

Table 4. *Cesium-oxygen distances* (Å)

Cs—O(1)	3.226 (2)	Cs—O(4 <sup>iv</sup> )	3.012 (2)
—O(2 <sup>i</sup> )	3.008 (2)	—O(5 <sup>v</sup> )	3.113 (2)
—O(2 <sup>ii</sup> )	3.019 (2)	—O(5 <sup>vi</sup> )	3.116 (2)
—O(3 <sup>i</sup> )	3.321 (2)	—O(6 <sup>iii</sup> )	3.684 (2)
—O(3 <sup>iii</sup> )	3.234 (2)		

## Symmetry code

(i)	$-\frac{1}{2} + x, -\frac{1}{2} - y, 2 - z$	(iv)	$\frac{3}{2} - x, -y, \frac{1}{2} + z$
(ii)	$1 - x, \frac{1}{2} + y, \frac{3}{2} - z$	(v)	$x, y, 1 + z$
(iii)	$1 - x, \frac{1}{2} + y, \frac{3}{2} - z$	(vi)	$\frac{1}{2} - x, -y, \frac{1}{2} + z$

agreement indices. Thus the precision of this data set is not great enough to reveal any angular dependence of  $f''$ .

The structure is nearly the same as that of ammonium hydrogen tartrate (van Bommel & Bijvoet, 1958). Bond distances (Fig. 1) and angles (Table 3) of the bitartrate ion are identical within about five times the estimated standard deviations. The Cs ion has nine O neighbors (Table 4), eight of them approximately at the corners of a square antiprism, with the ninth, O(6), at a slightly greater distance capping one square face (Fig. 2). A similar arrangement occurs in the ammonium salt, but with shorter distances from N to its O neighbors. Hydrogen bonds from O(1) to O(6), O(3) to O(6), and O(4) to O(5) are the same in the two salts, with corresponding O—O distances 2.58, 2.76 and 2.84 Å in the Cs salt and 2.55, 2.74 and 2.80 Å in the ammonium salt.

*Acta Cryst.* (1978), **A34**, 371–377

## Minimizing the Variance in Densities Evaluated by Fourier Synthesis

BY C. L. DAVIS, E. N. MASLEN AND J. N. VARGHESE

*Department of Physics, University of Western Australia, Nedlands, Western Australia*

(Received 29 September 1977; accepted 22 October 1977)

Criteria for minimizing the variance in the Fourier synthesis of scattering densities are established. Their value has been confirmed by applying them to a study of bonding electron density by difference Fourier methods. The procedure has conceptual and practical advantages when compared with other types of weighting for the coefficients in a difference synthesis. It is expected that the minimum-variance criterion will improve the interpretability of some other classes of Fourier synthesis, such as those used in the study of protein structures.

### Introduction

The techniques used for the study of the structure of materials by the analysis of diffraction data may be broadly classified as either Fourier-synthesis or least-

We thank R. D. Giauque and L. Y. Goda for the rubidium analysis; the potassium was analyzed by the Microchemical Analysis Laboratory of this University. This work was supported in part by the US Energy Research and Development Administration.

### References

- BOMMEL, A. J. VAN & BIJVOET, J. M. (1958). *Acta Cryst.* **11**, 61–70.
- CROMER, D. T. & LIBERMAN, D. (1970). *J. Chem. Phys.* **53**, 1891–1898.
- DONNAY, J. D. H. (1943). *Am. Mineral.* **28**, 313–328.
- DOYLE, P. A. & TURNER, P. S. (1968). *Acta Cryst.* **A24**, 390–397.
- HÉMON, Y. (1931). Thesis, Paris; *Strukturber.* (1937). **2**, 861.
- HÖNL, H. (1933). *Ann. Phys. (Leipzig)*, **18**, 625–655.
- JOHNSON, C. K. (1965). *ORTEP*. Report ORNL 3794. Oak Ridge National Laboratory, Tennessee.
- PHILLIPS, J. C., WLODAWER, A., GOODFELLOW, J. M., WATENPAUGH, K. D., SIEKER, L. C., JENSEN, L. H. & HODGSON, K. O. (1977). *Acta Cryst.* **A33**, 445–455.
- RAMASESHAN, S. & ABRAHAMS, S. C. (1975). *Anomalous Scattering*. Copenhagen: Munksgaard.
- RIETZ, R. R., ZALKIN, A., TEMPLETON, D. H., EDELSTEIN, N. M. & TEMPLETON, L. K. (1978). To be published.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.
- TEMPLETON, L. K. & TEMPLETON, D. H. (1973). *Abstracts, American Crystallographic Association Proceedings*, Ser. 2, Vol. 1, p. 143.

squares methods. The latter are more convenient for quantitative work on a wide range of structural problems but there is revived interest in quantitative studies by Fourier methods because of applications in the charge-density and protein-structure fields.

The parameters determined by least squares are coefficients which define a model for the structure. The analysis is carried out in reciprocal space, and the statistical procedures for optimal use of the data are established. Fourier methods mainly involve imaging of structural information in direct space. The statistical treatment of the data gives rise to anomalies which do not appear in the least-squares analysis. The spaces for the alternative methods are related by a linear transformation, and their formal equivalence has been demonstrated (Cochran, 1948*a,b*). This suggests the possibility of evaluating Fourier syntheses in a more nearly optimal way.

### Analysis

Our starting point is scattering theory based on the first Born approximation. The density of a component of the scattering material at a position  $\mathbf{r}$  in a crystal is given by

$$\rho(\mathbf{r}) = \frac{1}{V} \sum \mathbf{F}(\mathbf{S}) \exp(-i\mathbf{S} \cdot \mathbf{r}) \quad (1)$$

where  $V$  is the unit cell volume,  $\mathbf{S}$  is a reciprocal-lattice vector,  $\mathbf{F}(\mathbf{S})$  is the corresponding component of the structure factor. The summation extends over the whole of reciprocal space. We may write

$$\mathbf{F}(\mathbf{S}) = F \exp(i\alpha) \quad (2)$$

where  $F$  is real and non-negative and the  $\mathbf{S}$  dependence of  $F$  and  $\alpha$  is assumed. By Friedel's Law

$$\mathbf{F}(\mathbf{S}) = \mathbf{F}^*(-\mathbf{S}) \quad (3)$$

so that

$$\rho(\mathbf{r}) = \frac{1}{V} \sum F \cos(\mathbf{S} \cdot \mathbf{r} - \alpha). \quad (4)$$

In an experimental study of scattering density we have measurements or estimates  $F_e$  and  $\alpha_e$  for  $F$  and  $\alpha$  respectively. By analogy with (4) an estimate of the density is given by

$$\rho_e(\mathbf{r}) = \frac{1}{V} \sum F_e \cos(\mathbf{S} \cdot \mathbf{r} - \alpha_e). \quad (5)$$

Let

$$F_e = F + \delta F \quad (6)$$

and

$$\alpha_e = \alpha + \delta\alpha. \quad (7)$$

The exact values of  $\delta F$  and  $\delta\alpha$  are, of course, unknown. Our information is restricted to knowledge of their statistical behaviour.

The distribution function for the cosine terms in equation (5) cannot in all cases be treated as random,

since the Fourier components must combine to give a substantial positive contribution in the vicinity of a nuclear position. A statistical treatment thus relies on a random distribution for  $\delta F$  and  $\delta\alpha$ . For the moment it is assumed that there is negligible correlation between the  $\delta F$  values and the phases of the cosine terms, then the variance in the electron density is uniform over the cell and is given by\*

$$\sigma^2(\rho_e) = \frac{1}{V^2} \sum \sigma^2(F_e) + F_e^2 \sigma^2(\alpha_e) \quad (8)$$

where  $\sigma^2(F_e)$  and  $\sigma^2(\alpha_e)$  are the variance in  $F$  and  $\alpha$ , *i.e.* the expectation values for  $(\delta F)^2$  and  $(\delta\alpha)^2$  respectively. Equation (8) describes the variance in the density when the measurements used in (5) are treated as equally reliable. This leads to well known anomalies. The ideal summation (1) extends over an infinite set of lattice points which in practice is approximated by a finite set. As the finite set is extended so  $\sigma^2(\rho_e)$  extends, apparently without limit. This implies that for a Fourier map calculated from an infinite set of data points we should not be able to draw any conclusions about the structure!

### Optimum filtering

It is well known that the use of filtering can improve the signal-to-noise ratio in Fourier analysis of experimental data. The methods are standard where there are known relations between the Fourier coefficients of the signal. This is not generally true for diffraction data, and a variety of filtering functions have been proposed. Some of these are based on *ad hoc* procedures, such as truncation of the data at some maximum  $|\mathbf{S}|$  obtained by trial. In others, the Fourier coefficients have been assigned weights appropriate to a least-squares determination of structural parameters, but it has not been shown that this is an appropriate weighting for the calculation of scattering density.

We define a modulated experimental density

$$\rho_M(\mathbf{r}) = \frac{1}{V} \sum M F_e \cos(\mathbf{S} \cdot \mathbf{r} - \alpha_e) \quad (9)$$

where  $M$  is a modulating function which we seek to optimize.

We minimize a residual with the form of an estimate of the mean variance of the density in the unit cell

$$R = \langle \sigma^2(\rho_M) \rangle_V = \frac{1}{V} \int \langle [ \rho_M(\mathbf{r}) - \rho(\mathbf{r}) ]^2 \rangle d\mathbf{r}. \quad (10)$$

\* In deriving this expression we note that the products of the cosine terms average to zero, except for squared terms and those where the  $\mathbf{S}$  vectors differ only in sign.

It is shown in the Appendix that this is equivalent to minimizing

$$R = \frac{1}{V^2} \sum \langle |M\mathbf{F}_e - \mathbf{F}|^2 \rangle \quad (11)$$

where

$$\mathbf{F}_e = F_e \exp(i\alpha_e). \quad (12)$$

We now invoke the statistical behaviour of  $\delta F$  and  $\delta\alpha$  for the determination of  $M$ . A symmetrical distribution for  $\delta\alpha$  will be assumed, and this implies that  $M$  will be real, in which case

$$R = \frac{1}{V^2} \sum \langle M^2 F_e^2 - 2MFF_e \cos(\delta\alpha) + F^2 \rangle. \quad (13)$$

We now choose  $M$ , the optimum form for  $M$ , as a function with a fixed value for each ensemble average so that

$$(R)_{M=M} = \frac{1}{V^2} \sum M^2 \langle F_e^2 \rangle - 2M \langle FF_e \cos(\delta\alpha) \rangle + \langle F^2 \rangle$$

and hence

$$\begin{aligned} \left( \frac{\partial R}{\partial M} \right)_{M=M} &= \frac{2}{V^2} [M \langle F_e^2 \rangle - \langle F_e F \cos(\delta\alpha) \rangle] \\ &= 0 \end{aligned} \quad (14)$$

for each ensemble.

This is our basic equation, which has a simple geometrical interpretation, since it may be expressed, in the complex plane, as

$$M = \frac{\langle \mathbf{F}_e \cdot \mathbf{F} \rangle}{\langle \mathbf{F}_e \cdot \mathbf{F}_e \rangle}. \quad (15)$$

### Variance for $\rho_M(\mathbf{r})$

From the relations given above, the variance in the filtered electron density may be written as the expectation value

$$\sigma^2[\rho_M(\mathbf{r})] = \frac{1}{V^2} \sum \langle \mathbf{F} \cdot \mathbf{F} \rangle - \frac{\langle \mathbf{F}_e \cdot \mathbf{F} \rangle^2}{\langle \mathbf{F}_e \cdot \mathbf{F}_e \rangle}. \quad (16)$$

For those cases where  $\delta F$  and  $\delta\alpha$  are randomly distributed the variance is again uniform over the cell and is given by

$$\begin{aligned} \sigma^2(\rho_M) &= \frac{1}{V^2} \sum \left[ \sigma^2(F_e) + F_e^2 \sigma^2(\alpha_e) \right] \\ &\times \left\{ 1 - \frac{[\sigma^2(F_e) + F_e^2 \sigma^2(\alpha_e)]}{\langle F_e^2 \rangle} \right\}. \end{aligned} \quad (17)$$

A problem arises in assessing the effect on the variance when there is an approximate functional relation between  $\mathbf{F}_e - \mathbf{F}$  and  $\mathbf{F}$ . The relation between  $\delta F$ ,  $\delta\alpha$  and the phases of the cosine terms can no longer be treated as random. This has important consequences in practice because of the effect of errors in the scale of the  $\mathbf{F}_e$ . Let  $\delta K/K$  be the relative error in the scale factor for the data. In a well-conducted experiment the effect of this error on the structure factors is small, and this is also true for the density in the bulk of the synthesis. In the vicinity of a nuclear position  $\mathbf{r}_j$  the density  $\rho(\mathbf{r}_j)$  may be very large. Although the corresponding error  $\rho(\mathbf{r}_j)(\delta K/K)$  due to the scale factor is relatively small, its value may still be large compared with  $\sigma(\rho_M)$ .

The size of the error may be estimated if  $\sigma^2(K)$  is known; but, where the error function is dominated by a single large error of this type it is difficult to be sure that the requirements for validity of the central limit theorem (Cramér, 1947) are satisfied. The densities at nuclear positions are correspondingly uncertain.

A further difficulty arises because there are no observations of  $\mathbf{F}_e$  for  $S (=|\mathbf{S}|/4\pi)$  values greater than some maximum values  $S_{\max}$  determined by the experimental conditions. The effect is to convolute the density with the Fourier transform of a function which is unity for  $S < S_{\max}$  and zero elsewhere. Corrections for the termination of the Fourier series have been proposed (Cruickshank, 1949), but depend for their validity on the reliability of theoretical models for the structure.

Finally, it must be noted that special problems arise for non-centrosymmetric structures where the phase information is derived from a model for the structure, and not from experiment. As is shown by Maslen (1968) the bias of the phases towards those of the model precludes the use of the corresponding densities for quantitative work.

### Charge-density studies

X-ray diffraction difference Fourier syntheses are important because the residue after subtracting a set of unbonded atom densities depicts the redistribution of electron density by chemical bonding and packing forces. However, the data for large  $S$ , which are necessary for accurate determination of nuclear positions and vibration amplitudes, contain relatively less information on the bonding electron density than the low-angle data. Furthermore the high-angle reflections are relatively less accurate because of reduced form factors and attenuation by the temperature factors. Consequently, maps based on a subset of the data with  $S$  less than some maximum are often more readily interpreted than those where high-angle data are included. Such truncation is rather arbitrary and the minimum-variance criterion provides an approach to the problem which is nearer to ideal.

We are concerned only with the centrosymmetric case, with phases restricted to 0 or  $\pi$ .  $\mathbf{F}_e$  becomes  $\Delta\mathbf{F}_e = \mathbf{F}_o - \mathbf{F}_c$ , where  $\mathbf{F}_o$  and  $\mathbf{F}_c$  are observed and calculated structure factors respectively.  $\mathbf{F}$  becomes  $\Delta\mathbf{F} = \mathbf{F} - \mathbf{F}_c$  where  $\mathbf{F}$  is the true structure factor. Equation (15) then has the form

$$M = \frac{\langle \Delta\mathbf{F}_e \cdot \Delta\mathbf{F} \rangle}{\langle \Delta\mathbf{F}_e \cdot \Delta\mathbf{F}_e \rangle}. \quad (18)$$

Setting

$$\Delta\mathbf{F}_e = \Delta\mathbf{F} + \delta\Delta\mathbf{F} \quad (19)$$

and assuming that  $\delta\Delta\mathbf{F}$  is not correlated with  $\mathbf{F}$ , we obtain

$$\begin{aligned} M &= 1 - \frac{\langle \Delta\mathbf{F}_e \cdot \delta\Delta\mathbf{F} \rangle}{\langle \Delta\mathbf{F}_e \cdot \Delta\mathbf{F}_e \rangle} = 1 - \frac{\langle \delta\Delta\mathbf{F} \cdot \Delta\mathbf{F} \rangle}{\langle \Delta\mathbf{F}_e \cdot \Delta\mathbf{F}_e \rangle} \\ &= 1 - \frac{\sigma^2(\mathbf{F}_e)}{\langle (\Delta\mathbf{F}_e)^2 \rangle}. \end{aligned} \quad (20)$$

This is an expression for the optimum filtering function as it applies to difference syntheses for centrosymmetric structures. In principle it may be applied with individual reflections as the statistical ensembles, provided more than one observation per reflection is available. From equation (19), with the assumptions used to derive equation (20), we have

$$\langle (\Delta\mathbf{F}_e)^2 \rangle = (\Delta\mathbf{F})^2 + \sigma^2(\mathbf{F}_e). \quad (21)$$

Thus,  $M$  tends to unity for  $(\Delta\mathbf{F})^2 \gg \sigma^2(\mathbf{F}_e)$  and to zero for  $(\Delta\mathbf{F})^2 \rightarrow 0$ . Its use in practice is open to two objections. The first of these is the sensitivity of the denominator in the second term to noise, in the critical low  $(\Delta\mathbf{F})^2$  region. The second difficulty is that of finding an analytical description of the effect of the filtering function on the density.

The first of these problems may be avoided if we use a biased estimator for  $\langle (\Delta\mathbf{F}_e)^2 \rangle$  by substituting  $\langle (\Delta\mathbf{F}_e)^2 \rangle$  for  $(\Delta\mathbf{F})^2$ , which is unknown, in the right-hand side of (21), yielding

$$M = 1 - \frac{\sigma^2(\mathbf{F}_e)}{\langle (\Delta\mathbf{F}_e)^2 \rangle + \sigma^2(\mathbf{F}_e)}. \quad (22)$$

This obviously gives values which are too large, especially for low values of  $\langle (\Delta\mathbf{F}_e)^2 \rangle$ . Its virtue, apart from its satisfactory asymptotic behaviour, is that it avoids generating absurd negative values of  $M$  because of noise for low  $\langle (\Delta\mathbf{F}_e)^2 \rangle$  reflections.

#### $\chi$ -square weighting

To obtain an unbiased but less noisy estimate we note that if the  $\sigma^2$  values are reliable then the reduced  $\chi$ -square statistic

$$\chi^2 = \frac{\langle (\Delta\mathbf{F}_e)^2 \rangle}{\langle \sigma^2(\mathbf{F}_e) \rangle}, \quad (23)$$

which ideally is not less than unity, is less noisy if the averages are taken over more than one reflection. Since the same  $M$  value applies to all members of the ensemble the averages should be taken over reflections with similar  $\chi$ -square values.

We have then

$$M = 1 - \frac{1}{\chi^2}. \quad (24)$$

In a statistical sense this function is a measure of the information content of the data not accounted for by the model used to calculate the structure factors. It is effectively a signal-to-noise ratio, decreasing asymptotically to zero as the limit of resolution of the data, in the least-squares sense, is approached.

When the  $\mathbf{F}_c$  are structure factors for the unbonded-atoms model  $1 - (1/\chi^2)$  is a measure of the information on bonding density contained in the data. We expect it to vary with  $S$  as the contribution to bonding density varies with  $S$ , *i.e.* it will be near unity for low angles, and tend to zero for high angles where the atomic approximation is more reliable.

The use of the function

$$M(S) = 1 - \frac{1}{\chi^2}(S) \quad (25)$$

automatically provides a logical approach both to the choice of the experimental limit for data collection and to the calculation of standard deviations. Equation (17) becomes

$$\sigma^2(\Delta\rho_M) = \frac{1}{V^2} \sum M \sigma^2(\mathbf{F}_e). \quad (26)$$

It should be noted that this allows for the effect of  $M$  modulating the signal as well as the noise in the experiment.

The maximum value of  $S$  in the data collection should be large enough for  $M(S)$  to approach its asymptotic limit.  $\sigma^2(\Delta\rho_M)$  will simultaneously approach its asymptotic value. Data collected beyond this point will give no further contribution to the difference density and no increase in its standard deviation.

If a satisfactory model for the charge-density map could be devised  $\chi^2(S)$  would be one and  $M(S)$  zero for the full angle range. This would yield a flat difference map, instead of the map of noise obtained in the usual approach to difference-map evaluation.

The representation of  $M$  as a function of  $S$  has a further advantage. The signal in the difference map is given by the convolution of the Fourier transform of  $M(S)$  with the true difference density. The signal in a conventional difference map is similar except that  $M(S)$  is replaced by a step function which is unity for  $S$  less than the data cut-off, and zero elsewhere.

### Example

A set of accurate X-ray data has recently been collected for copper sulphate pentahydrate  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , for which neutron diffraction structural parameters have recently been determined (Bacon & Titterton, 1975). A best set of parameters, based on X-ray high-angle values for the heavier atoms and neutron values for the lighter atoms, has been derived. The experimental cut-off for the X-ray data is  $S_{\text{max}} = 1.07 \text{ \AA}^{-1}$ .  $\chi^2(S)$  ranges from 580 for the low-angle data to 1.47 for the outer reflections.  $M$ , calculated for intervals of  $S$  specified in an agreement analysis, is shown in Fig. 1. Difference syntheses for a section through the plane of one of the two crystallographically independent Cu atoms and its neighbouring water molecules are shown in Fig. 2.

The section of the density corresponding to the optimum form of  $M(S)$  [described by equation (25) and illustrated in Fig. 1] is shown in Fig. 2(a). This map resembles Fig. 2(b), the map with a data set truncated at  $S_{\text{max}} = 0.9 \text{ \AA}^{-1}$ . This has a total weight  $\int_0^\infty MS^2 dS$  which is very close to that of the optimum filtering function. Fig. 2(c) is the conventional difference synthesis. Fig. 2(d) was evaluated with data modulated by truncation at  $S_{\text{max}} = 0.7 \text{ \AA}^{-1}$ . It is obvious from inspection that the Fig. 2(a) map is a good compromise between resolution and sensitivity to noise.

Fig. 2(e) shows the difference map evaluated with  $M$  in the near-optimum form for individual reflections given by equation (22). Another synthesis was evaluated applying the expression

$$M''(S) = 1 - \frac{\langle \sigma^2(\mathbf{F}_e) \rangle_S}{\langle (\Delta \mathbf{F}_e)^2 + \sigma^2(\mathbf{F}_e) \rangle_S} \quad (27)$$

This is similar to equation (22) but the ensembles consist of all reflections in given intervals of  $S$  instead of individual reflections. This synthesis is not shown because it is virtually identical with Fig. 2(e). This confirms that taking ensembles as intervals of  $S$  is a satisfactory method of weighting difference densities for charge-density studies.

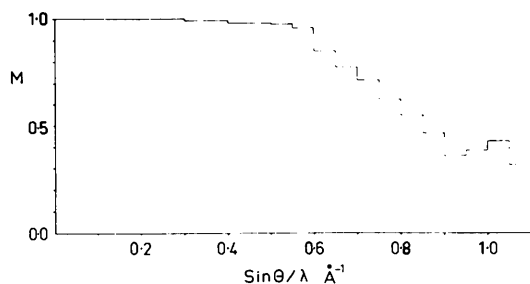


Fig. 1. The minimum-variance filtering function for X-ray data on copper sulphate pentahydrate.

The standard deviations in the electron density evaluated with equation (8) for Fig. 2(c) and equation (26) for Fig. 2(a) are  $0.09$ , and  $0.06$ ,  $e \text{ \AA}^{-3}$  respectively. There is a small reduction in height of the sharper features near the nuclear positions in the filtered map, but the features expected for the bonding electron density are defined to a higher level of significance. Furthermore, in so far as  $M(S)$  shown in Fig. 1 can be said to be approaching its asymptotic limit,  $\sigma^2(\Delta \rho_M)$  approaches a well-defined value determined by the accuracy of the observations rather than the data cut-off.

### Diffraction ripple

In addition to changing the resolution, modulation of the density may produce artefacts, in the form of ripple surrounding prominent features. It is desirable for ease of interpretation that this diffraction ripple should be small compared with the standard deviations. The magnitude of the ripple is determined by the shape of the Fourier transform of the modulating function. This is shown in Fig. 3 for (a) the optimum filtering function and (b), (c), (d) data truncated without further modulation at  $0.9 \text{ \AA}^{-1}$ , at the experimental cut-off ( $1.07 \text{ \AA}^{-1}$ ), and at  $S = 0.7 \text{ \AA}^{-1}$  respectively. The use of a logarithmic scale to enhance the structure of the ripple should be noted. The ripple is lower, relative to the height of the central maximum, for  $M(S)$  than for the truncated functions.

### Conclusions

The minimum-variance filtering function has conceptual and practical advantages in optimizing the difference syntheses for charge-density studies.

The usefulness of the minimum variance criterion will vary according to the nature of the problem, but there is an important case with a strong similarity to those in charge-density work.

In protein-structure analysis the weak reflections, which predominate at high Bragg angles, are determined with lower precision than the medium and high-intensity data, which predominate at low angles. Optimization of the Fourier maps is important in differentiating between functional groups with similar geometries, and in the refinement of large protein structures where least-squares methods are computationally prohibitive. It is probable that minimum-variance methods could usefully be applied.

We gratefully acknowledge the contributions of Dr A. H. White in collecting the intensity data for copper sulphate and of Professor G. E. Bacon who kindly provided advance information on the neutron structural parameters.

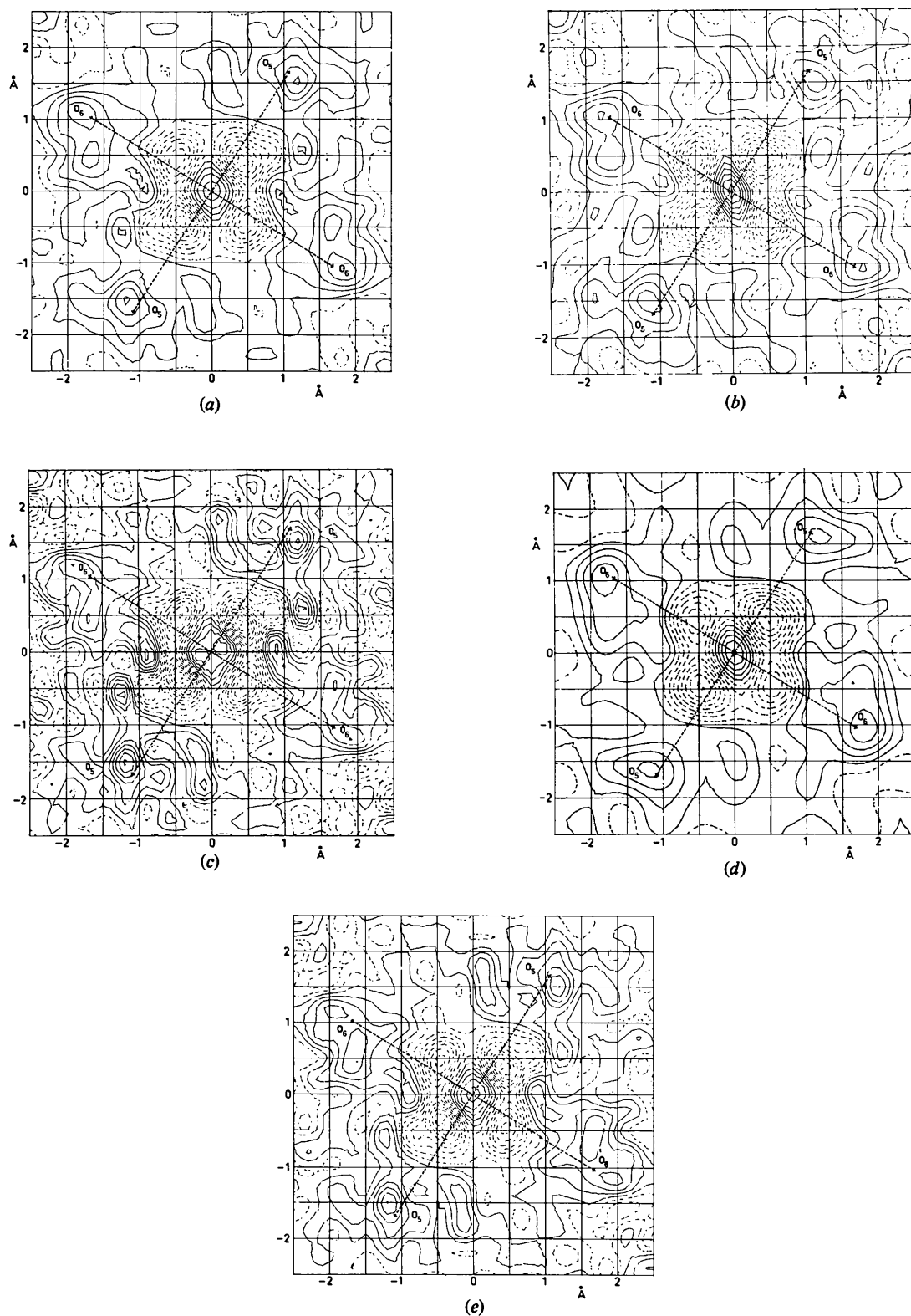


Fig. 2. Section of the X-ray difference density for copper sulphate pentahydrate through a Cu and four water oxygen atoms evaluated (a) with the optimum filtering function, (b) with  $S_{\max} = 0.9 \text{ \AA}^{-1}$ , (c) with the full data ( $S_{\max} = 1.07 \text{ \AA}^{-1}$ ). (d) with  $S_{\max} = 0.7 \text{ \AA}^{-1}$  and (e) with  $M$  as in equation (22) in the text. Contour interval  $0.1 e \text{ \AA}^{-3}$ .

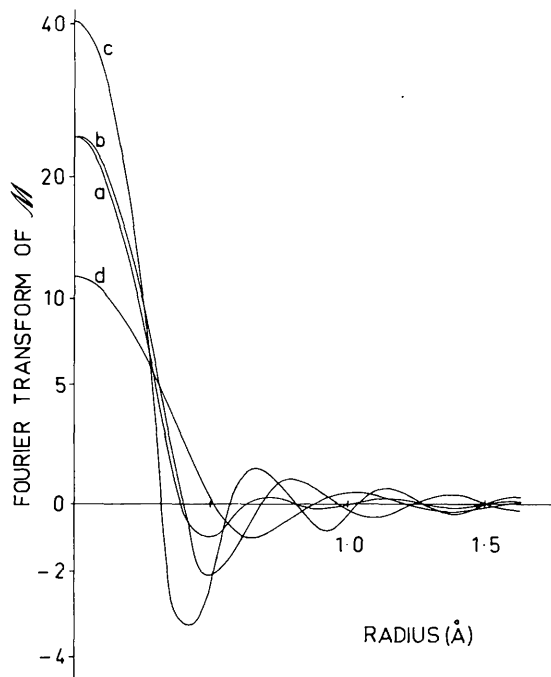


Fig. 3. The radial dependence of the Fourier transforms for (a) the optimum filtering function, (b) a step function with  $S_{\max} = 0.9 \text{ \AA}^{-1}$ , (c) step-function with  $S_{\max} = 1.07 \text{ \AA}^{-1}$  and (d)  $S_{\max} = 0.7 \text{ \AA}^{-1}$ . The logarithmic scale should be noted.

Financial assistance was provided by the Australian Research Grants Committee and the University of Western Australia Research Committee. One of us (CLD) acknowledges receipt of a Commonwealth Postgraduate Research Award.

## APPENDIX

The true density and filtered observed density are defined respectively by

$$\rho(\mathbf{r}) = \frac{1}{V} \sum \mathbf{F} \exp(-i\mathbf{S} \cdot \mathbf{r}) \quad (A.1)$$

and

$$\rho_M(\mathbf{r}) = \frac{1}{V} \sum M \mathbf{F}_e \exp(-i\mathbf{S} \cdot \mathbf{r}). \quad (A.2)$$

The variance of the filtered density can then be written as

$$\begin{aligned} \sigma^2[\rho_M(\mathbf{r})] &= \langle [\rho_M(\mathbf{r}) - \rho(\mathbf{r})]^2 \rangle \\ &= \left\langle \left[ \frac{1}{V} \sum \exp(-i\mathbf{S} \cdot \mathbf{r}) (M \mathbf{F}_e - \mathbf{F}) \right]^2 \right\rangle. \quad (A.3) \end{aligned}$$

Terms in the square of the indicated sum in the last expression can be separated into two classes: (i) products containing the factors  $\exp(i\mathbf{S} \cdot \mathbf{r})$  and  $\exp(-i\mathbf{S} \cdot \mathbf{r})$ , (ii) products containing the factors  $\exp(i\mathbf{S} \cdot \mathbf{r})$  and  $\exp(i\mathbf{S}' \cdot \mathbf{r})$ , where  $\mathbf{S}' \neq -\mathbf{S}$ . With this in mind and noting that Friedel's Law requires that  $\mathbf{F}(-\mathbf{S}) = \mathbf{F}^*(\mathbf{S})$  and  $M(-\mathbf{S}) \mathbf{F}_e(-\mathbf{S}) = M^*(\mathbf{S}) \mathbf{F}^*(\mathbf{S})$ , (A.3) can be written

$$\begin{aligned} \sigma^2[\rho_M(\mathbf{r})] &= \frac{1}{V^2} \langle \sum' |M \mathbf{F}_e - \mathbf{F}|^2 \rangle \\ &+ \frac{1}{V^2} \langle \sum \sum' \exp[i(\mathbf{S} + \mathbf{S}') \cdot \mathbf{r}] (M \mathbf{F}_e - \mathbf{F}) \\ &\times (M' \mathbf{F}_e' - \mathbf{F}') \rangle \quad (A.4) \end{aligned}$$

where  $\sum'$  means a restricted sum over all  $\mathbf{S}' \neq -\mathbf{S}$  and  $\mathbf{F}' = \mathbf{F}(\mathbf{S}')$  etc. The second expression in (A.4) is a superposition of periodic functions with different wavelengths. If the summation extends over a sufficiently large region of reciprocal space, this term will be negligible in comparison with the first. The variance therefore becomes

$$\sigma^2[\rho_M(\mathbf{r})] = \frac{1}{V^2} \langle \sum' |M \mathbf{F}_e - \mathbf{F}|^2 \rangle. \quad (A.5)$$

## References

- BACON, G. E. & TITTERTON, D. H. (1975). *Z. Kristallogr.* **141**, 330–341.  
 COCHRAN, W. (1948a). *Acta Cryst.* **1**, 138–142.  
 COCHRAN, W. (1948b). *Nature (London)*, **161**, 765.  
 CRAMÉR, H. (1947). *Mathematical Methods of Statistics*, pp. 213–218. Univ. Press.  
 CRUICKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 65–84.  
 MASLEN, E. N. (1968). *Acta Cryst.* **B24**, 1172–1175.