J. Appl. Cryst. (1998). **31**, 609–619

IFO: a Program for Image-Reconstruction-Type Calculation of Atomic Distribution Functions for Disordered Materials

V. Petkov[a]*† and R. Danev[b]

[a]*Fachbereich Physik, University of Rostock, Rostock 18051, Germany, and* [b]*Department of Solid State Physics, Sofia University, Sofia 1126, Bulgaria. E-mail: petkov@phys.uni-sofia.bg*

(*Received 14 October 1997; accepted 5 February 1998*)

Abstract

IFO is a newly developed computer program which transforms experimental structure factors for disordered materials to atomic distribution functions by employing an image-reconstruction-type technique. The transformation is carried out through a Monte Carlo search for an atomic distribution function which is a smooth, *i.e.* free of termination and other spurious ripples, real-space image of a given experimental structure factor. *IFO* has been tested on a number of data-sets and its efficiency has been demonstrated. The program is considered to be a useful tool for controlling and improving the quality of experimental structure functions for disordered materials.

1. Introduction

The structure of a disordered material such as glass, liquid, melt or gel is commonly described in terms of the radial distribution function (RDF), which is a directionally averaged representation of the atomic arrangement and gives the probability of finding an atom at a distance r from a reference atom (Klug & Alexander, 1974; Wagner, 1978). RDFs are usually obtained by diffraction experiments employing X-rays, neutrons or electrons in the following manner. A particular diffraction experiment is carried out and the spectrum obtained is subjected to appropriate corrections (Wagner, 1978; Thijsse, 1984; Muñoz *et al.*, 1988; Cockayne & McKenzie, 1988) in order to derive the so-called structure factor, $S(q)$, which is the only structure-dependent part of the recorded intensities. From the experimental $S(q)$ data the so-called reduced RDF, $G(r) = 4\pi r[\rho(r) - \rho_o]$, where $\rho(r)$ and ρ_o are the local and average atomic number densities, respectively, is obtained by carrying out a Fourier transformation as follows

$$G(r) = (2/\pi)\int_0^\infty q[S(q) - 1]\sin(qr)\,\mathrm{d}q. \qquad (1)$$

Here r is the radial distance and q the magnitude of the wave (scattering) vector. Once $G(r)$ has been obtained, some other frequently used RDFs, such as the pair distribution function $g(r) = [\rho(r)/\rho_o]$ or the total RDF = $4\pi r^2\rho(r)$, are readily derived. The generation of RDFs involves a numerical transformation of data from reciprocal [$S(q)$ data] into real space [$G(r)$ data] which should pose no problems provided all quantities on the right-hand side of (1) are exactly known. In practice, however, certain problems make direct application of the Fourier transformation of (1) rather inefficient. One of the problems stems from the fact that no diffraction experiment can yield $S(q)$ data covering the region from $q = 0$ to q approaching infinity. On one hand, $S(q)$ data at very low values of q are difficult to collect. On the other, even time-of-flight neutron diffraction experiments (Grimley *et al.*, 1990), wide-angle X-ray diffraction experiments employing hard X-rays (Poulsen *et al.*, 1995) and energy-dispersive X-ray diffraction (Petkov, 1995) yield $S(q)$ data at q values no higher than 30–40 Å^{-1}. Thus the set of $S(q)$ data available in practice always turns out to be confined to some finite region in reciprocal space. While the $S(q)$ data 'missing' at very low values of q can be, without much harm, derived by some wisely devised extrapolation towards $q = 0$ (Waseda, 1980), it is not possible to restore unambiguously the data 'missing' at higher q values. As predicted by the theory of Fourier transformations and as is shown by Warren & Mozzi (1975), the truncation of $S(q)$ data at some q_{max} value spoils the resolution of the corresponding RDF and, furthermore, gives rise to appreciable high-frequency false oscillations known as termination ripples.

Another problem is that the experimental $S(q)$ data at higher q values, even though extracted from diffraction spectra obtained using the most powerful radiation sources currently available (see Figs. 4 in Poulsen *et al.*, 1995; Jal *et al.*, 1991), are often of poor statistical accuracy. Thus, when the Fourier transformation of (1) is directly carried out the statistical noise in $S(q)$ propagates in the corresponding RDF and corrupts its fine features (see Figs. 1 and 2 presented later).

Yet another problem is that if some systematic errors are present in the experimental $S(q)$ data, considerable

† Permanent address: Department of Solid State Physics, Sofia University, Sofia 1126, Bulgaria.

spurious oscillations appear in the corresponding RDF obtained by a direct Fourier transformation (Klug & Alexander, 1974). These false oscillations may substantially distort the shape of RDFs and spoil the accuracy of the estimates of some important structural parameters, in particular those for the first coordination numbers and distances.

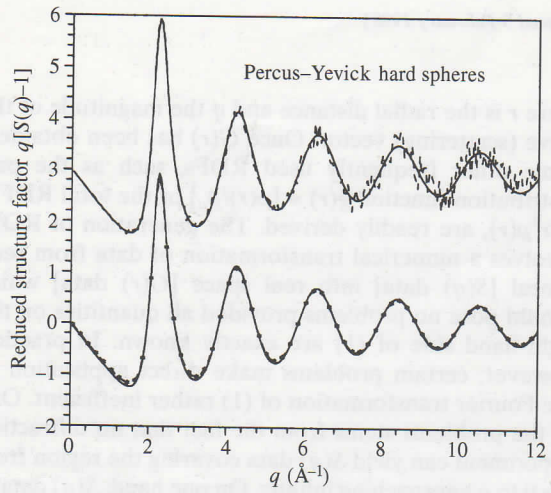

Fig. 1. Structure factors for Percus–Yevick hard-sphere liquid. Theoretical error-free data (full line); theoretical data with added statistical noise (broken line; upper part); reconstructed data (symbols; lower part).

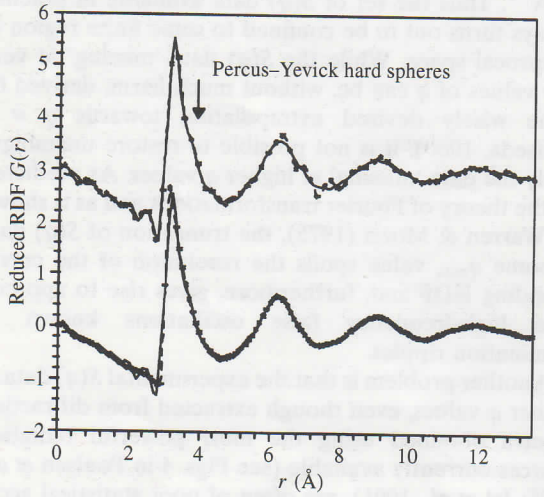

Fig. 2. Reduced RDFs for Percus–Yevick hard-sphere liquid obtained by: direct Fourier transformation of the error-free structure factor of Fig. 1 (thin full line); direct Fourier transformation of the noise-containing structure factor of Fig. 1 (symbols; upper part); the image-reconstruction technique presently considered (symbols; lower part). The nonphysical shoulder of the first peak in the conventionally derived $G(r)$ is marked by an arrow.

Several computational procedures have been worked out and put into practice in an attempt to solve the long-standing problems described above. To lessen the effects of statistical noise and $S(q)$ termination on RDFs a modifying function, $M(q)$, usually of the type $M(q) = \sin(\pi q/q_{max})/(\pi q/q_{max})$ or $M(q) = \exp(-\text{const}q^2)$, is often included in the direct Fourier transformation of (1). However, as shown by Thijsse (1984) the price for using a modification function is an additional obscuring of the fine features of RDFs (see also Figs. 11–13 presented later). To eliminate the systematic errors in $S(q)$ and their artifacts in the corresponding RDF, corrective procedures based on repeated direct Fourier transformations of artificially extended $S(q)$ data have been suggested (Kaplow *et al.*, 1965; Konnert & Karle, 1973). However, only the spurious oscillations at low r values are removed by these corrective procedures while the rest of the RDF data remain uncorrected (see Fig. 13 presented later). Other procedures aimed at improving the normalization of $S(q)$ data (Cumbrera *et al.*, 1995), obtaining smoothed RDFs by using analytical (Korsunsky & Naberukhin, 1980), numerical (Krylov & Vvedenskii, 1995) or computer-simulated (Howe *et al.*, 1996) approximations to the direct Fourier transformation of (1) have also been developed. None of them, however, has proved to be so efficient as to combat satisfactorily all the problems discussed so far. Thus it is still difficult to obtain reliable RDFs on the basis of $S(q)$ data that are incomplete (terminated at some q_{max} value), often, if not always, noisy at high q values and that possibly comprise some systematic errors. Generally speaking, similar tasks are frequently encountered in modern physics when a physical quantity (in our a case an RDF) is to be reconstructed from an underdetermined set of experimental data [in our case $S(q)$ data]. Practice has shown that such tasks can be quite successfully tackled by employing image-reconstruction-type techniques such as the maximum-entropy (MEM) method (Linden, 1995). In fact, MEM, or its modifications, is nowadays widely employed in applied crystallography (Müller & Hansen, 1994; Schotte *et al.*, 1995; Lin & Tsao, 1996; Schleger *et al.*, 1997) and some MEM applications to structure data of disordered materials are also known (Wei, 1986; Muñoz *et al.*, 1988; Jal *et al.*, 1991; Allen & Howe, 1992). However, no thoroughly tested, generally applicable and easy-to-use procedure for image-reconstruction-type processing of $S(q)$ and RDF data has been developed until now. It is the purpose of the present work to develop such a computational procedure, which, when necessary, could be applied with confidence to real data.

2. Theoretical background

The structure data for disordered materials are primarily obtained in reciprocal space as a set of $S(q_i)$ (usually

equidistant) data points, where the q_i's are the magnitudes of the scattering vectors at which diffracted intensities have been recorded. Accounting for the fact that the measured $S(q)$ data reflect the actual atomic arrangement of the disordered material being investigated, which, as already discussed, is usually described in terms of the RDF, one can rewrite (1) in the following two equivalent relations:

$$q_i[S(q_i) - 1] \equiv F(q_i) = \sum_{k=1}^{N} \sin(q_i r_k) G(r_k) \Delta r, \quad (2a)$$

$$F(q_i) = \sum_{k=1}^{N} T_{ik} G(r_k), \quad (2b)$$

where the Fourier integral has been approximated by a discrete summation over N properly selected data points r_k (usually $r_k = k\Delta r$; $k = 1, 2 \ldots N$; $\Delta r = \pi/q_{max}$), Δq is the sampling interval in reciprocal space and T_{ik} is a shorthand notation for the transformation matrix from $G(r_k)$ to $F(q_i)$. The main goal of the image-reconstruction technique is to restore the real RDF or, at least, to derive an RDF that is maximally close to the real one, *i.e.* that is free of termination and spurious ripples and other experimental or computational artifacts, from incomplete noisy experimental structure-factor data, which possibly comprise systematic errors, provided the sought $G(r_k)$ and the available $F(q_i)$ data are related by a nonlinear (Fourier sine) transformation (T_{ik}). To achieve this goal the theory of information is evoked. Since RDFs give nothing else but a time- and space-averaged picture of the atomic-scale structure of disordered materials, one may well consider them as statistical quantities. Formally, the information content of any statistical quantity regarded as a set of N positive data points $\{f_i\}$, which are to be determined, can be quantified by the information entropy, H, defined as

$$H = -\sum_{i=1}^{N} p_i \ln(p_i/b_i) - p_i + b_i, \quad (3)$$

where $p_i = f_i / \sum f_i$ and b_i is our prior knowledge about some features which $\{p_i\}$ (*i.e.* $\{f_i\}$) must satisfy (Skilling & Bryan, 1984; Skilling, 1988; Soper, 1990). It can be readily shown that our knowledge (information) about quantity $\{f_i\}$ is maximal when H is maximal. This information entropy lies at the root of the application of several image-reconstruction techniques, including MEM. Let us suppose that a set of experimental data $\{a_k\}$ that is related to $\{f_i\}$ through a transformation matrix T_{ik}, *i.e.*

$$a_k = \sum_{i=1}^{N} T_{ik} f_i, \quad (4)$$

is available. MEM selects a set of p_i's (or a corresponding set of $\{f_i\}$) for which H is maximal subject to a constraint that when this set is supplied to the right-hand side of (4) the data computed agree well with the experimentally measured data. Mathematically, this can be expressed by minimization of some functional, Q,

$$Q = \lambda \chi^2 - H, \quad (5)$$

where

$$\chi^2 = \sum_k (a_k^{exp} - a_k^{comp})^2 \quad (6)$$

and λ is a Lagrange-type multiplier predetermining how closely the derived $\{p_i\}$ (*i.e.* $\{f_i\}$) should reproduce the available experimental data $\{a_k\}$. The theoretical background and justification for the application of MEM to reconstruction of RDFs have been given by Wei (1986) and we will not repeat these details here. The application of MEM according to the approach outlined by Wei (1986), however, revealed that unreliable results may be obtained unless the experimental $S(q)$ data are good quality (Muñoz *et al.*, 1988). This drawback of MEM is partially due to the fact that when the experimental data are not good quality the respective Lagrange multiplier, λ, is chosen very close to zero. In such a case the minimization of the functional Q, *i.e.* the maximization of entropy H, yields a solution which closely reproduces our prior information, *i.e.* a set of $\{p_i\} \equiv \{b_i\}$. When, however, the prior information is scarce or not available at all, which occurs quite often in practice, the solution is the most disordered among all possible solutions, *i.e.* it is simply a set of uniformly distributed $\{p_i\}$ (Skilling & Bryan, 1984; Linden, 1995). Thus, when poor-quality $S(q)$ data are available MEM tends to yield uniform RDFs, *i.e.* RDFs whose characteristic features are smeared out, and thus produces not very reliable results. Since MEM, as it stands, does not take into account the underlying continuity ('smoothness') of the atomic arrangement in disordered materials but rather forces the reconstructed RDFs to be more or less featureless, Soper (1990) has argued that the information entropy, H, is not the function that most suitably reflects the physical information contained in RDFs. He postulated that the real RDFs being sought are most likely to be the least noisy (the smoothest) among all other possible solutions. To quantify the noise in the $G(r)$ data, Soper (1990) introduced a noise function I:

$$I = \sum_{k=1}^{N} [G''(r_k)]^2 / G'(r_k), \quad (7)$$

where $G''(r_k)$ and $G'(r_k)$ are, respectively, the second and first local derivatives of the reduced RDF, taking advantage of the fact that the derivatives of a function exaggerate the noise in that function.

Since the numerator of the function I is sensitive to the noise in the data while the denominator deemphasizes the importance of the noise in those regions where $G(r)$ is changing rapidly, one may obtain a smooth rather than flat (uniform) $G(r)$ by simply making I

minimal. Thus, according to Soper (1990), the functional Q to be minimized is

$$Q = \lambda\chi^2 + I. \tag{8}$$

We explored the approach of Soper (1990) and our test calculations showed that, when a noise function I of the type

$$I = \sum_{k=1}^{N}([-G(r_{k-2}) + 16G(r_{k-1}) - 30G(r_k) + 16G(r_{k+1}) - G(r_{k+2})]^2/\{\exp[\text{const}|G(r_k) - G(r_{k-1})|] + \exp[\text{const}|G(r_{k+2}) - G(r_k)|]\}) \tag{9}$$

is employed, quite reliable results are obtained. The good performance of this function is based on the suitable choice of its components. For the calculation of the local second derivative of $G(r)$ [the numerator of (9)] a five-data-point scheme is employed (Savitzky & Golay, 1964) which exactly accounts for the high-frequency noise in the data. If three- or seven-data-point schemes are employed, the noise in the data is under- or overestimated, respectively. The denominator is composed of two exponents whose factors are proportional to the local first derivative of $G(r)$. The exponents are necessary in order that the rapid changes in the shape of $G(r)$ are appropriately accounted for. If the local first derivative of $G(r)$ is not made a factor of exponents, the peaks in $G(r)$ turn out to be nonphysically rounded off. Because of the well balanced actions of the components of the noise function, one can both efficiently remove the high-frequency noise from $G(r)$ and preserve the physically relevant sharp features of $G(r)$. That is why the noise function, I, defined by (9) is implemented in the program *IFO*.

Furthermore, as already discussed, experimental $S(q)$ data often comprise some systematic errors that cause the appearance of false oscillating features in the corresponding RDF. Since the false oscillations usually involve correlations between a large number of data points in real space, the noise function I, designed for removing presumably high-frequency false features, may not be sensitive enough to them. In order to cope successfully with low-frequency artifacts in both reciprocal and real space, we introduced two additional constraints on the solution searched for by correspondingly adding two new terms to the functional Q to be minimized. The constraints involve all data points being considered and ensure a consistent overall behaviour of the reconstructed RDF and its Fourier associate $S(q)$. One of these constraints, $S1$, is based on the so-called integral 'sum rule' (Wagner, 1978),

$$\int_{q=0}^{\infty} q^2[S(q) - 1]\,\mathrm{d}q = 2\pi^2\rho_o \equiv S1^{\text{theor}}, \tag{10}$$

and concerns the experimental data in reciprocal space. For finite experimental $S(q)$ data it can be written as

$$\sum_{q=0}^{q_{\max}} q_i^2[S(q_i) - 1]\Delta q = 2\pi^2\rho_o \equiv S1^{\text{theor}}. \tag{11}$$

The other, $S2$, is based on the relation (Klug & Alexander, 1974)

$$\int_{r=0}^{\infty} \rho(r)4\pi r^2\,\mathrm{d}r = N^{\text{at}} - 1, \tag{12}$$

which can be readily reduced to the following 'sum rule' concerning finite real-space data:

$$\sum_{r=0}^{r_{\max}} r_i G(r_i)\Delta r = -1 \equiv S2^{\text{theor}}. \tag{13}$$

(Here N^{at} is the number of all atomic species in the disordered material.) As a result, the functional Q employed in the program *IFO* comprises the following terms

$$Q = \lambda\chi^2 + \mu I + \gamma|S1^{\text{theor}} - S1^{\text{calc}}| + \delta|S2^{\text{theor}} - S2^{\text{calc}}| \tag{14}$$

whose sum is to be minimized. Here it may be added that since (11) and (13) are only true for decided values of $q_{\max}$ and $r_{\max}$, respectively, the integral constraints $S1$ and $S2$ should be considered as relatively 'soft' and be used to direct the image-reconstruction process towards a solution consistent with them only. Thus, by selecting proper values for the Lagrange-type multipliers λ, μ, γ and δ, which are input parameters to the program *IFO*, one may obtain an RDF which reproduces as closely as necessary the experimental $S(q)$ data (due to the χ^2 term) and which is a smooth, *i.e.* a noise-free, function (due to the I term); it also behaves consistently with the 'sum rule' applicable to real-space data (due to the $S2$ term). The Fourier counterpart of such an RDF will be a structure factor that is as close to the original one as the quality of the experimental data allows (due to the χ^2 term), that does not introduce false oscillations in the corresponding $G(r)$ (due to the I and $S2$ terms) and that satisfies the respective 'sum rule' (due to the $S1$ term). Therefore, a minimization of the functional of (14), may, in principle, yield both a structure factor that is corrected for statistical and systematic errors and an RDF of improved quality. The efficiency of the image-reconstruction-type technique described here has been demonstrated by our test calculations discussed later.

There are several computational approaches that may be employed for the minimization of the functional Q presently considered. Since Q incorporates terms of data related by a nonlinear transformation and since a cycling between real and reciprocal space is inevitable with the minimization process, some iterative computational approach, as shown in practice (Press *et al.*, 1992), is appropriate. For the minimization of Q a Monte Carlo-type approach, that is essentially iterative and is not prone to get stuck into local 'false' minima, has been adopted in the program *IFO*. In the Monte Carlo

approach it is considered that there exists an ensemble of RDFs, $\{G(r)_m\}$, which encompasses all possible images of the available experimental data, each image occurring with some probability, p_m, defined as

$$p_m = \exp(-Q_m), \tag{15}$$

where Q_m is the value of the functional Q [(14)] for the respective image $G(r)_m$. Clearly, the $G(r)_m$ with the highest probability of occurring is the one for which the functional Q is minimal and it is exactly the solution searched for. Since the minimum of Q is usually rather flat, a number of $G(r)_m$'s have a high probability of occurring. As a result, the Monte Carlo approach yields a solution representing a weighted (ensemble) average of all such $G(r)_m$'s; with the weights are the corresponding probabilities p_m. The $G(r)_m$'s with a high probability of occurring are searched for by carrying out a random walk through configuration space of all possible $G(r)_m$'s. This random walk, known as an 'importance sampling' (Binder & Heerman, 1992), starts with the calculation of an RDF $G(r)_{\text{old}}$, exactly corresponding to the available set of experimental data $\{S(q_i)\}$, and of the corresponding value of the functional Q, Q_{old}. Then, one randomly selected data point of $G(r)_{\text{old}}$ is subjected to a small change (usually of the order of ± 0.05) and the value of the functional Q, Q_{new}, corresponding to this already modified RDF $G(r)_{\text{new}}$ is calculated. The change made is accepted if the following inequality,

$$\exp(-\alpha \Delta Q) > \eta, \tag{16}$$

where η is a random number uniformly distributed in the interval [0, 1] and $\Delta Q = Q_{\text{old}} - Q_{\text{new}}$ holds. The parameter α, which is an input parameter to the program *IFO*, is used here to renormalize the differences ΔQ, which are usually very small numbers, and to control the percentage of accepted changes. When a change is accepted, the just generated $G(r)_{\text{new}}$ is stored in a statistical sum and is further considered as the next $G(r)_{\text{old}}$. If a change is rejected, the current $G(r)_{\text{old}}$ is stored in the statistical sum and is again considered as $G(r)_{\text{old}}$. The process is repeated until a preset number of changes has been made and the functional Q has been minimized. Finally, an assembly-averaged RDF $G(r)$ is calculated on the basis of all $G(r)_m$'s which, together with their corresponding probabilities p_m, have been stored in the statistical sum

$$G(r) = \sum_{m=1}^{S} p_m G(r)_m \Big/ \sum_{m=1}^{S} p_m, \tag{17}$$

where S is the number of accepted changes. As usual, some number of $G(r)_m$'s sampled during the initial stages of the random walk are not included in the statistical sum defined by (17) in order that the solution is not influenced considerably by the starting data. Accordingly, accumulation of the statistical sum in the program *IFO* starts only when 10% of the preset number of trials to be accepted has really been achieved. As an option, one extra, rather stringent, constraint, requiring that the reconstructed $G(r)$ behaves as $-4\pi r\rho_o$ (Wagner, 1978) for values of r smaller than a preset small real-space distance, has also been included in the program *IFO*. It has been found that even a temporary switching on of this constraint in the course of the reconstruction process considerably reduces the number of Monte Carlo trials necessary to reach the global minimum of the functional Q and, furthermore, helps in removing some of the residual systematic errors in the $S(q)$ data which, if present, show as false oscillations near the origin of real space.

3. Test applications of *IFO*

To check the performance of the technique for image-reconstruction-type calculation of RDFs we applied the program *IFO* to a number of theoretical and experimental data-sets representing examples of some frequently occurring cases.

3.1. *Reconstruction of RDFs for a Percus–Yevick hard-sphere liquid*

A structure factor for a hypothetical hard-sphere-like liquid, having the atomic number density of liquid tin ($\rho_o = 0.035$ at. Å^{-3}), has been derived using the Percus–Yevick equation (Waseda, 1980). This is shown in Fig. 1, with the corresponding RDF $G(r)$, obtained by a direct Fourier transformation, shown in Fig. 2. Since, in order to mimic an experimental data-set, the $S(q)$ data have been deliberately calculated only up to $q_{\max} = 12$ Å^{-1}, some truncation ripples are seen in real space. No other false features are expected to be present in the original $G(r)$ data since the theoretically derived structure factor is inherently free from statistical and systematic errors. Such errors were introduced in the $S(q)$ data and the program *IFO* was then applied as follows. At first we simulated statistical noise in the $S(q)$ data by adding/subtracting a small number inversely proportional to the original $S(q)$ values. As one can see in Fig. 1 the noise we have introduced becomes more and more noticeable with the increase in q values, as usually occurs with experimental $S(q)$ data. One can also see in Fig. 2 that the statistical noise in the $S(q)$ data propagates in the $G(r)$ data obtained by a direct Fourier transformation and gives rise to some false features (see the region between 6 and 8 Å). Application of the image-reconstruction technique clearly restores the original $S(q)$ data in the finest detail (see Figs. 1 and 2) and correspondingly removes the false features from the $G(r)$ data. Besides, the amplitudes of the termination ripples are also reduced and the first peak in the reconstructed $G(r)$ does not exhibit a nonphysical shoulder on its high-r side. It may be added that the constraint requiring that

$G(r)$ behaves as $-4\pi r\rho_o$ at small values of r has not been applied during the particular reconstruction process carried out with the following values of the Lagrange parameters: $\lambda = 8$, $\mu = 5$, $\gamma = 2$ and $\delta = 2$.

Having ascertained that the image-reconstruction technique successfully copes with noisy $S(q)$ data, we went on with an evaluation of the case of $S(q)$ data comprising some systematic error. The original $S(q)$ data were systematically distorted by adding an oscillation having a short period in q. The imperfect $S(q)$ data obtained are shown in Fig. 3. and the corresponding RDF $G(r)$ in Fig. 4. As one can see in Fig. 4 the systematic error in the reciprocal-space data introduces pronounced false features in the real-space data. The period of oscillations in $G(r)$ shifts and the first peak of $G(r)$ appears split up. Application of the program *IFO* removes the systematic error from $S(q)$ (see Fig. 3) and the false features from $G(r)$ (see Fig. 4). The particular image-reconstruction process has been carried out with the use of the constraint requiring that $G(r)$ behaves as $-4\pi r\rho_o$ at small values of r and with the following values of the Lagrange parameters: $\lambda = 5$, $\mu = 50$, $\gamma = 10$ and $\delta = 10$.

One more case that occurs in practice is that of $S(q)$ data containing both statistical and systematic errors. This rather worse case was simulated by making the original $S(q)$ data noisy as already described and adding to them an oscillation slowly varying with q. The manipulated $S(q)$ data are shown in Fig. 5 and their direct Fourier transform, $G(r)$, in Fig. 6. As one can see in the figures, the combination of an imperfection slowly varying with q and high-frequency noise causes both pronounced false oscillations concentrated presumably at small values of r and high-frequency ripples modifying the shape of $G(r)$ at higher values of r [see the $G(r)$ data for values of r greater than 10 Å]. Application of the image-reconstruction technique restores the original $S(q)$ data (see Fig. 5) and removes all spurious features from the $G(r)$ data. To achieve this positive effect the constraint requiring that $G(r)$ data behave as $-4\pi r\rho_o$ at small values of r was applied and the following values of the Lagrange parameters were used: $\lambda = 5$, $\mu = 50$, $\gamma = 10$, $\delta = 10$. In summary, the test calculations carried out clearly demonstrated that the *IFO* program efficiently removes systematic errors of short and long period in q

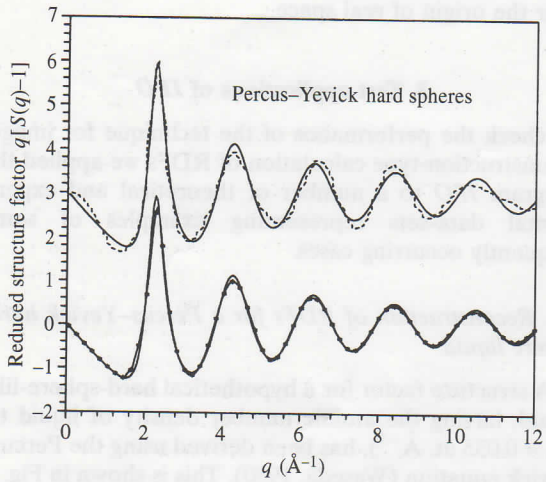

Fig. 3. Structure factors for Percus–Yevick hard-sphere liquid. Theoretical error-free data (full line); theoretical data with an added oscillation modelling a systematic error having a short period in q (full line; upper part); reconstructed data (full line with symbols; lower part).

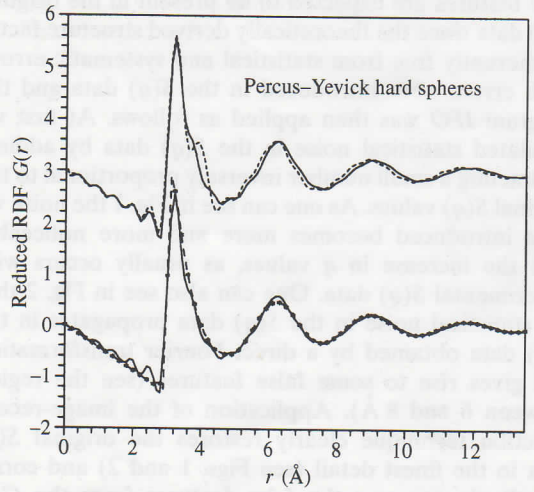

Fig. 4. Reduced RDFs for Percus–Yevick hard-sphere liquid obtained by: direct Fourier transformation of the error-free structure factor of Fig. 3 (full line); direct Fourier transformation of the deliberately distorted structure factor of Fig. 3 (broken line; upper part); the image-reconstruction technique (symbols; lower part).

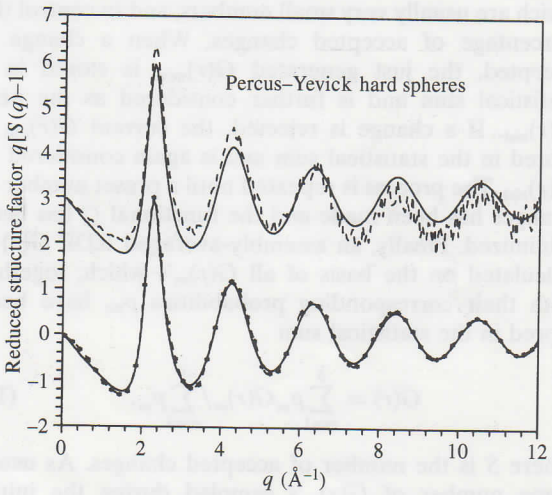

Fig. 5. Structure factors for Percus–Yevick hard-sphere liquid. Theoretical error-free data (full line); theoretical data with added statistical noise and an oscillation modelling a systematic error having a long period in q (broken line; upper part); reconstructed data (symbols; lower part).

and noise from $S(q)$ data and thus yields RDFs of improved quality.

3.2. *Reconstruction of RDFs for sol–gel-derived TiO_2 layers*

TiO_2 layers, produced by the sol–gel route, were subjected to electron diffraction experiments that yielded the structure factor given in Fig. 7. Detailed structural studies (Petkov, Holzhüter *et al.*, 1998) have shown that the particular TiO_2 layers are a mixture of amorphous TiO_2 and crystalline anatase, giving rise to the slowly and rapidly varying with q features in the experimental $S(q)$ data, respectively. Since the $S(q)$ data turned out to be rather noisy, mainly because the thickness of the layers investigated was only of the order of 500 Å, too many spurious sharp features appeared in the $G(r)$ data, which made interpretation difficult. As one can see in Fig. 8, application of the image-reconstruction technique successfully removed the statistical noise from the experimental $S(q)$ data and the corresponding false oscillations in the RDF $G(r)$; some physically relevant features of $G(r)$ were considerably sharpened [see the first peak in $G(r)$ positioned at approximately 2 Å]. At the same time, the sharp peaks in the original structure factor, due to the presence of a small amount of crystalline anatase, remained intact. This result illustrates well that, although smooth solutions are favoured by the image-reconstruction technique considered here, no physically relevant sharp features are removed from the experimental structure functions being reconstructed; this is due to the suitable choice of the noise function, I, implemented in the program *IFO*. The reconstructed $G(r)$ data have not been forced to behave as $-4\pi r\rho_o$ at small values of r since no reliable estimate for the atomic number density of the TiO_2 layers investigated could be obtained. The values of the Lagrange parameters used were: $\lambda = 10$, $\mu = 20$, $\gamma = 0$ and $\delta = 0$.

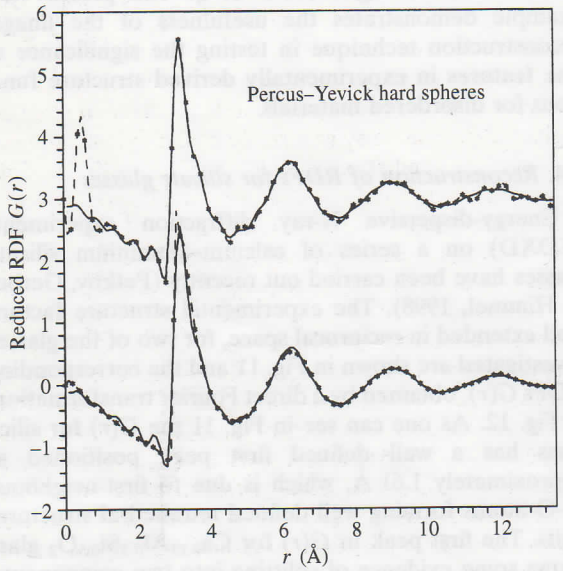

Fig. 6. Reduced RDFs for Percus–Yevick hard-sphere liquid obtained by: direct Fourier transformation of the error-free structure factor of Fig. 5 (thin full line); direct Fourier transformation of the imperfect structure factor of Fig. 5 (broken line with symbols; upper part); the image-reconstruction technique (broken line with symbols; lower part).

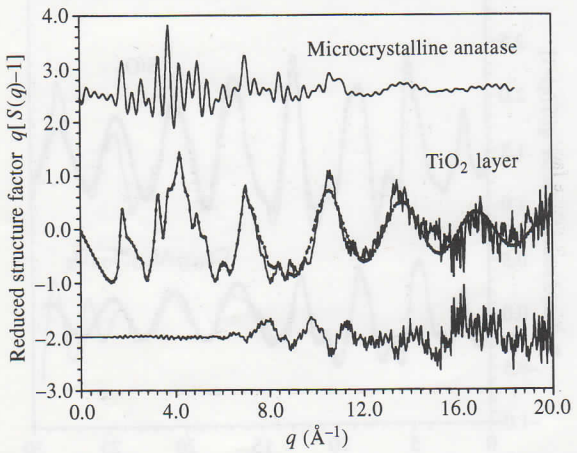

Fig. 7. Structure factors for TiO_2 layers (middle). Experimental data (thin full line); reconstructed data (symbols). The lower line gives the difference between the experimental and reconstructed data. For comparison, the upper line is a calculated structure factor for microcrystalline TiO_2.

3.3. *Reconstruction of the partial RDF $G_{Ni-Ni}(r)$ for $Ni_{81}B_{19}$ metallic glass*

Partial RDFs for $Ni_{81}B_{19}$ metallic glass have been obtained by neutron diffraction experiments making use of the isotopic substitution technique (Lamparter *et al.*,

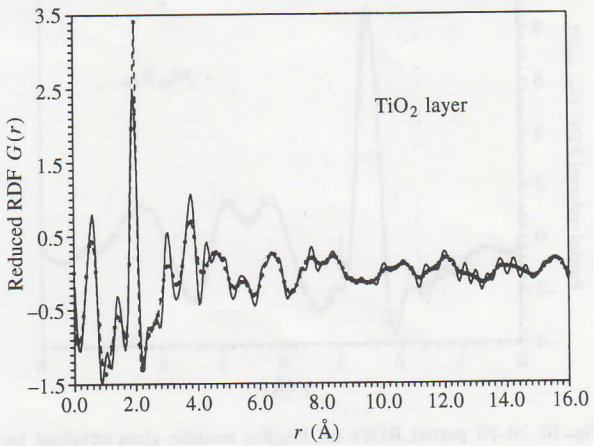

Fig. 8. Reduced RDFs for TiO_2 layers obtained by: direct Fourier transformation of the experimental data of Fig. 7 (thin full line); the image-reconstruction technique (broken line with symbols).

1982). Some features of the experimental data, in particular the presence of a small hump just after the first main peak in $G_{Ni-Ni}(r)$, however, have not been reproduced well by recent model calculations; this has inspired some discussion about their physical relevance (Sietsma & Thijsse, 1991; Ee *et al.*, 1998). To address this point we applied the image-reconstruction technique to the experimental $S_{Ni-Ni}(q)$ data shown in Fig. 9 with Lagrange parameters of the following values: $\lambda = 50$, $\mu = 5$, $\gamma = 2$ and $\delta = 2$. As one can see in Fig. 10 the reconstructed $G(r)$ does not exhibit any hump after its first main peak. Furthermore, the first peak in the reconstructed $G(r)$ is slightly increased in amplitude when compared to that in the $G(r)$ obtained by a direct Fourier transformation, which is in line with the predictions of molecular dynamics simulations (Ee *et al.*, 1998). These observations suggest that the small hump in $G_{Ni-Ni}(r)$ very likely originates from some residual error in the experimental structure-factor data. As one can see in Fig. 9, the errors likely to be present in $S_{Ni-Ni}(q)$ are of quite a small magnitude and are concentrated at high values of q. The present test example demonstrates the usefulness of the image-reconstruction technique in testing the significance of fine features in experimentally derived structure functions for disordered materials.

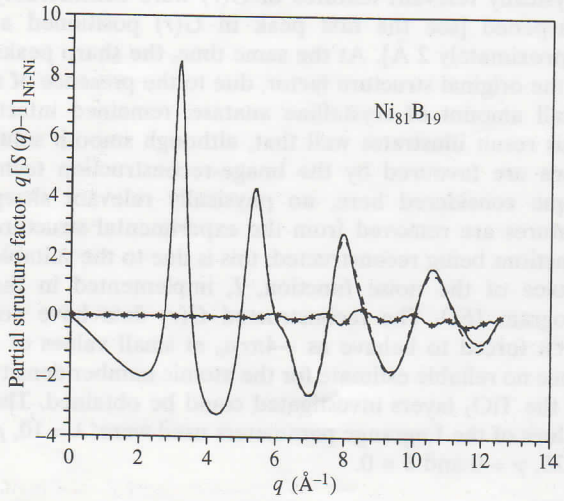

Fig. 9. Ni–Ni partial structure factors for $Ni_{81}B_{19}$ metallic glass. Experimental data (full line); reconstructed data (broken line); difference between the experimental and reconstructed data (full line with symbols).

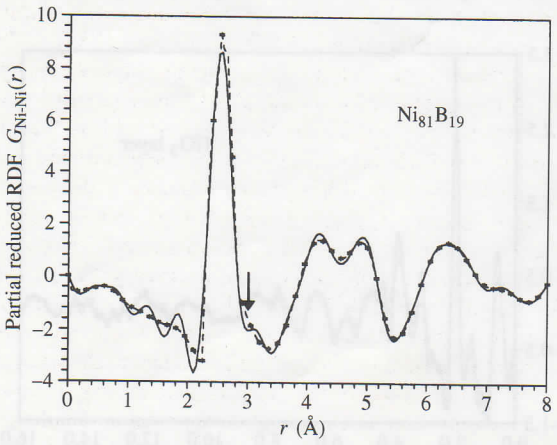

Fig. 10. Ni–Ni partial RDFs for $Ni_{81}B_{19}$ metallic glass obtained by: direct Fourier transformation of the experimental data of Fig. 9 (full line); the image-reconstruction technique (broken line with symbols). The position of the small hump under question (see text) is indicated by an arrow.

3.4. *Reconstruction of RDFs for silicate glasses*

Energy-dispersive X-ray diffraction experiments (EDXD) on a series of calcium–aluminium silicate glasses have been carried out recently (Petkov, Gerber & Himmel, 1998). The experimental structure factors, well extended in reciprocal space, for two of the glasses investigated are shown in Fig. 11 and the corresponding RDFs $G(r)$, obtained by a direct Fourier transformation, in Fig. 12. As one can see in Fig. 11 the $G(r)$ for silica glass has a well defined first peak positioned at approximately 1.61 Å, which is due to first-neighbour Si–O atoms forming well defined tetrahedral structural units. The first peak in $G(r)$ for $Ca_{0.17}Al_{0.5}Si_{0.5}O_2$ glass shows some evidence of splitting into two components, one due to Si–O first neighbours and another most probably due to Al–O first neighbours. Since the amplitude of the false oscillations in the experimental RDF data, although not very high, is of the order of the

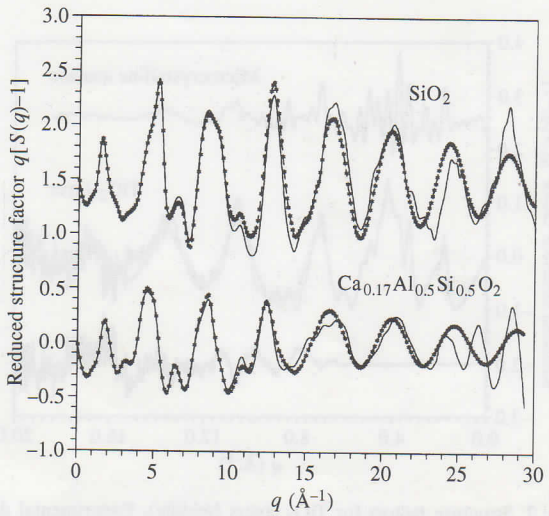

Fig. 11. Structure factors for silica and $Ca_{0.17}Al_{0.5}Si_{0.5}O_2$ glasses. Experimental data (thin full line); reconstructed data (broken line with symbols).

magnitude of the high-r component of the first peak in $G(r)$ for $Ca_{0.17}Al_{0.5}Si_{0.5}O_2$ glass, the question arises whether the observed splitting has a structural origin or is an experimental artifact. To answer this question we first applied some standard procedures for reducing the amplitudes of the false oscillations in $G(r)$ data. The results are shown in Fig. 13. As one can see in the figure, the inclusion of a modification function $M(q) = \exp(-\text{const}q^2)$ in the direct Fourier transformation reduces the amplitude of the false oscillations but the loss of resolution is so severe that the fine, physically relevant features of $G(r)$ are no longer recognizable. Application of the repetitive Fourier transformation correction procedure (Kaplow *et al.*, 1965) removes the false oscillations at small values of r only and, therefore, one is unable to discriminate between the false and real features of the RDF for $Ca_{0.17}Al_{0.5}Si_{0.5}O_2$ glass at higher values of r. By contrast, application of the image-reconstruction technique ($\lambda = 20$, $\mu = 50$, $\gamma = 5$ and $\delta = 5$) successfully removes the false oscillations from the $G(r)$ data without spoiling their resolution (see Fig. 12). As a result one can see that the splitting of the first peak in $G(r)$ for $Ca_{0.17}Al_{0.5}Si_{0.5}O_2$ glass is unlikely to be an experimental artifact and, hence, one can conclude that well defined Si–O and Al–O coordination units do exist in this disordered material. It is worth noting that the first Si–O coordination number has been estimated to be 3.76 (15) and 3.95 (15) when the first peak of the conventionally obtained and the reconstructed $G(r)$ for silica glass, respectively, have been fitted with Gaussian-like functions. This finding demonstrates the potential of the image-reconstruction process in improving the accuracy of all structure-relevant information contained in experimental structure functions for disordered materials.

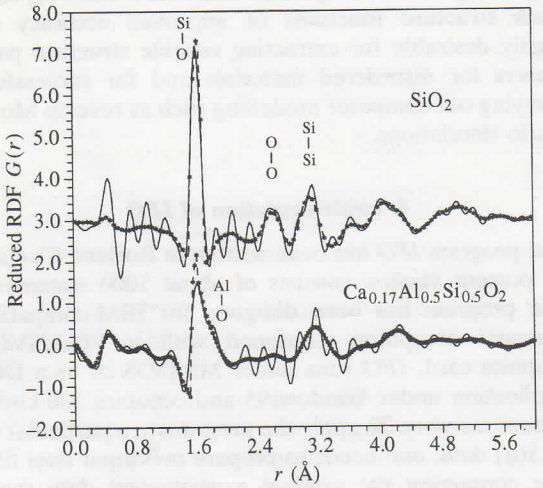

Fig. 12. Reduced RDFs for silica and $Ca_{0.17}Al_{0.5}Si_{0.5}O_2$ glasses obtained by: direct Fourier transformation of the experimental data of Fig. 11 (thin full line); the image-reconstruction technique (broken line with symbols).

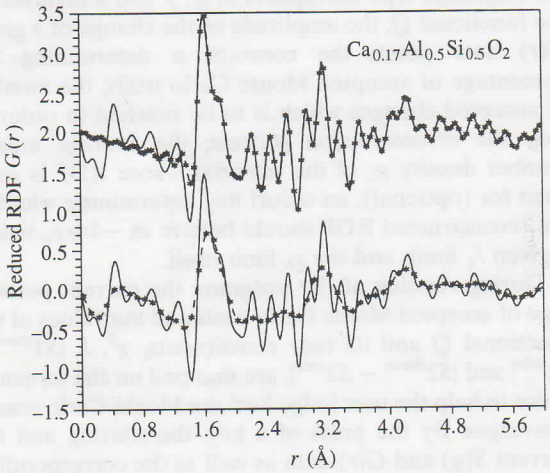

Fig. 13. Reduced RDFs for $Ca_{0.17}Al_{0.5}Si_{0.5}O_2$ glass obtained by: direct Fourier transformation of the experimental data of Fig. 11 (thin full line); application of a corrective procedure based on repetitive Fourier transformations (symbols; upper part); introduction of a modification function in the direct Fourier transformation (broken line with symbols; lower part).

3.5. *Reconstruction of $G_{Ge}(r)$ for $(Rb_2O)_{0.2}(GeO_2)_{0.8}$ glass*

Recently, anomalous wide-angle X-ray scattering (AWAXS) has become more and more popular thanks to the widespread availability of synchrotron radiation sources. AWAXS involves the measurement of a few diffraction spectra near the absorption edge of some of the atomic species constituting the material under study and a derivation of the so-called difference structure factor reflecting the atomic arrangement only around that species (Shevchik, 1977; Waseda, 1984). The difference structure factor is, however, highly susceptible to the quality of the individual total $S(q)$ and may turn out to be considerably distorted unless highly accurate AWAXS experiments have been carried out (Fishburn & Barton, 1995). As an example, a difference structure factor $S_{Ge}(q)$ of $(Rb_2O)_{0.2}(GeO_2)_{0.8}$ glass, obtained by recent AWAXS experiments carried out at Stanford Synchrotron Radiation Laboratory, California, USA (Stachs *et al.*, 1998), is shown in Fig. 14. As one can see in the figure, this structure factor is considerably distorted as a result of some deficiencies in the experimental and/or data-reduction procedures. As a result the corresponding RDF $G(r)$ exhibits large-amplitude spurious oscillations that render it almost unusable (see Fig. 15). Application of the image-reconstruction technique ($\lambda = 5$, $\mu = 50$, $\gamma = 10$ and $\delta = 10$), however, seems to have successfully removed most of the errors from the experimental data and yielded an RDF with quite a

reasonable behaviour (see Fig. 15). In fact, the reconstructed $S_{Ge}(q)$ and $G_{Ge}(r)$ are rather similar to those obtained by another independent AWAXS study on a glass of the same chemical composition (Price *et al.*, 1997). In order to achieve this result we forced the reconstructed RDF to behave strictly as $-4\pi r\rho_o$ at small values of r.

In general, our test calculations clearly demonstrated that the image-type reconstruction of RDFs has real advantages over the direct Fourier inversion of $S(q)$ data and may be applied with confidence to real data. However, the success of this approach in yielding reliable RDFs, even when the $S(q)$ data available are not quite accurate, does not lessen the importance of experimental efforts to improve the quality of the latter. In this respect, the program *IFO* could be used to control the quality of the experimental $S(q)$ data and reveal the presence of errors which are to be corrected by carrying out complementary measurements or adjusting the raw-data-reduction procedures. When all possible experimental improvements have been made *IFO* could be again applied for removing the small residual experimental and computational artifacts and improving the accuracy of the final RDF and $S(q)$ data. Such structure functions of improved accuracy are highly desirable for extracting reliable structure parameters for disordered materials and for successfully carrying out computer modelling such as reverse Monte Carlo simulations.

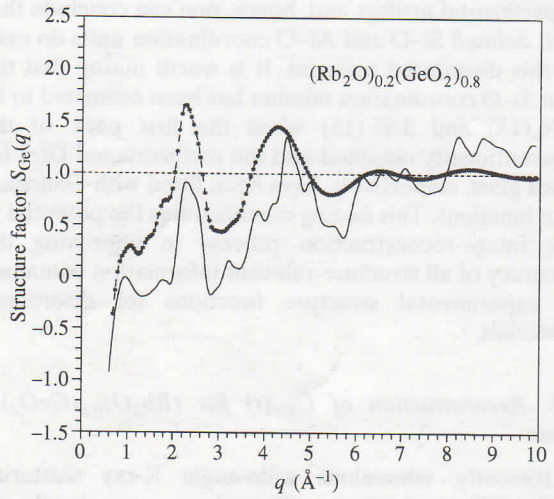

Fig. 14. Germanium difference structure factor $S_{Ge}(q)$ of $(Rb_2O)_{0.2}(GeO_2)_{0.8}$ glass. Experimental AWAXS data (thin full line); reconstructed data (broken line with symbols).

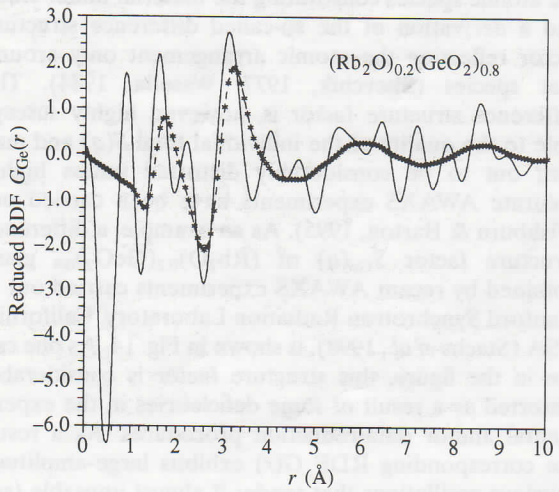

Fig. 15. Germanium difference reduced RDF $G_{Ge}(r)$ of $(Rb_2O)_{0.2}(GeO_2)_{0.8}$ glass obtained by: direct Fourier transformation of the experimental data of Fig. 14 (thin full line); the image-reconstruction technique (broken line with symbols).

4. Implementation of *IFO*

The program *IFO* has been written in Borland C++ and its present version consists of about 1000 statements. The program has been designed for IBM-compatible personal computers equipped with a VGA/SVGA graphics card. *IFO* runs under MS-DOS or as a DOS application under Windows95 and occupies 330 kbytes of core memory. To apply the program to a particular set of $S(q)$ data, one needs to prepare two input data files, one containing the original experimental data themselves [q, $S(q)$; two columns in free format] and the other containing the values of the parameters guiding the Monte Carlo sampling procedure. These parameters are as follows: the step, Δr, in which $G(r)$ is calculated (usually $\Delta r = \pi/q_{max}$); the maximal real-space distance, r_{max}, to which $G(r)$ is calculated (usually $r_{max} \leq \pi/\Delta q$); the Lagrange-type multipliers λ, μ, γ and δ involved in the functional Q; the amplitude of the change of a given $G(r)$ data point; the constant α determining the percentage of accepted Monte Carlo trials; the number of accepted changes which is to be reached in order to stop the reconstruction process; the average atomic number density ρ_o of the material whose RDF is searched for (optional); an on/off flag determining whether the reconstructed RDF should behave as $-4\pi r\rho_o$ below a given r_b limit; and the r_b limit itself.

During running of the program, the current percentage of accepted Monte Carlo trials and the values of the functional Q and its four constituents, χ^2, I, $|S1^{theor} - S1^{calc}|$ and $|S2^{theor} - S2^{calc}|$, are dumped on the screen in order to help the user judge how the Monte Carlo search converges. By the press of a key, the starting and the current $S(q)$ and $G(r)$ data as well as the corresponding differences $[S(q)_{start} - S(q)_{current}; G(r)_{start} - G(r)_{current}]$ may be displayed on the screen by means of built-in graphics routines. When a particular reconstruction process is complete, the program *IFO* automatically creates two files containing the $G(r)$ data calculated according to (17) and their Fourier associate, a corrected

$S(q)$ function, respectively. When necessary, the reconstruction process could be further continued by using these stored data files as input files. If proper values for the parameters guiding the Monte Carlo search have been selected a successful reconstruction of an RDF takes approximately half an hour on an IBM PC with 166 MHz Pentium processor. The most important of these parameters, the Lagrange-type multipliers λ, μ, γ and δ, are always automatically renormalized after having been supplied by the user so that their relative magnitudes, reflecting the actual quality of the $S(q)$ data being processed and the available prior information about the RDF being sought, are, in practice, involved in the functional Q. With this described architecture the program *IFO* is easy to use and no special knowledge of computers is required on the part of the user. A copy of the program *IFO* is available from the authors upon request (email petkov@phys.uni-sofia.bg) or may be downloaded from www.phys.uni-sofia.bg/dsspm/users/petkov/.

Thanks are due to the Alexander von Humboldt Foundation for financial support to one of the authors (VP). We are grateful to Professor P. Lamparter, MPI, Stuttgart, Germany, for kindly allowing us to reproduce his experimental data for $Ni_{81}B_{19}$ glass, and to Dr J. Sietsma, University of Delft, The Netherlands, Professor Th. Gerber and Dr B. Himmel, University of Rostock, Germany, for fruitful discussions.

References

Allen, D. A. & Howe, R. A. (1992). *J. Phys. Condens. Matter*, **4**, 6029–6038.

Binder, K. & Heerman, D. (1992). *Monte Carlo Simulations in Statistical Physics* (*Springer Series in Solid State Sciences*). Berlin: Springer-Verlag.

Cockayne, D. J. H. & McKenzie, D. R. (1988). *Acta Cryst.* A**44**, 870–880.

Cumbrera, F. L., Sanchez-Bajo, F. & Muñoz, A. (1995). *J. Appl. Cryst.* **28**, 408–415.

Ee, L., Thijsse, B. J. & Sietsma, J. (1998). *Phys. Rev. B*, **57**, 906–913.

Fishburn, J. R. & Barton, S. W. (1995). *Acta Cryst.* A**51**, 679–683.

Grimley, D. I., Wright, A. C. & Sinclair, R. N. (1990). *J. Non-Cryst. Solids*, **119**, 49–64.

Howe, M. A., McGreevy, R. L. & Pusztai, L. (1996). *Manual to MCGR Computer Program*.

Jal, J. F., Soper, A. K., Carmona, P. & Dupuy, J. (1991). *J. Phys. Condens. Matter*, **3**, 551–567.

Kaplow, R., Strong, L. & Averbach, B. (1965). *Phys. Rev. A*, **138**, 1336–1350.

Klug, H. & Alexander, L. (1974). *X-ray Diffraction Procedures for Polycrystalline and Amorphous Materials*. New York: Wiley.

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

Korsunsky, V. I. & Naberukhin, Yu. I. (1980). *Acta Cryst.* A**36**, 33–39.

Krylov, A. S. & Vvedenskii, A. V. (1995). *J. Non-Cryst. Solids*, **192&193**, 683–687.

Lamparter, P., Sperl, W., Steeb, S. & Bletry, J. (1982). *Z. Naturforsch. Teil A*, **37**, 1223–1234.

Lin, T.-L. & Tsao, C.-S. (1996). *J. Appl. Cryst.* **29**, 170–177.

Linden, W. (1995). *Appl. Phys.* A**60**, 155–165.

Müller, J. J. & Hansen, S. (1994). *J. Appl. Cryst.* **27**, 257–270.

Muñoz, A., Cumbrera, F. L. & Marquez, R. (1988). *J. Mater. Sci.* **23**, 2021–2028.

Petkov, V. (1995). *J. Non-Cryst. Solids*, **192&193**, 65–68.

Petkov, V., Gerber, Th. & Himmel, B. (1998). *J. Non-Cryst. Solids*. In the press.

Petkov, V., Holzhüter, G., Tröde, U., Gerber, Th. & Himmel, B. (1998). *J. Non-Cryst. Solids*. In the press.

Poulsen, H., Neuefeind, J., Neumann, H., Schneider, J. & Zeidler, M. (1995). *J. Non-Cryst. Solids*, **188**, 63–74.

Press, W. H., Teulkovsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in C*. Cambridge University Press.

Price, D., Ellison, J., Saboungi, M.-L., Hu, R., Egami, T. & Howells, W. (1997). *Phys. Rev. B*, **55**, 11249–11255.

Savitzky, A. & Golay, M. J. E. (1964). *Anal. Chem.* **36**, 1627–1640.

Schleger, P., Puig-Molina, A., Ressouche, E., Rutty, O. & Schweizer, J. (1997). *Acta Cryst.* A**53**, 426–435.

Schotte, K.-D., Schotte, U., Bleif, H.-J. & Papoular, R. J. (1995). *Acta Cryst.* A**51**, 739–746.

Shevchik, N. (1977). *Philos. Mag.* **35**, 805–815.

Sietsma, J. & Thijsse, B. J. (1991). *J. Non-Cryst. Solids*, **135**, 146–157.

Skilling, J. (1988). In *Maximum Entropy and Bayesian Methods in Science and Engineering*, Vol. 1, edited by G. Erickson & C. Smith. Dordrecht: Kluwer Academic Publishers.

Skilling, J. & Bryan, R. K. (1984). *Mon. Not. R. Astron. Soc.* **211**, 111–124.

Soper, A. K. (1990). *Inst. Phys. Conf. Ser.* **107**, 57–67.

Stachs, O., Gerber, Th. & Himmel, B. (1998). In preparation.

Thijsse, B. J. (1984). *J. Appl. Cryst.* **17**, 61–76.

Wagner, C. N. J. (1978). *J. Non-Cryst. Solids*, **31**, 1–40.

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

Waseda, Y. (1980). *The Structure of Non-Crystalline Materials*. New York: McGraw-Hill.

Waseda, Y. (1984). In *Novel Applications of Anomalous X-ray Scattering for Structural Characterization of Disordered Materials* (*Lecture Notes in Physics Series*), edited by H. Araki, J. Ehlers, K. Hepp, R. Kippenhahn, A. Weidenmüller & J. Zittarz. Berlin: Springer-Verlag.

Wei, W. (1986). *J. Non-Cryst. Solids*, **81**, 239–250.