance is hardly dependent on the slope of the interpolated background, k, which vanishes from the main term of equation (80) because of the symmetry of the range. Its main dependence on the mean background g is given by

$$\frac{\partial W}{\partial g} = -\frac{j^2}{L_0} + \frac{W}{L_0}$$

so that

$$\sigma_g^2(W) = \frac{1}{L_0^2}\sum_j (W^2 - 2Wj^2 + j^4)\sigma^2(g)$$

$$\sim \frac{R\sigma^2(g)}{L_0^2}\left[W^2 - \frac{WR^2}{6} + \frac{R^4}{80}\right], \tag{85}$$

where the variance of g is given by equations (134) to (136).

5.6. *The Fourier coefficients*

5.6.1. The normalized cosine Fourier coefficient A_m is evaluated as

$$A_m = \frac{\sum_j (I_j - G_j)\cos(2\pi mj/R)}{\sum_j (I_j - G_j)}, \tag{86}$$

and the corresponding sine component as

$$B_m = \frac{\sum_j (I_j - G_j)\sin(2\pi mj/R)}{\sum_j (I_j - G_j)}\,. \tag{87}$$

If the line profile is symmetrical the B_m's all vanish, and if the centroid is used as the working origin they will in any case be of smaller importance than the A_m's. Clearly

$$A_0=1,\ \sigma^2(A_0)=0\ ,\quad B_0=0,\ \sigma^2(B_0)=0\ . \tag{88}$$

For $m\neq 0$, the symmetry of the range and the assumed linearity of the background makes it possible to simplify equations (86) and (87) to

$$A_m = \frac{\sum_j I_j \cos(2\pi mj/R)}{\sum_j (I_j-g)} \tag{89}$$

and

$$B_m = \frac{\sum_j [I_j-(G_R-G_L)j/R]\sin(2\pi mj/R)}{\sum_j (I_j-g)}\ . \tag{90}$$

Differentiation gives

$$\frac{\partial A_m}{\partial I_j} = \frac{\cos(2\pi mj/R)}{L_0} - \frac{A_m}{L_0}\ , \tag{91}$$

$$\frac{\partial B_m}{\partial I_j} = \frac{\sin(2\pi mj/R)}{L_0} - \frac{B_m}{L_0}\ . \tag{92}$$

5.6.2. For fixed-time counting the variance of the cosine coefficients is therefore

$$\sigma^2_{\text{ft}}(A_m) = \frac{R}{TL_0^2}\sum_j [\cos(2\pi mj/R)-A_m]I_j \tag{93}$$

$$= \frac{R}{TL_0^2}\sum_j [\tfrac{1}{2}+\tfrac{1}{2}\cos(4\pi mj/R)-2A_m\cos(2\pi mj/R)+A_m^2]I_j$$

$$= \frac{R(L_0+I_g)}{2L_0^2T}+\frac{RA_{2m}}{2L_0T}+\frac{R(I_g-L_0)A_m^2}{L_0^2T}\ . \tag{94}$$

This takes on two limiting forms. For m small, $A_m\sim 1$, so that

$$\sigma^2_{\text{ft}}(A_m) \sim \frac{3RI_g}{2TL_0^2}\ , \tag{95}$$

and for m large, $A_m\sim 0$, so that

$$\sigma^2_{\text{ft}}(A_m) \sim \frac{R(L_0+I_g)}{2TL_0^2}\ . \tag{96}$$

One may also remark that equation (94) sets a lower limit for A_{2m}. Since the variance of A_m must be positive,

$$A_{2m} > (2A_m^2+1)(1-I_g/L_0)-2\ . \tag{97}$$

The importance of the background in determining the variance for small m should be noted. For the effect of variance of the background see paragraph 5.6.4.

The variance of the sine coefficients may be written down similarly:

$$\sigma^2_{\text{ft}}(B_m) = \frac{R}{TL_0^2}\sum_j [\sin(2\pi mj/R)-B_m]^2I_j \tag{98}$$

$$= \frac{R}{TL_0^2}\sum_j [\tfrac{1}{2}-\tfrac{1}{2}\cos(4\pi mj/R)-2B_m\sin(2\pi mj/R)+B_m^2]I_j$$

$$= \frac{R(L_0+I_g)}{2TL_0^2}-\frac{RA_{2m}}{2TL_0}+\frac{R(I_g-L_0)B_m^2}{TL_0^2} + \frac{(-)^m R^2(G_R-G_L)B_m}{\pi m TL_0^2}\ , \tag{99}$$

where the sum of $j\sin(2\pi mj/R)$ has been approximated by an integral. The peculiar contribution from the background slope should be noted. If it can be assumed that in practice the B_m's are all small, equation (99) reduces to

$$\sigma^2_{\text{ft}}(B_m) = \frac{R(L_0+I_g)}{2TL_0^2}-\frac{RA_{2m}}{2TL_0}\ . \tag{100}$$

For the effect of the variance of the background see paragraph 5.6.5.

5.6.3. The fixed-count and minimum-variance equivalents of equations (93) and (98) are

$$\sigma^2_{\text{fc}}(A_m) = \frac{R\langle I_j^{-1}\rangle}{TL_0^2}\sum_j [\cos(2\pi mj/R)-A_m]^2I_j^2\ , \tag{101}$$

$$\sigma^2_{\text{fc}}(B_m) = \frac{R\langle I_j^{-1}\rangle}{TL_0^2}\sum_j [\sin(2\pi mj/R)-B_m]^2I_j^2\ , \tag{102}$$

$$\sigma^2_{\min}(A_m) = \frac{1}{TL_0^2}\left\{\sum_j |\cos(2\pi mj/R)-A_m|I_j^{\frac{1}{2}}\right\}^2\ , \tag{103}$$

$$\sigma^2_{\min}(B_m) = \frac{1}{TL_0^2}\left\{\sum_j |\sin(2\pi mj/R)-B_m|I_j^{\frac{1}{2}}\right\}^2\ . \tag{104}$$

The minimum-variance expressions do not appear to be reducible to any more easily visualized form, but could easily be evaluated numerically by computer. The fixed-count expressions, equations (101) and (102) [and indeed (65) and (84)], could be expressed in terms of the A_m's and B_m's by writing

$$I_j^2 = \sum_{m,m'} [A_m\cos(2\pi mj/R)+B_m\sin(2\pi mj/R)] \times [A_{m'}\cos(2\pi m'j/R)+B_m\sin(2\pi m'j/R)] \tag{105}$$

and expressing the products of sines and cosines of m and m' as sines and cosines of the sums and differences $m+m'$ and $m-m'$. Since the variance of each parameter would involve all the Fourier coefficients, the usefulness of the result is somewhat doubtful, and it is probably better to rely on numerical evaluations of the equations. For the effect of background variance see paragraph 5.6.5.

5.6.4. The Fourier coefficients are not independent of one another, and for calculating the variance of, say, the particle-size distribution function from them by means of equation (1) it is necessary to know their covariances as well as their variances. The covariance of A_m with A_n is easily written down from the general expression for the covariance of any two functions of independent random variables x_j:

$$\text{cov}(F_1,F_2) = \sum_j \frac{\partial F_1}{\partial x_j}\frac{\partial F_2}{\partial x_j}\sigma^2(x_j)\ , \tag{106}$$

giving

$$\text{cov}(A_m,A_m) = \sum_j \frac{\partial A_m}{\partial I_j}\frac{\partial A_n}{\partial I_j}\sigma^2(I_j) \tag{107}$$

$$= \frac{1}{L_0^2}\sum_j [\cos(2\pi mj/R)-A_m] \times [\cos(2\pi nj/R)-A_n]\sigma^2(I_j) \tag{108}$$

$$= \frac{1}{L_0^2} \sum_j [\tfrac{1}{2} \cos\{2\pi(m+n)j/R\} + \tfrac{1}{2} \cos\{2\pi(m-n)j/R\} - A_m \cos(2\pi nj/R) - A_n \cos(2\pi mj/R) + A_n A_m] \sigma^2(I_j) . \quad (109)$$

This takes a simple form only for fixed-time counting, for which $\sigma^2(I_j) = I_j/\tau = RI_j/T$. In this case each term becomes a Fourier coefficient:

$$\mathrm{cov}_{\mathrm{ft}}(A_m, A_n) = \frac{R}{L_0 T} \left\{ \tfrac{1}{2} A_{m+n} + \tfrac{1}{2} A_{m-n} - A_n A_m + \frac{RgA_nA_m}{L_0} \right\} \quad (110)$$

where $m \neq n$. The covariance of the sine coefficients is similarly

$$\mathrm{cov}(B_m, B_n) = \frac{1}{L_0^2} \sum_j [\sin(2\pi mj/R) - B_m] \times [\sin(2\pi nj/R) - B_n] \sigma^2(I_j) \quad (111)$$

$$= \frac{1}{L_0^2} \sum_j [\tfrac{1}{2} \cos\{2\pi(m-n)j/R\} - \tfrac{1}{2} \cos\{2\pi(m+n)j/R\} - B_m \sin(2\pi nj/R) - B_n \sin(2\pi mj/R) + B_n B_m] \sigma^2(I_j) , \quad (112)$$

so that for fixed-time counting

$$\mathrm{cov}_{\mathrm{ft}}(B_m, B_n) = \frac{R}{L_0 T} \left\{ \tfrac{1}{2} A_{m-n} - \tfrac{1}{2} A_{m+n} - B_n B_m + \frac{RgB_mB_n}{L_0} - B_m \frac{G_R - G_L}{RL_0} \sum_j j \sin(2\pi nj/R) - B_n \frac{G_R - G_L}{RL_0} \sum_j j \sin(2\pi mj/R) \right\}$$

$$= \frac{R}{L_0 T} \left\{ \tfrac{1}{2} A_{m-n} - \tfrac{1}{2} A_{m+n} - B_m B_n + \frac{RgB_mB_n}{L_0} + \frac{(G_R - G_L)R}{2\pi L_0} \left[\frac{(-)^n B_m}{n} + \frac{(-)^m B_n}{m} \right] \right\}, \quad (113)$$

where again a slight approximation has been made in evaluating the sums multiplying the background slope. If it can be assumed that the B's are small in comparison with A's this reduces to

$$\mathrm{cov}_{\mathrm{ft}}(B_m, B_n) = \frac{R}{2L_0 T} \{A_{m-n} - A_{m+n}\} . \quad (114)$$

The covariance of a sine with a cosine coefficient is

$$\mathrm{cov}(A_m, B_n) = \frac{1}{L_0^2} \sum_j [\cos(2\pi mj/R) - A_m] \times [\sin(2\pi nj/R) - B_n] \sigma^2(I_j) \quad (115)$$

$$= \frac{1}{L_0^2} \sum_j [\tfrac{1}{2} \sin\{2\pi(m+n)j/R\} - \tfrac{1}{2} \sin\{2\pi(m-n)j/R\} - A_m \sin(2\pi nj/R) - B_n \cos(2\pi mj/R) + A_m B_n] \sigma^2(I_j) , \quad (116)$$

so that

$$\mathrm{cov}_{\mathrm{ft}}(A_m, B_n) = \frac{R}{L_0 T} \left\{ \tfrac{1}{2} B_{m+n} - \tfrac{1}{2} B_{m-n} - A_m B_n + \frac{(G_R - G_L)}{2\pi L_0} \cdot \frac{(-)^{m+n} n}{m^2 - n^2} - \frac{(G_R - G_L)R}{2\pi L_0} \cdot \frac{(-)^n A_m}{n} + \frac{RgA_mB_n}{L_0} \right\} \quad (117)$$

for fixed-time counting. This appears to be smaller than the other covariances, since the only terms that do not involve a B involve the background slope.

It should perhaps be stated explicitly that the above expressions are not valid if either m or n is zero. Covariances involving A_0 or B_0 are of course zero.

5.6.5. The derivatives of the Fourier coefficients with respect to the background observations are, from equations (89) and (90),

$$\frac{\partial A_m}{\partial G_R} = \frac{RA_m}{2L_0} , \quad (118)$$

$$\frac{\partial A_m}{\partial G_L} = \frac{RA_m}{2L_0} , \quad (119)$$

$$\frac{\partial B_m}{\partial G_R} = \frac{RB_m}{2L_0} - \frac{1}{RL_0} \sum_j j \sin(2\pi mj/R) \quad (120)$$

$$\sim \frac{RB_m}{2L_0} + \frac{(-)^m R}{2\pi L_0 m} , \quad (121)$$

$$\frac{\partial B_m}{\partial G_L} = \frac{RB_m}{2L_0} + \frac{1}{RL_0} \sum_j j \sin(2\pi mj/R) \quad (122)$$

$$\sim \frac{RB_m}{2L_0} - \frac{(-)^m R}{2\pi L_0 m} . \quad (123)$$

The contributions to the variance arising from the background are thus

$$\sigma^2(A_m) = \frac{R^2 A_m^2}{4L_0^2} [\sigma^2(G_R) + \sigma^2(G_L)] \quad (124)$$

and

$$\sigma^2(B_m) = \frac{R^2}{4L_0^2} \left[B_m^2 + \frac{1}{\pi^2 m^2} \right] [\sigma^2(G_R) + \sigma^2(G_L)] + \frac{(-)^m R^2 B_m}{2\pi L_0^2 m} [\sigma^2(G_R) - \sigma^2(G_L)] , \quad (125)$$

where the variances of G_R and G_L are as given in Appendix A. The covariances are similarly

$$\mathrm{cov}(A_m, A_n) = \frac{R^2 A_m A_n}{4L_0^2} [\sigma^2(G_R) + \sigma^2(G_L)] , \quad (126)$$

$$\mathrm{cov}(B_m, B_n) = \frac{R^2}{4L_0^2} \left[B_m B_n + \frac{(-)^{m+n}}{\pi^2 mn} \right] [\sigma^2(G_R) + \sigma^2(G_L)] + \frac{R^2}{4\pi L_0^2} \left[\frac{(-)^m B_n}{m} + \frac{(-)^n B_m}{n} \right] [\sigma^2(G_R) - \sigma^2(G_L)] , \quad (127)$$

$$\mathrm{cov}(A_m, B_n) = \frac{R^2 A_m}{4L_0^2} \left\{ B_n[\sigma^2(G_R)+\sigma^2(G_L)] + \frac{(-)^n}{\pi n} [\sigma^2(G_R)-\sigma^2(G_L)] \right\}. \quad (128)$$

The full variances and covariances are thus the sums of the corresponding pairs of equations from the groups (94)–(117) and (124)–(128).

This work was begun at the Georgia Institute of Technology, while the author was on leave from University College, Cardiff. I am indebted to Professor R. A. Young for many helpful discussions, and to Dr J. I. Langford for criticism of a draft of the paper.

APPENDIX A
Statistical variance of the interpolated background

If the background is determined from $\frac{1}{2}p$ observations at each end of the line, then effectively

$$g = \tfrac{1}{2}(G_L+G_R), \quad (129)$$

where G_L is the background on the left of the line and G_R that on the right. The calculation of the variances of G_L and G_R is trivial, giving

$$\sigma^2_{\mathrm{ft}}(G_L) = 2G_L/p\tau \quad (130)$$

for fixed-time counting, and

$$\sigma^2_{\mathrm{fc}}(G_L) = 2G_L^2/pc \quad (131)$$

for fixed-count timing. The variance of the mean background is then

$$\sigma^2(g) = \tfrac{1}{4}\sigma^2(G_L) + \tfrac{1}{4}\sigma^2(G_R) \quad (132)$$

$$= \frac{1}{2p\tau}(G_L+G_R) \quad (133)$$

so that

$$\sigma^2_{\mathrm{ft}}(g) = g/p\tau \quad (134)$$

for fixed-time counting, and

$$\sigma^2_{\mathrm{fc}}(g) = \frac{1}{2pc}(G_L^2+G_R^2) \quad (135)$$

for fixed-count timing. Unless the slope is extreme this will not differ essentially from

$$\sigma^2_{\mathrm{fc}}(g) = g^2/pc\,. \quad (136)$$

Similarly, the background slope is

$$k = (G_R-G_L)/R\,, \quad (137)$$

so that

$$\sigma^2(k) = \{\sigma^2(G_R)+\sigma^2(G_L)\}/R^2\,, \quad (138)$$

$$\sigma^2_{\mathrm{ft}}(k) = 4g/pR^2\tau \quad (139)$$

for fixed-time counting, and

$$\sigma^2_{\mathrm{fc}}(k) = 2\{G_L^2+G_R^2\}/pcR^2 \quad (140)$$

$$\sim 4g^2/pcR^2 \quad (141)$$

for fixed-count timing. The background variance at the jth step is to an appreciable extent a function of j. Since

$$G_j = g+kj \quad (142)$$

$$= \tfrac{1}{2}\{G_R+G_L\} + \frac{j}{R}\{G_R-G_L\}\,, \quad (143)$$

$$\sigma^2(G_j) = \left(\tfrac{1}{2}+\frac{j}{R}\right)^2\sigma^2(G_R) + \left(\tfrac{1}{2}-\frac{j}{R}\right)^2\sigma^2(G_L)$$

$$= \tfrac{1}{4}\{\sigma^2(G_R)+\sigma^2(G_L)\} + \frac{j}{R}\{\sigma^2(G_R)-\sigma^2(G_L)\} + \frac{j^2}{R^2}\{\sigma^2(G_R)+\sigma^2(G_L)\}\,. \quad (144)$$

For fixed-time counting this becomes

$$\sigma^2_{\mathrm{ft}}(G_j) = g/p\tau + 2kj/p\tau + 4gj^2/pR^2\tau\,, \quad (145)$$

and for fixed-count timing

$$\sigma^2_{\mathrm{fc}}(G_j) = \frac{1}{2pc}\{G_R^2+G_L^2\} + \frac{2j}{pcR}\{G_R^2-G_L^2\} + \frac{2j^2}{pcR^2}\{G_R^2+G_L^2\} \quad (146)$$

$$= (g^2+\tfrac{1}{4}R^2k^2)/pc + 4gkj/pc + 4(g^2+\tfrac{1}{4}R^2k^2)j^2/pcR^2 \quad (147)$$

$$\sim g^2/pc + 4gkj/pc + 4g^2j^2/pcR^2\,. \quad (148)$$

In either case the variance has a minimum (at $j = -R^2k/4g$ and $j \sim -R^2k/2g$ respectively), and is roughly twice as great near the ends of the range as it is near the middle.

The covariance of the mean background and the background slope is

$$\mathrm{cov}(g,k) = \frac{\partial g}{\partial G_L}\frac{\partial k}{\partial G_L}\sigma^2(G_L) + 0 + 0 + \frac{\partial g}{\partial G_R}\frac{\partial k}{\partial G_R}\sigma^2(G_R) \quad (149)$$

$$= \frac{1}{2R}\{\sigma^2(G_R)-\sigma^2(G_L)\} \quad (150)$$

and is thus small unless the background slope is large.

APPENDIX B
Effect of background slope on peak positions

The expressions derived by Wilson (1965) for the variance of the peak position apply to the observed line profile without correction for background. If a sloping background $g+kj$ is subtracted, the equation for the least-squares parabola becomes

$$I_j=(A-g)+(B\delta-k)j+C\delta^2j^2\,, \qquad (151)$$

with its peak displaced from $-B/2\delta C$ to

$$x_m=-(B-k/\delta)/2\delta C\,. \qquad (152)$$

The variance of its position is thus increased by

$$\sigma^2_{\rm ft}(k/2\delta^2C)\sim g/2pR^2\delta^4C^2\tau \qquad (153)$$

for fixed-time counting, and

$$\sigma^2_{\rm fc}(k/2\delta^2C)\sim g^2/2pcR^2\delta^4C^2 \qquad (154)$$

for fixed-count timing.

References

CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. Princeton: Princeton Univ. Press.
EASTABROOK, J. N. & HUGHES, J. W. (1952). *J. Sci. Instrum.* **30**, 317.
MACK, M. & SPIELBERG, N. (1958). *Spectrochim. Acta*, **12**, 169.
PIKE, E. R. & WILSON, A. J. C. (1959). *Brit. J. Appl. Phys.* **10**, 57.
SANDS, D. E. (1966). *Acta Cryst.* **21**, 868.
WILSON, A. J. C. (1965). *Brit. J. Appl. Phys.* **16**, 665.
WILSON, A. J. C., THOMSEN, J. S. & YAP, F. Y. (1965). *Appl. Phys. Letters*, **7**, 163.
ZEVIN, L. S., UMANSKIJ, M. M., KHEJKER, D. M. & & PANČENKO, JU. M. (1961). *Kristallografija*, **6**, 348.

Acta Cryst. (1967). **23**, 898

A New Method for Obtaining Phase Angles of Structure Factors of Non-Centrosymmetric Structures by the Use of Anomalous Dispersion

BY DIETRICH UNANGST, EBERHARD MÜLLER, JÜRGEN MÜLLER AND BERND KEINERT
Physikalisches Institut der Friedrich-Schiller-Universität, 69 *Jena*, *Max-Wien-Platz* 1, *Germany* (*DDR*)

(*Received* 23 *February* 1967)

A method is described for finding the phases of structure factors by using the ratio $|F|^2/|\bar{F}|^2$ instead of the Bijvoet difference $|F|^2-|\bar{F}|^2$. The influence of errors of measurement is pointed out and a comparison with the method of Bijvoet differences is made.

Introduction and theory

It is possible to obtain the phases of structure factors by making use of the effects of anomalous dispersion. The method is based on experimental measurements of the Bijvoet differences $|F|^2-|\bar{F}|^2$ between pairs of inverse reflexions hkl and $\bar{h}\bar{k}\bar{l}$ produced by the imaginary component $\Delta f''$ of the scattering factor (Ramachandran, 1964). In order to avoid difficulties and errors of measurement in the determination of the scale factor and the absorption-correction a new method is proposed, which uses the ratios $|F|^2/|\bar{F}|^2$ instead of the differences $|F|^2-|\bar{F}|^2$. If the irradiated volume of the crystal has an inversion centre, there is an immediate connexion between the direct measurable intensities S, $\bar{S}$, and the ratio $|F|^2/|\bar{F}|^2$. This connexion is given by $|F|^2/|\bar{F}|^2=S/\bar{S}$. Supposing the crystal contains only one kind of atom, for which the imaginary component of the scattering factor is significant, whereas for all the other atoms this component is negligible, we get by simple geometrical considerations (Fig. 1) the following equation (Unangst, 1965):

$$\frac{S(\lambda_m)}{\bar{S}(\lambda_m)}=\frac{|F_N(\lambda_m)|^2}{|\bar{F}_N(\lambda_m)|^2}=1-\frac{4\dfrac{|F_P^0|}{|F_N^0|f_P^0}\Delta f''(\lambda_m)\sin\alpha}{1+2\dfrac{|F_P^0|}{|F_N^0|f_P^0}[\Delta f'(\lambda_m)\cos\alpha+\Delta f''(\lambda_m)\sin\alpha]+\left[\dfrac{|F_P^0|}{|F_N^0|f_P^0}\right]^2[\Delta f'^2(\lambda_m)+\Delta f''^2(\lambda_m)]}\,. \qquad (1)$$

Taking the following abbreviations

$$a_m=\frac{\Delta f''(\lambda_m)}{f_P^0}\qquad b_m=\frac{\Delta f'(\lambda_m)}{f_P^0}\qquad c_m=a_m^2+b_m^2$$

$$f_P(\lambda_m)=f_P^0+\Delta f'(\lambda_m)+i\Delta f''(\lambda_m)\qquad v=\frac{|F_P^0|}{|F_N^0|}$$

$$\eta_m=\frac{|F_N(\lambda_m)|^2+|\bar{F}_N(\lambda_m)|^2}{|F_N(\lambda_m)|^2-|\bar{F}_N(\lambda_m)|^2}=\frac{S(\lambda_m)+\bar{S}(\lambda_m)}{S(\lambda_m)-\bar{S}(\lambda_m)}$$

we get

$$\eta_m=\frac{1+2vb_m\cos\alpha+v^2c_m}{-2va_m\sin\alpha}$$

or

$$v=\frac{-1}{c_m}\Big[a_m\eta_m\sin\alpha+b_m\cos\alpha \pm\sqrt{(a_m\eta_m\sin\alpha+b_m\cos\alpha)^2-c_m}\Big]\,. \qquad (2)$$

The ambiguity in respect of the sign of the square root can be resolved by using the independence of λ of the ratio $v=|F_P^0|/|F_N^0|$. From this it follows that

$$\text{sign}\sqrt{\ }\neq\text{sign}(a_m\eta_m\sin\alpha+b_m\cos\alpha)\,.$$