Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 14 November 2000
Accepted 3 January 2001

Phase determination *via* Sayre-type equations with anomalous-scattering data

Jeffrey Roach, Pascal Retailleau and Charles W. Carter Jr*

Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, North Carolina, USA. Correspondence e-mail: carter@med.unc.edu

The necessary background for the analysis of complex-valued electron-density maps is established. Various systems of structure-factor equations of convolutional type akin to Sayre's squaring method equations are tested for agreement on the real and imaginary parts of the electron density as well as approximations thereof. A system of convolutional structure-factor equations holding in a complex-valued electron density generated by two atom types is developed. The scope of application of these equations is determined and it is shown that the equations provide a method of extrapolating high-resolution phases from a low-resolution base phase set without introducing further model bias. Additional applications to phase refinement are explored.

1. Introduction

Although the Sayre equations (Sayre, 1952) represent the theoretical cornerstone of the so-called direct methods, their application to phase determination in actual macromolecular structures has been relatively limited (*cf.* Sayre, 1974; Zhang & Main, 1990; Shiono & Woolfson, 1991; Refaat *et al.*, 1995). Geometrically, the Sayre equations are derived from the assumption that the electron density consists of the sum of a number of smaller, identical and non-interpenetrating electron densities. This 'atomicity', however reasonable in principle, necessarily limits the scope of application of these equations. Generalizations of the Sayre equations have been developed that partially overcome the equi-atom requirement (*cf.* Woolfson, 1958; Rothbauer, 1980, 2000).

Experimental phasing techniques such as the multiple-wavelength anomalous-dispersion method (MAD) (Hendrickson, 1991) are predicated on the existence of anomalously scattering atoms. Typically, these atoms will severely violate the equi-atom assumption of the Sayre equations, consequently limiting their effectiveness. In some cases, anomalous-scattering information from a single wavelength (SAD) in conjunction with direct methods (or other techniques) is sufficient for structure determination. The series of five papers by Fan *et al.* beginning with Fan, Han, Qian & Yao (1984) and Fan, Han & Qian, (1984) as well as Wang (1985) and Brodersen *et al.* (2000) investigate this problem. Direct methods that exploit the additional complex structure of the electron density induced by anomalous scattering should prove useful in this pursuit.

In this work, convolutional structure-factor equations are developed that are a more appropriate representation of the situation commonly encountered in the use of MAD phasing in macromolecular crystallography. Based on the generalization of the Woolfson approach, these equations assume that the complex-valued density maps are generated by two atom types. One atom type is intended to model large anomalously scattering atoms in the heavy-atom substructure and the other meant to represent the lighter atoms. The resolution range where the new system of equations holds is delineated and the capabilities of the system for phase extrapolation and refinement are investigated.

In particular, it is shown that given a reasonably correct set of experimental phases the new system of equations can extrapolate phases to high resolution of comparable quality to experimental phases. These extrapolated phases have the additional advantage of limiting model bias that would be introduced in model-based phase extension.

We conclude that convolutional structure-factor equations modelling complex electron-density maps afford accurate extrapolations – in some cases superior to those based on model coordinates. Such extrapolations should be of considerable value in determining unbiased phases for high-resolution data measured to extend the resolution of a structure. Refinement of phases is less tractable. The residual of the convolutional structure-factor equations alone is insufficient to support significant phase refinement of large structures at typical resolutions. Nevertheless, the ability to model complex-valued density provides new insights into the disappointing performance of phase refinements based on minimizing the least-squares magnitude of the residual of convolution-type structure-factor equations.

1.1. Complex-valued electron density

It is well known that in the absence of anomalous scattering the structure factors satisfy a particular type of Hermitian symmetry described by Friedel's law, *viz*

$$F_{-\mathbf{h}} = \bar{F}_{\mathbf{h}}.$$

This observation implies that the electron density takes only real values. More generally, however, the presence of anomalous scattering introduces complex-valued corrections into the electron density and consequently the electron density is composed of a real part and a significantly weaker imaginary part.

The structure factors of the real and imaginary electron density, $F^R_{\mathbf{h}}$ and $F^I_{\mathbf{h}}$ respectively, are calculated from the available structure factors as follows:

$$F^R_{\mathbf{h}} = \frac{F_{\mathbf{h}} + \bar{F}_{-\mathbf{h}}}{2} \qquad F^I_{\mathbf{h}} = \frac{F_{\mathbf{h}} - \bar{F}_{-\mathbf{h}}}{2i}.$$

Note that if the Friedel law holds then all the structure factors of the imaginary part vanish and the structure factors of the real part are equal to the structure factors of the whole electron density.

In general, the imaginary anomalous corrections will be smaller than the real scattering, thus the imaginary density will be significantly weaker than the real density. However, the imaginary density has certain properties of interest to the study of Sayre-type equations. In particular, the relatively large imaginary anomalous corrections will be associated with the strongest anomalous scatterers. Therefore, the imaginary density will concentrate around these atoms. Furthermore, these atoms will typically be of a similar type and be positioned distant enough from one another to be separated at relatively low resolution, thus providing an electron density highly suited to the equal-atomicity assumptions of Sayre-type equations.

Two limitations remain. First, a complete set of structure factors, both $F_{\mathbf{h}}$ and $F_{-\mathbf{h}}$, is necessary to calculate the structure factors of the real and imaginary parts of the electron density. Second, the phases of the imaginary part are not directly related to the phases of the real part and therefore information gained from the imaginary electron density will be difficult to apply to the real, and more important, part of the electron density. In light of these observations, in the sequel we develop an approximation to the electron density requiring phases for only one of a Friedel-related pair of reflections. This approximation has the property that the phases of its real part are directly related to the phases of its imaginary part and potentially affords a means of estimating the phases of the real map from the phases of the imaginary map.

1.2. Friedel-average and Bijvoet-difference maps

Although the anomalous-scattering data gives the amplitudes $|F_{\mathbf{h}}|$ and $|F_{-\mathbf{h}}|$ for both halves of reciprocal space determined by the Friedel Hermitian symmetry, typically phases are determined, experimentally or otherwise, for only one asymmetric unit, *i.e.* $\varphi_{\mathbf{h}}$. The following construction (Kraut, 1968) provides a method of simulating this situation when both $\varphi_{\mathbf{h}}$ and $\varphi_{-\mathbf{h}}$ are known. Let $\varphi^0_{\mathbf{h}}$ be defined for any $\mathbf{h}$ in reciprocal space by

$$\varphi^0_{\mathbf{h}} = \begin{cases} \varphi_{\mathbf{h}} & \text{for } \varphi_{\mathbf{h}} = \varphi_{-\mathbf{h}} = \pi \\ (\varphi_{\mathbf{h}} - \varphi_{-\mathbf{h}})/2 & \text{otherwise.} \end{cases}$$

Note that angle $\pi = -\pi$, thus, in either case, $\varphi^0_{-\mathbf{h}} = -\varphi^0_{\mathbf{h}}$. Consequently, one half of the approximate phases determines the other half.

Consider an electron density $\tilde{\rho}$ defined by the structure factors

$$\tilde{F}_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(i\varphi^0_{\mathbf{h}}). \tag{1}$$

This density approximates the true electron density. Like the true density, it does not satisfy the Friedel law and thus it takes both real and imaginary values. The imaginary part of this electron density is the Bijvoet-difference Fourier as defined by Kraut (1968). The structure factors of this map for any $\mathbf{h}$ are

$$\tilde{F}^I_{\mathbf{h}} = (1/2i)[|F_{\mathbf{h}}| - |F_{-\mathbf{h}}|] \exp(i\varphi^0_{\mathbf{h}}).$$

Note that the i in the denominator of the Bijvoet-difference structure factor causes the well known, though under-appreciated, phase shift of $-\pi/2$. The structure factors of the real part of this map for any $\mathbf{h}$ are

$$\tilde{F}^R_{\mathbf{h}} = \tfrac{1}{2}[|F_{\mathbf{h}}| + |F_{-\mathbf{h}}|] \exp(i\varphi^0_{\mathbf{h}}).$$

For lack of more compelling terminology, the real part will be called the Friedel-average Fourier.

1.3. Approximation quality

In order to use these approximate maps intelligently, their precise nature as approximations must be determined. Note that the structure factors of the two maps are related by a particular rotation:

$$F_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(i\varphi_{\mathbf{h}}) = |F_{\mathbf{h}}| \exp[i(\varphi_{\mathbf{h}} - \varphi^0_{\mathbf{h}} + \varphi^0_{\mathbf{h}})] = \tilde{F}_{\mathbf{h}} \exp(i\varepsilon_{\mathbf{h}}),$$

where

$$\varepsilon_{\mathbf{h}} = \begin{cases} 0 & \text{when } \varphi_{\mathbf{h}} = \varphi_{-\mathbf{h}} \\ (\varphi_{\mathbf{h}} + \varphi_{-\mathbf{h}})/2 & \text{otherwise.} \end{cases}$$

These equations determine a complex-valued map with the following real and imaginary parts (here $\mathbf{h}^+$ denotes one particular half of the reciprocal space determined by Friedel's law):

$$\begin{aligned}
\tilde{\rho}(x) = {} & \sum_{\mathbf{h}^+}[(\Re F_{\mathbf{h}} + \Re F_{-\mathbf{h}}) \cos(2\pi\mathbf{h}\cdot\mathbf{x}) \\
& - (\Im F_{\mathbf{h}} - \Im F_{-\mathbf{h}}) \sin(2\pi\mathbf{h}\cdot\mathbf{x})] \cos(\varepsilon_{\mathbf{h}}) \\
& + \sum_{\mathbf{h}^+}[(\Im F_{\mathbf{h}} + \Im F_{-\mathbf{h}}) \cos(2\pi\mathbf{h}\cdot\mathbf{x}) \\
& + (\Re F_{\mathbf{h}} - \Re F_{-\mathbf{h}}) \sin(2\pi\mathbf{h}\cdot\mathbf{x})] \sin(\varepsilon_{\mathbf{h}}) \\
\Im\tilde{\rho}(x) = {} & \sum_{\mathbf{h}^+}[(\Im F_{\mathbf{h}} + \Im F_{-\mathbf{h}}) \cos(2\pi\mathbf{h}\cdot\mathbf{x}) \\
& + (\Re F_{\mathbf{h}} - \Re F_{-\mathbf{h}}) \sin(2\pi\mathbf{h}\cdot\mathbf{x})] \cos(\varepsilon_{\mathbf{h}}) \\
& + \sum_{\mathbf{h}^+}[(\Re F_{\mathbf{h}} + \Re F_{-\mathbf{h}}) \cos(2\pi\mathbf{h}\cdot\mathbf{x}) \\
& - (\Im F_{\mathbf{h}} - \Im F_{-\mathbf{h}}) \sin(2\pi\mathbf{h}\cdot\mathbf{x})] \sin(-\varepsilon_{\mathbf{h}}).
\end{aligned} \tag{2}$$

Note the real and imaginary parts of the true density:

$$\begin{aligned}\Re\rho(x) &= \sum_{\mathbf{h}^+}[(\Re F_{\mathbf{h}} + \Re F_{-\mathbf{h}})\cos(2\pi\mathbf{h}\cdot\mathbf{x}) \\ &\quad - (\Im F_{\mathbf{h}} - \Im F_{-\mathbf{h}})\sin(2\pi\mathbf{h}\cdot\mathbf{x})] \\ \Im\rho(x) &= \sum_{\mathbf{h}^+}[(\Im F_{\mathbf{h}} + \Im F_{-\mathbf{h}})\cos(2\pi\mathbf{h}\cdot\mathbf{x}) \\ &\quad + (\Re F_{\mathbf{h}} - \Re F_{-\mathbf{h}})\sin(2\pi\mathbf{h}\cdot\mathbf{x})]\end{aligned} \tag{3}$$

are intertwined in (2) with respect to the sine and cosine of $\varepsilon_{\mathbf{h}}$. When $\varepsilon_{\mathbf{h}}$ is close to zero, the value of $\cos\varepsilon_{\mathbf{h}}$ is close to one and the value of $\sin(\pm\varepsilon_{\mathbf{h}})$ is close to zero. Therefore, when all the $\varepsilon_{\mathbf{h}}$ are close to zero, the approximation will most closely resemble the true complex-valued electron density. In particular, the Friedel average will approach the real part and the Bijvoet difference will approach the imaginary part.

Recall, however, that the imaginary density is in general weaker than the real density. Therefore, even for small values of $\varepsilon_{\mathbf{h}}$ the presence of the real part of the map in the Bijvoet difference can cause significant corruption. This is not a concern with the Friedel-average map where the weak imaginary density is further diminished by the values of $\sin(\pm\varepsilon_{\mathbf{h}})$. Since the $\{\varepsilon_{\mathbf{h}}\}$ are the phase differences between the true and approximate phases, they are potentially useful in estimating the difference between the true imaginary map and the Bijvoet-difference map.

These observations are illustrated by an example. Consider a test system based on the atomic coordinates for tetragonal hen egg white lysozyme (Protein Data Bank ID 193L) where all the sulfur atoms are replaced by selenium. The map correlation coefficient between the true real map and the Friedel-average approximation is 0.9931 whereas the map correlation coefficient between the true imaginary map and the Bijvoet-difference approximation is only 0.4617. Although the agreement between the imaginary density and the Bijvoet-difference Fourier appears very weak, bear in mind that the Bijvoet-difference Fourier does, in fact, locate the large anomalous scatterers and at reasonable contour levels the maps of the imaginary density and the Bijvoet-difference Fourier look very similar, *cf.* Fig. 1. Therefore, it is important to recognize that the Bijvoet-difference map is a good approximation to the imaginary density around the large anomalous scatterers; however, throughout the comparatively larger volume of the entire unit cell the approximation is of a lower quality.

2. Convolutional structure-factor equations

The original convolutional structure-factor equations, the Sayre equations (Sayre, 1952), stem from the assumption that the electron density consists of a finite number of identical non-overlapping atoms. They take an appealingly simple formulation:

$$F_{\mathbf{h}} = a_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} \qquad a_{\mathbf{h}} \text{ real}$$

and have led to extensive theoretical development; the sign and triplet relationships and tangent formula all make use of the Sayre equations. Further use is being made in connection with bulk solvent correction (Guo *et al.*, 2000).

Unfortunately, the equi-atom structure assumption can be somewhat restrictive. In this regard, Woolfson (1958) developed a generalized system of convolutional structure-factor equations that take into account two different atom types. This system,

$$F_{\mathbf{h}} = b_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} + c_{\mathbf{h}} \sum_{\mathbf{l}} \sum_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}}, \qquad b_{\mathbf{h}}, c_{\mathbf{h}} \text{ real},$$

proved to be well suited for polypeptide crystals containing some heavy atoms (Shiono & Woolfson, 1991). Note that, if all $c_{\mathbf{h}}$ vanish, the system of equations becomes equivalent to that of Sayre.

Although the Woolfson equations have been generally applied to macromolecular crystals containing metals, even structures consisting of only C, N and O atoms can benefit. In general, the O atoms corresponding to solvent water will have significantly larger temperature (Debye–Waller) factors than O atoms corresponding to a protein. The difference in temperature factors will introduce, loosely speaking, two atom types: a generic protein-atom type and a generic solvent-atom type. In this case, the two-atom-type approximation can be used to more accurately represent the distinction between solvent and protein.

Since both systems of equations described above are derived from ideal circumstances, it is not clear exactly how well they hold with actual structures. One test of agreement is to consider the correlation between the map of a known structure and the map determined from the right-hand side of the equation calculated with the correct structure factors.

Consider the real density of the model system derived from lysozyme where all the sulfur atoms are replaced by selenium. The agreement of the two systems of equations, described as a map correlation coefficient, were determined at various resolutions from 1.5 to 3.0 Å. Results are summarized in Fig. 2(*a*). In these calculations, the $a_{\mathbf{h}}$ parameters of the Sayre equations were determined based on a carbon atom with average temperature factor. The $b_{\mathbf{h}}$ and $c_{\mathbf{h}}$ parameters of the Woolfson equations were determined using a temperature-factor-averaged carbon atom and a temperature-averaged selenium atom. The convolutions were calculated up to the given resolution with fast Fourier transforms.

It is clear from the figure that strong agreement in the equations for the real density requires both multiple atom types and relatively high (greater than 2.0 Å) resolution.

The imaginary density much more closely satisfies the hypotheses of the convolutional structure-factor equations. In particular, in this case the selenium atoms are the only significant anomalous scatterers and they are reasonably separated in the unit cell. Although it is necessary only to consider the Sayre equations, the agreement of the equations on true imaginary density and Bijvoet-difference Fourier differs. The great advantage of the Bijvoet-difference Fourier is that correctly determined phases could be used to determine correct phases for the Friedel average, a close approximation to the real density. The correlation coefficients for the two maps are found in Fig. 2(*b*). Parameters for the Sayre equa-

tions are based on the imaginary anomalous corrections associated with selenium for Cu $K\alpha$ radiation.

Unlike the Bijvoet-difference Fourier, the imaginary density has strong agreement at all resolutions. The weak agreement in the Bijvoet-difference Fourier indicates that the $a_{\mathbf{h}}$ values determined for the imaginary part are not well suited for the Bijvoet-difference Fourier. Although it might be possible to determine more suitable equation parameters, the problem remains that there are many fewer non-zero structure factors for the Bijvoet-difference map than for the imaginary map. This is because Friedel-related pairs with the same amplitude and different phases determine a structure factor of 0 in the Bijvoet difference. In the imaginary density, structure factors of 0 occur only if the amplitudes of the Friedel-related pairs are the same and the phases differ by π.

3. Extrapolation of phases *via* convolutional structure-factor equations

In light of the previous section, it is possible to define a set of convolutional structure-factor equations intended to model a complex density that consists of two types of atoms: heavy atoms, *e.g.* selenium, with large imaginary anomalous corrections and lighter carbon-like atoms with very slight imaginary anomalous corrections. Such a system of equations differs from the two-atom system of Woolfson in the fact that the parameters are complex numbers reflecting the complex scattering of the two atom types.

Our first application of the system:

$$F_{\mathbf{h}} = b_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} + c_{\mathbf{h}} \sum_{\mathbf{l}} \sum_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \qquad (4)$$

with $b_{\mathbf{h}}$, $c_{\mathbf{h}}$ complex will be to choose phases to accompany known amplitudes lying above a fixed low-resolution threshold. It is assumed that the phases of resolution lower than the threshold are essentially correct.

Let R be the set of reflections $\mathbf{h}$ within a particular resolution sphere. Consider the expression

$$G_{\mathbf{h}} = b_{\mathbf{h}} \sum_{\mathbf{k}\in R} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} + c_{\mathbf{h}} \sum_{\mathbf{l}\in R} \sum_{\mathbf{k}\in R} F_{\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}}. \qquad (5)$$

Provided $\mathbf{h}$ is not far outside R, the phase of $G_{\mathbf{h}}$ together with the known amplitude $|F_{\mathbf{h}}|$ should be a reasonable approximation to $F_{\mathbf{h}}$. These extrapolated structure factors produce a complex-valued electron density approximating the true density.

Of course, the further $\mathbf{h}$ lies outside R, the more susceptible this approximation becomes to truncation errors, *i.e.* due to $F_{\mathbf{k}}$ excluded from the summations when $\mathbf{k} \notin R$, indicating that stepwise extrapolation is more likely to lead to quality phases. However, improvements gained by reducing the truncation error must be balanced by the cumulative error introduced at each step.

This procedure should be compared with the practice of extending the resolution of a structure solved at lower resolution by measuring structure-factor amplitudes at higher resolution and generating phases from the lower-resolution model. The essential difference is that the new phases are derived from a basis phase set and not an atomic model. In fact, good experimental phases should be preferable as a basis for phase extrapolation. Unlike experimental phases, phases constructed from a pre-existing lower-resolution model will be biased toward the original model. Therefore, if the initial low-resolution phases are experimentally determined, extrapolation will produce higher-resolution phases without model bias. We illustrate in §3.2 that one effect of such a model bias is to omit density around solvent molecules not in the low-resolution model.

Table 1
Extrapolation from model phases with errors in °.

Resolution	2.0–1.5 Å	2.5–1.5 Å	3.0–1.5 Å
∞–4.0 Å	1.267	1.267	1.267
4.0–3.0 Å	0.775	0.775	0.775
3.0–2.0 Å	1.606	28.022	52.659
2.0–1.7 Å	26.019	36.080	48.296
1.7–1.5 Å	25.972	37.697	51.444
∞–1.5 Å	15.372	29.478	44.237
Map correlation	0.986	0.956	0.885

3.1. Extrapolation from a model structure

The model lysozyme system where the sulfur atoms are replaced by selenium provides the first test of the extrapolation *via* the set of convolutional structure-factor equations (4). Assume that the amplitudes to 1.5 Å are given. The basis phase set will be the phases of the $\tilde{\rho}$ approximation below a given lower-resolution limit, *e.g.* 2.0 Å. The convolutional structure-factor equations (2) give approximations to the true phases.

Proceeding in 0.25 Å increments, the extrapolation is applied from three low-resolution thresholds: 2.0, 2.5 and 3.0 Å. It is readily apparent from the phase errors and map correlation coefficients of the real map (Table 1) that error introduced in the extrapolation remains reasonable even from as far as 3.0 Å. The small phase errors occurring at low resolution due to the approximation $\tilde{\rho}$, *cf.* equation (1), provide a good basis for extrapolation. Figs. 3(*a*)–(*c*) give an impression of the general degradation occurring in the electron density when extrapolating from larger distances. For the most part, the maps are quite reasonable. One particularly interesting aspect of the extrapolation is that it appears to intensify the heavy-atom effects, in this case diminishing the density of the smaller atoms surrounding the selenium.

3.2. Extrapolation from experimental data

Extrapolations from experimental data are of significantly more interest than extrapolation from an essentially correct model system. An example is provided by the TrpRs-5′AMP complex (Retailleau *et al.*, 2001). A set of 2.0 Å model-independent SAD phases produced by *SHARP* (La Fortelle & Bricogne, 1997) and processed by *Solomon* (Abrahams & Leslie, 1996) is used as a starting point for generating approximate phases for measured amplitudes between 2.0 and 1.7 Å from the convolutional structure-factor equations (4).

Table 2
Extrapolation from experimental phases, errors in °.

Resolution	Model extension	Extrapolation
∞–3.0 Å	24.275	24.275
3.0–2.5 Å	23.507	23.507
2.5–2.0 Å	24.590	24.590
2.0–1.85 Å	46.272	42.814
1.85–1.7 Å	49.484	39.700
2.0–1.7 Å	47.713	41.417
Map correlation	0.945	0.945

The phases produced by the convolutional structure factors are intended to be used to construct a high-resolution model when none exists. A second set of phases between 2.0 and 1.7 Å was calculated for comparison from a 2.9 Å model (Doublié *et al.*, 1995). These two phase sets are compared against the final high-resolution coordinate-refined structure (Retailleau *et al.*, 2001).

Phase errors and map correlation coefficients *versus* the final model are tabulated in Table 2. Note that the phases extrapolated using equations (4) are significantly better: in fact, by nearly 10° in the highest-resolution shell. Because the structure factors generating these maps differ only at high resolution, the physical differences apparent in the respective maps will be somewhat subtle and are expressed most dramatically through molecules such as bound solvent. Of particular interest are glycerol molecules which do not occur in the low-resolution model. Figs. 4(*a*)–(*d*) show the map extended from the low-resolution model as compared with the extrapolation on two particular glycerol molecules. A similar effect occurs at other large solvent molecules, in particular the sulfates shown in Figs. 4(*e*)–(*h*). In both cases, the extrapolation using the convolutional structure-factor equations produces a fuller more accurate density.

4. Refinement *via* convolutional structure-factor equations

The next application is to determine whether refining the phases by enforcing agreement in the convolutional structure-factor equations can actually improve the phases. Consider the least-squares magnitude D of the difference between the left-hand side and right-hand side of the system of equations (4):

$$D = \sum_{\mathbf{h}} \left| F_{\mathbf{h}} - b_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} - c_{\mathbf{h}} \sum_{\mathbf{l}} \sum_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \right|^2$$

for $b_{\mathbf{h}}$, $c_{\mathbf{h}}$ complex. Refinement would adjust phases for the structure factors of known amplitude so as to minimize the size of D.

In order for this sort of refinement to work, there must be a strong correlation between the size of D and the phase error; otherwise, reducing D will make little improvement on the phases. To test this correlation, 21 structures based on the lysozyme model were generated with phase errors randomly distributed according to a normal distribution with mean 0 and standard deviation ranging from 20 to 40°. The correlation between the D values and mean phase error of the 21 test systems is 0.9753, indicating a strong linear dependence, *cf.* Fig. 5.

Although these results appear impressive, it is important to bear in mind exactly what they establish: *viz* random phase errors introduced in a correct structure weaken the agreement in the convolutional structure-factor equations. The more worthwhile question, does reducing D bring an incorrect structure closer to the correct structure?, has, as yet, been studied only inadequately.

The method of phase refinement by minimizing the least-squares residual of a system of convolutional structure-factor equations was proposed by Sayre (1972) as a method of bridging the gap between initial phasing and final coordinate refinement. Because Sayre's squaring method equations represent an equi-atom approximation to the electron density, this method was not intended to provide intensive phase refinement. Surprisingly, the initial results of Sayre (1972) demonstrated that refinement by enforcing the Sayre equations significantly reduced the mean phase error for computer-generated protein-like equi-atom structures of 100 atoms. The local refinement method of these results has been drawn into question in the paper by Chen & Su (2000), where simulated annealing produces the correct phases up to certain space-group-related limitations for several computer-generated equi-atom systems of approximately 100 atoms. Although these latter systems are based on actual molecules, applications to real proteins have been quite modest.

Refinements of 2Zn pig insulin (Zhang & Main, 1990), a protein of 806 non-hydrogen atoms, show that the Sayre equations alone cannot improve the mean phase error of 3.0 Å MIR phases. Further, using the Sayre equations to incrementally extend and refine phases to 2.0 Å results in a mean phase error considerably worse than that corresponding to 2.0 Å MIR phases. Nevertheless, the Sayre equations provide a complement to density-modification refinement. In fact, the same process of incremental extension and refinement *via* the Sayre equations accompanied with solvent flattening and histogram matching produces a set of 2.0 Å phases slightly more accurate than the experimental phases.

The first protein submitted to this type of refinement was rubredoxin containing 424 non-hydrogen atoms (Sayre, 1974). Extending an initial 2.5 Å experimental phase set to 1.5 Å *via* the Sayre equations and refining with respect to the Sayre equations resulted in a mean phase difference of 46.6° from a coordinate-refined structure. Although this phase difference seems large, the very promising aspect of the study was that the method contributed to correctly identifying several non-polar side chains that were known to be incorrect in the model and hence a source of model bias in coordinate refinement. In light of the results on 2Zn pig insulin, however, this probably did not account entirely for the substantial phase difference. Nor is it clear that the success with rubredoxin resulted from phase refinement and not simply from phase extrapolation, which we have shown in §3 to be quite accurate.

One essential obstacle to this type of refinement is that the convolutional structure-factor equations hold only

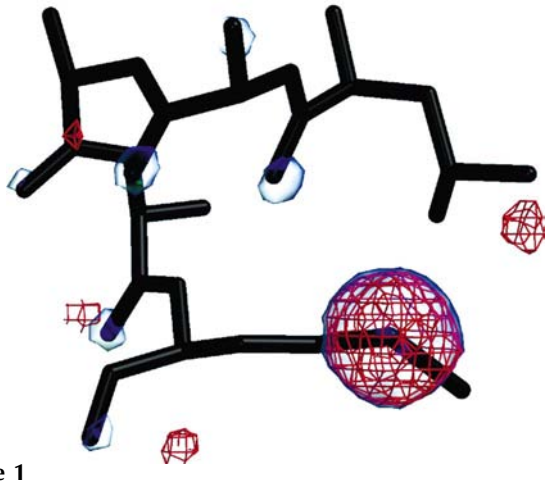

Figure 1
The Bijvoet-difference Fourier shown in red approximates the true imaginary density shown as the blue surface. The two maps are contoured at the same absolute level. Although the map correlation coefficient between the maps is only 0.4617, the Bijvoet difference accurately determines the location of the large anomalous scatterers. Figure produced by *Bobscript* (Esnouf, 1997) and *Raster3D* (Merritt & Bacon, 1997).

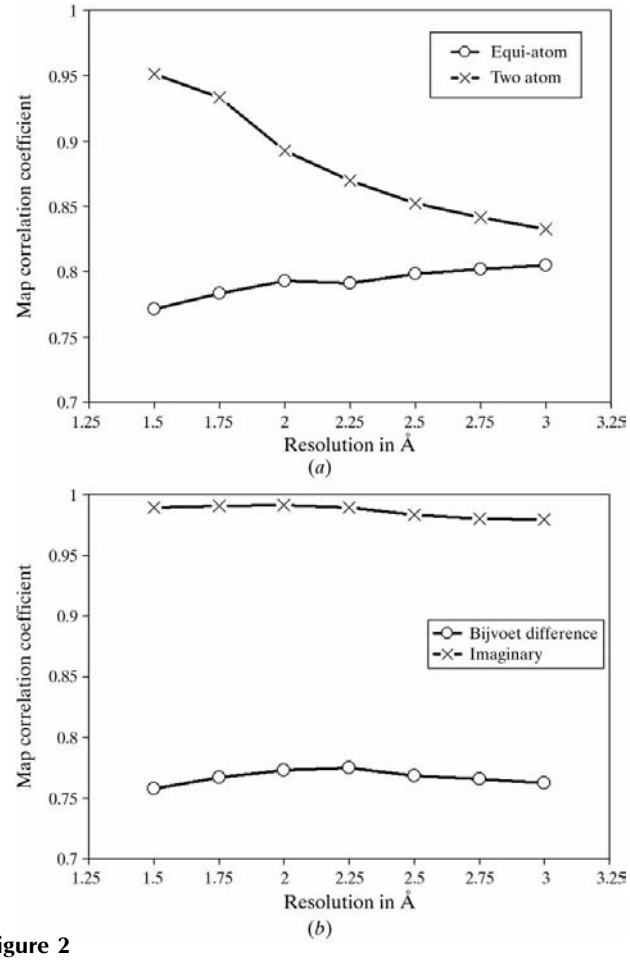

Figure 2
(*a*) The real density map is best modelled by two atom types at high resolution. As resolution goes down neither convolutional structure-factor equation provides a good model for the density. (*b*) The single-atom-type convolutional structure-factor equations hold well on the imaginary density at all tested resolutions. The Bijvoet-difference Fourier is not nearly as well modelled.

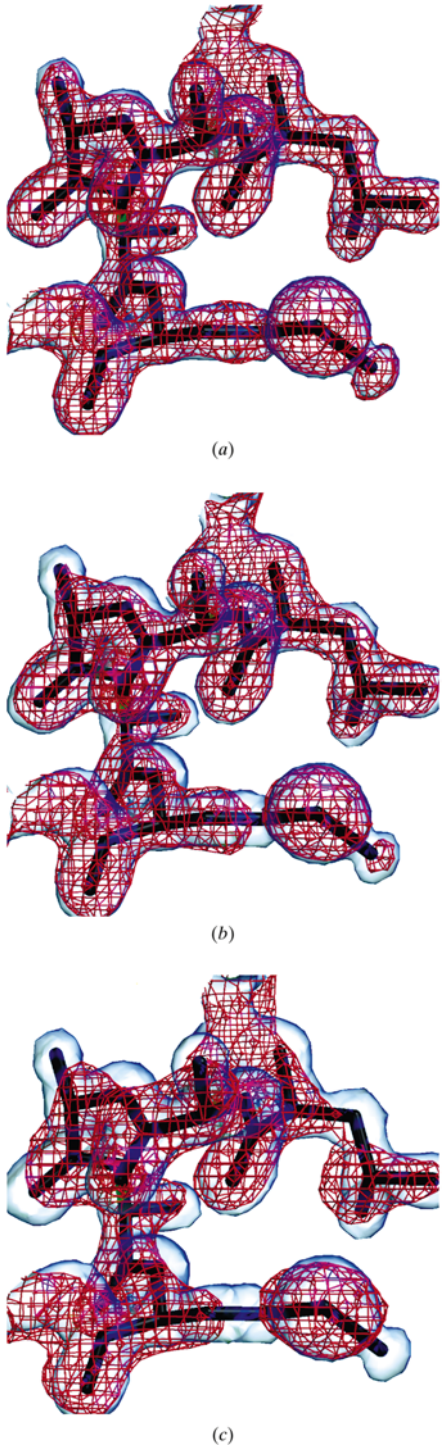

Figure 3
Phase extrapolation based on complex-valued convolutional structure-factor equations for a model system. The true density is represented by the blue surface. Extrapolation based on the convolutional structure-factor equations is shown in red. (*a*) Extrapolation from 2.0 Å provides an excellent approximation to the true density. Note the furrows that are beginning to form around the selenium in the lower right-hand corner. The extrapolation intensifies some of the effects of finite resolution. (*b*) Extrapolation from 2.5 Å suffers somewhat from the disruptive effects of finite resolution, in particular around the selenium in the lower right-hand corner. However, the map is of high quality overall. (*c*) Extrapolation from 3.0 Å loses density around the smaller atoms surrounding the heavy atom on the lower right-hand side. Figure produced by *Bobscript* (Esnouf, 1997) and *Raster3D* (Merritt & Bacon, 1997).

approximately for real structures. Therefore, the correct phase set may not necessarily have the minimum value of D among all possible phase sets. For example, the extrapolation structures of the lysozyme-based model all have lower D values than the correct structure. In this regard, it is possible to 'over-refine' by introducing features into the map that are artifacts of the convolutional structure-factor equation hypotheses, *e.g.* limited atom types.

The existence of false minima is not an entirely unfamiliar situation. Certainly any optimization method will have to effectively distinguish false minima from true minima. Generally, however, it is assumed that the correct structure occupies a local minimum; therefore, if the starting point of the refinement is close enough to the true minimum, the false minima will play no role. The final example reveals that this need not always be the case.

Consider the lysozyme test system generated with 40° mean phase error. This system is used as a starting point for steepest-descent refinement (see Appendix *A*) of two different objective functions: the D value above, denoted D_C, and the analogous D value applied to the real part of the map only, denoted D_R. Note that for D_R the values of $b_{\mathbf{h}}$ and $c_{\mathbf{h}}$ are real. Each D value and its corresponding mean phase error are calculated after each cycle and displayed graphically in Fig. 6. Within 20 cycles of refinement, both objective functions have decreased dramatically. Yet the mean phase error in either case has changed by less than 2°. Apparently, in both cases, the local minimum closest to the initial phase set is not the correct solution.

Although neither refinement finds the correct solution, it should be emphasized that the behavior of the refinement of the complex-valued density differs significantly from the behavior of the refinement of its real part. After 20 cycles, the

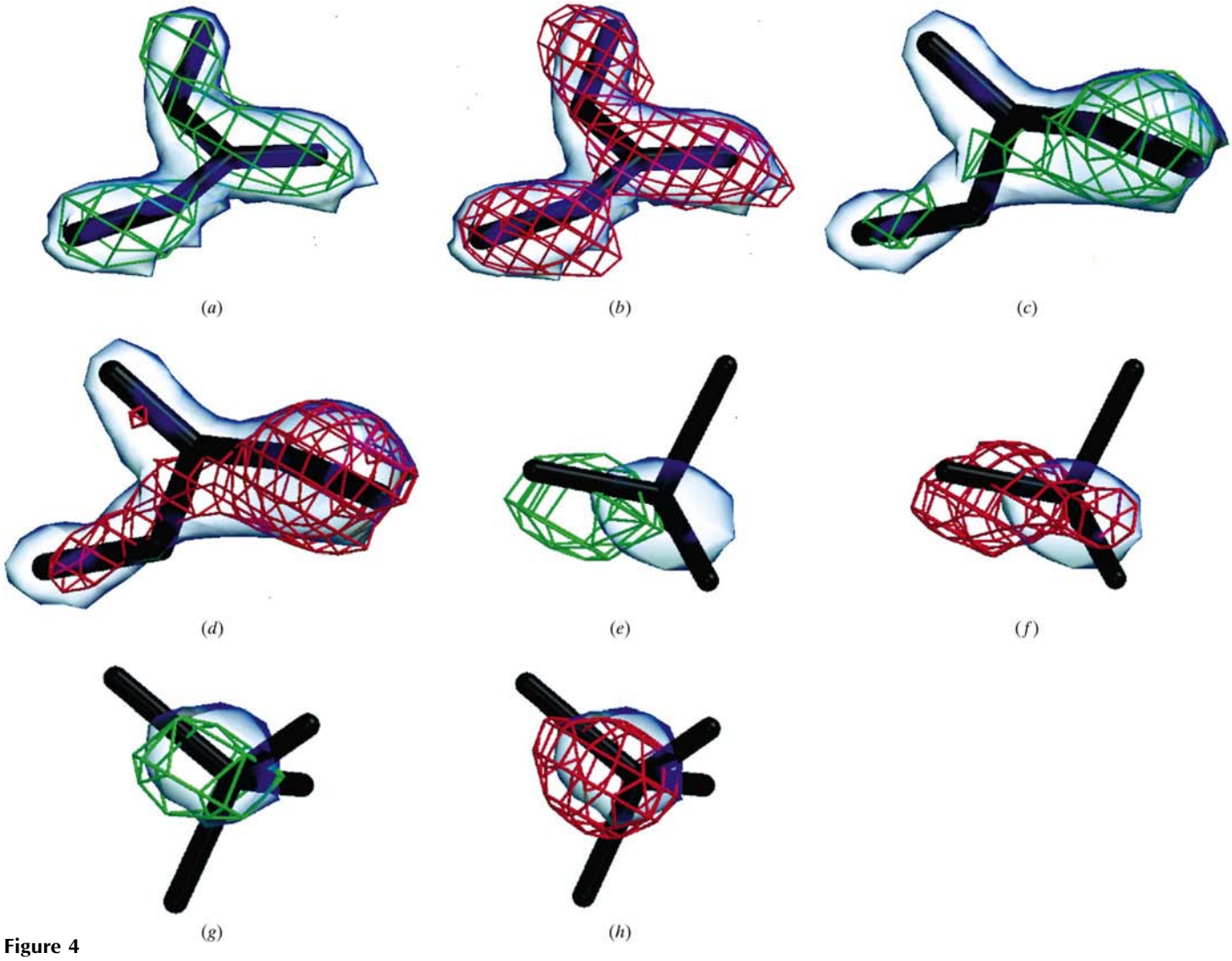

Figure 4
Extrapolation from experimental phases. Low-resolution model-based extension shown in green (*a*, *c*, *e* and *g*). Extrapolation based on convolutional structure-factor equations shown in red (*b*, *d*, *f* and *h*). The coordinate-refined map is represented by the blue surface. (*a*) Failure to connect the density in the glycerol molecule. (*b*) Reasonable glycerol structure predicted. (*c*) Most of the glycerol structure lost. (*d*) An improved density compared to (*c*). (*e*) Reasonable sulfate density. (*f*) More complete sulfate density predicted. Note both red and green densities position the sulfate differently from the coordinate-refined model. (*g*) Weak sulfate density. (*h*) Extrapolation produces fuller density. Figure produced by *Bobscript* (Esnouf, 1997) and *Raster3D* (Merritt & Bacon, 1997).

refinement of the real electron density has yet to reach a stable point. In fact, after 40 cycles the refinement has reduced D_R to 3.055×10^{10}, well below that of the correct structure where $D_R = 3.158 \times 10^{10}$. At this point, the mean phase error has risen approximately 6.8° from the starting point. Therefore, there exists an incorrect structure with a much lower D_R than that of the correct solution, indicating that D_R alone is insufficient to distinguish the correct structure from false structures.

Although the refinement of the complex-valued electron density also fails to gain significant phase improvement, the sensitivity of D_C to an incorrect local minimum contrasts markedly with the pathological behavior of D_R. The mean phase error decreases monotonically to a local minimum approximately 0.4° below the starting point after 19 cycles. This behavior suggests that a more sophisticated algorithm taking advantage of additional information could avoid the local minimum and refine closer to the correct solutions. In the case of the Sayre equations, Refaat *et al.* (1995) have demonstrated that any phase refinement at all with resolutions lower than 1.5 Å requires more delicate treatment. The use of unitary structure factors as well as a number of additional constraints have lead to an improvement in the phase error by 3 to 4°.

A particularly promising modification is to use an explicit estimation of $\{\varepsilon_{\mathbf{h}}\}$ to coordinate refinement of the Bijvoet difference *via* the Sayre equations with refinement of the real electron density using the Woolfson generalization.

Any of these techniques could be applied *mutatis mutandis* to the system of equations described here and modest improvement could be expected. However, the potential for the new system of equations to complement density-modification routines such as solvent flattening and histogram matching is great. Given the work of Zhang and Main and more recent work, *e.g.* Foadi *et al.* (2000), where good results are obtained by using the Sayre equations in conjunction with density modification, a system of equations that represents more accurately the existence of multiple atom types and anomalous scattering will probably lead to significant phase improvement for a wider class of problems.

5. Conclusions

It has been shown that the real and imaginary parts of a complex-valued electron density have distinct structure and any convolutional structure-factor equations must model them differently. The real part of the density does not closely satisfy the hypotheses of Sayre-type equations; thus additional atom types and relatively high resolution are needed. In contrast, the imaginary part of the density is highly suited for modelling by Sayre-type equations. However, the imaginary map is significantly weaker than the real map. Consequently, any system of convolutional structure-factor equations intended to model a complex-valued electron density will necessarily be dominated by the real part and therefore it is difficult to take advantage of the highly suitable imaginary electron density in applications to the overall complex density.

An alternative is to make use of some relation between the phases of the two maps. Unfortunately, the phases of the real map and the phases of the imaginary map are not directly related. In response to this problem, the Bijvoet-difference Fourier of Kraut was introduced. This map, an approximation to the imaginary map, has the additional property that its phases are directly related to the phases of a very close approximation to the real part. We have shown that the Bijvoet-difference Fourier is composed in a non-trivial manner of both the real map and the imaginary map. Moreover, this intertwining together with the relative weakness of the map makes the Bijvoet-difference Fourier somewhat unsuited for analysis *via* convolutional structure-factor equations. Further study of the possible estimation and use of $\varepsilon_{\mathbf{h}}$ is warranted in this regard.

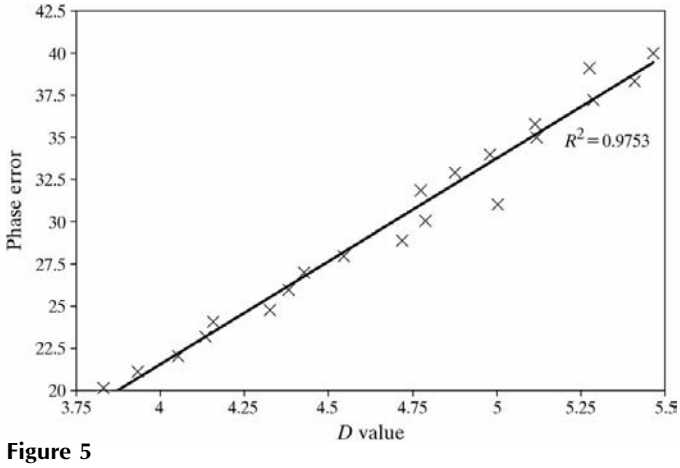

Figure 5
A strong linear correlation between increasing the phase error of a correct structure and the corresponding increase in D value.

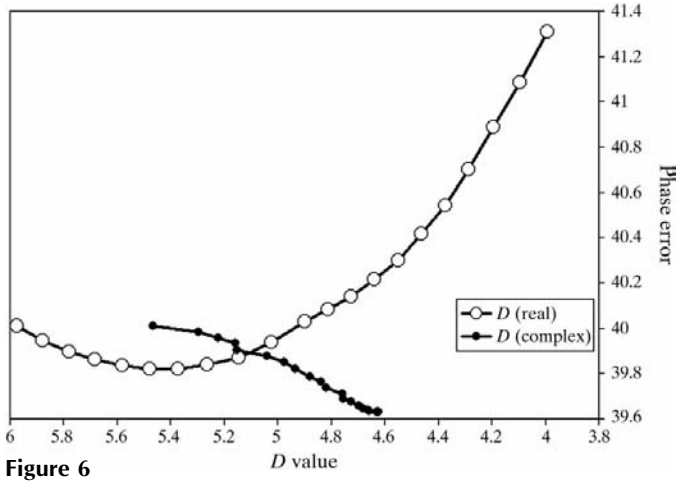

Figure 6
The refinement of D(real) fails to reach a stable point after 20 cycles. During this period, the refinement improves the mean phase error by less than one degree initially and then increases the mean phase error. The refinement of D(complex) behaves more sensibly, reaching a local minimum and stopping. However, it fails to improve the phase error significantly.

A variation on a system of equations of Woolfson has been developed to represent a system composed of two atom types, as an approximation to commonly encountered MAD and SAD phasing experiments. One atom type corresponds to a large real scattering contribution and a modest imaginary anomalous correction. The second atom type corresponds to a smaller real scattering and a negligible imaginary correction. The new system of equations has been used to derive phases for a set of known high-resolution structure-factor amplitudes from a lower-resolution base phase set. This method provides for model-independent extrapolation to higher resolution. Furthermore, if the low-resolution base phase set is itself experimentally determined, the high-resolution extrapolated phases are relatively free from model bias.

Despite the surprising accuracy of phase extrapolation by this method, it has been shown that the system of convolutional structure-factor equations is insufficient as a sole means of refinement. Although the inclusion of information from the complex structure of the density improves the refinement path, only slight improvement in the phases was obtained in the test case. Significant improvement, however, has been documented using density-modification techniques in conjunction with convolutional structure-factor equations that represent more drastic approximations to the electron density, *e.g.* the Sayre equations. Therefore, it is likely that the new system of equations will be an improved component in a suite of complementary refinement techniques.

APPENDIX *A*
Refinement computation formulae

The volume of computation necessary to successfully apply these methods can be prohibitive if the calculations are performed naïvely. In this Appendix, formulae analogous to those found in Sayre & Toupin (1975) and Refaat *et al.* (1995) are established. The principal difference between this Appendix and previous work is that the convolutional structure-factor equations are based on Woolfson (1958) rather than Sayre (1952); however, the assumption that the electron density is complex valued and, consequently, that Friedel's law fails, introduces additional pathology.

Consider the objective function

$$D = \sum_{\mathbf{h}} \left| F_{\mathbf{h}} - b_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} - c_{\mathbf{h}} \sum_{\mathbf{l}} \sum_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \right|^2 .$$

It is well known that the gradient of this objective function, ∇D, can be expressed in terms of the Jacobian of a particular system of functions of the phases $\varphi_{\mathbf{h}}$. For example, let

$$d_{\mathbf{h}} = F_{\mathbf{h}} - b_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}} - c_{\mathbf{h}} \sum_{\mathbf{l}} \sum_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}}$$

for some reflection $\mathbf{h}$. Thus, with $\bar{z}$ denoting the complex conjugate of z, D can be written as

$$D = \sum_{\mathbf{h}} d_{\mathbf{h}} \bar{d}_{\mathbf{h}}.$$

The gradient is then

$$\begin{aligned}(\nabla D)_{\mathbf{m}} = \frac{\partial D}{\partial \varphi_{\mathbf{m}}} &= \sum_{\mathbf{h}} \frac{\partial d_{\mathbf{h}}}{\partial \varphi_{\mathbf{m}}} \bar{d}_{\mathbf{h}} + \sum_{\mathbf{h}} d_{\mathbf{h}} \frac{\partial \bar{d}_{\mathbf{h}}}{\partial \varphi_{\mathbf{m}}} \\ &= 2\Re\left\{ \sum_{\mathbf{h}} \frac{\partial d_{\mathbf{h}}}{\partial \varphi_{\mathbf{m}}} \bar{d}_{\mathbf{h}} \right\} \\ &= 2\Re\left\{ J^T \bar{d}_{\mathbf{h}} \right\}\end{aligned}$$

for the Jacobian J given by

$$J_{\mathbf{h},\mathbf{m}} = \frac{\partial d_{\mathbf{h}}}{\partial \varphi_{\mathbf{m}}}.$$

The Jacobian can also be used, as in Sayre & Toupin (1975), to construct a first-order approximation to the Hessian.

It will be shown that the Jacobian of a particular system decomposes into a sum of matrix representations of elementary vector operations. These elementary vector operations can be calculated quickly using the fast Fourier transform.

Consider the so-called Hadamard product $\times$ of two vectors $\mathbf{u}$ and $\mathbf{v}$: $(\mathbf{u} \times \mathbf{v})_{\mathbf{h}} = \mathbf{u}_{\mathbf{h}} \mathbf{v}_{\mathbf{h}}$. Fixing the vector $\mathbf{u}$, we see that the vector operation corresponding to the Hadamard product with $\mathbf{u}$ is given by multiplying the vector $\mathbf{v}$ with a matrix U that has the values $\mathbf{u}_{\mathbf{h}}$ along the diagonal and zero elsewhere. Explicitly, we have $U\mathbf{v} = \mathbf{u} \times \mathbf{v}$. Such a matrix will be called a matrix representation of the operation. Another common vector operation is the inversion $\sigma(\mathbf{v})_{\mathbf{h}} = \mathbf{v}_{-\mathbf{h}}$. The matrix representing the inversion operation will be denoted T. The two other vector operations necessary in the Jacobian decomposition are convolution $\mathbf{u} * \mathbf{v}$ and cross correlation $[\mathbf{u}, \mathbf{v}]$. Convolution and cross correlation are given by

$$\begin{aligned}\mathbf{u} * \mathbf{v} &= \sum_{\mathbf{k}} \mathbf{u}_{\mathbf{k}} \mathbf{v}_{\mathbf{h}-\mathbf{k}} \\ [\mathbf{u}, \mathbf{v}] &= \sum_{\mathbf{k}} \mathbf{u}_{\mathbf{k}} \mathbf{v}_{\mathbf{h}+\mathbf{k}}.\end{aligned}$$

Each of these operations can be calculated quickly using the fast Fourier transform.

The Jacobian is given by the formula

$$\frac{\partial d_{\mathbf{h}}}{\varphi_{\mathbf{m}}} = i\delta(\mathbf{h}-\mathbf{m}) F_{\mathbf{h}} - 2ib_{\mathbf{h}} F_{\mathbf{m}} F_{\mathbf{h}-\mathbf{m}} - 3ic_{\mathbf{h}} F_{\mathbf{m}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}-\mathbf{m}},$$

where $\delta(\mathbf{x}) = 1$ if and only if $\mathbf{x} = 0$. Therefore, if B, C and F denote the diagonal matrices with $b_{\mathbf{h}}$, $c_{\mathbf{h}}$ and $F_{\mathbf{h}}$ along the diagonal, respectively, and I denotes the identity matrix, we have that

$$J^T \mathbf{v} = iF[I - 2TX_1 - 3TX_2]\mathbf{v},$$

where X_1 and X_2 are matrix representations of particular cross correlations. If $\mathbf{f}$ denotes the vector of structure factors $F_{\mathbf{h}}$ and $\mathbf{k} = \mathbf{f} * \mathbf{f}$, then

$$\begin{aligned}(X_1\mathbf{v})_{\mathbf{m}} &= [B\mathbf{v}, \mathbf{f}]_{\mathbf{m}} \\ (X_2\mathbf{v})_{\mathbf{m}} &= [C\mathbf{v}, \mathbf{k}]_{\mathbf{m}}.\end{aligned}$$

Consequently, the action of J^T on some vector $\mathbf{v}$ is given by the operations:

$$J^T \mathbf{v} = i\mathbf{f} \times \{\mathbf{v} - 2\sigma([\mathbf{b} \times \mathbf{v}, \mathbf{f}]) - 3\sigma([\mathbf{c} \times \mathbf{v}, \mathbf{k}])\}.$$

In particular, if $\mathbf{v_h} = \bar{d}_{\mathbf{h}}$, the above gives the gradient of D. Thus, storing only three vectors **b**, **c**, **f**, the gradient can be calculated with standard operations.

It is interesting to point out that when Friedel's law holds the cross correlation and inversion operations are no longer necessary and can be replaced by equivalent convolutions as in Sayre & Toupin (1975).

We would like to thank David Sayre and Richard Toupin for helpful discussions throughout the course of this work. After this work was in the press, the authors became aware of the work of Hendrickson & Sheriff (1987) concerning the Bijvoet-difference Fourier. In particular, equation (2) of §1.3 is closely related to equation (8) of the theoretical analysis of Hendrickson & Sheriff (1987).

References

Abrahams, J. & Leslie, A. (1996). *Acta Cryst.* D**52**, 30–42.
Brodersen, D., La Fortelle, E. de, Vonrhein, C., Bricogne, G., Nyborg, J. & Kjeldgaard, M. (2000). *Acta Cryst.* D**56**, 431–441.
Chen, Y. & Su, W. (2000). *Acta Cryst.* A**56**, 127–131.
Doublié, S., Bricogne, G., Gilmore, C. & Carter, C. Jr (1995). *Structure*, **3**, 17–31.
Esnouf, R. M. (1997). *J. Mol. Graphics*, **15**, 132–134.
Fan, H.-F., Han, F.-S. & Qian, J.-Z. (1984). *Acta Cryst.* A**40**, 495–498.
Fan, H.-F., Han, F.-S., Qian, J.-Z. & Yao, J.-X. (1984). *Acta Cryst.* A**40**, 489–495.
Foadi, J., Woolfson, M., Dodson, E., Wilson, K., Yao, J. & Zheng, C. (2000). *Acta Cryst.* D**56**, 1137–1147.
Guo, D., Blessing, H. & Langs, D. (2000). *Acta Cryst.* D**56**, 451–457.
Hendrickson, W. (1991). *Science*, **254**, 51–58.
Hendrickson, W. A. & Sheriff, S. (1987). *Acta Cryst.* A**43**, 121–125.
Kraut, J. (1968). *J. Mol. Biol.* **35**, 511–512.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **277**, 173–207.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Refaat, L., Tate, C. & Woolfson, M. (1995). *Acta Cryst.* D**51**, 1036–1040.
Retailleau, P., Yin, Y., Huang, X., Hu, M., Roach, J., Bricogne, G., Roversi, P., Vonrhein, C., Blanc, E., Sweet, R. & Carter, C. Jr (2001). *Acta Cryst.* Submitted.
Rothbauer, R. (1980). *Acta Cryst.* A**36**, 27–32.
Rothbauer, R. (2000). *Z. Kristallogr.* **215**, 157–168.
Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.
Sayre, D. (1972). *Acta Cryst.* A**28**, 210–212.
Sayre, D. (1974). *Acta Cryst.* A**30**, 180–184.
Sayre, D. & Toupin, R. (1975). *Acta Cryst.* A**31**, S20.
Shiono, M. & Woolfson, M. (1991). *Acta Cryst.* A**47**, 526–533.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.
Woolfson, M. (1958). *Acta Cryst.* **11**, 277–283.
Zhang, K. & Main, P. (1990). *Acta Cryst.* A**46**, 377–381.