Acta Cryst. (1979). A35, 598–603

An Improved Transformation Procedure for Radial Distribution Function Analysis

By R. Lovell, G. R. Mitchell and A. H. Windle

Department of Metallurgy and Materials Science, Pembroke Street, Cambridge, England

(*Received* 21 *November* 1978; *accepted* 2 *March* 1979)

Abstract

Various methods of reducing the effect of termination error on radial distribution functions are briefly reviewed. A new approach is introduced in which the RDF is calculated by generating the Fourier transform at predetermined points. The resulting sampled transform is relatively free of termination ripple and the spacing of the points reflects the true resolution of the method for the particular, experimentally limited, $s_{\rm max}$. The optimum choice of sampling points is determined in relation to special data terminations such as at a zero value of the interference functions or at a peak or trough position. The effectiveness of the sampled transform routine in reducing termination error is demonstrated by applying it to prematurely truncated interference functions derived from scattering data from atactic polystyrene. The advantage of the sampled transform approach to RDF analysis is that it prevents a termination discontinuity in the interference function, such as is often unavoidable even when special care is taken to apply precise corrections to the data, from causing an obscuring ripple on the RDF. In fact, in the extreme, it enables generation of useful RDF's of glassy polymers from data which have been neither corrected nor normalized.

1. Introduction

In this paper we propose a method of calculating sampled radial distribution functions (RDF) from scattering data from amorphous materials. Its object is the minimization of the termination errors due to the finite range of the scattering vector, $s(=4\pi\sin\theta/\lambda)$, accessible by experiment. The method is illustrated with X-ray diffraction data for atactic polystyrene but it is of general applicability.

For X-rays, the RDF is calculated from fully corrected intensity data, $I(s)$, by the formula (Warren, 1969):

$$G(r)=4\pi r[\rho(r)-\rho_0]=\frac{2}{\pi}\int\limits_0^\infty sZ(s)\sin rs\,{\rm d}s,\qquad(1)$$

where $Z(s)=[kI(s)-\sum f^2(s)]/g^2(s)$ is the interference function, k is the scaling factor (to electron units), $\sum f^2(s)$ is the sum of independent scattering from one unit of composition and $g^2(s)$ is a sharpening function, usually $[\sum f(s)]^2$.

In practice the integral must have finite limits, $s_{\rm min}$ and $s_{\rm max}$ $(<4\pi/\lambda)$, which gives rise to errors in the RDF. The error due to loss of data between zero and $s_{\rm min}$ (typically in our experiments $s_{\rm min}=0\cdot25$ Å^{-1}) corresponds to variations of electron density with wavelengths in real space > 20 Å. These variations are normally examined in reciprocal space. For most materials, the first features in $sZ(s)$ are at about $s=1$ Å^{-1} or greater, and few errors are introduced into the RDF by extrapolating $sZ(s)$ smoothly to zero at $s=0$ (Cargill, 1971).

Termination at $s_{\rm max}$ causes two errors in the RDF:

(i) loss of resolution, since wavelengths in real space $<2\pi/s_{\rm max}$ are lost,

(ii) spurious ripples of wavelength $2\pi/s_{\rm max}$ caused by the discontinuity in $sZ(s)$ at $s_{\rm max}$.

These ripples are due to sharp peaks in the true RDF being convoluted with the cosine transform of the cut-off function, $h(s)$ (Waser & Schomaker, 1953; Warren & Mozzi, 1975). The cut-off function is given by

$$\begin{aligned}h(s)&=1,\quad 0<|s|<s_{\rm max}\\&=0,\quad |s|>s_{\rm max}\end{aligned}$$

and its cosine transform is

$$\frac{1}{\pi}\int\limits_0^\infty h(s)\cos rs\,{\rm d}s=\frac{\sin rs_{\rm max}}{\pi r}=\frac{s_{\rm max}}{\pi}\,{\rm sinc}\,rs_{\rm max},$$

which has its main peak at the origin and subsidiary peaks at $rs_{\rm max}=5\pi/2, 9\pi/2, 13\pi/2$, *etc.*

In polymers, the first and second nearest-neighbour distances are fixed by covalent bonds and give sharp peaks at about 1·5 and 2·5 Å. These are the usual source of termination ripple. However, faulty data correction or scaling (normalization) can also give a prominent peak or trough below 1 Å and hence be a source of ripple (Kaplow, Strong & Averbach, 1965).

0567-7394/79/040598-06$01.00

2. Established procedures which minimize termination error

2.1. *Subtraction of theoretical scattering*

In (1), the independent scattering, $\sum f^2(s)$, is subtracted from the scaled intensity since it is due to scattering within the constituent atoms, which is not usually of interest and would also give a large peak at very low r and consequently a marked termination ripple. {The peak is at the origin in $[\rho(r)-\rho_0]$ but displaced to low r in the function $4\pi r[\rho(r)-\rho_0]$ which is the RDF.} Subtraction of $\sum f^2(s)$ reduces the amplitude of $sZ(s)$ at $s_{\max}$ and hence reduces the ripple.

A further modification of this is to include known interatomic distances in the theoretical scattering (Konnert & Karle, 1973). This will further reduce $sZ(s)$ at $s_{\max}$. We have developed a similar procedure for polymers (Waring, Mitchell & Windle, 1979).

2.2. *Electronic RDF*

The sharpening factor $g^2(s)$ decays at high s and hence dividing by it causes the data at high s to be enhanced. Not only is this data usually the least reliable, but any discontinuity at $s_{\max}$ is increased by sharpening, giving an increased ripple.

Finbak (1949) proposed that the electronic RDF, calculated by taking $g^2(s) = 1$, was more reliable in studies of atomic liquids. With $g^2(s)$ constant, $sZ(s)$ decays more rapidly and termination is less of a problem. Sharply defined interatomic distances are, of course, broadened by the electron distribution around each atom.

This method was in fact applied in an early RDF analysis of polymers by Bjørnhaug, Ellefsen & Tønnesen (1954).

2.3. *Damping functions (convergence factors)*

This is the usual method of dealing with the finite range of data and was originally proposed by Bragg & West (1930). In (1), $sZ(s)$ is replaced by $sZ(s)M(s)$, where $M(s)$ decays monotonically from 1 at $s = 0$ to 0·1 or less at $s = s_{\max}$. Waser & Schomaker (1953) have analysed a number of different functions, all of which broaden the peaks in the RDF whilst reducing the spurious ripple. Since $M(s)$ and $g^2(s)$ are of somewhat similar shape, the ratio $M(s)/g^2(s)$ is often approximately constant over the range 0 to $s_{\max}$ and hence the resultant RDF is quite similar to the electronic RDF.

The two most common damping functions are the artificial temperature factor due to Bragg & West (1930);

$$M(s) = \exp\left(-\frac{2\cdot303\,s^2}{s_{\max}^2}\right),$$

and the σ termination function (Lanczos, 1966) applied to RDF's by Lorch (1969);

$$M(s) = \frac{s_{\max}}{\pi s}\sin\left(\frac{\pi s}{s_{\max}}\right).$$

The effects of these two functions have been compared by Wright & Leadbetter (1976) who favour the latter.

Any damping function must reduce the resolution below that inherent in the limited range of data.

2.4. *Extension of data*

Various methods of approximating the experimentally inaccessible data have been suggested. The simplest, though rarely used, is to extrapolate $sZ(s)$ smoothly to zero beyond $s_{\max}$. The form of the extrapolation would be determined by the shape of $sZ(s)$ near $s_{\max}$.

d'Anjou & Sanz (1978) have used the more easily applied method due to Shevchik (1972) where a damped sine wave is fitted by least squares to the last two oscillations of $sZ(s)$ and then continued to approximately $2s_{\max}$.

More involved methods using forward and back transformation have been used by Kaplow, Strong & Averbach (1965) [recently applied to polymers by Wang & Yeh (1978) and by Gupta & Yeh (1978)] and by Hansen, Knudsen & Carneiro (1975). In both cases the procedure is based on the assumption of a featureless RDF below $r = 1$ Å, although the second group of authors also assume a particular shape for the first two peaks in the RDF.

2.5. *Termination at $sZ(s) = 0$*

If the upper limit of integration is taken as the last point at which $sZ(s)$ crosses zero, then the discontinuity in $sZ(s)$ is removed, though discontinuities in its first and higher derivatives may still remain. As will be shown later, this gives a more rapidly decaying ripple and hence the RDF is improved. This termination point has been used by Wignall, Rothon, Longman & Woodward (1977).

3. Reduction of termination error by sampling

The treatments described above have taken little account of the fact that the resolution in the RDF is limited to $\pi/s_{\max}$, indeed the transform has often been calculated at much more closely spaced points. Dixon, Wright & Hutchinson (1977) have stressed the limited resolution and have used a fast Fourier transform (FFT) to calculate the RDF at points $\pi/s_{\max}$ apart. However, as will be shown later, these points ($r = n\pi/s_{\max}$) will frequently lie near the peaks and troughs of the termination ripple. In general, FFT's sample at twice this spacing (*i.e.* $2\pi/s_{\max}$, the period of the ripple) which would remove all evidence of the ripple. This, of course, reduces the resolution but will not necessarily diminish the errors.

We now develop a systematic treatment of sampled transforms which minimize errors while optimizing resolution.

The nature of termination error is first illustrated by considering some simple, though rather unrealistic, forms of $sZ(s)$ [$= x(s)$] and comparing their infinite transform, $X_0(r)$, and finite transform, $X_1(r)$:

$$\frac{\pi}{2}G(r) = X_o(r) = \int_0^{\infty} x(s) \sin rs \, ds,$$

$$X_1(r) = \int_0^{s_{\max}} x(s) \sin rs \, ds.$$

We then present a general treatment and show how this leads to a general method of sampling the transform.

3.1. *Error for various forms of x(s)*

3.1.1. *Constant:* $x(s) = 1$. The function to be transformed can approximate to this at high s if the normalization is performed badly.

The infinite and finite transforms are given by:

$$X_0(r) = \int_0^{\infty} \sin rs \, ds = \frac{1}{r},$$

$$X_1(r) = \int_0^{s_{\max}} \sin rs \, ds = \frac{1}{r}(1 - \cos rs_{\max}).$$

These are plotted in Fig. 1. It can be seen that X_1 oscillates as a damped cosine wave about X_0, *i.e.* the termination error is:

$$X_2(r) = X_0 - X_1 = \frac{1}{r}\cos rs_{\max}.$$

The two transforms are therefore identical at the points given by $r = (2n + 1)\pi/2s_{\max}$.

3.1.2. *Sine wave:* $x(s) = \sin as$. This illustration is closer to a well normalized interference function. Its infinite transform is:

$$X_0(r) = \int_0^{\infty} \sin as \sin rs \, ds$$

$$= \tfrac{1}{2}\int_0^{\infty} [\cos(a - r)s - \cos(a + r)s] \, ds$$

$$= \frac{\pi}{2}[\delta(a - r) - \delta(a + r)],$$

whereas the finite transform is

$$X_1(r) = \int_0^{s_{\max}} \sin as \sin rs \, ds$$

$$= \frac{s_{\max}}{2}[\operatorname{sinc}(a - r)s_{\max} - \operatorname{sinc}(a + r)s_{\max}].$$

This is plotted in Fig. 2.

As mentioned earlier, the effect of termination is to convolute the ideal transform (in this case δ-functions) with $(s_{\max} \operatorname{sinc} rs_{\max})/\pi$, the cosine transform of the cut-off function, $h(s)$.

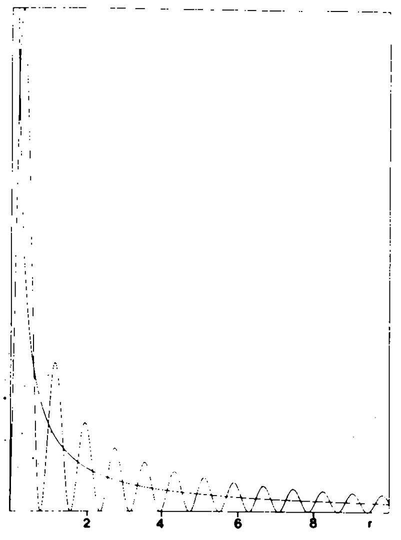


Fig. 1. The infinite and finite ($s_{\max} = 8{\cdot}0$) transform of $x(s) = 1{\cdot}0$.

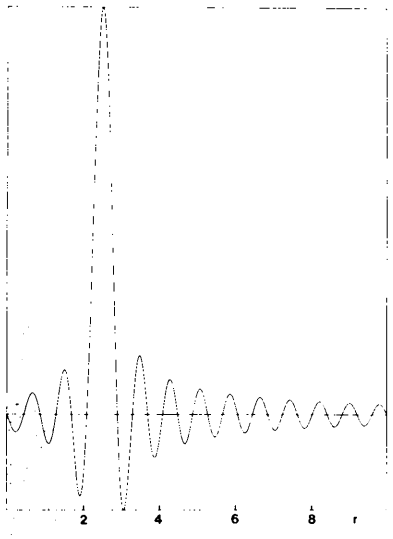


Fig. 2. The finite ($s_{\max} = 8{\cdot}0$) transform of $x(s) = \sin as$.

The expression for X_1 can be rewritten:

$$X_1(r) = -\left(\frac{r \sin as_{\max}}{r^2 - a^2}\right)\cos rs_{\max} + \left(\frac{a \cos as_{\max}}{r^2 - a^2}\right)\sin rs_{\max}. \qquad (2)$$

Thus, in general, there are two components: a cosine ripple decaying as $1/r$ at large r and a sine ripple decaying as $1/r^2$ at large r. The form of the finite transform simplifies for particular termination points:

(*a*) *Termination at a node, i.e.* $as_{max} = m\pi$. The first term in (2) is now identically zero, whereas the second (sine) term has zeros at

$$r = n\pi/s_{\max},$$

except for $n = m$, *i.e.* $r = a$.

(*b*) *Termination at a maximum or minimum, i.e.* $as_{max} = (2m + 1)\pi/2$. The second term is now identically zero and the first (cosine) term has zeros at

$$r = (2n + 1)\pi/2s_{\max},$$

except for $n = m$, *i.e.* $r = a$.

Termination at an arbitrary point will give finite sine and cosine components. The sine will dominate in the region from 0 to a, and the cosine above a.

3.1.3. *General function.* The termination error is given by:

$$X_2(r) = \int_{s_{\max}}^{\infty} x(s) \sin rs \, ds.$$

By assuming that $x(s)$ and all its derivatives are zero at $s = \infty$ and carrying out repeated partial integration, $X_2(r)$ may be expanded in the series (Erdélyi, 1956):

$$X_2(r) = x(s_{\max})\cos rs_{\max}/r - x'(s_{\max})\sin rs_{\max}/r^2 - x''(s_{\max})\cos rs_{\max}/r^3 + x'''(s_{\max})\sin rs_{\max}/r^4, \textit{ etc.} \qquad (3)$$

The assumption that all derivatives tend to zero must be satisfied by any real interference function since the peaks in the true RDF must have finite width.

From (3), we can see that the general termination error has cosine components dependent on the zeroth and even derivatives at $s_{\max}$ and sine components dependent on the odd derivatives at $s_{\max}$. However, some of these may be zero for particular termination points:

(*a*) *Termination at a node:* $x(s_{max}) = 0$. The first non-zero term in (3) is now a sine ripple decaying as $1/r^2$:

$$X_2(r) = -x'(s_{\max}) \sin rs_{\max}/r^2 - \ldots,$$

with the first cosine term decaying as $1/r^3$. The error at large r is therefore likely to be negligible. This is the justification for the method referred to in § 2.5.

(*b*) *Termination at a maximum or minimum:* $x'(s_{max}) = 0$. The first two non-zero terms are now cosine ripples decaying as $1/r$ and $1/r^3$:

$$X_2(r) = x(s_{\max}) \cos rs_{\max}/r - x''(s_{\max}) \cos rs_{\max}/r^3 + \ldots,$$

with the first sine term decaying as $1/r^4$. Hence, particularly at large r, the termination error is very close to zero at values of r corresponding to the zeros of $\cos rs_{\max}$, *i.e.*

$$r = (2n + 1)\pi/2s_{\max}.$$

3.1.4. *Decay of ripple in radial density function.* The decay rates at large r, derived above, all refer to the radial distribution function, $4\pi r[\rho(r) - \rho_0]$. It is, however, often preferable to plot the radial density function, $4\pi r^2[\rho(r) - \rho_0] = rG(r)$, particularly when looking at intermolecular order in the 5–30 Å range. In this latter case, the residual ripple will be correspondingly more prominent at large r, and that due to a finite value at $s_{\max}$ will maintain a constant amplitude.

3.2. *Best sampling points*

We can conclude from the above treatment that:

(i) the RDF with termination error oscillates about the true RDF with a period $2\pi/s_{\max}$,

(ii) for termination at a node, the error is predominantly a sine ripple,

(iii) for termination at a maximum or minimum, the error is predominantly a cosine ripple.

Hence, if we can choose to terminate at one of these and sample the transform at the zeros of the dominant term, we will minimize the errors. These two possible schemes are shown in Table 1. As can be seen, the second method gives the more rapidly decaying error. In practice, we may need to discard too much of the data (say > 20%) in order to terminate at one of these special points. In this case, we suggest that the best procedure is to sample at the zeros of the cosine ripple, $r = (2n + 1)\pi/2s_{\max}$, and minimize the sine ripple by averaging over three points:

$$X(r) = \tfrac{1}{4}[X(r_{i-1}) + 2X(r_i) + X(r_{i+1})]. \qquad (4)$$

Table 1. *Comparison of errors at different sampling points*

Terminated at	Sampled at	Approximate error in $X(r)$
Node $x(s_{\max}) = 0$	$r = n\pi/s_{\max}$	$\pm x''(s_{\max})/r^3$
Maximum or minimum $x'(s_{\max}) = 0$	$r = (2n + 1)\pi/2s_{\max}$	$\pm x'''(s_{\max})/r^4$
Arbitrary point	$r = n\pi/s_{\max}$	$\pm x(s_{\max})/r$
	$r = (2n + 1)\pi/2s_{\max}$	$\pm x'(s_{\max})/r^2$

4. Examples of sampled transforms

In this section the utility of sampled transforms in minimizing termination error is demonstrated with scattering data from atactic polystyrene.

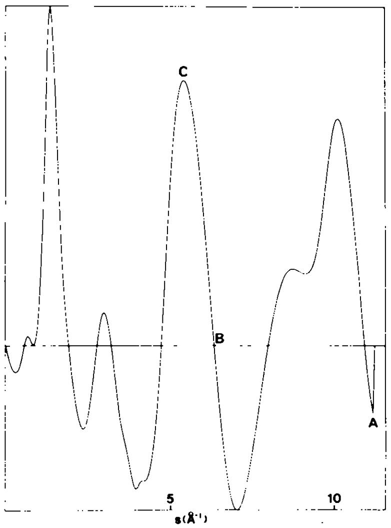
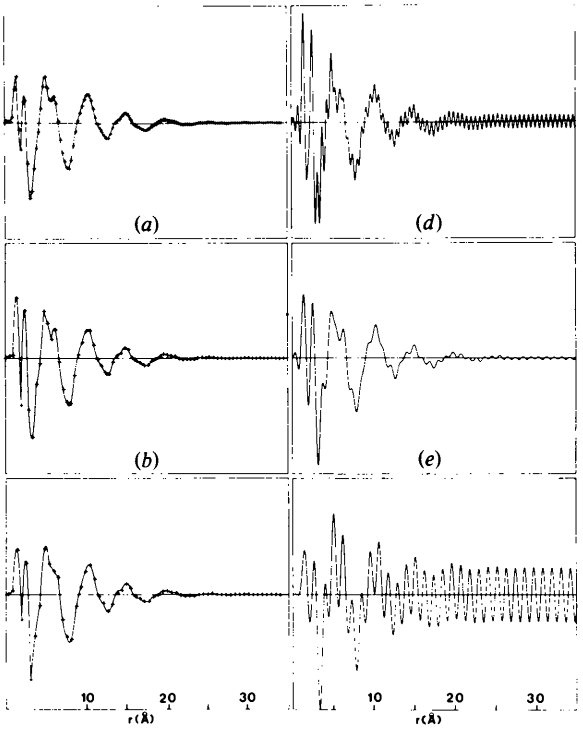


Fig. 3. Interference function for atactic polystyrene.



(c) (f)

Fig. 4. Transforms of the interference function of Fig. 3. (a), (b), (c) are sampled radial density functions for the termination points A, B and C (Fig. 3), whereas (d), (e) and (f) are the corresponding conventional radial density functions (+ sampled points). Vertical axis is $4\pi r^2[\rho(r) - \rho_0]$.

Fig. 3 is an interference function calculated from these data with (1), with $g^2(s) = [\sum f(s)]^2$. Marked on the diagram are three termination points A, B, C. A corresponds to the experimental $s_{\max}$ for these data, while B and C are premature terminations chosen to coincide with a node and a maximum.

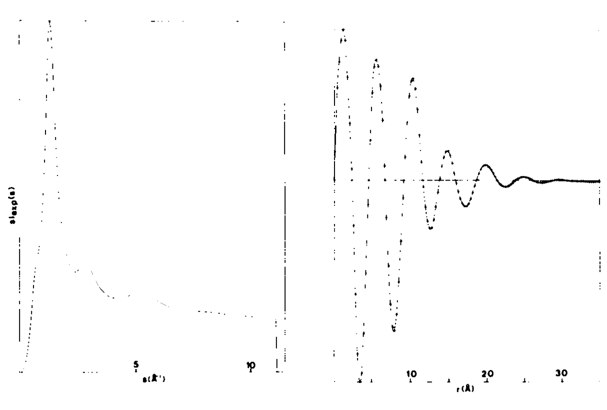
The radial density functions $\{4\pi r^2[\rho(r) - \rho_0]\}$ derived from the interference function for these termination points are shown in Fig. 4. The transforms are arranged in pairs; the left-hand member is discretely sampled at the appropriate points, the right-hand member is a continuous (conventional) transform. The transforms were calculated by Filon's (1928) method.

The three sampled transforms (Fig. 4a,b,c) are similar except for suggestions of a lower resolution corresponding to the smaller $s_{\max}$'s. The actual reduction in resolution due to premature termination at around $s_{\max}/2$ is surprisingly small. In this case it probably reflects the fact that the resolution of the RDF derived from termination A, which is at a general point is reduced by the necessary averaging (equation 4). There are no discernible termination errors, even at low r values.

The magnitude and period of the termination errors in the continuous transforms (Fig. 4d,e,f) clearly varies with the termination point. Termination at C (Fig. 4f) generates a transform dominated by the termination ripple, which is not apparent in the sampled version (Fig. 4c). The magnitude of the oscillation is reduced for termination at A and B, particularly for B. However these RDF's still contain considerable termination errors which are not visible in the corresponding sampled transforms.

5. General summary

The modification of the RDF analysis procedure by the incorporation of sampled transforms greatly reduces the problems associated with termination ripple. The



(a) (b)

Fig. 5. (a) Uncorrected intensity function from atactic polystyrene. (b) The sampled transform of (a).

only ways in which misleading peaks can appear in the RDF's produced in this way is as the result of bad data or of poor normalization. In the latter case, the peak(s) will be confined to the region below 1 Å.

Fig. 5 shows an extreme case. The intensity function, $sI_{\exp}(s)$, smoothed but neither corrected nor normalized, is shown in Fig. 5(*a*). This would, with a traditional transform, give a RDF swamped by termination ripple. The result of applying a sampled transform is shown in Fig. 5(*b*). Beyond 5 Å, the RDF is very similar to that generated from properly corrected and normalized data. At low values of r, however, it is dominated by the very large peak [centred on the origin in $\rho(r) - \rho_0$] which represents the distribution of vector lengths between electrons within the same atom. This peak obscures the first and second nearest-neighbour peaks. For polymers, these are, of course, fixed by covalent bonds and are of limited interest.

In demonstrating the feasibility of transforming unnormalized data, we are not suggesting that this would necessarily be a sensible approach. The point is that carrying out corrections for absorption, double scattering and the Compton component, to the degree of precision required for completely satisfactory normalization, is an exacting and time consuming task. Small uncertainties in the correction factors can lead to significant termination problems with conventional transforms. The corrections are discussed in more detail elsewhere (Waring, Mitchell & Windle, 1979).

By using sampled transforms, the stringent requirements of perfect normalization can be considerably relaxed. Also the fact that the first and second nearest neighbours give very sharp peaks in atomic (sharpened) RDF's for polymers means that the interference function still has appreciable amplitude at $s_{\max}$. Consequently, even with perfect normalization, termination error can still be troublesome unless sampled transforms are employed.

We suggest that the use of sampled transforms is the most appropriate method of minimizing the effect of termination error on RDF's. It is straightforward; it does not require judgement, either as to a correctly smoothed RDF prior to repeated forward and backward transformation as in the method of Kaplow, Strong & Averbach (1965), or as to the best form of a damping function to be applied to the interference function. In addition, the points generated on a RDF by the sampled transform method are at the spacing appropriate to the resolution as determined by the experimentally limited $s_{\max}$.

The authors acknowledge useful discussions with Dr A. Colebrooke at the inception of this work and with Mr J. R. Waring, Dr I. Voigt-Martin and Dr D. J. Ewins. They also thank Professor R. W. K. Honeycombe for enabling the work to take part in his Department and the Science Research Council for Grant GR/A 13387.

References

Anjou, A. d' & Sanz, F. (1978). *J. Non-Cryst. Solids*, **28**, 319–326.

Bjørnhaug, A., Ellefsen, Ø. & Tønnesen, B. A. (1954). *J. Polym. Sci.* **12**, 621–632.

Bragg, W. L. & West, J. (1930). *Philos. Mag.* **10**, 823–841.

Cargill, G. S. (1971). *J. Appl. Cryst.* **4**, 277–283.

Dixon, M., Wright, A. C. & Hutchinson, P. (1977). *Nucl. Instrum. Methods*, **143**, 379–383.

Erdélyi, A. (1956). *Asymptotic Expansions*, p. 47. New York: Dover.

Filon, L. N. G. (1928). *Proc. R. Soc. (Edinburgh)*, **49**, 38–47.

Finbak, C. (1949). *Acta Chem. Scand.* **3**, 1279–1292.

Gupta, M. R. & Yeh, G. S. Y. (1978). *J. Macromol. Sci. Phys.* **15**, 119–137.

Hansen, F. Y., Knudsen, T. S. & Carneiro, K. (1975). *J. Chem. Phys.* **62**, 1556–1565.

Kaplow, R., Strong, S. L. & Averbach, B. L. (1965). *Phys. Rev. Sect. A*, **138**, 1336–1345.

Konnert, J. H. & Karle, J. (1973). *Acta Cryst.* A**29**, 702–710.

Lanczos, C. (1966). *Discourse on Fourier Series*, p. 65. London: Oliver & Boyd.

Lorch, E. (1969). *J. Phys. C*, **2**, 229–237.

Shevchik, N. J. (1972). PhD Thesis, Harvard Univ.

Wang, C. S. & Yeh, G. S. Y. (1978). *J. Macromol. Sci. Phys.* **15**, 107–118.

Waring, J. R., Mitchell, G. R. & Windle, A. H. (1979). *J. Mater. Sci.* Submitted.

Warren, B. E. (1969). *X-ray Diffraction*, p. 137. Reading, Mass: Addison-Wesley.

Warren, B. E. & Mozzi, R. L. (1975). *J. Appl. Cryst.* **8**, 674–677.

Waser, J. & Schomaker, V. (1953). *Rev. Mod. Phys.* **25**, 671–690.

Wignall, G. D., Rothon, R. N., Longman, G. W. & Woodward, G. R. (1977). *J. Mater. Sci.* **12**, 1039–1049.

Wright, A. C. & Leadbetter, A. J. (1976). *Phys. Chem. Glasses*, **17**, 122–145.