COOPER, M. J. & NATHANS, R. (1968*a*). *Acta Cryst.* A**24**, 481.
COOPER, M. J. & NATHANS, R. (1968*b*). *Acta Cryst.* A**24**, 619.
COOPER, M. J. & ROUSE, K. D. (1968). *Acta Cryst.* A**24**, 405.
LOWDE, R. D. (1954). *Proc. Roy. Soc.* A**221**, 206.
ROUSE, K. D. & COOPER, M. J. (1969). *Acta Cryst.* A**25**, 615.
ROUSE, K. D. & COOPER, M. J. (1970). *Acta Cryst.* A**26**, 457.
SEEGER, R. J. & TELLER, E. (1942). *Phys. Rev.* **62**, 37.
WALLER, I. & FROMAN, P. O. (1952). *Ark. Fys.* **4**, 183.
WILLIS, B. T. M. (1970). *Acta Cryst.* A**26**, 396.

Acta Cryst. (1971). A**27**, 157

Normal Probability Plot Analysis of Error in Measured and Derived Quantities and Standard Deviations

BY S. C. ABRAHAMS AND E. T. KEVE*

Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey, U.S.A.

(*Received* 25 *May* 1970)

Normal probability plot analysis is applied to independent sets of crystallographic structure factor measurements (F) and the derived coordinates (p). Differences between corresponding pairs of structure factors (ΔF) in the two sets are examined in terms of their pooled standard deviations (σF) by plotting the ordered statistic $\delta m = \Delta F/\sigma F$ against the expected normal distribution. Differences between pairs of coordinates (Δp) are similarly examined in a $\delta p = \Delta p/\sigma p$ half-normal probability plot. Both plots result in linear arrays of unit slope and zero intercept, for normal error distribution in the experiment and the model and correctly assigned standard deviations. Analysis of departures from this ideal, especially when both plots are considered together, provides detailed information of the kinds of error in δm and in δp. By inference, the kinds of error in F and σF as well as in p and σp can be deduced. The normal probability plot $\delta R = |F_{meas}| - |F_{calc}|/\sigma F_{meas}$ should ideally also be linear, with unit slope and zero intercept. Deviations from ideal provide considerably more information than the conventional R values. Analysis of δR in combination with δm plots allows further specification of the error distribution. Examples using these plots are given and discussed, based both on real and on simulated data.

Introduction

The association of a given measured quantity with a reliable estimate of the uncertainties in that quantity is of fundamental importance. The uncertainties in derived quantities, however, have often been given greater importance than those assigned to observed quantities. Propagation of error theory shows that estimates of these two kinds of uncertainty, later referred to as the derived and the assigned standard deviations respectively, are functionally related (for a discussion, see Birge, 1939). The additional dependence of the least-squares refined parameters on the weights of the observations, obtained from the assigned standard deviations, was implied by Gauss (1809). The assigned standard deviations are therefore critical, and should be estimated with care. Both these uncertainties should be capable of satisfying stringent *post facto* tests for validity.

The increasing use in structural crystallography of diffractometers to measure structure factors (F_{meas}) has stimulated attempts at assigning experimental standard deviations (σF_{meas}) to these quantities. Thus, Busing & Levy (1957) proposed an expression for σF_{meas} that included both counting statistics and an empirical term proportional to the net count in the reflection. Similar proposals have subsequently been made by others. An objective assessment of σF_{meas} on the same scale as the F_{meas} is, however, possible only from the sum of all the independent variances entering the measurement. A method for evaluating these variances has been given (Abrahams, 1964). The assigned σF_{meas} magnitudes are shown by model-dependent indicators to be close to their absolute scale (Abrahams, 1969).

In any model-independent procedure for testing the assigned standard deviations, it is necessary to measure at least two independent sets of F_{meas}. The availability of two independent data sets, which need not be complete, allows the validity both of the derived as well as the assigned standard deviations to be tested by methods developed below. In addition, duplicate measurements on two different crystals of the material under study considerably increase the chances that the crystallographic results reported are indeed typical of that material.

The results from a recent structural investigation (Keve, Abrahams & Bernstein, 1970), and a parallel simulated case, are used as examples to illustrate the

* Present address: Mullard Research Laboratories, Redhill, Surrey, England.

test procedures both on the assigned σF_{meas} and on the derived σp_i, where p_i is the ith parameter obtained in a least-squares fit to a theoretical model.

The probability plot

The distribution of any set of magnitudes may be compared with any assumed distribution in a probability plot. Consider the quantities $F(1)_i$, $F(2)_i$ each of which is an independent measure of the same ith structure factor, obtained in two separate experiments. Let us assign standard deviations $\sigma F(1)_i$ and $\sigma F(2)_i$ to these quantities. The statistic δm_i is defined as

$$\delta m_i = [F(1)_i - KF(2)_i]/[\sigma^2 F(1)_i + K^2\sigma^2 F(2)_i]^{1/2} \quad (1)$$

where K is a scale factor between the two sets of measurements such that

$$\sum_{i=1}^{j} (\delta m_i)^2 = \text{a minimum} \quad (2)$$

and there is a total of j structure factors. The distribution of the δm_i is Gaussian if the $F(1)_i$, $F(2)_i$ contain only random error and if the $\sigma F(1)_i$, $\sigma F(2)_i$ are correct. In general, this ideal is not met and deviations can be examined with great sensitivity by means of the probability plot. In constructing the normal probability plot the collection of δm_i is rearranged in order of increasing magnitude, *i.e.* as *order statistics*. If these order statistics are plotted against x_i, the values (quantiles) expected for a normal distribution, the resulting probability plot will be a straight line of slope 1 that passes through the origin. Values of x_i are readily obtained from the normal probability function

$$P(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} \exp\{-\alpha^2/2\} d\alpha \quad (3)$$

The ith value of $P(x)$ for the j-ordered statistics is given by $|(j-2i+1)/j|$. The sign of x is positive for $i > j/2$, negative for $i < j/2$. A convenient tabulation of $P(x)$ is found in the National Bureau of Standards' *Tables of Normal Probability Functions* (1953).*

The probability plot is entirely general, and any basis distribution may be chosen, such as exponential or gamma: if the measured distribution of δm_i is identical to that chosen, the resulting plot will remain linear with a slope of unity and zero intercept. Normal distributions only are considered here. A general discussion of probability plotting methods for data analysis has been given by Wilk & Gnanadesikan (1968).

The information revealed by a normal probability plot displaying all the δm_i simultaneously considerably extends that contained in the simple interexperimental agreement factor

$$R_{1,2} = \sum_{i=1}^{j} [F(1)_i - KF(2)_i] / \sum_{i=1}^{j} [F(1)_i + KF(2)_i]/2$$

or that in an examination only of the largest ith individual differences.

Properties of the normal probability plot of δm

The normal probability plot is sensitive to individual δm_i containing gross error, and results in such points lying far from the ideal linear distribution. The ranking process in forming order statistics necessarily places the largest δm_i at the extremes of the array. The density of points on this array, projected on the observation axis, represents the distribution of the δm_i. Hence, in judging a normal probability plot for linearity and slope as is discussed later, it is important to give greatest weight to the central portion, containing the majority of the data, and least to the sparsely populated outermost portion of the array. Departure of an individual δm_i from the remainder of the array is less important than overall trends. The significance of such departure may be found by standard means, such as the χ^2 test.

A linear normal probability plot with slope different from unity may be an indication of uniform misestimation in σm_i. Thus a slope of 0·25 would be obtained with data for which the assigned standard deviations were too large, on average, by a factor of 4. A linear normal probability plot based on a value of K derived from equation (2) should have zero intercept.

Numerous plots of quantiles of a wide variety of probability distributions against quantiles of normal, uniform, and exponential probability functions have been computed by Chambers & Fowlkes (1967). In

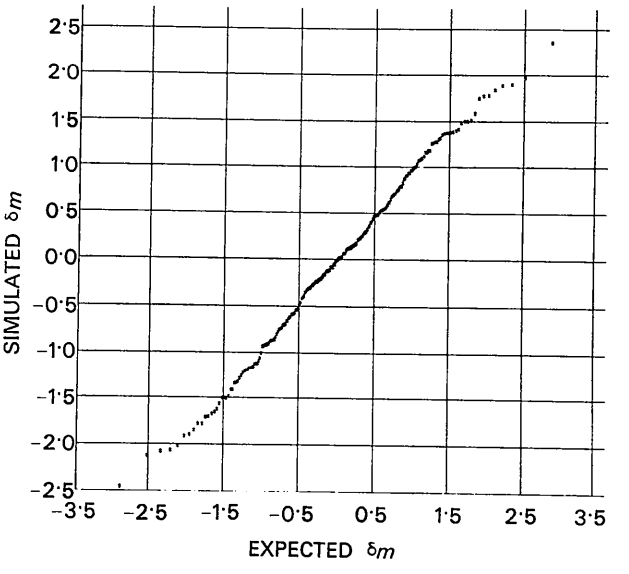

Fig. 1. Normal probability plot of 254 simulated δm_i [equation (1)] with normally distributed random error. For one set, the error is 0·05 F_{EF}, for the other it is 0·04 F_{EF}.

* For example, with $j=254$, the extreme δm_i with $i=1$ and 254 have $P(x)_1 = 253/254$, $x_1 = -2{\cdot}883$ and $P(x)_{254} = 253/254$, $x_{254} = 2{\cdot}883$: the next pair of δm_i with $i=2$ or 253 have $P(x)_{2,253} = 251/254$ and $x_{2,253} = \mp 2{\cdot}518$, *etc.*

general, although various concave–convex, convex, and convex–concave shaped probability plots may be obtained, a true match of distributions gives a function clearly recognizable as linear.

Simulated example and resulting normal probability plot of δm

As a standard with which the real example given below may be judged, two sets of measured structure factors containing random error only are considered in the simulated example of this section. Let:

$$F(1)_R = F_{EF} + N_1 \,.\, \sigma_1 F_R \qquad (4)$$

and

$$F(2)_R = F_{EF} + N_2 \,.\, \sigma_2 F_R \,. \qquad (5)$$

The values of F_{EF} are derived from the final atomic position coordinates and temperature coefficients of $BaCoF_4$ (Keve, Abrahams & Bernstein, 1970), using the atomic scattering factors and dispersion corrections of that paper, as error free magnitudes. N_1 and N_2 are two sets of computer generated (Chambers, 1968) random standard normal deviates. The assumed standard deviations due to random error alone are $\sigma_1 F_R = 0{\cdot}05\ F_{EF}$, $\sigma_2 F_R = 0{\cdot}04\ F_{EF}$. The resulting interexperimental agreement factor is 0·049, close to the expected value.

In this simulated example, it is instructive to note that the δm_i do not involve the values of $F(1)_R$ and $F(2)_R$, but have values given by $(0{\cdot}05N_1 - 0{\cdot}04N_2)/0{\cdot}064$, equation (1). A display of the δm_i from the present simulation is given in Fig. 1 on a normal probability plot. The various small breaks in the array, and the departure from linearity at the extremes, should be regarded as of no significance and to be expected if N_1 and N_2 are indeed random standard normal deviates.

Calculation and plotting of this and all subsequent normal or half-normal probability plots has been performed using DeChaine's (1967) *GNORMS* plotting routine.

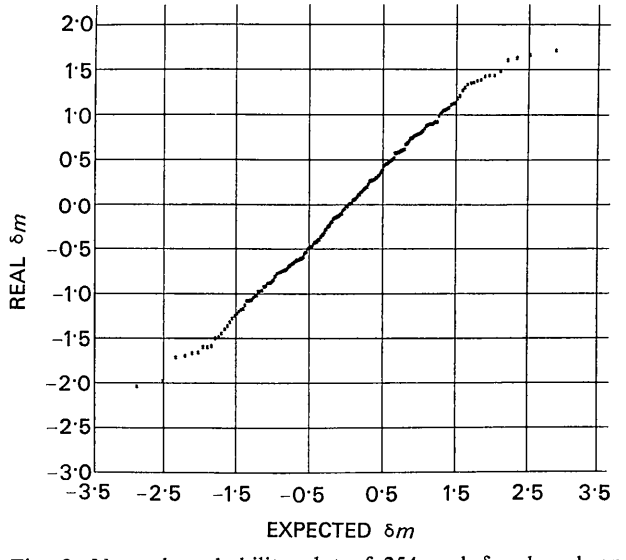

Fig. 2. Normal probability plot of 254 real δm_i based on measurement of $BaCoF_4$ structure factors using two different crystals.

Normal probability plot of δm for real case

Two independent sets of F_{meas} were measured on $BaCoF_4$ (Keve, Abrahams & Bernstein, 1970) under experimental conditions that were essentially identical except for the size and regularity of the polyhedral crystal specimens studied. The experimental values of σF_{meas} were assigned from equation 6

$$\sigma^2 F^2_{meas} = V(\text{counting statistics}) + V(\text{form}) + 4 \sum_n k_n^2 F^4_{meas} \qquad (6)$$

where V(counting statistics) is given by the Poisson distribution in the arrival of quanta at the counter and V(form) is given by equation (7)

$$V(\text{form}) = \sum_{j=1}^{J} \left(F^2_{meas_j} - \frac{1}{J} \sum_{j=1}^{J} F^2_{meas_j} \right)^2 \Big/ (J-1) \qquad (7)$$

which measures the variation among the J-crystallographically equivalent members of a form. The percentage error estimated in the nth recognized variable is given by k_n (see Abrahams, 1969, for details).

The first two terms of equation 6 are in general different for corresponding F_{meas} measured on each crystal. However, in evaluating the third term, each component is identical except for that relating to the estimated error due to differences between the measured and actual crystal dimensions. For crystal 1, this component has value $207 \times 10^{-4}\ F^4_{meas}$; for crystal 2, the corresponding value is $169 \times 10^{-4}\ F^4_{meas}$. The complete value of $4 \sum_i k_n^2 F^4_{meas}$ is $267 \times 10^{-4}\ F^4_{meas}$ for crystal 1 and $229 \times 10^{-4}\ F^4_{meas}$ for crystal 2. It may be noted that the variance given by equation (7) is necessarily partly due to the variation in absorption for different reflections within a form. To the extent that this effect is also contained in the third term of equation (6), σF_{meas} is overestimated.

Values of δm_i for the 254 structure factors common to the measurements made on crystal 1 and crystal 2 were obtained after scaling [equation (2)]. The resulting normal probability plot, shown in Fig. 2, does not depart appreciably from that expected for a normal distribution of δm_i (compare with Fig. 1). The linear array, in the plot, has slope of 0·8 and passes through the origin.

Fig. 2 thus allows the conclusion to be drawn that the δm_i have a normal random distribution. The slope of the plot in Fig. 2 of 0·8 is interpretable either on the basis that $\sigma \Delta F_{meas}[= \{\sigma^2 F_{meas}(1) + K^2 \sigma^2 F_{meas}(2)\}^{1/2}]$ is overestimated on average by 20 per cent, or that $\Delta F_{meas}[= \{F_{meas}(1) - K F_{meas}(2)\}]$ is underestimated by

about 20 per cent, or a partial combination of both effects. It is assumed that errors in $\sigma\Delta F_{\text{meas}}$ or in ΔF_{meas} due to error in K are negligible. In the absence of additional information, the third of these interpretations would be chosen as most likely, with an equal distribution of the two effects *i.e.* that σF_{meas} is generally too large by about 11 per cent (see 'Normal probability plot comparison of model with measurement' for additional information).

This purely external or experimental assessment of the validity of the assigned standard deviations may be compared with the quasi-external $\mathscr{U}_\alpha$ indicator (Abrahams, 1969) where

$$\mathscr{U}_\alpha=[\sum_j(\mathscr{D}^2F_i/\sigma^2F_{\text{meas}_i})/\chi^2_{j-v,\alpha}]^{1/2}$$

and

$$\mathscr{D}F_i=|F_{\text{meas}_i}|-|F_{\text{calc}_i}|,$$

$\chi^2_{j-v,\alpha}$ is the χ^2 distribution magnitude at the α-significance level, and the model for which $\sum(\mathscr{D}^2F_i/\sigma^2F_{\text{meas}_i})$ is minimized contains v variables. Parallel series of least-squares refinements carried out on all the F_{meas} (1) and F_{meas} (2) resulted in values for the standard deviation of an observation of unit weight S of 0·887 for crystal 1 and 0·899 for crystal 2. For both refinement series 36 parameters (including the factor relating F_{meas} to the absolute scale) were varied: for crystal 1 there were 386 and for crystal 2 there were 877 independent F_{meas}. At the one per cent confidence level, the range of $\mathscr{U}_{0\cdot01}$ is 0·088, 0·982 for crystal 1 and 0·846, 0·959 for crystal 2.

The conclusion may be inferred from these $\mathscr{U}_{0\cdot01}$ ranges that the σF_{meas} are overestimated. The probability that the error in σF_{meas} lies between 1·8 and 19·2 per cent for crystal 1, and between 4·1 and 15·4 per cent for crystal 2, is 98 per cent in each case. The agreement between this conclusion and that drawn from the normal plot (Fig. 2) is highly satisfactory. Hence, for investigations in which a single set of independent measurements only have been made, a measure of the average error in σF_{meas} can be derived by calculation of $\mathscr{U}_\alpha$, although this statistic is necessarily less revealing than the normal probability plot.

Half-normal probability plot

In cases for which the sign of δ_i is redundant, the full-normal probability plot must be replaced by the half-normal probability plot, corresponding to an assumed standard half-normal distribution of δ_i. An important example arises in comparisons of pairs of experimentally determined sets of positional coordinates. It is equally valid to use either the stated coordinate or a transformed coordinate. Hence if the position coordinate xyz is transformed, for example, to $\frac{1}{2}-x$, $\frac{1}{2}+y$, $\frac{1}{2}-z$, the sign of δ_i for x and z would appear reversed. Of course, only corresponding pairs of coordinates may be compared.

The half-normal probability plot is constructed by ordering the $|\delta_i|$ from zero to maximum value. The ith value of $P(x)$ for the j-ordered statistics is given by $(2i+1)/2j$, and the resulting value of x is obtained from $P(x)$ [equation (3)]. Half-normal probability plots are more prone to ambiguity of interpretation than normal probability plots, since the modulus of the parent distribution is now used. A half-normal probability plot close to linear with slope of unity and zero intercept may be generally interpreted as due to a correct match between measured and assumed distribution, with correctly estimated standard deviations. It may be noted that, unlike the normal probability plot in which the sets of F_{meas} to be compared should first be placed on a common scale [equation (2)], the parameter values in the half-normal probability plots discussed below are absolute. Hence, non-zero intercepts in such cases indicate the presence of systematic error.

Simulated example and resulting half-normal probability plot of δp

The structure factor magnitudes and standard deviations generated by equations (4) and (5) were used in a standard least-squares process with Busing, Martin & Levy's (1962) *ORFLS* program, to obtain refined position coordinates and temperature coefficients for the simulated model of $BaCoF_4$. The statistic δp_i was then formed where

$$\delta p_i=||p(1)_i|-|p(2)_i||/\{\sigma^2p(1)_i+\sigma^2p(2)_i\}^{1/2} \quad (8)$$

$p(1)_i$, $p(2)_i$ are the final parameters obtained with the $F(1)_R$ and $F(2)_R$ structure factors, and the $\sigma^2p(1)_i$,

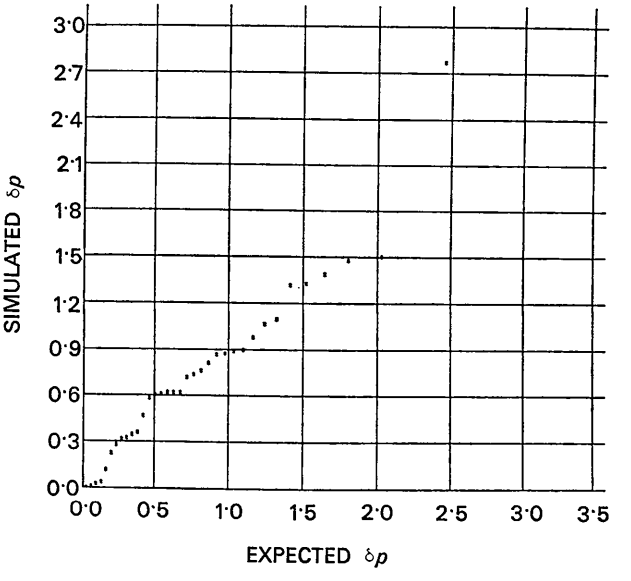

Fig. 3. Half-normal probability plot of 35 simulated δp_i [equation (8)] derived by refining the two sets of F_R containing the random errors used for Fig. 1.

$\sigma^2 p(2)_i$ are the associated variances in these parameters.

Assuming a standard half-normal distribution, the ordered δp_i are shown plotted as a half-normal probability plot in Fig. 3. The apparent departures from linearity are assumed to be without significance and provide a standard for judging subsequent half-normal plots. It may also be seen that the slope is close to unity as is to be expected.

Structure factors calculated on the basis of the refined parameter sets $p(1)$ and $p(2)$ have accuracy indicators given in Table 1 for comparison with the experimental results in the next section.

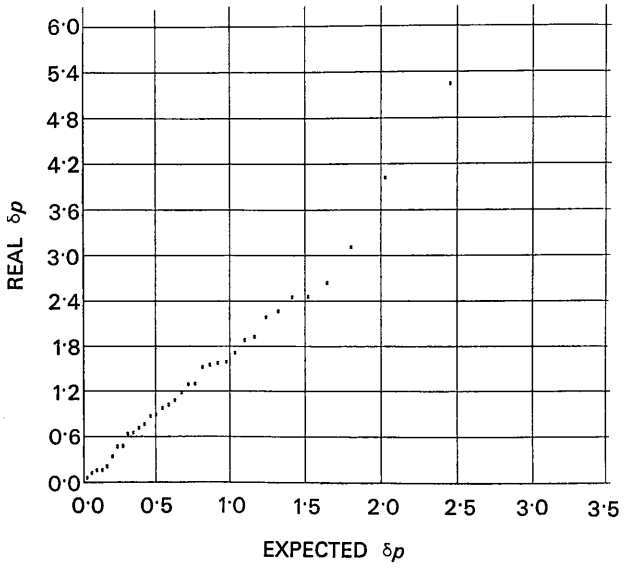

Fig. 4. Half-normal probability plot of 35 real δp_i derived from measurement on $BaCoF_4$ as in Fig. 2.

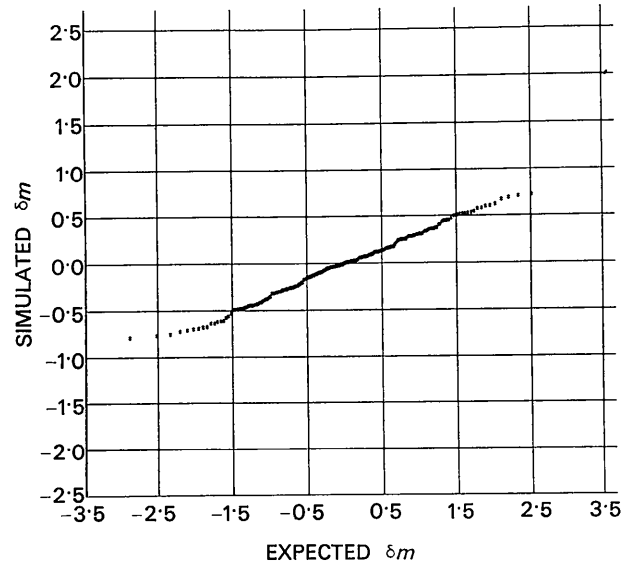

Fig. 5. Normal probability plot for 254 simulated δm_i based on a combination of random and systematic error.

Table 1. *Indications of accuracy in simulated structure factors containing random error*

	R	wR	S	$\mathscr{U}_{0.01}$
$F(1)_R$ set	0·0389	0·0462	0·998	0·888, 1·163
$F(2)_R$ set	0·0333	0·0392	1·061	0·943, 1·237

Half-normal probability plot of δp for real case

The two independent sets of structure factors measured from crystals 1 and 2 refined to give the final parameters $p(1)_i$ and $p(2)_i$ of Keve, Abrahams & Bernstein (1970). Table 2 contains accuracy indicators for these refinements. Fig. 4 gives the resulting half-normal probability plot.

Table 2. *Indicators of accuracy in measured structure factors of* $BaCoF_4$

	R	wR	S	$\mathscr{U}_{0.01}$
$F_{meas}(1)$	0·0749	0·0924	0·887	0·808, 0·982
$F_{meas}(2)$	0·0584	0·0850	0·899	0·846, 0·959

The ordered array of real δp_i in Fig. 4 is as close to linear as the array of simulated δp_i in Fig. 3 obtained with random error only present in the data. The array, which points at zero, has a slope of about 1·75 indicating either that the Δp_i are too large, the σp_i too small, or that both conditions are applicable. The only way in which the Δp_i can be too large is if bias is present in the data such as to cause most $p(1)_i$ parameter values to differ from the true (but unknown) value in a sense opposite to that in which the corresponding $p(2)_i$ value differs. The half-normal probability plot for such a case cannot be linear, hence the plot of Fig. 4 provides strong evidence for excluding such bias.

The alternative hypothesis, that σp_i is too small, may be accounted for if both F_{meas} (1) and F_{meas} (2) contain error components that vary systematically with parameters in the model. In this case, both sets of data are necessarily fitted by least-squares to the resulting calculated structure factors better than is warranted by the accuracy of the data, thus giving sets of σp_i that are too small. If this alternative hypothesis is correct, and if the distribution of error between $\sigma p(1)_i$ and $\sigma p(2)_i$ is equal, then both sets of standard deviations are too small by 75 per cent.

In the absence of a half-normal probability plot, it would be necessary to regard the largest Δp_i of $5{\cdot}26 p_i$, corresponding to the data in Fig. 4, as highly significant. For 35 p_i parameters (excluding the scale factor) the chance of such a large Δp_i occurring is only 7×10^{-6}. However, application of the factor 1·75 to all the σp_i results in a half-normal probability plot with

an essentially normal distribution of unit variance. The largest Δp_i is in consequence reduced to 2·97 σp_i, for which the chance of occurrence is 10 per cent. Thus, the entirely false conclusion that the largest values of Δp_i represent highly significant differences is averted by means of the probability plot.

Case of simulated structure factor sets containing similar bias

Measured sets of structure factors, in general, contain both random and systematic error. The influence of different combinations of these errors may conveniently be investigated by simulation. In this section, brief consideration is given to the simple case of two sets of structure factors, each of which contains similar but not identical systematic bias both as the sole source of error (F_S) and also in combination with a random normal error ($F_{S,R}$).

The structure factors are given explicitly by equations 9 to 12.

$$F(1)_S = F_{EF} + 0{\cdot}3\, F_{EF}\,(\sin\theta_{hkl})/(\sin\theta_{\max}) \quad (9)$$

$$F(2)_S = F_{EF} + 0{\cdot}2\, F_{EF}\,(\sin\theta_{hkl})/(\sin\theta_{\max}) \quad (10)$$

$$F(1)_{S,R} = F(1)_S + 0{\cdot}05\, N_1 F_{EF} \quad (11)$$

$$F(2)_{S,R} = F(2)_S + 0{\cdot}04\, N_2 F_{EF}\,. \quad (12)$$

(See equations 4 and 5). The systematic bias is closely related to an artificial isotropic temperature factor. The standard deviation of $F(1)_S$ is taken as $\sigma F_S = 0{\cdot}15\, F_{EF}$, and of $F(2)_S$ as 0·10 F_{EF} from equations (9) and (10). The standard deviation of $F_{S,R}$ includes σF_S and σF_R defined previously.

With only systematic bias present, the interexperimental agreement factor $R_{1,2}$ is 0·0153. The normal probability plot (not shown) of δm_i is slightly concave–convex, passing through zero and with a slope of about 0·11. This small slope is primarily due to the difference terms in equation (1) being too small. With both systematic and random error present, $R_{1,2}$ for the two $F_{S,R}$ sets is 0·0452. The resulting normal probability plot of δm_i is shown in Fig. 5. The normal error component has essentially straightened the plot, which should be regarded as linear by comparison with Fig. 1. The small slope is primarily governed by the nearly parallel biases present in the data sets.

Refinement of the $F_{S,R}$ sets, and the F_S sets, results in half-normal probability plots of δp containing distinct structure with associated departure from linearity. The simulated bias correlates strongly with the β_{ii} parameters and very weakly with the remaining parameters, causing a separation between these two types of parameter on the half-normal plot.

In several simulations of error in which the σp plot was not linear, the corresponding values of σp_i were entirely unreliable in terms of the departures of the p_i from the true (known) values. Conversely, a linear δp_i array of slope $\mathscr{S}(\delta p)$ in the simulations corresponded to departures of p_i that are acceptable in terms of $\mathscr{S}(\delta p)\cdot\sigma p$.

Diagnostics for systematic error from essentially linear normal probability plots

Markedly nonlinear δm and δp plots are necessarily caused by systematic error. It is assumed throughout this paper that the basis distribution is Gaussian. The presence of most kinds of systematic error is hence readily recognized, although the scatter inherent in plots drawn from a limited number of data may obscure small departures from linearity. Certain types of systematic error result in apparently linear δm and δp plots. These are errors that have approximately equal

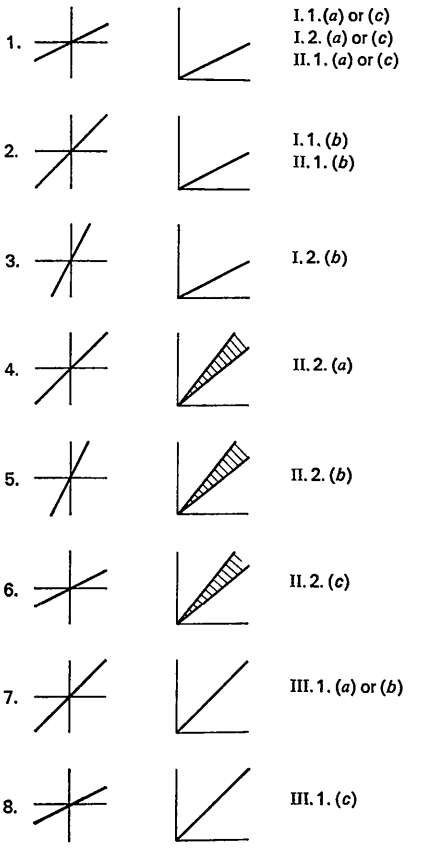

Fig. 6. Idealized linear normal (left) and half-normal (center) probability plots expected for various combinations of systematic and random error and case (right) given in the Appendix. Abscissae represent expected δ values, ordinates experimental values.

1. $S \gg \Delta S$, $\sigma F >$ correct value
2. $\Delta S \ll \Delta R$, $\sigma F = \sigma F_R$ only
3. $S \gg R$, $\Delta S \simeq \Delta R$, $\sigma F = \sigma F_R$ only
4. $S \approx R$, $\Delta S \approx \Delta R$, σF correct
5. $S \approx R$, $\Delta S \approx \Delta R$, $\sigma F = \sigma F_R$ only
6. $S \approx R$, $\Delta S \approx \Delta R$, $\sigma F >$ correct value
7. $S \ll R$, $\sigma F = \sigma F_R =$ correct value
8. $S \ll R$, $\sigma F >$ correct value.

but low correlation with all parameters varied in the model.

Ideally, the assigned σF_{meas} should combine an accurate estimate of the variance due to random error with the mean square errors due to systematic bias (case *a*). The real case will deviate from ideal toward one of two extremes. At one, σF_{meas} consists of estimates of the variances due to random error only (case *b*). At the other, σF_{meas} contains an overestimate of systematic error in addition to an estimate of the variance due to random error (case *c*). Evaluation of $\mathscr{U}_x$ may indicate which of the above cases is applicable.

It is assumed that corresponding pairs of structure factors that are independently measured and correctly scaled may be represented by equations (13) and (14).

$$F(1)=|\mathbf{F}_{EF}+\mathbf{S}(1)+\mathbf{R}(1)| \qquad (13)$$

$$F(2)=|\mathbf{F}_{EF}+\mathbf{S}(2)+\mathbf{R}(2)| \qquad (14)$$

where S(1), S(2), R(1), R(2) are vectors giving the magnitudes and sense of the systematic (*S*) and random (*R*) components of the error associated with the true structure factor. Detailed analysis of combinations of these vectors (see Appendix) lead to apparently linear δm and δp plots of varying slopes. By matching the experimental δm and δp plots to one of the idealized pairs shown in Fig. 6, the relative magnitudes of the random and systematic errors present may be diagnosed.

Normal probability plot comparison of model with measurement

In previous sections, probability plot analysis has been applied to comparisons of independent sets of measurements in δm plots, and to independent sets of derived parameters in δp plots. The probability plot may also be applied to a comparison of measured and calculated structure factors (δR plots). It is emphasized that the δR plot by itself is not as powerful as, nor is it a substitute for, the δm and δp plots obtainable from independent measurement sets. The δR plot in combination with δm and δp plots can however give additional information: it is also considerably more convenient to use than are the columns of $\mathscr{D}F/\sigma F_{\text{meas}}$ (printed by many least-squares refinement programs) and is greatly to be preferred to the conventional *R* value.

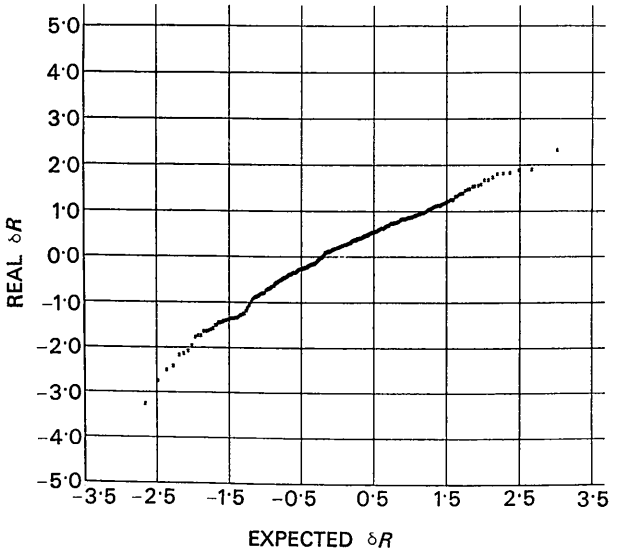

Fig. 7. Normal probability plot of 386 δR_i, based on measurement of $BaCoF_4$ crystal 1 structure factors.

The δR plot is constructed similarly to the δm normal probability plot: the statistic $\delta R_i=\mathscr{D}F_i/\sigma F_{\text{meas}_i}$, where the least-squares refined scale factor is used to place F_{meas} and σF_{meas} on the F_{calc} scale, is ordered and plotted against x_i [equation (3)]. If the F_{meas} contain no systematic error, the σF_{meas} are correctly estimated, and the trial structure is essentially correct, then the method of least-squares will result in a linear δR plot with zero intercept and a slope of unity. Deviations from any of these plot characteristics indicate a failure in one or more of the assumptions.

An example of a δR plot provided in Fig. 7, based on the published data for crystal 1 of $BaCoF_4$, is briefly analysed. This array is close to linear, except for a few data points, with a slope of 0·8 and intercept of $\frac{1}{4}\,\delta R$. The near linearity of the array corresponds to an absence of systematic error in the F_{meas} (1) of the kind that cannot be absorbed by the model. It is noted that the slope is consistent with the result of Table 2 and of Fig. 2, and that the intercept is not significantly different from zero, in terms of the least squares error of 1·1 % in the scale factor.

Combination of the δR with the δm plot provides further information. The common slope of 0·8 indicates the σF_{meas} (1) to be most-likely overestimated by about 20 per cent, and hence the earlier interpretation that part of the deviation of $\mathscr{S}(\delta m)$ from unity should be attributed to errors in ΔF_{meas} (see *Normal probability plot of δm for real case*) is unnecessary. An additional conclusion is that there is no essential difference between any F_{meas} $(1)_i$ and F_{calc_i} for this data set and model.

Discussion

The implicit assumption is always made that a single crystal specimen of a given substance is indeed representative of the intrinsic properties of that material, so long as an investigation is confined to that one specimen alone. It is only possible to examine this assumption by studying the properties of numerous samples. Where the necessary measurements are long and tedious, the study of at least two specimens enables a partial test of this assumption to be made. The information thus generated then allows several other hypotheses to be tested.

Comparison of sets of F_{meas} from two different crystals, in a normal probability plot, immediately shows

if the δm_i are normally distributed. A normal distribution is consistent with the hypothesis that both crystals generate the same population of F_{meas} and that the σF_{meas} are correctly estimated. Conversely, a nonlinear δm plot would arise from a distribution other than normal, allowing the inference to be made that either the two data sets differ systematically or the σF_{meas} are systematically misestimated. Apparent or real differences in the data sets may be caused by one or more of the following:

1. Unrecognized errors introduced in measurement.
2. Physical differences such as the degree of primary or secondary extinction present.
3. Chemical differences resulting from failure of the assumption that the crystals are indeed representative of the given material.

The possibility that cause (1) is operative should always be considered and positively eliminated before proceeding further. For two sets of F_{meas} drawn from the same statistical population, the δm plot must have zero intercept after application of equation (2) to derive K. If the intercept is not zero, the data sets cannot be related by a single scale factor and hence cannot come from the same normal population.

Some of the differences due to cause (2) may be corrected for in the model. Remaining significant differences of unrecognized origin result in a nonlinear δp plot. Suspected errors sources may be tested by examination of the δp plot after each attempted correction. Similarly, subsets of either F_{meas} or p_i may be examined, using δm or δp probability plots of the subsets. Thus, the presence of suspected symmetry elements among the set of F_{meas} is detectable from probability plots of appropriate subsets of F_{meas}: likewise, the possibility that the δp_i associated with the position coordinates belong to a different population from those associated with the thermal coefficients may also be tested directly on subset plots.

Departures from unity in the slopes of *linear* plots of δm and δp are discussed in the Appendix and are summarized in Fig. 6.

Underestimation of derived standard deviations has also been found in the recent I. U. Cr. single crystal intensity project, which undertook, in part, to refine the structural parameters of D(+)-tartaric acid for each of seventeen independent sets of F_{meas} (Hamilton & Abrahams, 1969). Comparison of the agreement among individual parameters and the weighted mean with the standard deviations derived by the method of least-squares showed that in this project the derived σp_i for all coordinates are about 1·95 too small on average. This I.U.Cr. project result, together with the findings obtained with the $BaCoF_4$ data, strongly suggests that underestimation of derived standard deviations is not uncommon.

Derived standard deviations should be regarded as suspect unless *post facto* tests for validity have been applied. Replicate measurement and analysis of normal probability plots provide a stringent test. Once it has been determined that both crystals are indeed characteristic of the material under study, the two data sets may be combined to produce a final set of p_i and σp_i which are superior to those derived from either individual set of F_{meas}.

Probability plot analysis may also be used in comparing data sets measured with two different instruments but with the same crystal, thus providing information on the validity of the experimental procedures. An extension of this approach is use of two instruments and two crystals, producing four data sets.

The design of an optimum diffractometer experiment must now be modified in light of the above analyses. Since use of two crystals is demonstrated to be highly advantageous it becomes preferable to measure not a full set of F_{meas} on one crystal alone, but instead, a half set on one crystal and the remaining half plus an overlapping set on a second crystal. The overlapping set should randomly sample reciprocal space and should contain at least one hundred F_{meas}.

The application of probability plots to measured and derived quantities and associated standard deviations, although restricted in the present paper to structure factors and structural parameters, is entirely general and may be used for the analysis of error in physical measurement of all kinds.

It is a pleasure to thank Miss I. G. Abrahamson for suggesting the use of normal probability plots to us, and for valuable discussions on the properties of such plots, and Dr Walter C. Hamilton for numerous illuminating comments on this manuscript.

APPENDIX

Synthesis of diagnostic probability plots

Various combinations of the vector pairs S(1), S(2) and **R**(1), **R**(2) of equations (13) and (14) are considered in this Appendix as the basis for the section *Diagnostics for systematic error from essentially linear normal probability plots.*

The magnitude $\Delta S = \langle |\mathbf{S}(1) - \mathbf{S}(2)| \rangle$ and $\Delta R = \langle |\mathbf{R}(1) - \mathbf{R}(2)| \rangle$ represent the modal values of the vector differences. The relative magnitudes of S and R are considered in three main groups. Within each group, the relative magnitudes of ΔS and ΔR are considered for each of the cases (*a*), (*b*) and (*c*) defined in the above section. The arguments in later groups are parallel to those given in detail for group I. 1 below. The numbering corresponds to the labeling on Fig. 6. All symbols represent modal values.

I. *Systematic error large compared with random error*

1. For ΔS very small compared with ΔR, *i.e.* S(1) is essentially equal to S(2), equations (13) and (14) give ΔF approximately equal to ΔR, thus ΔF is essentially independent of the systematic error.

(*a*) σF contains the components of error due to **R** and **S**, hence $\Delta F/\sigma F \simeq \Delta R/\sigma F \ll 1$ and $\mathscr{S}(\delta m)$ is much less than 1. For the parameters obtained by least-squares refinement with the $F(1)$ and $F(2)$ structure factor sets, as ΔS becomes small, Δp approaches Δp_R (where p_R is the parameter value obtained for random error present only). However, σp is much greater than σp_R since the systematic error is large and not taken up by variation of the parameters in the model. Hence $\Delta p/\sigma p$ is much smaller than $\Delta p_R/\sigma p_R$ which is unity, and $\mathscr{S}(\delta p)$ is therefore much less than 1.

(*b*) with $\sigma F = \sigma F_R$, $\Delta F/\sigma F \simeq \Delta R/\sigma F_R \simeq 1$ and hence $\mathscr{S}(\delta m) \simeq 1$. However, $\Delta p/\sigma p$ is small as in case I. 1 (*a*). Hence $\mathscr{S}(\delta p) < 1$.

(*c*) with $\sigma F > \sigma F_S > \sigma F_R$, $\Delta F/\sigma F < \Delta R/\sigma F_R$ or less than 1. Hence $\mathscr{S}(\delta m) < 1$. $\Delta p/\sigma p$ is small as in case I.1. (*a*). Hence $\mathscr{S}(\delta p) < 1$.

2. For $\Delta S \simeq \Delta R$, $\Delta F > \Delta R$. For (*a*), (*b*), and (*c*), $\Delta p > \Delta p_R$, but $\sigma p \gg \sigma p_R$, hence $\mathscr{S}(\delta p) < 1$.

(*a*) $\sigma F \gg \sigma R$ and $\Delta F > \Delta R$, hence $\mathscr{S}(\delta m < 1$
(*b*) $\sigma F = \sigma R$ but $\Delta F > \Delta R$, hence $\mathscr{S}(\delta m) > 1$
(*c*) $\sigma F \gg \sigma R$ and $\Delta F > \Delta R$, hence $\mathscr{S}(\delta m) < 1$

3. For $\Delta S \gg \Delta R$, a linear δm plot cannot be obtained and need not be considered further in this section.

II. *Systematic error comparable with random error*

1. When $\Delta S \ll \Delta R$, $\Delta F \simeq \Delta R$. For (*a*), (*b*) and (*c*) $\Delta p \simeq \Delta p_R$, but $\sigma p > \sigma p_R$, hence $\mathscr{S}(\delta p) < 1$.
(*a*) $\sigma F \simeq \sqrt{2}\sigma R$ and $\Delta F \simeq \Delta R$, hence $\mathscr{S}(\delta m) < 1$
(*b*) $\sigma F \simeq \sigma R$ and $\Delta F \simeq \Delta R$, hence $\mathscr{S}(\delta m) \simeq 1$
(*c*) $\sigma F > \sqrt{2}\sigma R$ and $\Delta F \simeq \Delta R$, hence $\mathscr{S}(\delta m) < 1$.

2. For $\Delta S \simeq \Delta R$, in all three cases, $\Delta p > \Delta p_R$ and $\sigma p > \sigma p_R$, hence $\mathscr{S}(\delta p)$ is indeterminate.

(*a*) $\sigma F > \sigma R$ and $\Delta F > \Delta R$, for correct σF, $\mathscr{S}(\delta m) \simeq 1$
(*b*) $\sigma F \simeq \sigma R$ but $\Delta F > \Delta R$, hence $\mathscr{S}(\delta m) > 1$
(*c*) $\sigma F > \sqrt{2}\sigma R$ and $\Delta F > \Delta R$, hence $\mathscr{S}(\delta m) < 1$.

III. *Systematic error small compared with random error*

1. If $\Delta S \ll \Delta R$, then $\Delta F \simeq \Delta R$. In each case $\Delta p \simeq \Delta p_R$ and $\sigma p \simeq \sigma p_R$, hence $\mathscr{S}(\delta p) \simeq 1$.
(*a*) and (*b*) $\sigma F \simeq \sigma R$ and $\Delta F \simeq \Delta R$, hence $\mathscr{S}(\delta m) \simeq 1$
(*c*) $\sigma F > \sigma R$ and $\Delta F \simeq \Delta R$, hence $\mathscr{S}(\delta m) > 1$.

References

ABRAHAMS, S. C. (1964). *Acta Cryst.* **17**, 1327.

ABRAHAMS, S. C. (1969). *Acta Cryst.* A**25**, 165.

BIRGE, R. T. (1939). *Amer. Phys. Teach.* **7**, 351.

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

BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). ORNL Report TM-305, modified by W. C. HAMILTON and by B. B. CETLIN.

CHAMBERS, J. M. (1968). Private communication.

CHAMBERS, E. & FOWLKES, E. B. (1967). Unpublished Report, *A Dictionary of Distribution*, Bell Telephone Laboratories, Murray Hill, N. J. Available from Dr Fowlkes.

DECHAINE, T. L. (1967). Unpublished Report, *Half-normal, Full-normal and Gamma Labeling*, Bell Telephone Laboratories, Murray Hill, N. J. Available from Dr DeChaine.

GAUSS, C. F. (1809). *Theoria Motus Corporum Coelestium*, Hamburg.

HAMILTON, W. C. & ABRAHAMS, S. C. (1970). *Acta Cryst.* A**26**, 18.

KEVE, E. T., ABRAHAMS, S. C. & BERNSTEIN, J. L. (1970). *J. Chem. Phys.* **53**, 3279.

National Bureau of Standards (1953). *Applied Mathematics Series*, No. 23, U.S. Government Printing Office, Washington, D. C.

WILK, M. B. & GNANADESIKAN, R. (1968). *Biometrika*, **55**, 1.