Acta Cryst. (1975). A31, 388

A comment on Hawthorne's paper *Refinement of merohedrally twinned crystals.*

By K. SUDARSANAN, *Georgia Institute of Technology, Atlanta, Georgia* 30332, *U.S.A.*

(*Received* 7 *October* 1974; *accepted* 6 *December* 1974)

Comment is made on the paper by Hawthorne, F. C. [*Acta Cryst.* (1974). A30, 603–604].

In the procedure outlined by Hawthorne (1974) for the refinement of crystal structure by the least-squares method with intensities from merohedrally twinned crystals, the structure factors calculated for the two crystals in the twin are summed. By introducing linear constraints in the least-squares refinement program, as stated by the author, the newly generated parameters are used to calculate additional structure-factor components to be added to the original. But, instead, it is the intensities calculated for the two crystals in the twin which have to be added to get the total intensity. The program has to be rewritten to minimize $w\{I_o-(I_1+I_2)\}^2$ where I_o is the observed intensity, I_1 and I_2 are the calculated intensities due to the two crystals in the twin and w is the weight assigned for each observed intensity.

Reference

HAWTHORNE, F. C. (1974). *Acta Cryst.* A30, 603–604.

Acta Cryst. (1975). A31, 388

Efficiency in Fourier phase refinement for protein crystal structures.

By DOUGLAS M. COLLINS, *Department of Chemistry, Texas A&M University College Station, Texas* 77843, *U.S.A.*

(*Received* 8 *August* 1974; *accepted* 24 *December* 1974)

Fourier phase refinement for macromolecular crystal structures can be accomplished using electron density maps sampled at intervals of half the minimum interplanar spacing for which diffraction data have been measured in that direction. Order-of-magnitude economy in computation is thus gained as compared to use of customary sampling rates.

Fourier phase refinement is a much discussed method for improvement of the crystallographic description of large biological molecules [see, *e.g.*, Barrett & Zwick (1971) and references cited therein]. Use of this method certainly will become more widespread as the real cost of computing decreases and computer programs based on Fast Fourier Transform algorithims become freely available. Notwithstanding the present speed and efficiency of Fourier inversion, repetitive calculations for crystals with large asymmetric units still require a most substantial investment of computational resources. In the interest of further economy we have examined the proposition that the sampling criterion of Lipson & Cochran (1966) must be satisfied in Fourier phase-refinement calculations, as implied by Barrett & Zwick (1971).

The conceptual cornerstone of these refinement calculations is the idea that a mediocre electron density map (*i.e.* a map based on experimental structure amplitudes and estimated phases) can be modified in a reasonable way to provide a density function rather more similar to the true density function than was the original (Hoppe & Gassmann, 1968). This being so, it is a simple matter to transform the modified density function and obtain improved phases. This concept was first presented by Sayre (1952) who gave the relationships

$$\varrho(\mathbf{r}) \sim \varrho^2(\mathbf{r}), \tag{1a}$$

and consequently,

$$F_{\mathbf{h}} = \frac{1}{V}\theta_{\mathbf{h}} \sum_{\mathbf{K}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}}, \tag{1b}$$

where V is the unit cell volume and θ is a shape function. The development of these equations has been directed principally toward use of the latter in *ab initio* estimation of phases for crystal structures of moderate complexity. In fact, (1*b*) and similar equations have been so successful in application that many crystallographers have come to think of (1*b*) as a fundamental relationship among structure

Table 1. *Figures of merit and convergence during phase refinement*

All numbers are given in pairs; the upper is for case I, the lower for case II as described in the text.

		Iteration				
		0	1	2	3	4
Average change in phase angle (°)*	I	—	32 (37)	17 (20)	6 (8)	2 (4)
	II	—	32 (37)	16 (19)	5 (8)	2 (4)
R†	I	—	0·28	0·21	0·19	0·19
	II	—	0·26	0·21	0·21	0·22
Average error in phase angle (°)*	I	46 (46)	26 (32)	17 (24)	15 (22)	15 (22)
	II	46 (46)	24 (29)	16 (23)	16 (23)	17 (24)

* Modulus-weighted averages are given; the corresponding unweighted averages are in parentheses.
† R is calculated from $R=\sum||F_{\text{exp}}|-|F_{\text{calc}}||/\sum|F_{\text{exp}}|$.

factors while relegating (1*a*) to the status of but an interesting corollary.

Let us turn that about and focus on the following form of (1*a*):

$$\varrho(\mathbf{r})=f[\varrho_{\text{exp}}(\mathbf{r})], \tag{2}$$

where f may be any reasonable function that will change an experimental electron density map into a better representation of the true structure. Suppose that $f(\varrho_{\text{exp}})$ gives an accurate representation of the correct structure at experimental resolution. In this case it is clear that to obtain accurate structure factors upon Fourier inversion of ϱ, the grid upon which it is represented need be no finer than required by Shannon's (1949) sampling theorem. On the other hand, the sampling rate in ϱ would have to be doubled (Gold & Rader, 1969) in each direction to obtain an overlap-free transform of $\varrho(\mathbf{r})=\varrho^2_{\text{exp}}(\mathbf{r})$.

Although we have no interest in transforming $\varrho(\mathbf{r})=\varrho^2_{\text{exp}}(\mathbf{r})$, we can see that the need for finer sampling in this case arises from the (artificial) increase of resolution and sharpness in direct space. Thus informed, we consider that an efficient modification function may increase resolution and sharpness at low density levels, but must leave the highest levels essentially unchanged. It may be possible then, to obtain from the minimally sampled density function, structure factors effectively free from overlap. Alternatively, in terms of Sayre's (1951) hypothetical atoms, an ideally modified density function, among other things, will be comparable to a collection of hypothetical scattering centers whose transforms have about the same range in reciprocal space as do the experimental data.

For an empirical test, an artificial structure of 17 single-bonded carbon atoms was constructed in space group $P1$ with $a=b=c=16$ Å, and $\alpha=\beta=\gamma=90°$. The atoms were assigned isotropic thermal parameters of 10 Å^2 and theoretical structure factors, $|F_t|\exp i\varphi_t$, were calculated to a minimum interplanar spacing of 2·0 Å. To give this test some likeness of a protein problem, fixed random errors were introduced into $|F_t|$ to give $|F_{\text{exp}}|$ such that $\sum||F_t|-|F_{\text{exp}}||/\sum|F_t|=0{\cdot}12$; similarly, errors were added to φ_t to give starting phases in error by an average of 46·4°. With ϱ_{exp} always scaled to have a maximum value of 1·0 e Å^{-3}, the following scheme of modification was used.

$$\varrho_{\text{calc}}(\mathbf{r})=\begin{cases} 3\varrho^2_{\text{exp}}(\mathbf{r})-2\varrho^3_{\text{exp}}(\mathbf{r});\ \varrho_{\text{exp}}>0 \\ 0;\ \varrho_{\text{exp}}\leq 0 \end{cases} \tag{3}$$

and an updated ϱ_{exp} is the transform of $|F_{\text{exp}}|\exp i\varphi_{\text{calc}}$.

Parallel procedures were carried through using two sampling rates in ϱ. Case I was based on maps of 16 divisions per cell edge to meet the Shannon (1949) criterion, and case II was based on maps of 32 divisions per cell edge. Improved phases were developed in four iterations for both cases and the results are given in Table 1.

It is clear that case II does not give better results. Moreover, the slightly better results for case I are not factitious but arise from the inability of direct-space phase-refinement methods to account properly for series-termination error. In particular, certain areas of any experimental electron density map must be negative because of series-termination effects, yet they are set to some non-negative value. The situation is clear if we consider the δ function at the origin, its discrete transform truncated at order 4, then transformed onto grids of eight per period and 16 per period. In the first instance, the density function is positive at the origin and zero elsewhere. In the second, the density function will quite correctly show negative regions; this effect can be minimized only through use of a grid of minimal fineness. Of course, for a real protein structure, series-termination errors are trifling at worst (if we use F_{h} rather than E_{h}), so we shall base our preference for case I solely on the fact that it requires less calculation by an order of magnitude than does case II.

Evidently Barrett & Zwick (1971) were wrong to imply that the costly sampling criterion of Lipson & Cochran (1966) must generally be observed in Fourier phase-refinement calculations.* A well chosen density-modification function allows these calculations to be carried out quite satisfactorily using maps based on a grid of minimal fineness.

Our practical interest is in the structure of staphylococcal nuclease (Arnone *et al.*, 1971) and phase extension and refinement for a 1·5 Å (minimum interplanar spacing) data set. We estimate that for this case, the outlined calculations, as based on Shannon's (1949) criterion, will require 80 min on an IBM 360/65 computer at a total cost of $160. To double the sampling rate in each direction would increase the required time by 560 minutes and the cost by $1120.

The computer programs used for this work were modified versions of the programs by Hubbard, Quicksall & Jacobson (1971). This work was supported in part by a grant of the National Institute of General Medical Sciences.

* Lipson & Cochran (1966) state that 'the number of points at which the electron density should be sampled in any one direction should be *about* three times the highest index observed in that direction' (our italics). They also state, *pari passu*, that 'it is necessary to ensure that the various sets of spectra do not overlap – that is, for the highest index q observed, the $(n-q)$th and higher orders should be negligible'. It is essential to recognize that this latter general requirement leads to the numerical rule of thumb first quoted only in the context in which it is given, that is, in the context of normal high-resolution studies of small molecules. In a broad sense then, it would be correct to say that the sampling criterion of Lipson & Cochran (1966) *always* must be observed in phase-refinement calculations, but in the same broad sense Barrett & Zwick's (1971) statement 'It is important to note that ϱ_E must be sampled more finely than its transform if the inverse Fourier operation is to yield accurate E_{h}'s' is incorrect. Let it be observed that Lipson & Cochran's (1966) general requirement is, in effect, a restatement of Shannon's (1949) sampling theorem.

References

ARNONE, A., BIER, C. J., COTTON, F. A., DAY, V. W., HAZEN, E. E. JR, RICHARDSON, D. C., RICHARDSON, J. S. & YONATH, A. (1971). *J. Biol. Chem.* **246**, 2302–2316.

BARRETT, A. N. & ZWICK, M. (1971). *Acta Cryst.* A**27**, 6–11.

GOLD, B. & RADER, C. M. (1969). *Digital Processing of Signals*, p. 166. New York: McGraw-Hill.

HOPPE, W. & GASSMAN, J. (1968). *Acta Cryst.* B**24**, 97–107.

HUBBARD, C. R., QUICKSALL, C. O. & JACOBSON, R. A. (1971). *The Fast Fourier Algorithm and the Programs ALFF, ALFFDP, ALFFPROJ, ALFFT and FRIEDEL.* Ames, Iowa: Ames Laboratory Report.

LIPSON, H. & COCHRAN, W. (1966). *The Crystalline State.* Vol. III. *The Determination of Crystal Structures*, p. 102. London: Bell.

SAYRE, D. (1951). *Acta Cryst.* **4**, 362–367.

SAYRE, D. (1952). *Acta Cryst.* **5**, 60–65.

SHANNON, C. E. (1949). *Proc. Inst. Radio Eng.* **37**, 10–21.