Acta Cryst. (1976). A32, 859

On the Contribution of Short-Term Fluctuations to the Variance of the Measured Intensity

By J. Sygusch* and T. Cyr†

Department of Chemistry, University of Montreal, Montreal, Quebec, Canada

(*Received* 25 *July* 1975; *accepted* 20 *March* 1976)

The contribution of short-term fluctuations to the variance of the net intensity is shown to increase the variance above usual Poisson statistics and to be dependent upon two factors: the instrumental stability and the reflexion profile or time spent measuring a given data point. Simple formulae accounting for the increase in variance from normal Poisson statistics are presented.

Introduction

The statistical treatment of short-term fluctuations or drift during the course of a measurement of a reflexion has received very little theoretical attention. It has been recognized experimentally that fluctuations will affect the counting statistics and increase the variance in excess of that predicted by ordinary Poisson statistics (Schulz & Huber, 1971*a*,*b*). To account for the increase in variance, Schulz & Huber have modified the ordinary Poisson variance by a simple multiplicative factor, c, that is,

$$\sigma^2(k)=ck\,, \tag{1}$$

where c is greater than one and k represents the number of counts received in a given time interval.

Other authors have suggested the empirical modification to the variance (Busing & Levy, 1957; Peterson & Levy, 1957; Corfield, Doedens & Ibers, 1967; Stout & Jensen, 1968, p. 456; McCandlish, Stout & Andrews, 1975),

$$\sigma^2(k)=k+c'k^2\,, \tag{2}$$

where c' is a constant which can be related to the estimated value of the instrumental instability. In this note we propose a theoretical treatment of this problem together with the derivation of simple formulae which can be used to account for the increase in variance from normal Poisson statistics.

Theory

Fluctuations or variations of the intensity of the diffracted beam can be attributed to various sources. Of these, fluctuations of the incident beam and of the crystal orientation are probably the most important. To facilitate the subsequent analysis, it is advantageous to assume that fluctuations may be decomposed into two independent components. The first represents a slow or long-term drift which is considered to be constant during the measurement of a reflexion and can be completely determined by repeated measurement of reference reflexions suitably spaced in time. The second component represents short-term fluctuations occurring during the measurement of a reflexion and will be described as a continuous, stationary stochastic process. Such a process is realized when the nature of the fluctuations is the same from one measurement period to the next. This is normally found to be true provided that the crystal is firmly mounted and the power supply is well regulated.

The effect of short-term fluctuations cannot be separated from the statistical process of diffraction. In the absence of any fluctuations, this process can be characterized by the Poisson parameter,†

$$\lambda=\int_{t_1}^{t_2} I_0 f(t)\,\mathrm{d}t\,, \tag{3}$$

where I_0 is the incident beam intensity per unit time interval, $f(t)$ is the probability that a photon is diffracted at time t, and the integration is made over the time interval, (t_1,t_2), of the measurement of the reflexion.

The effect of random fluctuations on I_0 and $f(t)$ results in λ becoming a random variable. In the absence of any *a priori* model, fluctuations of the incident beam, I_0, and of crystal orientation affecting $f(t)$ cannot be distinguished from a single process occurring in the incident beam only. Provided that maximum noise amplitudes, $I_n(t)$, are at least one order of magnitude smaller than I_0, $I_n(t)$ can be incorporated into equation (3) in an additive manner:

$$\lambda=\int_{t_1}^{t_2} [I_0+I_n(t)]f(t)\,\mathrm{d}t \tag{4}$$

where the long-term drift is constant during the measurement. The equation (4) represents a stochastic

* Present address: Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada, T6G 2H7.

† Present address: National Research Council of Canada, Ottawa, Ontario, Canada.

† A Poisson process is represented by a frequency distribution, $\varphi_p(k)=e^{-\lambda}\lambda^k/k!$ where k is the number of counts measured in a given time interval.

integral whose statistics generally depend on the size of the interval, (t_1, t_2). The subsequent counting statistics must now be based upon a double stochastic Poisson process (Cox, 1955; Bartlett, 1963). The frequency distribution of this process is

$$\varphi_p(k) = E_\lambda\{e^{-\lambda}\lambda^k/k!\}\,, \tag{5}$$

where k is the number of counts received in the time interval, (t_1, t_2), and $E_\lambda\{\quad\}$ is the expectation operator with respect to the frequency distribution of λ. If it is assumed that short-term fluctuations are zero on the average, the expected value of the number of counts received in time interval (t_1, t_2) is

$$E(k) = E_\lambda\{\lambda\} = I_0 \int_{t_1}^{t_2} f(t)\mathrm{d}t \tag{6}$$

with

$$\sigma^2(k) = E_\lambda\{\lambda\} + \sigma^2\{\lambda\}\,, \tag{7}$$

where

$$\sigma^2\{\lambda\} = E_\lambda\{\lambda^2\} - (E_\lambda\{\lambda\})^2\,. \tag{8}$$

Then the effect of short-term fluctuations or noise is to broaden the variance $\sigma^2(k)$, as has been found experimentally by Schulz & Huber and shown in equation (1). It should be noted that the above derivation is free from assumptions of the type of probability distribution representing the short-term fluctuation. Furthermore, should the short-term fluctuations or noise be correlated, the ordinary computed sample variance will only in the limit of a large number of observations be a measure of $\sigma^2(k)$ or, put more precisely, the sample variance estimator is only asymptotically unbiased (Cox & Lewis, 1966).

Explicitly, the variance $\sigma^2(k)$ for correlated noise becomes

$$\sigma^2(k) = E(k) + \xi^2 \int_{t_1}^{t_2}\int_{t_1}^{t_2} f(u_1)f(u_2)\varrho(u_2-u_1)\mathrm{d}u_2\mathrm{d}u_1\,, \tag{9}$$

where $\varrho(u_2-u_1)$ is the autocorrelation function defined by

$$\varrho(u_2-u_1) = E_\lambda\{I_n(u_1)I_n(u_2)\}/\xi^2 \tag{10}$$

with

$$\xi^2 = E_\lambda\{I_n(u_1)I_n(u_1)\}\,.$$

Since the noise is stationary, ξ^2 will be independent of time.

To ascertain the importance of the contribution of noise to the variance $\sigma^2(k)$, the following two simple models will be considered:

$$\varrho(\tau) = \begin{cases} 1 \text{ for } \tau = 0 \\ 0 \text{ for } \tau \neq 0 \end{cases} \tag{10a}$$

and

$$\varrho(\tau) = 1 \text{ for all values } \tau\,. \tag{10b}$$

The variances for white noise [based on equation (10a)] and highly correlated noise [based on equation (10b)] become respectively

$$\sigma^2_{\mathrm{WHITE}}(k) = E(k) + (\xi/I_0)^2 \int_{t_1}^{t_2} [g(u)]^2\mathrm{d}u \tag{10c}$$

and

$$\sigma^2_{\mathrm{CORR}}(k) = E(k) + (\xi/I_0)^2\Big[\int_{t_1}^{t_2} g(u)\mathrm{d}u\Big]^2 \tag{10d}$$

with

$$g(t) = I_0 f(t)\,.$$

The quantities $g(t)$ and ξ/I_0 are what are commonly referred to as the reflexion profile and the instrument instability factor respectively.

If a value of 0·01 for ξ/I_0 is employed, together with a constant profile $g(t)$ in time (*i.e.* a step scan) of 1000 c.p.s. and time limits of 10 s, the numerical values of equations (10c) and (10d) are respectively

$$\sigma^2_{\mathrm{WHITE}}(k) = 11000(\text{c.p.s.})^2$$

and

$$\sigma^2_{\mathrm{CORR}}(k) = 20000(\text{c.p.s.})^2\,.$$

Clearly it is important to determine the type of noise present in the measurement system.

Experiment

The noise autocorrelation function, $\varrho(u_2-u_1)$ of equation (10), for a Picker FACS-I diffractometer is shown in Fig. 1. The crystal and detector were held stationary throughout the experiment at a setting corresponding to a peak of a high-intensity reflexion. Counts were repeatedly accumulated for one second every three seconds, for a period of two and one-half hours at a time of the day corresponding to high laboratory activity. The following procedure was used to eliminate the very slow fluctuations of periods greater than 2000 s and corresponding to the periodic operation of the air-conditioning system and other heavy equipment: The observed time series was first numerically filtered using a first-order recursive low-pass filter (Otnes & Enochson, 1972) and the new time series was then subtracted from the observed series. The autocorrelation function for the resultant series corresponds to the curve having a filter factor α set equal to 0·998. It is immediately apparent that the noise in the $\alpha = 0{\cdot}998$ series is highly correlated. In other words, noise having a period of less than 2000 s is definitely correlated and cannot be removed by repeated intensity remeasurement within this time period [see discussion after equation (8)].

Power supplies for X-ray generators are essentially low-pass filters of line noise. To remove the power-line noise, the original time series was filtered using a filter factor α of 0·67 chosen to match the power-supply parameters. This low-pass time series was then subtracted from the original series and the autocorrelation function of the resultant series computed. This is shown in Fig. 1 as a rather scattered horizontal trace. We conclude that when line noise is removed, *i.e.* by good voltage regulation either by the diffractometer generator and/or through use of a stable power line, the remaining noise is uncorrelated. It is then reasonable to suppose that only white noise remains and it will be so assumed in the following discussion.*

Discussion

The contribution of the short-term fluctuations in equation (10c) to the variance $\sigma^2(k)$ is determined by what has be defined as the beam instability parameter ξ/I_0 and by the profile $g(t)$ of the reflexion considered. It should be noted that the quantities ξ/I_0 and $g(t)$ are both experimentally measurable. Typical values of beam instability range from 0·005 to 0·025 in this laboratory. As the reflexion profile $g(t)$ is a function of the measurement technique employed, the formulae for the variance of the net intensity are derived below for the more commonly employed scan techniques. The net intensity is defined by

$$N = T - \frac{t_T}{t_B} B , \tag{11}$$

where T is the total measured intensity, t_T is the measurement time of the intensity, B is the background and t_B is the background sampling time.

* A further consequence of poor voltage regulation is that reflexions measured close in time will be correlated, with the implication that the weight matrix used in the construction of the normal equations is no longer diagonal.

(*a*) *Step scan*

The profile $g(t)$ for each step of the scan will be a function of the type

$$g(t) = T/t_T , \tag{12}$$

where the quantities, T and t_T, refer to a single step. The variance $\sigma^2(N)$ then is expressed for n steps by

$$\sigma^2(N) = \sum_{i=1}^{n} T_i + (\xi/I_0)^2 T_i^2/t_T + (nt_T/t_B)^2[B + (\xi/I_0)^2(B^2/t_B)] . \tag{13}$$

If the ratio of intensity measurement time t_T to the background measurement t_B is a constant, then for reflexions possessing the same number of counts for both peak and background, those clases of reflexions which have been measured over the greater time period will have the smaller variance.

(*b*) *Continuous scans*

A simple Gaussian model will be assumed to describe, in the absence of an integrating geometry, the reflexion profile due to a single wavelength such as $K\alpha_1$.* To take $K\alpha_1\alpha_2$ splitting into account the reflexion profile will be represented by a double Gaussian profile,

* It should be noted that the final observed profile is the result of many convolution operations (Ladell, 1965) which in the limit will yield a Gaussian profile (Papoulis, 1965, p. 233).

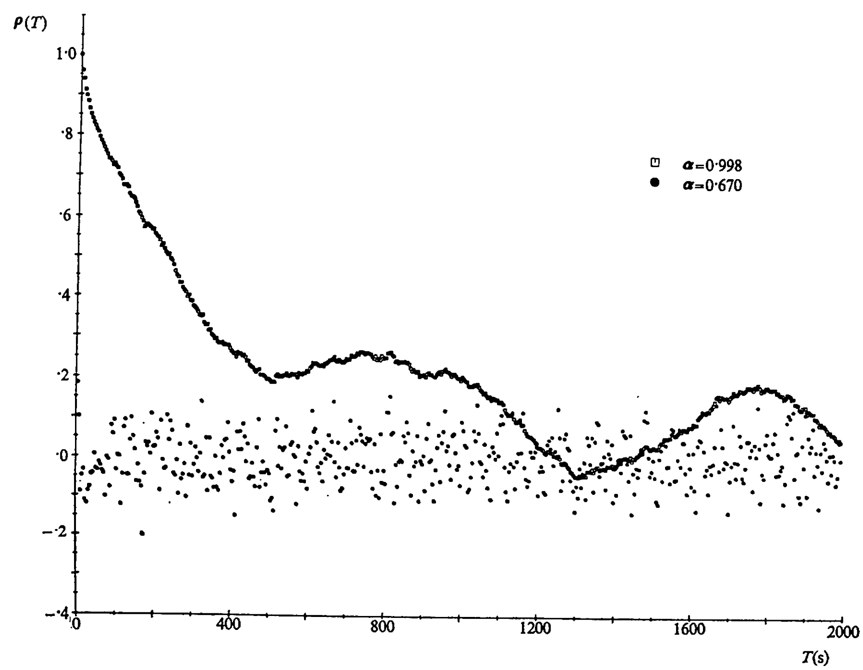

Fig. 1. Plot of the autocorrelation function p as a function of time T. The curve □, corresponding to filter factor $\alpha = 0·995$, represents a time series filtered to remove noise with periods greater than 2000 s. The curve ●, corresponding to filter factor $\alpha = 0·67$, represents a time series filtered to simulate a stable power supply.

$$g(t)=\tfrac{2}{3}\frac{T}{h\sqrt{2\pi}}\exp\left(\frac{-t^2}{2h^2}\right)+\tfrac{1}{3}\frac{T}{h\sqrt{2\pi}}\exp\left(\frac{-(t-\Delta)^2}{2h^2}\right), \quad (14)$$

where T is the total measured intensity, h is the half width in time and Δ is the separation in time of the $K\alpha_1$ and $K\alpha_2$ peaks. If truncation errors due to finite scanning limits are small, the variance of the net intensity becomes

$$\sigma^2(N)=T+\left(\frac{\xi}{I_0}\right)^2\frac{T^2}{2h\sqrt{\pi}}\left[\tfrac{5}{9}+\tfrac{4}{9}\exp\left(\frac{-\Delta^2}{4h^2}\right)\right]+\left(\frac{t_T}{t_B}\right)^2\left[B+\left(\frac{\xi}{I_0}\right)^2\frac{B^2}{t_B}\right]. \quad (15)$$

It is apparent from the above equation that the variance of a given reflexion will depend upon its profile as seen by the detector. A smaller half width, h, increases the sensitivity of a given reflexion to short-term fluctuations. Although the $K\alpha_1\alpha_2$ splitting decreases the contribution of short-term fluctuations to the variance by as much as $\frac{4}{9}$, there exists no maximum or minimum in the second term of equation (15), the term which represents the contributions of the short-term fluctuations to the variance of the peaks such that a compromise could be achieved between $K\alpha_1\alpha_2$ separation and half width. The variance for a reflexion with small half width will always be greater than a reflexion with a larger half width, provided all other quantities in equation (15) are equal and scanning limits do not truncate the profile. It is assumed that the detector aperture is large enough to accommodate the $K\alpha_1\alpha_2$ splitting in order for the above treatment to apply to a continuous ω scan.

Conclusion

The contributions of the short-term fluctuations to the variance of a reflexion are dependent upon two factors: the instrumental stability and the reflexion profile. The greater the time spent in measuring a given reflexion the smaller the contribution made by the short-term fluctuations. For continuous scans, the greater the half width in time of a given reflexion the smaller will be the contribution of short-term fluctuations to its variance. When $K\alpha_1\alpha_2$ splitting becomes significant, the increase in the variance due to short-term fluctuations will be less for high-angle reflexions than for comparable low-angle reflexions. In order for these conclusions to be valid the power supply of the generator must be well stabilized and the crystal firmly mounted.

We would like to thank Drs F. Brisse and G. Williams for their encouragement and time and Dr U. Maag of the Département d'Informatique for his helpful discussion and the National Research Council of Canada for the granting of a Scholarship (J.S.).

References

BARTLETT, M. S. (1963). *J. Roy. Stat. Soc.* **B25**, 129–164.

BUSING, W. A. & LEVY, H. A. (1957). *J. Chem. Phys.* **26**, 563–568.

CORFIELD, P. W. R., DOEDENS, R. J. & IBERS, J. A. (1967). *Inorg. Chem.* **6**, 197–204.

COX, D. R. (1955). *J. Roy. Stat. Soc.* **B17**, 129–164.

COX, D. R. & LEWIS, P. A. W. (1966). *The Statistical Analysis of Series of Events*, p. 87. London: Methuen.

LADELL, J. (1965). *Trans. Amer. Cryst. Assoc.* **1**, 89.

MCCANDLISH, L. E., STOUT, G. H. & ANDREWS, L. C. (1975). *Acta Cryst.* **A31**, 245–249.

OTNES, R. K. & ENOCHSON, L. (1972). *Digital Time Series Analysis*, p. 69. New York: Wiley–Interscience.

PAPOULIS, A. (1965). *Fourier Integral.* New York: McGraw. Hill.

PETERSON, S. W. & LEVY, H. A. (1957). *Acta Cryst.* **10**, 70–76.

SCHULZ, H. & HUBER, P. J. (1971*a*). *Acta Cryst.* **A27**, 536–539.

SCHULZ, H. & HUBER, P. J. (1971*b*). *Acta Cryst.* **A27**, 540–544.

STOUT, G. H. & JENSEN, L. H. (1968). *X-ray Structure Determination.* New York: Macmillan.