Two referees have given our paper their careful attention, and we acknowledge their constructive and helpful comments. One of us (AKL) also thanks St John's College, Cambridge, and the Cavendish Laboratory Trust Fund for support.

References

ABLES, J. G. (1974). *Astron. Astrophys. Suppl. Ser.* **15**, 383-393.
BRICOGNE, G. (1983). Paper presented at Medical Foundation of Buffalo (NY) School on Direct Methods and Macromolecular Crystallography, and private communication.
BRICOGNE, G. (1984*a*). *Acta Cryst.* A**40**, 410-445.
BRICOGNE, G. (1984*b*). *Methods and Applications in Crystallographic Computing*, edited by T. ASHIDA & S. HALL. Oxford Univ. Press. in the press.
BRITTEN, P. L. & COLLINS, D. M. (1982). *Acta Cryst.* A**38**, 129-132.
BRYAN, R. K. (1980). PhD thesis. Univ. of Cambridge, England.
BRYAN, R. K., BANSAL, M., FOLKHARD, W., NAVE, C. & MARVIN, D. A. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 4728-4731.
BURCH, S. F., GULL, S. F. & SKILLING, J. (1983). *Comput. Graph. Image Process.* **23**, 113-128.
BURG, J. P. (1967). Paper presented at 37th Meeting of the Society of Exploration Geophysicists, Oklahoma City.
BURG, J. P. (1972). *Geophysics*, **38**, 375-376.
COLLINS, D. M. (1982). *Nature (London)*, **298**, 49-51.
COLLINS, D. M. & MAHAR, M. C. (1983). *Acta Cryst.* A**39**, 777-783.
DANIELL, G. J. & GULL, S. F. (1980). *Proc. Inst. Electr. Eng.* **127**E, 170-172.
DIAMOND, R. (1963). *Acta Cryst.* **16**, 627-639.
FRIEDEN, B. R. (1972). *J. Opt. Soc. Am.* **62**, 511-518.
FRIEDEN, B. R. (1983). *J. Opt. Soc. Am.* **73**, 927-938.
GERCHBERG, R. W. (1974). *Opt. Acta*, **21**, 709-720.
GULL, S. F. (1983). Paper presented at the 3rd Workshop on Maximum Entropy, Univ. of Wyoming.
GULL, S. F. & DANIELL, G. J. (1978). *Nature (London)*, **272**, 686-690.
HARKER, D. & KASPER, J. S. (1948). *Acta Cryst.* **1**, 70-75.
HOPPE, W. & GASSMANN, J. (1968). *Acta Cryst.* B**24**, 97-107.
HOSOYA, S. & TOKONAMI, M. (1976). *Acta Cryst.* **23**, 18-25.
JAYNES, E. T. (1957*a*). *Phys. Rev.* **106**, 620-630.
JAYNES, E. T. (1957*b*). *Phys. Rev.* **108**, 171-190.
JAYNES, E. T. (1968). *IEEE Trans. Syst. Sci. Cybern.* SSC-**4**, 227-241.
JAYNES, E. T. (1983). (Collected Works.) *Papers on Probability, Statistics, and Statistical Physics*, edited by R. D. ROSENKRANTZ. Dordrecht, Holland: Reidel.
KARLE, J. & HAUPTMAN, H. (1950). *Acta Cryst.* **3**, 181-187.
KEMP, M. C. (1980). *Med. Radionuclide Imaging (IAEA, Vienna)*, **1**, 313-323.
KIKUCHI, R. & SOFFER, B. H. (1977). *J. Opt. Soc. Am.* **67**, 1656-1665.
LIVESEY, A. K. (1984). PhD thesis, Univ. of Cambridge, England. Submitted.
NARAYAN, R. & NITYANANDA, R. (1982). *Acta Cryst.* A**38**, 122-128.
NAVAZA, J., CASTELLANO, E. E. & TSOUCARIS, G. (1983). *Acta Cryst.* A**39**, 622-630.
NITYANANDA, R. & NARAYAN, R. (1982). *J. Astrophys. Astron.* **3**, 419-450.
PAPOULIS, A. (1975). *IEEE Trans. Circuits Syst.* CAS-**22**, 735-742.
PIRO, O. E. (1983). *Acta Cryst.* A**39**, 61-67.
RANGO, C. DE, TSOUCARIS, G. & ZELWER, C. (1974). *Acta Cryst.* A**30**, 342-353.
SAXTON, W. O. (1980). In *Computer Processing of Electron Microscope Images*, edited by P. W. HAWKES. *Topics in Current Physics* 13, pp. 35-87. Berlin: Springer-Verlag.
SCOTT, P. F. (1981). *Mon. Notices R. Astron. Soc.* **194**, 23P-29P.
SHANNON, C. E. (1948). *Bell Syst. Tech. J.* **27**, 379-423, 623-656.
SHORE, J. E. & JOHNSON, R. W. (1980). *IEEE Trans. Inf. Theory*, IT-**26**, 26-37; correction: *IEEE Trans.* (1983). IT-**29**, 942-943.
SIBISI, S. (1983). *Nature (London)*, **301**, 134-136.
SIBISI, S., SKILLING, J., BRERETON, R. G., LAUE, E. D. & STAUNTON, J. (1984). *Nature (London)*, **311**, 446-447.
SKILLING, J. (1981). Paper presented at 1st Workshop on Maximum Entropy Estimation and Data Analysis, Univ. of Wyoming.
SKILLING, J. (1983). Paper presented at Opt. Soc. Am. meeting on Signal REcovery and Synthesis with Incomplete Information and Partial Constraints, Incline Village, Nevada.
SKILLING, J. & GULL, S. F. (1984). *SIAM (Soc. Ind. Appl. Math.) Proc. Am. Math. Soc.* **14**, 167-189.
SKILLING, J., STRONG, A. W. & BENNETT, K. (1979). *Mon. Notices R. Astron. Soc.* **187**, 145-152.
STEENSTRUP, S. (1984). *Acta Cryst.* A**40**, 66-67.
TIKHONOV, A. & ARSENIN, V. (1977). *Solutions of Ill-posed Problems.* London: Wiley.
TSOUCARIS, G. (1970). *Acta Cryst.* A**26**, 492-499.
ULRYCH, T. J. & BISHOP, T. N. (1975). *Rev. Geophys. Space Phys.* **13**, 183-200.
WILKINS, S. W. (1983). *Acta Cryst.* A**39**, 892-895, 896-898.
WILKINS, S. W., VARGHESE, J. N. & LEHMANN, M. S. (1983). *Acta Cryst.* A**39**, 47-60; erratum: *Acta Cryst.* (1983), A**39**, 594.

Acta Cryst. (1985). A**41**, 122-128

Round-off Errors in Inter-experimental Comparisons

By Robin Taylor and Olga Kennard

Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England

(*Received 6 April* 1984; *accepted* 17 *September* 1984)

Abstract

Two independent determinations of the same structure may be compared by means of statistical techniques such as normal probability plots and χ^2 hypothesis tests. Computer simulations show that errors may arise in the application of these techniques if rounded estimates of structural parameters and their e.s.d.s are used in the calculations. Round-off errors are particularly serious in goodness-of-fit hypothesis tests, since they increase the probability of making type I errors, *i.e.* falsely rejecting the null hypothesis.

0108-7673/85/020122-07$01.50

I. Introduction

It sometimes happens that the same structure is determined independently by two different research groups. In this situation, the two sets of results may be compared by means of statistical techniques such as normal or half-normal probability plots (Abrahams & Keve, 1971; Hamilton & Abrahams, 1972) and χ^2 hypothesis tests (Hamilton, 1969). These techniques are used to examine the distribution of the quantities δ_i, defined as

$$\delta_i = (p_{i,1} - p_{i,2})/(\sigma_{i,1}^2 + \sigma_{i,2}^2)^{1/2}, \quad (1)$$

where $p_{i,j}$ is the value of the ith parameter (*e.g.* atomic coordinate, bond length, temperature factor) in the jth structure determination and $\sigma_{i,j}$ is its e.s.d. The purpose of this paper is to discuss some difficulties that may arise when the δ_i are calculated from rounded estimates of $p_{i,j}$ and $\sigma_{i,j}$. This will be the case if the inter-experimental comparison is based on published data, since atomic coordinates and their e.s.d.s are usually reported to only four or five decimal places.

Intuitively, we may expect that round-off errors in the numerator and denominator of (1) will tend to increase the standard deviation and/or kurtosis of the δ_i distribution. In fact, the effects of rounding are more subtle, as may be seen from a simple example. Suppose that two determinations of the same structure are published, the fractional atomic coordinates in each being reported to four decimal places. Suppose, further, that the atomic coordinate e.s.d.s are all quoted as 0·0003 in the first paper and 0·0004 in the second. If δ_i values are calculated from (1) using these rounded quantities, the denominator will always be $(0{\cdot}0003^2 + 0{\cdot}0004^2)^{1/2}$ *i.e.* 0·0005. The numerator can only take values 0·0, ±0·0001, ±0·0002, *etc.* Consequently, the calculated values of δ_i can only be 0·0, ±0·2, ±0·4, ..., *i.e.* the distribution will appear to be discontinuous. Application of normal probability plots or χ^2 hypothesis tests may then lead to incorrect conclusions regarding the normality of the error distribution. In particular, the δ_i distribution may appear to be significantly non-normal, even if the underlying distribution of experimental errors is normal.

In this paper, we use the method of computer simulation to investigate the effects of rounding on (*a*) the shape and standard deviation of the δ_i distribution, (*b*) the appearance of normal probability plots, and (*c*) the validity of hypothesis tests. We begin by defining some mathematical symbols and describing the general procedure used in the computer simulations.

II. Definition of mathematical symbols

Throughout the paper, the symbol (e) indicates an exact quantity or a statistic calculated from exact quantities (strictly speaking, 'exact' in this context means 'as exact as can be accommodated in IBM double-precision arithmetic, *i.e.* in a 56-bit mantissa word'). Conversely, (r) denotes a quantity rounded to four decimal places, or a statistic calculated from such rounded quantities. Thus, $p(e)_{i,j}$ is the exact value of the ith parameter in the jth structure determination and $\sigma(e)_{i,j}$ is the exact value of its e.s.d.; $p(r)_{i,j}$ and $\sigma(r)_{i,j}$ are the corresponding rounded values. The weighted deviations $\delta(e)_i$ and $\delta(r)_i$ are then calculated from the formulae

$$\delta(e)_i = [p(e)_{i,1} - p(e)_{i,2}]/[\sigma^2(e)_{i,1} + \sigma^2(e)_{i,2}]^{1/2} \quad (2)$$

$$\delta(r)_i = [p(r)_{i,1} - p(r)_{i,2}]/[\sigma^2(r)_{i,1} + \sigma^2(r)_{i,2}]^{1/2}. \quad (3)$$

III. Generation of δ_i values by simulation

The following procedure was used to generate artificial simulated values of $\delta(e)_i$ and $\delta(r)_i$. E.s.d.s $\sigma(e)_{i,1}$ and $\sigma(e)_{i,2}$ were selected at random from a uniform distribution in the range $\sigma_{\min} - \sigma_{\max}$ (typical values would be $\sigma_{\min} = 0{\cdot}00005$, $\sigma_{\max} = 0{\cdot}00025$). The parameters $p(e)_{i,j}$ ($j = 1, 2$) were drawn at random from normal distributions with mean = 1, standard deviation = $\sigma(e)_{i,j}$. The simulated value of $\delta(e)_i$ was then calculated from (2). Values of $p(r)_{i,j}$ and $\sigma(r)_{i,j}$ were obtained by rounding the corresponding exact quantities to four decimal places, and $\delta(r)_i$ calculated from (3). All simulations were performed with the aid of pseudo-random number generators written by the Numerical Algorithms Group (*NAG Fortran Library Manual*, 1983).

IV. Effect of round-off errors on δ_i distribution

In our first simulation we generated 10 000 values of $\delta(e)_i$ and $\delta(r)_i$, using the simulation parameters $\sigma_{\min} = 0{\cdot}00005$ and $\sigma_{\max} = 0{\cdot}00015$ [thus, the only possible value of $\sigma(r)_{i,j}$ was 0·0001]. The resulting distributions of $\delta(e)_i$ and $\delta(r)_i$ are shown in Figs. 1(*a*)–(*b*) and the sample means and variances are given in the first line of Table 1. Kolmogorov–Smirnov and χ^2 goodness-of-fit tests (Siegel, 1956) indicate that the $\delta(e)_i$ distribution is not significantly different from a normal distribution with zero mean and unit variance. This confirms that the pseudo-random number generators used in our simulation program are satisfactory. The $\delta(r)_i$ distribution consists of only 13 discrete values and therefore shows pronounced departures from normality (*e.g.* there are no observations in the range $0{\cdot}0 < \delta(r)_i < 0{\cdot}7$, whereas about 2580 would be expected if the distribution was normal). However, the variance of the $\delta(r)_i$ values (= 1·166) is quite close to that of the underlying δ_i population (= 1). Several other simulations were performed with different values of $\sigma_{\min}$ and $\sigma_{\max}$. Results are summarized in Table 1 and some of the $\delta(r)_i$

Table 1. *Details of simulated $\delta(e)_i$ and $\delta(r)_i$ distributions*

Simulation parameters		Possible values of	$\delta(e)_i$ distribution		$\delta(r)_i$ distribution		Number of discrete
$\sigma_{\min}$	$\sigma_{\max}$	$\sigma(r)_{ij}(\times 10^4)$	Mean	Variance	Mean	Variance	values in $\delta(r)_i$ distribution
0·00005	0·00015	1	−0·006	1·002	−0·010	1·166	13
0·00005	0·00025	1, 2	−0·009	1·009	−0·010	1·108	42
0·00005	0·00035	1, 2, 3	0·010	0·978	0·012	1·024	102
0·00005	0·00045	1, 2, 3, 4	0·018	0·982	0·015	1·019	185
0·00005	0·00055	1, 2, 3, 4, 5	−0·005	0·992	−0·006	1·008	343
0·00005	0·00105	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	0·007	1·005	0·006	1·013	1536
0·00015	0·00025	2	−0·010	1·011	−0·011	1·047	25
0·00025	0·00055	3, 4, 5	−0·005	0·996	−0·006	1·004	202
0·00045	0·00055	5	0·011	1·020	0·012	1·026	54
0·00085	0·00105	9, 10	0·007	1·007	0·007	1·010	246

distributions are shown in Figs. 1(*c*)–(*g*). The $\delta(e)_i$ distributions are not shown as they were all similar to Fig. 1(*a*). As would be expected, the worst effects of rounding are confined to simulations in which $(\sigma_{\min}+\sigma_{\max})/2$ is small. Presumably, this is because errors due to rounding become relatively less important as $\sigma(r)_{ij}$ increases along the series 0·0001, 0·0002, 0·0003, ...

Table 1 shows that the sample variances of the $\delta(r)_i$ distributions are all quite close to the ideal value of unity. This is an encouraging result, because estimation of the standard deviation of the δ_i is often a major objective of inter-experimental comparisons. The variances of $\delta(r)_i$ distributions were further investigated as follows. A sample of 30 simulated $\delta(e)_i$ and $\delta(r)_i$ values was generated (simulation parameters: $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00025$) and the $\delta(e)_i$ and $\delta(r)_i$ sample variances calculated. After 2500 repetitions of this procedure, the distributions of the sample variances were as shown in Figs. 2(*a*)–(*b*). The mean values of the distributions are 1·002 for $\sigma^2[\delta(e)_i]$ and 1·084 for $\sigma^2[\delta(r)_i]$; the values of $\sigma^2[\delta(r)_i]$ show slightly more variation than those of $\sigma^2[\delta(e)_i]$. We conclude that estimates of δ_i variances based on rounded parameters tend to be too large, and less precise than those based on exact parameters. However, they are likely to be adequate for most purposes.

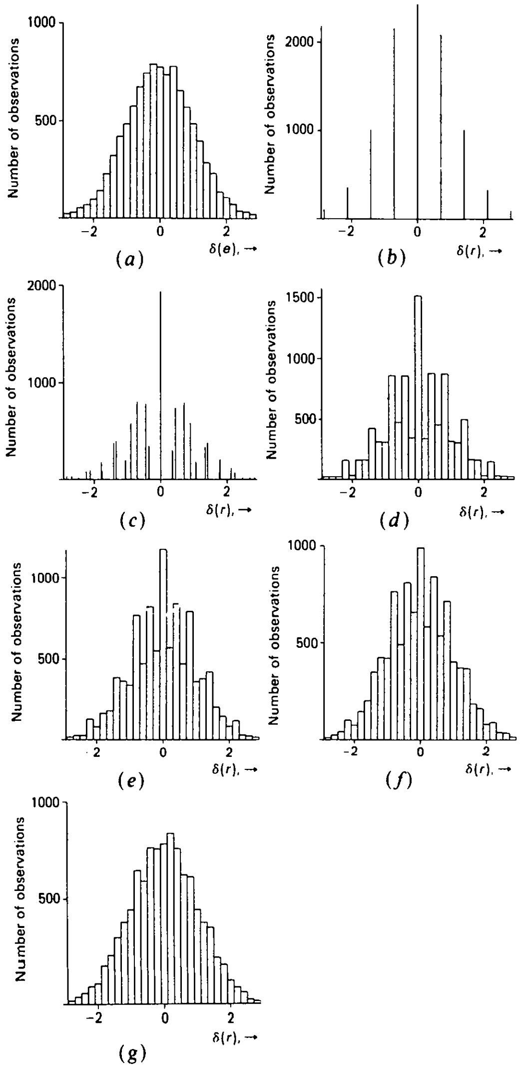

Fig. 1. Simulated distributions of: (*a*) $\delta(e)_i$, $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00015$; (*b*) $\delta(r)_i$, $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00015$; (*c*) $\delta(r)_i$, $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00025$; (*d*) $\delta(r)_i$, $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00035$; (*e*) $\delta(r)_i$, $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00045$; (*f*) $\delta(r)_i$, $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00055$; (*g*) $\delta(r)_i$, $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00105$. All distributions are shown as conventional histograms except for (*b*) and (*c*), where each discrete value in the distribution is represented by a separate bar. All distributions are shown in the range $-2{\cdot}9<\delta_i<2{\cdot}9$; few observations fall outside this range.

V. Effect of round-off errors on normal probability plots

In normal probability plotting, the n observed δ_i values are ranked in order of increasing magnitude and plotted against the ranked δ_i values expected for a sample of size n from a normal distribution with zero mean and unit variance. A straight-line plot suggests that the observed δ_i are normally distributed (Abrahams & Keve, 1971). Fig. 3(*a*) shows a normal probability plot based on a simulated sample of 30 $\delta(r)_i$ values (simulation parameters: $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00015$). For comparison the corresponding $\delta(e)_i$ plot is given in Fig. 3(*b*). The $\delta(r)_i$ plot has a staircase-like appearance with sudden steps upwards corresponding to discontinuities in the $\delta(r)_i$ distribution. Round-off errors therefore have a profound effect on this normal probability plot. However, further simulations suggest that the effects of rounding diminish rapidly as the parameters σ_{min} and σ_{max} are increased. For example, Figs. 3(*c*)–(*f*) show $\delta(r)_i$ and $\delta(e)_i$ plots produced by two simulations with parameters $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00025$ and $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00035$. The departures from linearity due to rounding are relatively small compared with the background scatter due to random sampling effects.

VI. Effect of round-off errors on hypothesis tests

Normal probability plotting is an excellent technique for obtaining an overview of the δ_i distribution. However, for more quantitative work it may be necessary to use statistical hypothesis tests. Such tests will generally focus on one or both of the following questions. (*a*) Is the δ_i distribution normal? (*b*) Does it have unit variance? A variety of tests may be used and we consider here the effects of rounding on three of them.

VI-1. *Hamilton's* χ^2 *test*

If experimental errors are normally distributed and the $\sigma_{i,j}$ are accurate estimates of the standard deviations of these errors, then each δ_i will be a random variable from a normal distribution with zero mean and unit variance. Consequently, the sum of squares of any n δ_i values will follow a χ^2 distribution with n degrees of freedom. The statistic

$$Q(e) = \sum_{i=1}^{n} \delta(e)_i^2 \qquad (4)$$

may therefore be used as the basis of a hypothesis test (Hamilton, 1969). If $Q(e)$ exceeds the tabulated value of $\chi^2_{n,\alpha}$ we may reject the null hypothesis that the δ_i are normally distributed with zero mean and unit variance. The test will be at the $100(1-\alpha)\%$ confidence level, *i.e.* the probability of falsely rejecting the null hypothesis (a so-called 'type I' error; Hamilton, 1964) will be α. If, however, the Q statistic is calculated from rounded quantities

$$Q(r) = \sum_{i=1}^{n} \delta(r)_i^2, \qquad (5)$$

we may anticipate that the probability of making a type I error will be inflated, *i.e.* a test at the formal confidence level of $100(1-\alpha)\%$ will really be at some lower level $100(1-\alpha')\%$, where $\alpha' > \alpha$.

The effect of round-off errors on the distribution of Q was investigated as follows. A sample of 40 simulated $\delta(e)_i$ and $\delta(r)_i$ values was generated using the simulation parameters $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00015$. The statistics $Q(e)$ and $Q(r)$ were calculated and tested against the tabulated value of $\chi^2_{40,0{\cdot}05} (= 55{\cdot}76)$. This procedure was repeated 2500 times and a count made of the number of significant values of $Q(e)$ and $Q(r)$ obtained. Results are summarized in

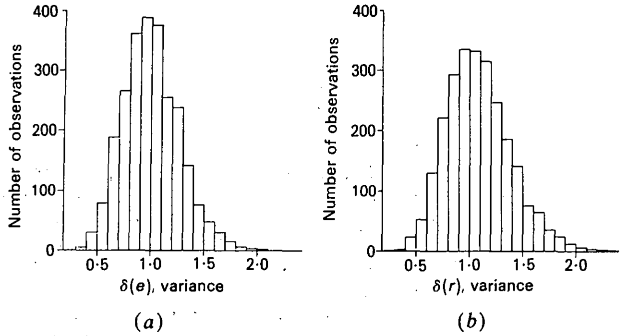

Fig. 2. Distribution of sample variances of (*a*) $\delta(e)_i$ and (*b*) $\delta(r)_i$ samples.

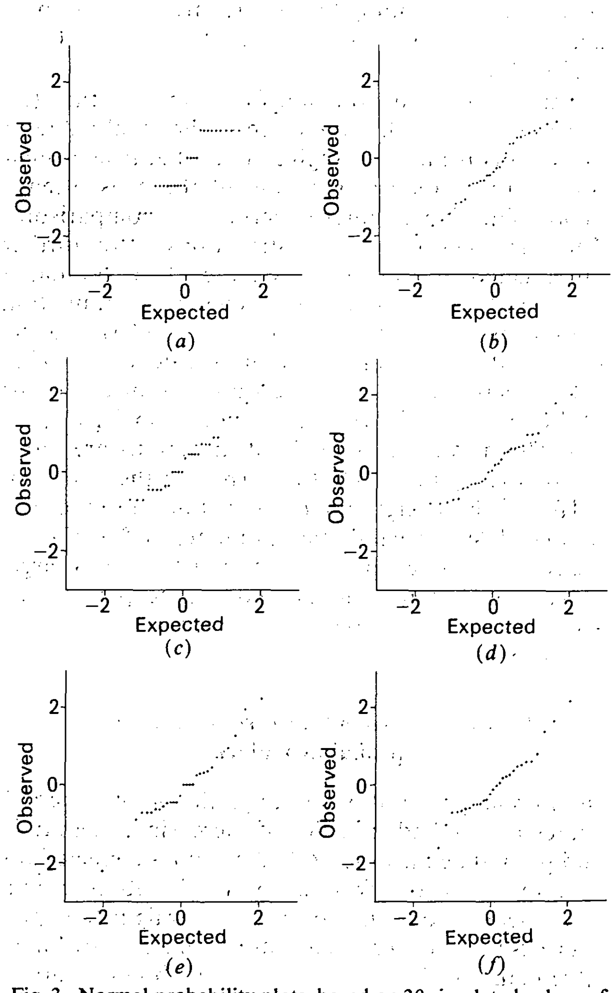

Fig. 3. Normal probability plots, based on 30 simulated values of: (*a*) $\delta(r)_i$, $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00015$; (*b*) $\delta(e)_i$, $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00015$; (*c*) $\delta(r)_i$, $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00025$; (*d*) $\delta(e)_i$, $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00025$; (*e*) $\delta(r)_i$, $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00035$; (*f*) $\delta(e)_i$, $\sigma_{min} = 0{\cdot}00005$, $\sigma_{max} = 0{\cdot}00035$.

the first line of Table 2. They show that 5·5% of the $Q(e)$ values were significant - close to the theoretical proportion of 5% - but 19·8% of the $Q(r)$ values. Thus, the true confidence level of the test based on rounded parameters was about 80%, appreciably smaller than the formal confidence level of 95%. The simulation was repeated with various values of σ_{min} and σ_{max}, and the results are summarized in Table 2. They show that round-off errors can seriously jeopardize the validity of the test when $(\sigma_{min}+\sigma_{max})/2$ is small. Slightly more of the $Q(e)$ values were significant that would have been expected. This is probably due to minor defects in the pseudo-random number generators used in our simulation program, or to residual round-off errors caused by the finite word length of the computer.

Table 2. *Effect of round-off errors on Hamilton's* χ^2 *test*

Simulation parameters		Number of significant results obtained		Percentage of significant results obtained	
σ_{min}	σ_{max}	Exact	Rounded	Exact	Rounded
0·00005	0·00015	138	496	5·5	19·8
0·00005	0·00025	134	301	5·4	12·0
0·00005	0·00035	128	228	5·1	9·1
0·00005	0·00045	130	180	5·2	7·2
0·00005	0·00055	143	196	5·7	7·8
0·00005	0·00105	134	156	5·4	6·2
0·00015	0·00025	129	188	5·2	7·5
0·00025	0·00055	138	147	5·5	5·9
0·00045	0·00055	136	145	5·4	5·8
0·00085	0·00105	126	134	5·0	5·4

VI-2. χ^2 *Goodness-of-fit test*

The above procedure tests the combined null hypothesis that the δ_i distribution is (a) normal, with (b) zero mean and (c) unit variance. If any part of the hypothesis is false, the test is likely to give a significant result. However, it may be desired to test one or two parts of the hypothesis separately, *e.g.* that the δ_i distribution is normal, with zero mean and unknown standard deviation σ_δ. One way of doing this is to perform a χ^2 goodness-of-fit test (Siegel, 1956). The sample standard deviation of the δ_i is used to estimate σ_δ. The test is then based on the statistic

$$C=\sum_{k=1}^{N}(f_k-F_k)^2/F_k, \qquad (6)$$

where f_k is the observed number of δ_i values falling in a particular range and F_k is the number expected to fall in that range if the δ_i distribution is normal with zero mean and standard deviation σ_δ. The summation is over N ranges, which may be chosen more or less arbitrarily [but see Siegel (1956) for a discussion of this point]. If the test is based on exact $\delta(e)_i$ values - *i.e.* the statistic $C(e)$ is calculated - the null hypothesis that the δ_i distribution is normal with zero mean can be rejected if $C(e)$ exceeds the tabulated value of $\chi^2_{N-2,\alpha}$. The probability of making a type I error will be approximately α (the test is not exact). However, if $\delta(r)_i$ values are used - *i.e.* $C(r)$ is calculated - we may again expect that the true probability of making a type I error will be greater than α.

The effects of rounding on the χ^2 goodness-of-fit test were investigated as follows. A sample of 40 simulated $\delta(e)_i$ and $\delta(r)_i$ values was generated, using the simulation parameters $\sigma_{min}=0{\cdot}00005$, $\sigma_{max}=0{\cdot}00015$. The sample standard deviation of the $\delta(e)_i$ was used to estimate σ_δ, and a count was made of the number of $\delta(e)_i$ values falling in each of the following ranges: $\delta(e)_i<-1{\cdot}0\sigma_\delta$; $-1{\cdot}0\sigma_\delta\le\delta(e)_i<-0{\cdot}4\sigma_\delta$; $-0{\cdot}4\sigma_\delta\le\delta(e)_i<0$; $0\le\delta(e)_i<0{\cdot}4\sigma_\delta$; $0{\cdot}4\sigma_\delta\le\delta(e)_i<1{\cdot}0\sigma_\delta$; $1{\cdot}0\sigma_\delta\le\delta(e)_i$. The number of observations expected to fall in each of these ranges was calculated from the standard normal curve [*e.g.* the number expected in the third range was calculated as $40(0{\cdot}5000-0{\cdot}3446)=6{\cdot}22$, 0·5000 being the area under the standard normal curve to the left of $z=0$ and 0·3446 being the area to the left of $z=-0{\cdot}4$]. The statistic $C(e)$ was thus calculated, and tested against the tabulated value of $\chi^2_{4,0\cdot05}(=9{\cdot}49)$. The value of $C(r)$ was determined in an exactly analogous fashion and similarly tested against $\chi^2_{4,0\cdot05}$. The complete procedure was repeated 2500 times and a count made of the number of significant values of $C(e)$ and $C(r)$ obtained. Results are summarized in Table 3, together with those obtained from similar simulations with different values of σ_{min} and σ_{max}.

Table 3. *Effect of round-off errors on* χ^2 *goodness-of-fit test*

Simulation parameters		Number of significant results obtained		Percentage of significant results obtained	
σ_{min}	σ_{max}	Exact	Rounded	Exact	Rounded
0·00005	0·00015	152	2260	6·1	90·4
0·00005	0·00025	153	1322	6·1	52·9
0·00005	0·00035	144	720	5·8	28·8
0·00005	0·00045	132	443	5·3	17·7
0·00005	0·00055	141	348	5·6	13·9
0·00005	0·00105	157	217	6·3	8·7
0·00015	0·00025	143	1197	5·7	47·9
0·00025	0·00055	166	286	6·6	11·4
0·00045	0·00055	149	254	6·0	10·2
0·00085	0·00105	150	180	6·0	7·2

Table 3 shows that the proportion of significant $C(e)$ values was slightly greater than the theoretical proportion (*i.e.* 5%) in all of the simulations performed. This probably reflects the approximate nature of the test and the minor defects in the simulation procedure commented upon above (see VI-1). The number of $C(r)$ values exceeding $\chi^2_{4,0\cdot05}$ was far greater than 5% in many of the simulations. For example, when $\sigma_{min}=0{\cdot}00005$ and $\chi_{max}=0{\cdot}00035$, some 28·8% of the $C(r)$ values were *formally* significant at the 95% confidence level, *i.e.* the *true* confidence level of the test was about 71%. The poor performance of the $C(r)$ statistic in this series of simulations is partly due to an unfortunate choice of ranges. Many of the $\delta(r)_i$ values will be exactly equal to zero [corresponding to $p(r)_{i,1}=p(r)_{i,2}$] and this inflates the number of observations in the range $0\le\delta(r)_i<0{\cdot}4\sigma_\delta$ relative to the number in the adjoining range $-0{\cdot}4\sigma_\delta\le\delta(r)_i<0$. This is particularly serious in the simulation with $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00015$, since the two lowest possible values of $|\delta(r)_i|$ are 0 and 0·7071 [$=0{\cdot}0001(0{\cdot}0001^2+0{\cdot}0001^2)^{-1/2}$]. Thus, the number of $\delta(r)_i$ values falling in the range $-0{\cdot}4\sigma_\delta\le\delta(r)_i<0$ is *always* equal to zero. The 'expected' number, calculated from the standard normal probability curve, is 6·22 (see preceding paragraph). Thus, this range *on its own* contributes 6·22 towards the value of $C(r)$, accounting for the disastrous effect that round-off errors have in this simulation (see Table 3). Some improvement can therefore be effected by altering the ranges used in calculating $C(r)$. For example, the simulation was repeated using the ranges: $\delta(r)_i<-0{\cdot}8\sigma_\delta$; $-0{\cdot}8\sigma_\delta\le\delta(r)_i<-0{\cdot}2\sigma_\delta$; $-0{\cdot}2\sigma_\delta\le\delta(r)_i<0{\cdot}2\sigma_\delta$; $0{\cdot}2\sigma_\delta\le\delta(r)_i<0{\cdot}8\sigma_\delta$; $0{\cdot}8\sigma_\delta\le\delta(r)_i$. The percentage of $C(r)$ values formally significant at the 95% confidence level was thereby reduced to 49·8%.

The χ^2 goodness-of-fit test can be used successfully with rounded parameters if the expected frequencies [*i.e.* the F_k of (6)] are determined by simulation. For example, Fig. 1(*b*) shows the $\delta(r)_i$ distribution that would be expected if all of the $\sigma(r)_{i,j}$ were equal to 0·0001 (corresponding to $\sigma_{\min}=0{\cdot}00005$, $\sigma_{\max}=0{\cdot}00015$) and the underlying δ_i distribution was normal with zero mean and unit variance (see § IV). Some 2739 of the 10 000 observations in Fig. 1(*b*) lie between −0·2 and +0·2, so in a random sample of 40 $\delta(r)_i$ values we would expect 10·96 (=2739 × 40/10 000) observations in this range. By way of comparison, the expected frequency calculated from the standard normal probability curve is 6·34. In a simulation to test the performance of the $C(r)$ statistic when the expected frequencies are estimated from Fig. 1(*b*), we found that 4·5% of the $C(r)$ values were formally significant at the 95% confidence level (*c.f.* the theoretical value of 5%).

VI-3. *Kolmogorov–Smirnov goodness-of-fit test*

The hypothesis that the δ_i are normally distributed with zero mean and unknown standard deviation, σ_δ, may be tested by another type of goodness-of-fit test, the Kolmogorov–Smirnov test (Siegel, 1956). As before, the sample standard deviation of the observed δ_i is used to estimate σ_δ. The test is then based on the statistic

$$D=\text{maximum}\,|P(\delta)-p(\delta)|. \quad (7)$$

For any value δ, $p(\delta)$ is the observed proportion of δ_i values that are less than or equal to δ, *i.e.* $p(\delta)$ is the observed cumulative probability distribution of the δ_i. $P(\delta)$ is the theoretical cumulative probability distribution for a normal population with zero mean and standard deviation σ_δ. Critical values of D for various sample sizes are tabulated in statistical texts (*e.g.* Siegel, 1956) but they are slightly conservative if, as here, the variance of the theoretical population is estimated from the sample (Massey, 1951).

Table 4. *Effect of round-off errors on Kolmogorov–Smirnov test*

Simulation parameters		Number of significant results obtained		Percentage of significant results obtained	
$\sigma_{\min}$	$\sigma_{\max}$	Exact	Rounded	Exact	Rounded
0·00005	0·00015	93	962	3·7	38·5
0·00005	0·00025	87	362	3·5	14·5
0·00005	0·00035	85	228	3·4	9·1
0·00005	0·00045	88	150	3·5	6·0
0·00005	0·00055	69	100	2·8	4·0
0·00005	0·00105	99	89	4·0	3·6
0·00015	0·00025	109	317	4·4	12·7
0·00025	0·00055	103	129	4·1	5·2
0·00045	0·00055	113	147	4·5	5·9
0·00085	0·00105	93	94	3·7	3·8

The effect of round-off errors on the Kolmogorov–Smirnov test was investigated by simulations analogous to those described above, *i.e.* samples of 40 $\delta(e)_i$ and $\delta(r)_i$ values were generated, statistics $D(e)$ and $D(r)$ were calculated using the exact and rounded values, respectively, and both were tested for significance against the tabulated value of $D_{40,0{\cdot}05}$ (=0·215). Results are summarized in Table 4. The percentage of significant $D(e)$ values is less than the ideal 5% in all of the simulations, confirming that the test is slightly conservative. The percentage of significant $D(r)$ values tends to be too large, particularly when $(\sigma_{\min}+\sigma_{\max})/2$ is small. However, the true confidence levels of the tests based on $D(r)$ are much closer to the formal confidence level of 95% than was the case for the χ^2 goodness-of-fit tests (see Table 3).

VII. Conclusions

Inter-experimental comparisons based on published data necessarily use rounded estimates of atomic coordinates, temperature factors, e.s.d.s, *etc.* By far the most important consequence is that the distribution of the weighted parameter differences [*i.e.* the δ_i of (1)] appears to be discontinuous. This does not seriously impair the accuracy and precision with which the standard deviation of the δ_i can be estimated. Nor, in general, does it have an appreciable effect on the appearance of normal probability plots. However, round-off errors can seriously jeopardize the validity of various hypothesis tests by increasing the probability of falsely rejecting the null hypothesis. The severity of the problem depends on the particular type of hypothesis test used and the method of computation employed. In unfavourable circumstances the validity of the test may be totally nullified.

In reporting the results of crystal-structure analyses it is conventional to follow each atomic coordinate and temperature factor by its e.s.d. Usually this is given in brackets in the units of the least significant digit of the coordinate or temperature factor. Our results suggest that serious round-off errors in inter-experimental comparisons only occur when this number in brackets is small, *i.e.* 1, 2 and, possibly, 3. Thus, the problems discussed in this paper can be overcome if an extra digit of significance is reported in these cases.

Olga Kennard is a member of the external staff of the Medical Research Council.

References

ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* **A27**, 157-165.
HAMILTON, W. C. (1964). *Statistics in Physical Science*, pp. 46-47. New York: Ronald Press.
HAMILTON, W. C. (1969). *Acta Cryst.* **A25**, 194-206.
HAMILTON, W. C. & ABRAHAMS, S. C. (1972). *Acta Cryst.* **A28**, 215-218.
MASSEY, F. J. JR (1951). *J. Am. Stat. Assoc.* **46**, 68-78.
NAG Fortran Library Manual (1983). Mark 10, Vol. 5.
SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* Tokyo: McGraw-Hill Kogakusha.

Acta Cryst. (1985). **A41**, 128-133

Simulation of Multiple Diffraction

BY Y. SOEJIMA AND A. OKAZAKI

Department of Physics, Kyushu University, Fukuoka 812, *Japan*

AND T. MATSUMOTO

Department of Earth Sciences, Kanazawa University, Kanazawa 920, *Japan*

(*Received* 27 *April* 1984; *accepted* 1 *August* 1984)

Abstract

The multiple diffraction of X-rays and neutrons is discussed on the basis of the kinematical theory; a program for the simulation of ψ scanning and λ scanning is developed, where the influence of the wavelength width of incident beams on the Ewald construction is properly taken into account. The effect of higher-order diffraction (n-beam interaction, $n > 3$) is treated as the sum of those of $(n-2)$ pairs of relevant double diffractions (three-beam interactions). Applications are made for some examples for which experimental data are available; it is shown that the results are in very good agreement with experiment. This suggests that the kinematical approach is appropriate. The simulation is useful in planning ψ-scanning experiments for precise structure determination and for examining experimental data.

1. Introduction

Since Renninger (1937) showed the phenomenon of double diffraction (*Umweganregung*) in the pattern of ψ scanning of 222 of diamond, the importance of the effect on structure determinations has often been discussed. In the early days, discussions were mainly concerned with possible errors in space-group determination because, as mentioned by Lipson & Cochran (1953), the effect can interfere with the detection of glide planes and screw axes. Later, in connection with precise structure determination, the effect has been considered more generally (Coppens, 1968; Panke & Wölfel, 1968): its effect on general reflections has been considered, to improve the accuracy of the intensity.

The ψ-scanning experiment, on the other hand, is not very easy even at present, particularly with specimens in special environments such as low or high temperature; it also requires a long machine time. If a computer simulation is available for double (*i.e.* multiple) diffraction, it will therefore be very useful in planning the ψ-scanning experiment. Moreover, in some cases, the experiment will be replaced by such a simulation, in part at least. Examination of the results of the ψ-scanning data by comparison with the simulation will also be worth doing.

In the following, a computer simulation based on the kinematical theory will be developed for ψ scanning with monochromatic incident beams and for λ scanning with white beams. The results will be applied to some examples and compared with experiments.

0108-7673/85/020128-06$01.50