Acta Cryst. (1996). A**52**, 629–633

Discrete Hilbert Transforms and the Phase Problem in Crystallography

A. F. Mishnev

Latvian Institute of Organic Synthesis, LV 1006 Riga, Latvia. E-mail: mishnevs@osi.lanet.lv

(Received 21 December 1995; accepted 17 February 1996)

Abstract

With analogy to optical-image reconstruction theory, the role of 'in-between' structure factors (*i.e.* those having half-integral Miller indices) has been investigated in the context of the crystallographic phase problem. It is shown that in-between structure factors can be incorporated into the phasing process by means of the autocorrelation function (the Patterson for a single unit cell). Three-dimensional discrete Hilbert transforms (DHTs) that are obtained express the in-between structure factors in terms of standard ones. When the casual Fourier transform condition is satisfied, a DHT applied to an intensity function provides twofold intensity oversampling in the reciprocal space.

1. Introduction

It has long been recognized that, if the magnitudes of the Fourier transform of a crystal unit cell could be observed continuously in reciprocal space (*i.e.* not only at the discrete Bragg locations), then the solution to the phase problem would be greatly facilitated (Sayre, 1952; Buerger, 1960; Sayre, 1991). This situation is, in fact, met in optics, when the whole Fourier transform pattern is obtained. The literature on image recovery in two dimensions is rich with the contributions of many authors on this topic (Stark, 1987; Bates & Mnyama, 1986, and references therein). Indeed, any normally two-dimensional image having a finite envelope can be uniquely reconstructed (Hayes, 1982; Bates, 1982). In fact, the whole Fourier transform pattern is not needed. Only magnitudes at in-between (having half-spacing) and normal sampling points are important. Magnitudes at any other point can then be calculated by means of Shannon's sampling theorem (Shannon, 1949; Jerri, 1977).

In its application to crystallography, Sayre (1952) pointed out that knowledge of the intensity at half-integral Miller indices is equivalent to having the signs of structure factors in the case of a centrosymmetric crystal structure. A few years earlier, Boyes-Watson, Davidson & Perutz (1947) reported the determination of the signs of centrosymmetric zone structure factors by observation of intensity at non-integral Miller indices in the process of haemoglobin shrinkage with the variation of water content in the crystal. Later, Millane (1986, 1990) and Kim & Hayes (1991) directly transferred optical image reconstruction techniques to the field of crystallography.

The importance of in-between magnitudes in general image-reconstruction theory stimulated this work to solve the phase problem in crystallography. Recently, discrete Hilbert transforms (DHTs) have been derived for the crystallographic structure factor (SF) in the one-dimensional case (Mishnev, 1993). It was shown that DHTs are capable of expressing the SF at in-between points (*i.e.* having half-integral Miller indices) in terms of standard ones. In the present paper, DHTs are generalized to three dimensions (3D) and applied to both the SF and intensity.

DHTs obtained for SFs allow in-between structure factors $F(h/2, k/2, l/2)$ to be calculated using the normal $F(h, k, l)$. Trial phase estimates normally available from direct methods (DM) for smaller problems (Woolfson, 1987) and isomorphous replacement or anomalous-dispersion methods for macromolecules (Blundell & Johnson, 1978) can be used to calculate in-between intensities $|F(h/2, k/2, l/2|^2$. A new formula is derived here for the convenient computation of the autocorrelation function (*i.e.* the Patterson for a single unit cell) from in-between and normal intensities. Compared with the normal Patterson, the autocorrelation function (ACF) should be free of the peak superposition inherent to the Patterson and thus more easily interpretable. Properties of the ACF and its relation to the Patterson function give rise to new criteria for the quality of the phase set. Introduction of the ACF into the phasing process enables one to employ its physical properties (non-negativity, boundaries in space and magnitude) and thus to construct iterative ACF modification algorithms. Possibly, deconvolution of the ACF rather than the Patterson (Harrison, 1990) might give better results as the ACF is dependent on phases whereas the Patterson is not. Furthermore, DHT's applied to the intensity function in three-dimensions and the formula for in-between intensities in terms of normal ones are obtained. These formulae hold for atomic distributions, having ACF projections concentrated mainly in the I and III quadrants of the plane, thus producing crystallographic oversampling in reciprocal space.

2. 3D sampling theorems

The sampling theorems for the crystallographic SF and intensity are defined as follows

$$F(\mathbf{s}) = \int_0^a \int_0^b \int_0^c \rho(\mathbf{r}) \exp(2\pi i \mathbf{r}\mathbf{s})\, \mathrm{d}v \qquad (1)$$

$$I(\mathbf{s}) = \int_{-a}^{a} \int_{-b}^{b} \int_{-c}^{c} A(\mathbf{u}) \exp(2\pi i \mathbf{u}\mathbf{s})\, \mathrm{d}\mathbf{u}, \qquad (2)$$

where s is a continuous 3D variable in reciprocal space and $A(\mathbf{u})$ is the ACF for a single unit cell. From (1) and (2), it can be seen that the SF and intensity spectra are limited to the unit cell and twice the unit cell, respectively.

The sampling theorem (Shannon, 1949; Jerri, 1977) states that a band-limited function can be specified exactly at any point by its sampled values, taken at regular intervals, which do not exceed Nyquist spacing. Inspection of (1) and (2) shows that both $F(\mathbf{s})$ and $I(\mathbf{s})$ can be considered as band-limited functions, whose Fourier transforms vanish outside the unit cell. The Nyquist spacing, or maximal-allowed interval between samples, in our case is 1/2 for all three variables. Thus, the sampling theorem for $F(\mathbf{s})$ and $I(\mathbf{s})$ has one and the same form, and for the SF it is given by Brillouin (1962) and Petersen & Middleton (1962):

$$F(s_1, s_2, s_3) = \sum_{h,k,l} F(h/2, k/2, l/2) \times \operatorname{sinc}(2s_1 - h) \operatorname{sinc}(2s_2 - k) \operatorname{sinc}(2s_3 - l) \qquad (3a)$$

with the notation $\operatorname{sinc}(x) = \sin(x)/x$. The summation is taken over all positive and negative integers h, k, l.

By setting s_3 to integral-valued L, we obtain the corresponding sampling theorem for reciprocal-space sections

$$F(s_1, s_2, L) = \sum_{h,k} F(h/2, k/2, L) \times \operatorname{sinc}(2s_1 - h) \operatorname{sinc}(2s_2 - k). \qquad (3b)$$

Likewise, a linear sampling theorem can be obtained by setting s_2 and s_3 to integral K and L, respectively:

$$F(s_1, K, L) = \sum_h F(h/2, K, L) \operatorname{sinc}(2s_1 - h). \qquad (3c)$$

From (3*a*)–(3*c*), one can see that interpolation of the SF at an arbitrary point in the reciprocal space requires the knowledge of both in-between and normal samples.

3. 3D discrete Hilbert transforms for the structure factor

It is well known that if a complex function of a real variable $f(x)$ has a Fourier transform $F(y)$ that vanishes for negative argument (causal Fourier transform), $f(x)$ satisfies the Hilbert transform (Toll, 1956; Wu & Ohmura, 1962)

$$f(x) = (1/\pi j) P \int_{-\infty}^{\infty} f(y)/(y - x)\, \mathrm{d}y,$$

where P denotes the Cauchy principal value and $j = (-1)^{1/2}$. By definition, the SF satisfies the causal Fourier transform condition and one can apply Hilbert transforms to the sampling expansion (3*a*). From Papoulis (1968) and Kramer (1973), the Hilbert transform of the sinc function

$$f(s) = [\sin \pi(2s - h)]/\pi(2s - h)$$

is

$$g(s) = -[1 - \cos \pi(2s - h)]/\pi(2s - h)$$

and, after application of the transformation to (3*a*)–(3*c*) for each variable, we obtain the following discrete Hilbert transform expressions:

$$F(s_1, K, L) = (-1/j) \sum_h F(h/2, K, L) \times [1 - \cos \pi(2s_1 - h)]/[\pi(2s_1 - h)], \qquad (4a)$$

{further with the notation $\operatorname{cosc}(x) = [1 - \cos(x)]/x$}

$$F(s_1, s_2, L) = -\sum_{h,k} F(h/2, k/2, L) \times \operatorname{cosc}(2s_1 - h) \operatorname{cosc}(2s_2 - k) \qquad (4b)$$

and

$$F(s_1, s_2, s_3) = (1/j) \sum_{h,k,l} F(h/2, k/2, l/2) \times \operatorname{cosc}(2s_1 - h) \operatorname{cosc}(2s_2 - k) \times \operatorname{cosc}(2s_3 - l). \qquad (4c)$$

Now, by setting $s_1 = h'/2$ with h' integer in (4*a*), one obtains

$$F(h'/2, K, L) = (-1/j) {\sum_h}' F(h/2, K, L) \times [1 - (-1)^{h'-h}]/[\pi(h' - h)], \qquad (5a)$$

where the prime on the summation sign indicates that terms with $h' = h$ are omitted. In the same way, using (4*b*)–(4*c*), the following discrete Hilbert transforms for two and three variables can be derived:

$$F(h'/2, k'/2, L) = -{\sum_{h,k}}' F(h/2, k/2, L) \times [1 - (-1)^{h'-h}]/[\pi(h' - h)] \times [1 - (-1)^{k'-k}]/[\pi(k' - k)] \qquad (5b)$$

and

Table 1. *Test of the DHT (equation 5c) for the cycloalanine structure*

$F(\mathrm{s}) = A(\mathrm{s}) + iB(\mathrm{s})$.

s_1	s_2	s_3	$A(\mathrm{s})$ True	$A(\mathrm{s})$ From (5*c*)	$B(\mathrm{s})$ True	$B(\mathrm{s})$ From (5*c*)
0.5	0.5	0.5	−18.10	−17.28	−11.5	−13.35
0.5	1.5	4.5	−2.87	−3.05	−3.02	−2.98
1.5	4.5	1.5	−14.23	−14.57	7.98	8.05
1.5	6.5	5.5	3.56	3.40	−4.65	−4.43
2.5	1.5	4.5	3.37	3.42	0.70	0.70
2.5	6.5	4.5	2.01	2.16	0.17	0.23
3.5	0.5	2.5	−6.58	−6.27	−2.29	−1.98
3.5	2.5	4.5	5.18	4.98	−6.37	−6.36
4.5	1.5	3.5	−12.02	−11.57	4.21	4.14
4.5	6.5	4.5	−2.72	−3.16	−0.58	−0.55
5.5	1.5	5.5	−1.39	−1.40	4.83	5.48
5.5	5.5	5.5	−4.58	−5.09	−3.62	−3.79
6.5	2.5	4.5	4.88	4.81	−4.22	−4.01
6.5	5.5	6.5	4.53	4.77	−0.43	−0.38

$$\begin{aligned} F(h'/2, k'/2, l'/2) = {} & (1/j) \sum_{h,k,l}{}' F(h/2, k/2, l/2) \\ & \times [1 - (-1)^{h'-h}]/[\pi(h' - h)] \\ & \times [1 - (-1)^{k'-k}]/[\pi(k' - k)] \\ & \times [1 - (-1)^{l'-l}]/[\pi(l' - l)]. \qquad (5c) \end{aligned}$$

Suppose that h', k', l' on the left-hand side of (5*c*) are odd, then, on the right-hand side, terms with odd h, k, l will cancel and only terms with even h, k, l will remain. Thus, (5*c*) gives the value of the in-between SF $F(H + 1/2, K + 1/2, L + 1/2)$ in terms of normal ones. In turn, (5*a*) and (5*b*) express $F(H + 1/2, K, L)$ and $F(H + 1/2, K + 1/2, L)$, respectively, in terms of SFs, having integral Miller indices. It should be noted that (5*a*) and (5*c*) are discrete Hilbert transforms in a classical sense as they link the real part of the SF with the imaginary parts of the other SFs and *vice versa*. However, (5*b*) expresses the real part of the SF through the real parts of the others. The same is true for the imaginary parts in (5*b*).

To verify the validity of the formulae obtained, the triclinic crystal structure of cyclo-L-alanyl-L-alanine (Sletten, 1970), having cell parameters $a = 5.1552$, $b = 8.0596$, $c = 4.6698$, $\alpha = 103.155$, $\beta = 103.680$, $\gamma = 97.587°$, space group $P1$, was used. SFs were calculated using real atomic form factors for ten non-H atoms, all treated as C atoms. Table 1 shows the real and imaginary parts of selected in-between SFs calculated from the atomic coordinates and by means of (5*c*).

4. Autocorrelation function

Having the estimates of in-between SFs, and hence intensities, we can incorporate them into the phasing process by means of the ACF. Unlike the normal Patterson, the ACF is not periodic and dies when interatomic vectors exceed maximal distances between atoms in a single unit cell.

As follows from the intensity definition (2), the ACF is the inverse Fourier transform of the intensity function

$$A(\mathbf{u}) = \int \int_{-\infty}^{\infty} \int I(\mathbf{s}) \exp(-2\pi i \mathbf{u}\mathbf{s})\, \mathrm{d}\mathbf{s}. \qquad (6)$$

By substituting the sampling expansion for $I(\mathbf{s})$,

$$\begin{aligned} I(s_1, s_2, s_3) = {} & \sum_{h,k,l} I(h/2, k/2, l/2) \operatorname{sinc}(2s_1 - h) \\ & \times \operatorname{sinc}(2s_2 - k) \operatorname{sinc}(2s_3 - l) \qquad (7) \end{aligned}$$

into (6) and performing the integration over the sinc functions, we obtain

$$\begin{aligned} A(u, v, w) = {} & (1/8V) \sum_{h,k,l} I(h/2, k/2, l/2) \\ & \times \cos \pi(hu + kv + lw). \qquad (8) \end{aligned}$$

Expression (8) gives a convenient way of calculating the ACF instead of by numerical integration by means of (6). Formula (8) looks similar to the Patterson function although (8) also contains in-between intensities. The fundamental difference between the ACF and the Patterson is that the former is dependent on phases whereas the Patterson is not. This feature of the ACF makes it useful in treating the phase problem.

Let us separate the ACF into phase-independent and phase-dependent parts as follows:

$$A(u, v, w) = (1/8)P(u, v, w) + D(u, v, w), \qquad (9)$$

where $P(u, v, w)$ is the conventional Patterson function and

$$\begin{aligned} D(u, v, w) = {} & (1/8V) \sum I(h/2, k/2, l/2) \\ & \times \cos \pi(hu + kv + lw) \qquad (10) \end{aligned}$$

is dependent on the phases of the normal SFs. The function $D(\mathbf{u})$ assumes either positive or negative values. For correct phases, it suppresses peaks arising from interactions between atoms in the central and surrounding unit cells and improves the first term in (9) for interactions within the cell. Neither $A(\mathbf{u})$ nor $D(\mathbf{u})$ has a direct correlation with $P(\mathbf{u})$. However, in so far as the role of $D(\mathbf{u})$ is to compensate positive Patterson peaks, there would be some certain correlation between $P(\mathbf{u})$ and $|D(\mathbf{u})|$. Moreover, a good phase set should produce an ACF map with a minimal amount of negativity and also with small ACF values at grid points where $P(\mathbf{u})$ is small. To test these considerations as potential figures of merit for the goodness of phase sets, the cycloalanine structure was used again. Table 2 shows the correlation coefficient (CC) between $P(\mathbf{u})$ and $|D(\mathbf{u})|$, the amount (SN) of negativity of $A(\mathbf{u})$ and SD, the sum of the $A(\mathbf{u})$ values at the points where the Patterson is close to zero, for phase sets with different absolute mean phase errors $\langle |\Delta\varphi| \rangle$. It can be seen that

Table 2. *Figures of merit (FOMs) for a good phase set based on the ACF and the* $D(\mathbf{u})$ *function*

CC is the correlation coefficient between $P(\mathbf{u})$ and $|D(\mathbf{u})|$, SN is the negative density of $A(\mathbf{u})$ and SD is the density in the zero-valued Patterson regions.

$\langle\lvert\Delta\varphi\rvert\rangle$ (°)	CC	SN	SD
0.0	0.731	−206.9	297.8
40	0.604	−8380.7	7735.4
60	0.516	−10523.4	11707.0
90	0.507	−10267.8	10112.3

Table 3. *Test of DHT (equation 15) for intensity*

s_1	s_2	L	$I(s_1, s_2, L)$ True	$I(s_1, s_2, L)$ From (15)
−0.5	5.5	0	127.6	134.6
5.5	2.5	0	47.8	46.5
−3.5	4.5	1	44.2	45.7
6.5	−6.5	1	82.6	85.2
−5.5	−4.5	2	126.7	131.2
0.5	0.5	2	877.3	881.5
−6.5	0.5	3	17.7	13.6
2.5	5.5	3	89.6	93.0
−5.5	−4.5	4	53.3	51.1
1.5	6.5	4	54.9	58.7
−4.5	1.5	5	175.2	170.5
2.5	5.5	5	36.0	33.5
4.5	6.5	6	198.7	197.2
6.5	−3.5	6	110.3	106.0

all three figures show a fairly good indication of the correct phase set.

5. 3D discrete Hilbert transforms for intensity

Experience in application of the DHT to SFs will now be useful in deriving the desired expression for intensity. Although mathematically the Hilbert transform can be obtained for any suitable function, intensity is a real function and its imaginary part has no physical meaning. Thus, the only reasonable form of the DHT for intensity is that equivalent to (5*b*). For simplicity of analysis, we start with the expression for the two-dimensional case. Then the DHT for intensity reads

$$\begin{aligned} I(h'/2, k'/2) = &-\sum_{h,k}{}' I(h/2, k/2) \\ &\times [1-(-1)^{h'-h}]/[\pi(h'-h)] \\ &\times [1-(-1)^{k'-k}]/[\pi(k'-k)], \quad (11) \end{aligned}$$

which can be rewritten in a clearer form:

$$\begin{aligned} &I(h'+1/2, k'+1/2) \\ &\quad = -(1/\pi^2)\sum_{h,k}{}' I(h,k)[(h'-h+1/2)(k'-k+1/2)]^{-1}, \end{aligned} \quad (12)$$

where h, h', k and k' all have integral values. It should be noted that the same expression can be obtained by direct discretization of 2D integral Hilbert transforms, as derived by Nieto-Vesperinas (1980). Now we have to answer the question how general is this formal expression, in so far as the 2D intensity function must satisfy the causal Fourier transform condition? According to the latter, its Fourier transform must vanish for negative arguments. The definition for 2D intensity is

$$I(s_1, s_2) = \int_{-a}^{a}\int_{-b}^{b} A(u,v)\exp[2\pi i(us_1 + vs_2)]\,\mathrm{d}u\,\mathrm{d}v. \quad (13)$$

The ACF $A(u, v)$ normally has spread to both positive and negative arguments (u, v). Using the centrosymmetric property of the ACF, we can rearrange the integral (13) as follows:

$$\begin{aligned} I(s_1, s_2) = &\,2\int_0^a\int_0^b A(u,v)\cos[2\pi i(us_1 + vs_2)]\,\mathrm{d}u\,\mathrm{d}v \\ &+ 2\int_0^a\int_0^b A(u,-v)\cos[2\pi i(us_1 - vs_2)]\,\mathrm{d}u\,\mathrm{d}v. \end{aligned} \quad (14)$$

From (14), it can be seen that for a general atomic distribution it is impossible to satisfy the causal Fourier transform condition. However, when the atomic arrangement is such that its ACF is localized in the I and III quadrants of the plane, the second integral in (14) vanishes and the DHT for intensity must hold. Another alternative when the causal Fourier transform condition for intensity can be satisfied is the case when the electron density $\rho(x, y)$ can be factorized in the form $\rho(x)\rho(y)$. By introducing the third Miller index as a parameter in (12), the 3D DHT for intensity can be obtained:

$$\begin{aligned} &I(h'+1/2, k'+1/2, L) \\ &\quad = -(1/\pi^2)\sum_{h,k}{}' I(h,k,L) \\ &\qquad \times [(h'-h+1/2)(k'-k+1/2)]^{-1}. \end{aligned} \quad (15)$$

To verify the validity of (15), we used a 3D model structure with six C atoms having coordinates $(x, y, z) = (0.1, 0.1, 0.4)$; $(0.2, 0.3, 0.3)$; $(0.4, 0.4, 0.2)$; $(0.5, 0.6, 0.7)$; $(0.7, 0.7, 0.6)$; $(0.8, 0.9, 0.9)$. The atoms have been assumed to be Gaussian. The reconstruction has been carried out with the origin peak removed from $A(u, v, w)$. After the reconstruction, the subtracted intensity was restored. The results of the recovery of intensity at half-integral Miller indices for this model structure are given in Table 3, which shows the in-between intensities calculated directly and by the use of DHT (15) for some selected reflections.

6. Discussion

So far, this research is mainly theoretical. The formulae have yet to be checked with real diffraction data.

Two properties of a general type, causality and bandlimitedness, applied to the structure factor, allowed the intensities to be brought into the phasing process by means of the autocorrelation function. When the causality condition is satisfied for the intensity function, the DHT led to a twofold intensity oversampling in reciprocal space. The causality in our case was expressed in projection. This condition is realistic only for some linear chain molecules.

Model numerical calculations on the ACF suggest that new figures of merit (FOMs) for judging the quality of the phase set can be constructed. Reliable FOMs, able to recognize the best set in multisolution direct methods when applied to large structures, are a great necessity. Likewise in protein crystallography, FOMs to characterize the quality of electron-density maps, in the process of phase extension and refinement, need to be fast and easy to calculate. Two efficient but rather time-consuming FOMs have been developed recently (Gilmore, Henderson & Bricogne, 1991; Mishnev & Woolfson, 1994). The first one is based on the evaluation of the log-likelihood gain, based upon entropy-maximization procedures. The second FOM exploits the expected characteristics of an electron-density map for proteins. The FOMs proposed in this paper are built on a different principle, being related to the Patterson function.

There are also other ways of fitting the ACF to the Patterson function in zero-valued regions or by applying a minimum-negativity constraint to the ACF. Least-square procedures for phase refinement could be designed on that basis.

The author is grateful to the Royal Society for partial support of this work.

References

Bates, R. H. T. (1982). *Optik*, **61**, 247–262.

Bates, R. H. T. & Mnyama, D. (1986). *Adv. Electron. Electron Phys.* **67**, 1–64.

Blundell, T. L. & Johnson, L. N. (1978). *Protein Crystallography*. New York: Academic Press.

Boyes-Watson, J., Davidson, E. & Perutz, M. F. (1947). *Proc. R. Soc. London Ser. A*, **191**, 83–132.

Brillouin, L. (1962). *Science and Information Theory*. New York: Academic Press.

Buerger, M. J. (1960). *Crystal-Structure Analysis*. New York/London: Wiley.

Gilmore, C. J., Henderson, A. N. & Bricogne, G. (1991). *Acta Cryst.* A**47**, 842–846.

Harrison, R. W. (1990). *Acta Cryst.* A**46**, 606–619.

Hayes, M. H. (1982). *IEEE Trans. Acoust. Speech Signal Process.* **30**, 140–154.

Jerri, A. (1977). *Proc. IEEE*, **65**, 1565–1596.

Kim, W. & Hayes, M. H. (1991). *Proc. IEEE*, **91**, 1765–1768.

Kramer, H. P. (1973). *J. Math. Anal. Appl.* **44**, 275–287.

Millane, R. P. (1986). *Biophys. J.* **49**, 60–62.

Millane, R. P. (1990). *J. Opt. Soc. Am.* **7**, 394–411.

Mishnev, A. F. (1993). *Acta Cryst.* A**49**, 159–161.

Mishnev, A. F. & Woolfson, M. M. (1994). *Acta Cryst.* D**50**, 842–846.

Nieto-Vesperinas, M. (1980). *Optik*, **56**, 377–384.

Papoulis, A. (1968). *Systems and Transforms with Applications in Optics*. New York: McGraw-Hill.

Petersen, D. P. & Middleton, D. (1962). *Inf. Control*, **5**, 279–323.

Sayre, D. (1952). *Acta Cryst.* **5**, 843.

Sayre, D. (1991). *Direct Methods of Solving Crystal Structures*, edited by H. Shenk, pp. 353–356. New York: Plenum Press.

Shannon, C. E. (1949). *Proc. IRE*, **37**, 10–21.

Sletten, E. (1970). *J. Am. Chem. Soc.* **92**, 172–176.

Stark, H. (1987). Editor. *Image Recovery: Theory and Application*. Orlando: Academic Press.

Toll, J. S. (1956). *Phys. Rev.* **104**, 1760–1770.

Woolfson, M. M. (1987). *Acta Cryst.* A**43**, 593–612.

Wu, T.-Y. & Ohmura, T. (1962). *Quantum Theory of Scattering*. New York: Prentice-Hall.