Determination of Particle-Size Distribution Functions from Small-Angle Scattering Data by Means of the Indirect Transformation Method

BY O. GLATTER

Institut für Physikalische Chemie der Universität Graz, Heinrichstrasse 28, A-8010 Graz, Austria and Institut für Röntgenfeinstrukturforschung der österreichischen Akademie der Wissenschaften und des Forschungszentrums Graz, Steyrergasse 17, A-8010 Graz, Austria

(Received 25 May 1979; accepted 25 July 1979)

Abstract
Small-angle scattering data from polydisperse systems can be evaluated under the assumption that all particles have the same shape and that the size distribution depends only on one linear size parameter R. The shape of the particles is assumed to be known a priori. The corresponding size distribution function for the number of particles \( D_n(R) \) or for the volume \( D_v(R) \) can be computed from the smeared, unsmoothed scattering data by the indirect transformation method restricting the range of definition of the \( D(R) \) functions to a finite range \( R_{\text{min}} \leq R \leq R_{\text{max}} \). \( R_{\text{min}} \) may be equal to zero, \( R_{\text{max}} \) is limited by the sampling theorem of the Fourier transformation. The resolution in real space is given by the distance of the knots of the B spline functions approximating the size distribution function. The propagated statistical error band in real space can be computed using the inverted matrix of the normal equations. The method gives satisfactory results even in those cases where the shape of the particles is not known exactly, and is superior to analytical methods if the termination effect is critical.

Introduction
The scattered intensity of X-rays or neutrons from a dilute system of particles of identical shape but having a variety of sizes, can be regarded as approximately the integral of the intensities scattered by the individual particles. The total scattered intensity, \( I(h) \), the normalized scattered intensity, \( \Phi(hR) \), of a single particle of size \( R \) (sometimes called the shape factor), and the particle size distribution function, \( D_n(R) \), that states the number of particles defined by the size parameter \( R \), are related by the integral

\[
I(h) = \int_0^\infty D_n(R) \Phi(hR) dR \tag{1a}
\]

(Vonk, 1976).

Here, \( h = 4\pi\lambda_0^{-1} \sin \theta \), \( \lambda_0 \) is the wavelength of the monochromatic radiation, \( 2\theta \) is the scattering angle and \( m(R) \) is the integral over the excess scattering length density over the surrounding medium of a particle of size \( R \). If we want to determine the function \( D_n(R) \) (this attempt is sometimes called the solution of the inverse scattering problem), we have to assume that the shape of the particles is known a priori, giving a well defined shape factor, \( \Phi(hR) \). The parameter \( R \) can be any characteristic dimension of the particle (e.g. the radius if the particle is a sphere).

If we are interested in the size distribution function, \( D_v(R) \), the volume of all particles defined by the size parameter \( R \), we have to use a similar integral

\[
I(h) = \text{constant} \int_0^{R_{\text{max}}} D_v(R) R^{-3} m^2(R) \Phi(hR) dR. \tag{1b}
\]

The mathematical description of the method and the quality of the numerical results are nearly identical for \( D_n(R) \) and \( D_v(R) \). Therefore, the formulae and results for \( D_v(R) \) are not presented in the following.

In (1) we have assumed infinite dilution, no multiple scattering, the same shape for all particles, random orientation for non-spherical particles, monochromatic radiation and pinhole collimation. Concentration effects are usually removed by concentration series and extrapolation to zero concentration (Kratky, 1963). The wavelength effect and the collimation effect can be described by the following equation:

\[
I_{\text{exp}}(h) = \left[ 2 \int_0^\infty D_n(R) m^2(R) \Phi(hR) dR \right]^{\frac{1}{2}} \int_0^{\infty} dx Q(x) P(t) W(\lambda') \tag{2}
\]

(GLatter, 1977a), where \( P(t) \) is the weighting function for the slit-length effect, \( Q(x) \) is the weighting function for the slit-width effect, \( \lambda' = \lambda / \lambda_0 \) and \( w(\lambda') \) is the wavelength distribution of the polychromatic radiation.

The wavelength profile should be measured very carefully because any polychromatic effect could be interpreted as polydispersity. Generally speaking there is no way to decide from the experimental scattering data whether they originate from a monodisperse or a polydisperse system, from monochromatic or polychromatic radiation. Additional a priori information is necessary to answer these questions.

The integrals (1) and (2) can be solved analytically by an inverse transformation like the Fourier transformation; the type of transformation depends on the assumed particle shape. The equations can be solved numerically assuming the analytical form of \( D(R, c_0) \)
and adjusting the parameters $c_i$ in reciprocal space. Again, one has to make an assumption concerning the particle shape.

**Analytical methods**

An explicit analytical solution for the special case of spherical particles was derived by Roess (1946) and Riseman (1952). These methods involve an integration over the scattering curve from zero to infinity. If it is possible to determine a Guinier region (Guinier & Fournet, 1955) and the Porod constant (Porod, 1951) with sufficient accuracy, such transformations can be carried out without strong termination effects (Letcher & Schmidt, 1966; Brill & Schmidt, 1968; Brill, Weil & Schmidt, 1968).

Similar transformations have since been developed for different types of particles: long cylinders (Fedorova, 1977), thin spherical shells (Fedorova & Emelyanov, 1977), thin cylinders (Schmidt, Emelyanov & Fedorova, 1978), thin circular discs (Schmidt, Fedorova & Emelyanov, 1978), and flat sheets (Pringle & Schmidt, 1977). A general analytical method for calculating particle dimension distributions from scattering data was published recently by Fedorova & Schmidt (1978). The size distribution can be calculated using a Titchmarsh transformation if the shape factor of the particle can be expressed as a square of a Bessel function of the first kind of order $v$. None of these transformations makes assumptions for the solution of (1), but the solution function is very sensitive to the extrapolated parts of $I(h)$.

Equation (2) has been solved only for negligible wavelength and slit width effects.

**Numerical methods – linear models**

Several simple methods approximate the function $D_\sigma(R)$ with a special two-parameter function (for example the so-called log normal distribution): Roess & Shull (1947), Hosemann (1951), Mittelbach & Porod (1965), Mittelbach (1965), Harkness, Gould & Hren (1969), Neillson (1973), Plesil & Baldrian (1976).

The parameter estimation is comparatively easy for such two-parameter models but the size distribution is biased to a considerable extent by the particular model that was assumed for the size distribution.

More general approaches permit a larger number of parameters to be estimated. The methods of Hendricks, Schelten & Schmatz (1974), Plavnik Kozhevnikov & Shishkin (1976), Vonk (1976) and the method presented in this paper belong to this group.

**Indirect transformation method**

A general method for data evaluation of small-angle scattering data from monodisperse systems applying constrained weighted least-squares techniques for parameter estimation (called the indirect transformation method) has been described by the author in detail recently (Glatter, 1977a, b). This method can easily be adapted to the problem of polydisperse systems.

The approach starts with a linear expansion for the size distribution function $D_\sigma(R)$:

$$D_\sigma(R) = \sum_{v=1}^{N} c_v \varphi_v(R), \quad (3)$$

where $N$ is the number of basis functions $\varphi_v(R)$. This number is usually not larger than 30. Cubic $B$ splines are used as basis functions (Greville, 1969; Schelten & Hossfeld, 1971; Glatter, 1977a, b). They are defined as convolution products of a step function and represent curves with minimum second derivative. They overlap only with six neighbours as they differ from zero only in the range of four knots. The expansion coefficients are the unknowns.

As a priori information we use the assumption that the size distribution function is a band-limited function, i.e.

$$D_\sigma(R) \neq 0 \quad \text{only for} \quad R_{\text{min}} \leq R \leq R_{\text{max}}, \quad (4)$$

where $R_{\text{max}}$ may be equal to zero. The linear expansion (3) leads to the approximation curve in reciprocal space, $I_\lambda(h)$,

$$I_\lambda(h) = \sum_{v=1}^{N} c_v \Psi_v(h), \quad (5)$$

where the $\Psi_v(h)$ are the $B$ splines transformed into reciprocal space using (1),

$$\Psi_v(h) = \int_{R_{\text{min}}}^{R_{\text{max}}} \varphi_v(R) R^2 \Phi(hR) dR. \quad (6)$$

The approximation to the uncorrected data point, $I_\lambda(h)$, is given by a similar formula

$$\tilde{I}_\lambda(h) = \sum_{v=1}^{N} c_v \chi_v(h), \quad (7)$$

where $\chi_v(h)$ results from transformation of the basis functions $\varphi_v(R)$ according to (2):

$$\chi_v(h) = 2 \int_{0}^{x} dR \int_{0}^{x} d\alpha \int_{0}^{x} dt \int_{0}^{x} dx \Phi(x) P(t) \times W(\alpha) \varphi_v(R) R^2 \Phi(\beta). \quad (8)$$

The coefficients $c_v$ can be determined by the constrained least-squares condition

$$L + \lambda N_c \sum_{i=1}^{M} \left[ I_\exp(h_i) - \sum_{v=1}^{N} c_v \chi_v(h_i) \right]^2 \sigma(h_i)^2$$

$$+ \lambda \sum_{v=1}^{N-1} (c_v + c_{v+1} - 2c_v)^2 = \text{Min.}, \quad (9)$$

where $M$ is the number of data points and $\sigma(h_i)$ is the variance of the $i$th point. The constraint $N_c$ is necessary to avoid artificial oscillations. The optimum value of the stabilizing parameter $\lambda$ can be determined by the point of inflexion method (for details see Glatter, 1977a).

The solution of the least-squares condition (9) leads
to the so-called normal equations which can be solved using a conventional matrix inversion routine. The inverted matrix provides the basis for the error propagation analysis (for details see Glatter, 1977b).

All the necessary computations are performed by the extended Fortran program system *ITP-79* which allows either the evaluation of the size distribution function \( D_n(R) \) of polydisperse systems or the computation of the distance distribution functions of monodisperse systems.

* For further details of the program system *ITP-79* and for program listings, please contact the author.

The resolution of details in \( D_n(R) \) is limited by the distance DRB of the knots of the B splines:

\[
DRB = \frac{R_{\text{max}} - R_{\text{min}}}{N}.
\]  

(10)

The maximum value \( R_{\text{max}} \) is limited by the sampling theorem (Bracewell, 1965). The maximum dimension \( D \) of the particles is correlated with the size parameter by the relation

\[
D = \alpha R
\]  

(11)

and from this

\[
R_{\text{max}} \leq \frac{\pi}{2h_1};
\]  

(12)

\( h_1 \) is the scattering angle of the first data point. There exists no similar relation between DRB and \( h_2 \), where \( h_2 \) is the scattering angle of the last data point. However, the maximum number of coefficients, \( N \), and the distance of the knots, DRB, are limited by the

---

Fig. 1(a) — Scattered intensity of polydisperse system of spheres (the corresponding size distribution is shown in (b)). ●●● Data points, smeared with an infinitely long slit function \( P(t) \), statistical error proportional to the square root of the intensity. Termination angles: \( h_1 = 0.01 \), \( h_2 = 0.2 \). (b) Size distribution of (a) — Exact distribution function, ○○○○ size distribution computed from the exact scattering curve without smearing effect and statistical noise, × × × size distribution computed from data points in (a).

Fig. 2(a) — Smeared scattered intensity (infinitely long slit) from a double Gaussian distribution of spheres (b). ●●● Data points with a standard deviation proportional to the intensity; termination angles: \( h_1 = 0.01 \), \( h_2 = 0.2 \). (b) —— Exact double Gaussian size distribution of (a); ○○○ Size distribution and propagated error band resulting from data points in (a).
termination effect \( (h_2) \), the statistical error \( \sigma(h) \) and the smearing effect \( \{P(t), Q(x) \text{ and } W(\lambda)\} \). This limitation cannot be expressed analytically but it leads to insurmountable stability problems signalled by a breakdown of the point-of-inflexion method for the determination of \( \lambda_{opt} \).

**Numerical results**

The method has been tested by simulations, which have been performed in the following way by the program system. The theoretical scattering intensity is computed assuming an arbitrary but well defined size distribution. This scattering function is smeared according to geometry and wavelength effects, distorted by statistical fluctuations from a random-number generator and terminated at the scattering angles \( h_1 \) and \( h_2 \). These simulated data points are subjected to the evaluation procedure, and the results are compared with the starting function \( D(R) \). Such simulations show clearly the limits of the method and possible systematic artefacts. Some typical results from a very large number of simulations are reported in the following; the units of \( R \) and \( h \) are Å and Å\(^{-1}\), respectively.

The theoretical scattering curve, \( I(h) \), and the simulated data, \( \tilde{I}(h) \), starting from a system of spheres with a polynomial size distribution \( D(R) \), are shown in Fig. 1(a). \( \tilde{I}(h) \) corresponds to an infinitely long slit function \( P(t) \) and an error band proportional to the square root of the smeared intensity. The termination angles are \( h_1 = 0.01 \) and \( h_2 = 0.2 \). The size distribution functions resulting from the smeared scattering curve with and without statistical error are shown in Fig. 1(b) together with the starting function \( (R_{min} = 0, R_{max} = 300) \).

A similar test for a double Gaussian distribution of spheres is shown in Fig. 2(a), (b). The geometrical effect represents again an infinitely long slit. The data points shown in Fig. 2(a) are generated with an error band proportional to the square root of the intensity. The resulting \( D(R) \) function is plotted together with the propagated statistical error band and the starting function in Fig. 2(b). A better resolution of the profile cannot be expected, because the termination effect at \( h_1 = 0.01 \) corresponds to an allowed \( R_{max} \) of about 160, while the actual \( R_{max} \) used for evaluation is 250.

The resolution of details can also be tested by a delta function, this means evaluating the scattering function of a monodisperse system as a polydisperse system. The scattering function of a sphere with \( R = 100 \) was used for this test. The results for the exact scattering function without collimation effect and statistical error are shown together with the results from smeared data with \( 5\% \) statistics in Fig. 3. A delta function can never be

![Fig. 3(a) --- Theoretical scattering function I(h) of a sphere with radius R = 100. --- Solution function \( I_s(h) \) approximating I(h) by a polydisperse system of spheres according to (5). \( h_1 = 0.01, h_2 = 0.2 \). The corresponding \( D(R) \) function is shown in (b). (b) Interpretation of a monodisperse system of spheres with radius \( R = 100 \) (delta function at \( R = 100 \)) as a polydisperse system of spheres. \( R_{min} = 20, R_{max} = 200, N = 20 \). --- \( D(R) \) computed from the theoretical scattering function. \( \bigcirc \bigcirc \bigcirc \) \( D(R) \) computed from a smeared scattering function \( \tilde{I}(h) \) (slit length effect) with \( 5\% \), statistical noise.

![Fig. 4. --- Exact polynomial size distribution \( D(R) \) of prolate ellipsoids \( (R, R, 3R) \). \( \bigcirc \bigcirc \bigcirc \) \( D(R) \) computed from \( I(h) \) \( (h_1 = 0.0, h_2 = 0.2) \) with exact shape factor \( \Phi(hR) \). \( R_{max} = 60, N = 20 \); \( \triangle \triangle \triangle \) \( D(R) \) computed from \( I(h) \) with the shape factor of spheres. \( R_{max} = 120, N = 20. \)
described exactly by a limited number of spline functions. But the results with \( DRB = 10 \) \((N = 20)\) are very reasonable. The oscillations around zero indicate a very sharp distribution and allow for a good approximation in reciprocal space. Of course, negative regions of \( D(R) \) have no real physical meaning.

The examples in Fig. 1–3 are computed with the exact shape factor \( \Phi(hR) \). In practice it will usually not be possible to get such perfect a priori information. Thus it is important to perform a test with a wrong shape factor. The same polynomial size distribution function as in Fig 1(b) is used for this purpose. The particles are prolate ellipsoids with axes \( R, R \) and \( 3R \). The corresponding scattering function is interpreted as a polydisperse system of spheres with radius \( R \). The resulting size distribution function \( D(R) \) has a similar shape to the starting function but is shifted to larger \( R \) values (Fig. 4). This shift is caused by the long axis of the ellipsoids.

This and other similar simulations suggest that a wrong shape factor may cause some deformation of the true profile but may not prevent a useful application of the method even to mixtures of particles of different size and shape. However, a good approximation of the scattering function does not indicate a correct shape factor, as it is not possible to determine the size distribution and the particle shape of a polydisperse system simultaneously from a small-angle scattering experiment.

Conclusion

The numerical tests show that the indirect transformation method can be successfully applied to the evaluation of size distribution functions from small-angle scattering data of polydisperse systems. The method allows the evaluation of smeared and unsmoothed data, takes into account the statistical accuracy and gives a propagated error band for the size distribution function. Negative values in the distribution function are not suppressed as is the case in the method of Hendricks, Schelten & Schmatz (1974), but such values are necessary for a satisfactory approximation of scattering curves of very narrow size distributions. Imperfect a priori information on the shape factor leads to quantitative errors in the size distribution function. If, for example, a distribution of prolate ellipsoids is interpreted as a distribution of spheres the method still provides a distribution similar in shape but slightly shifted to larger diameters, taking into account the larger distances in the ellipsoids. In any case, such imperfect a priori information on the shape factor, which is typical for practical applications, does not cause a numerical breakdown of the method.

The comparison of the results from this method and from the method of Vonk (1976) for a few test examples shows no drastic differences. The Vonk method needs an extrapolation to zero angle if \( h \) is too large. The logarithmic scaling in real space used in the latter method is advantageous for profiles centered near \( R = 0 \), because of the narrow spacing in this region. On the other hand, this narrow spacing may introduce stability problems. The wide spacing at large \( R \) values prevents the description of details in this range. The methods of Hendricks, Schelten & Schmatz (1974) and Plavnik, Kozhevnikov & Shishkin (1976) have not been tested.

The analytical method of Fedorova & Schmidt (1978) does not take into account the statistical error and requires extrapolations to zero angle and to the large angles. The latter method will be suitable for data without smoothing problems and with small termination effects as no assumptions besides the a priori information on the shape of the particles are required.

References