Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 17 November 2008
Accepted 16 March 2009

Relationship between the atomic pair distribution function and small-angle scattering: implications for modeling of nanoparticles

Christopher L. Farrow[a] and Simon J. L. Billinge[a,b]*

[a]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, USA, and [b]Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, NY 11973, USA. Correspondence e-mail: sb2896@columbia.edu

The relationship between the equations used in the atomic pair distribution function (PDF) method and those commonly used in small-angle-scattering (SAS) analyses is explicitly shown. The origin of the sloping baseline, $-4\pi r \rho_0$, in PDFs of bulk materials is identified as originating from the SAS intensity that is neglected in PDF measurements. The nonlinear baseline in nanoparticles has the same origin, and contains information about the shape and size of the nanoparticles.

1. Introduction

The atomic pair distribution function (PDF) analysis of X-ray and neutron powder diffraction is growing in popularity with the advent of nanoscience and nanotechnology. The technique is more than 70 years old (Debye & Menke, 1930; Warren, 1990; Egami & Billinge, 2003) and was originally applied almost exclusively to the study of glass and amorphous structures (Warren, 1934; Warren *et al.*, 1936; Franklin, 1950, 1951; Wright, 1998). However, the approach is proving powerful in solving structure on the nanoscale (Billinge & Kanatzidis, 2004), where traditional crystallographic methods break down (Billinge & Levin, 2007). In particular, the study of the structure of discrete nanoparticles using the PDF method has recently become a focus (McKenzie *et al.*, 1992; Zhang *et al.*, 2003; Gateshki *et al.*, 2004; Gilbert *et al.*, 2004; Page *et al.*, 2004; Korsunskiy & Neder, 2005; Petkov *et al.*, 2005; Bedford *et al.*, 2007; Ehm *et al.*, 2007; Masadeh *et al.*, 2007; Pradhan *et al.*, 2007). The convergence of this new need with the availability of powerful sources of high-energy synchrotron X-rays and spallation neutrons and fast computing is greatly expanding the power and applicability of the method.

The PDF, $G(r)$, is defined both as a function of the real-space pair density, $\rho(r)$, and the reciprocal-space scattering, $F(Q) = Q[S(Q) - 1]$, as follows:

$$G(r) = (2/\pi) \int_0^\infty F(Q) \sin(Qr)\, \mathrm{d}Q \tag{1}$$

and

$$G(r) = 4\pi r[\rho(r) - \rho_0], \tag{2}$$

where ρ_0 is the number density of the material (Kaplow *et al.*, 1965, 1968; Klug & Alexander, 1974; Johnson *et al.*, 1982; Soper & Silver, 1982; Korsunskii, 1985; Wright, 1985; Nanao *et al.*, 1987; Warren, 1990; Egami *et al.*, 1991; Billinge *et al.*, 1994; Petkov *et al.*, 1998; Petkov & Danev, 1998; Keen, 2001; Tucker *et al.*, 2001; Egami & Billinge, 2003; Soper, 2007; Neder *et al.*, 2007; Billinge, 2008; Dinnebier & Billinge, 2008; Gilbert, 2008). Equation (2) in this form works well for bulk materials, but the negatively sloping baseline, $-4\pi r \rho_0$, is no longer valid when the PDF is calculated from models of finite-sized objects such as discrete nanoparticles (Korsunskiy & Neder, 2005). Motivated by the need for a rigorous definition of the form of this baseline we rederive these equations here. We show the correct form of equation (2) in a number of cases of practical interest such as discrete nanoparticles and nanoparticle PDFs calculated from bulk models. The important distinction is provided by the small-angle-scattering (SAS) intensity and we explicitly relate the commonly used PDF functions with commonly used results from SAS. We also show that the widely used pair of definitions for the PDF above are actually incompatible with each other and that the definition given in equation (1) does not give rise to equation (2) but rather to $R(r)/r$, where $R(r)$ is the radial distribution function. This work therefore resolves a long-standing ambiguity in the PDF literature.

Compared to the PDF of a bulk sample (Levashov *et al.*, 2007), the PDF of a nanoparticle is attenuated with increasing r by a function that is related to the form of the nanoparticle (Guinier, 1963). For simple shapes, such as spheroids, spherical shells, rods and discs, this nanoparticle form factor can be computed analytically (Rayleigh, 1914; Glatter & Kratky, 1982; Thorpe & Lei, 2007; Gilbert, 2008) and integral equations exist for more complex shapes (Kodama *et al.*, 2006). This lends itself to a simple nanoparticle-modeling procedure where the nanoparticle PDF is calculated from the PDF of a bulk-phase analogue by multiplying by the assumed nanoparticle form factor (Guinier, 1963; Qiu *et al.*, 2005; Masadeh *et al.*, 2007). The approach is successful for extracting precise quantitative structural information about the crystalline core of nanoparticles, including defects and size-dependent bond lengths (Masadeh *et al.*, 2007), and this functionality has been incorporated in the latest version of the PDF modeling software *PDFgui* (Farrow *et al.*, 2007). However, the method is not applicable when the nanoparticle structure has no bulk-phase analogue. This is the case in general, for example in nanoparticles with surface modifications (Zhang *et al.*, 2003) and inhomogeneous compositions such as core–shell nanoparticles (Liz-Marzán *et al.*, 1996). In these cases, models of discrete nanoparticles must be applied. As we discuss below, this results in an ambiguity about the precise form of the measured correlation function, and therefore how to calculate it. Currently, this is dealt with quite successfully in an *ad hoc* way, as for example in Korsunskiy & Neder (2005), Neder & Korsunskii (2005), Korsunskiy *et al.* (2007) and Neder *et al.* (2007). This paper presents a rigorous definition of the form of the baseline in terms of the nanoparticle form factor.

In §2 we rederive the equations giving rise to the PDF to show the precise relationship between the measured correlation function in an X-ray or neutron total-scattering experiment and the underlying model. In §3 we make explicit the link between the commonly used PDF and small-angle scattering equations. This has implications for calculating PDFs from discrete nanoparticle models for quantitative comparison with data, which are discussed in §4. In §5 we discuss the conditions under which the PDF can be calculated in real space. §6 presents a brief summary.

2. Derivation of the PDF equations

To understand the precise relationship between the commonly used PDF equations, nanoparticle structures and small-angle scattering, we rederive the PDF equations from the beginning since subtle details of the derivation that are often overlooked have a significant impact on the discussion presented here. Furthermore, the full derivation is not reproduced even in many textbooks on the subject (Warren, 1990; Klug & Alexander, 1974; Egami & Billinge, 2003; Dinnebier & Billinge, 2008) and so these subtleties are not widely appreciated in the community.

We start from the scattering amplitude from a set of i atoms at points $\mathbf{r}_i$ in the kinematical limit:

$$\begin{aligned}\psi(\mathbf{Q}) &= \sum_i f_i(Q)\exp(i\mathbf{Q}\cdot\mathbf{r}_i)\\ &= \sum_i \psi_i.\end{aligned}\tag{3}$$

If the scattering from these atoms were totally incoherent the total intensity would be the sum of the intensities from each atom,

$$\begin{aligned}I_{\rm inc} &= \sum_i \psi_i^*\psi_i\\ &= \sum_i f_i^*(Q)f_i(Q)\\ &= \sum_\alpha N_\alpha f_\alpha^*(Q)f_\alpha(Q)\\ &= N\sum_\alpha c_\alpha f_\alpha^*(Q)f_\alpha(Q)\\ &= N\langle f^2\rangle,\end{aligned}\tag{4}$$

where the sum over α is now over the different species of atoms in the sample with N being the total number of atoms, N_α being the number of atoms of type α and where the concentration of species α is $c_\alpha = N_\alpha/N$. The asterisk indicates the complex conjugate. Similarly, we can define the sample-averaged scattering power, $\langle f\rangle = \sum_\alpha c_\alpha f_\alpha$ and

$$\begin{aligned}\langle f\rangle^2 &= (1/N^2)\sum_{ij} f_j^* f_i\\ &= \sum_{\alpha\beta} c_\alpha c_\beta f_\alpha^* f_\beta.\end{aligned}\tag{5}$$

Here we have dropped the Q dependence of the atomic scattering factors to simplify the notation, but the f's are understood to retain their Q dependence.

The full coherent scattering intensity is given by $\psi^*\psi$, which is

$$\begin{aligned}I_{\rm c} &= \sum_i\sum_j f_j^* f_i \exp[i\mathbf{Q}\cdot(\mathbf{r}_i-\mathbf{r}_j)]\\ &= \sum_{i,j} f_j^* f_i\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij}).\end{aligned}\tag{6}$$

We can separate out the self-scattering, $i = j$, for which $\mathbf{r}_{ij} = 0$:

$$\begin{aligned}I_{\rm c} &= \sum_i f_i^* f_i + \sum_{i\neq j} f_j^* f_i\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij})\\ &= N\langle f^2\rangle + \sum_{i\neq j} f_j^* f_i\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij}),\end{aligned}\tag{7}$$

where we have used equation (4), resulting in an expression for the discrete scattering intensity for $i \neq j$ as

$$\begin{aligned}I_{\rm d} &= I_{\rm c} - N\langle f^2\rangle\\ &= \sum_{i\neq j} f_j^* f_i\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij}).\end{aligned}\tag{8}$$

We want an expression for the total scattering structure function, $S(\mathbf{Q})$, which is defined as $I_{\rm c}/(N\langle f\rangle^2) - \langle(f-\langle f\rangle)^2\rangle/\langle f\rangle^2$. The second term in this definition is the Laue monotonic diffuse scattering that comes about because of the imperfect cancellation of intensity at the destructive interference condition when atomic sites are occupied by atoms of different scattering strength. It results in a monotonic incoherent background even in the case of perfectly coherent scattering.

To get $S(\mathbf{Q})$ from equation (7) we therefore must normalize by the total number of scatterers, N,

$$I_{\rm c}/N = \langle f^2\rangle + (1/N)\sum_{i\neq j} f_j^* f_i\exp(i\mathbf{Q}\cdot\mathbf{r}_{ij}).\tag{9}$$

Subtracting the normalized self-scattering term to get

$$(I_{\rm c}/N) - \langle f^2 \rangle = (1/N) \sum_{i \neq j} f_j^* f_i \exp(i\mathbf{Q} \cdot \mathbf{r}_{ij}) \qquad (10)$$

and then normalizing by $\langle f \rangle^2$, we obtain

$$I_{\rm c}/(N\langle f \rangle^2) - \langle f^2 \rangle / \langle f \rangle^2 = [1/(N\langle f \rangle^2)] \sum_{i \neq j} f_j^* f_i \exp(i\mathbf{Q} \cdot \mathbf{r}_{ij}). \qquad (11)$$

Thus,

$$\begin{aligned} S(\mathbf{Q}) - 1 &= I_{\rm c}/(N\langle f \rangle^2) - \langle f^2 \rangle / \langle f \rangle^2 \\ &= I_{\rm d}/(N\langle f \rangle^2) \\ &= 1/(N\langle f \rangle^2) \sum_{i \neq j} f_j^* f_i \exp(i\mathbf{Q} \cdot \mathbf{r}_{ij}). \qquad (12) \end{aligned}$$

This expression yields precisely $S(\mathbf{Q}) - 1$ in terms of scattering from atoms in our sample.

For an isotropic sample, *e.g.* a powder of crystals or nanoparticles, we assume there to be a crystallite with every orientation with equal probability and we can take an orientational average. Place the $\mathbf{Q}$ along z so that we can express $\mathbf{Q} \cdot \mathbf{r}_{ij} = Qr_{ij} \cos\theta$. Then the orientational averaging means that θ takes all values with equal probability. The sample-averaged intensity for a pair of atoms will therefore be

$$\begin{aligned} \overline{\exp(i\mathbf{Q} \cdot \mathbf{r}_{ij})} &= \frac{\int_0^{2\pi} {\rm d}\varphi \int_0^{\pi} {\rm d}\theta \exp(iQr_{ij} \cos\theta) r_{ij}^2 \sin\theta}{\int_0^{2\pi} {\rm d}\varphi \int_0^{\pi} {\rm d}\theta \, r_{ij}^2 \sin\theta} \\ &= \frac{-2\pi r_{ij}^2 \big[\exp(iQr_{ij}\cos\theta)\big]_0^{\pi}}{4\pi r_{ij}^2 iQr_{ij}} \\ &= \frac{\big[\exp(iQr_{ij}) - \exp(-iQr_{ij})\big]}{2iQr_{ij}} \\ &= \frac{\sin(Qr_{ij})}{Qr_{ij}}. \qquad (13) \end{aligned}$$

Using this in equation (6) gives the average coherent scattering intensity, as expressed originally by Debye (1915). From this we get the total scattering structure function for an isotropic sample,

$$S(Q) - 1 = [1/(N\langle f \rangle^2)] \textstyle\sum_{i \neq j} f_j^* f_i [\sin(Qr_{ij})/Qr_{ij}]. \qquad (14)$$

Thus, the reduced total scattering structure function, $F(Q) = Q[S(Q) - 1]$, is

$$F(Q) = [1/(N\langle f \rangle^2)] \sum_{i \neq j} f_j^* f_i [\sin(Qr_{ij})/r_{ij}]. \qquad (15)$$

For completeness we describe the removal of the Q dependence of the X-ray form factors. This is a standard result and this paragraph can be skipped without losing the thread of the derivation. The form factors are assumed to be isotropic so depend only on Q and not $\mathbf{Q}$, which is a good approximation for scattering from core electrons in particular. Write $f(Q) = f(0)\tilde{f}(Q)$, where $\tilde{f}(Q)$ has value 1 at $Q = 0$ and contains the Q dependence of the form factor and $f(0) \simeq Z$, where Z is the atomic number that scales the form factor. The Morningstar–Warren approximation (Warren *et al.*, 1936) is that the Q-dependent part of the form factors can be well approximated by an average Q dependence, $\overline{\tilde{f}(Q)} = (1/N_{\rm species}) \sum_\alpha c_\alpha \tilde{f}_\alpha(Q)$. In this case the Q dependence, $\overline{\tilde{f}(Q)}^2$, comes out of the double sums in equation (15) on the top and the bottom and cancels out. The fs that remain are Q independent, and are normally replaced by the atomic number (modified by any anomalous scattering factors). The same result holds for neutron scattering where the fs are replaced by coherent neutron scattering lengths, b. These have no Q dependence and therefore the approximate method for removing the Q dependence is not needed.

Now we want to consider the inverse Fourier transform of $F(Q)$. Because $F(Q)$ is an odd function, we use the sine-Fourier transform,

$$f(r) = (2/\pi) \int_0^\infty F(Q) \sin(Qr) \, {\rm d}Q. \qquad (16)$$

We choose the $2/\pi$ prefactor so that the direct sine transform has a prefactor of 1. This is precisely the definition of the PDF in equation (1). From this we get

$$\begin{aligned} f(r) &= \frac{2}{\pi} \int_0^\infty \frac{1}{N\langle f \rangle^2} \sum_{i \neq j} f_j^* f_i \frac{\sin(Qr_{ij})}{r_{ij}} \sin(Qr) \, {\rm d}Q \\ &= \frac{2}{\pi N\langle f \rangle^2} \sum_{i \neq j} \frac{f_j^* f_i}{r_{ij}} \int_0^\infty \sin(Qr_{ij}) \sin(Qr) \, {\rm d}Q \\ &= \frac{1}{N\langle f \rangle^2} \sum_{i \neq j} \frac{f_j^* f_i}{r_{ij}} [\delta(r - r_{ij}) - \delta(r + r_{ij})] \\ &= \frac{1}{rN\langle f \rangle^2} \sum_{i \neq j} f_j^* f_i [\delta(r - r_{ij}) - \delta(r + r_{ij})], \qquad (17) \end{aligned}$$

which, if we confine ourselves to the positive axis only, is

$$f(r) = [1/(rN\langle f \rangle^2)] \sum_{i \neq j} f_j^* f_i \delta(r - r_{ij}). \qquad (18)$$

We can interpret $f(r)$ in terms of the radial distribution function (RDF). The RDF, denoted $R(r)$, is defined for an elemental system such that for an arbitrary atom i at the origin, $R_i(r)\,{\rm d}r$ gives the number of atoms in a shell of thickness ${\rm d}r$ at a distance r from that atom and the total RDF is the average of the partial RDFs over each atom taken at the origin. Thus, the integral of the RDF between two bounds, a and b, gives the number of atomic pairs per atom with separation within those bounds, N_{ab}. By inspection we see that equation (18) yields this behavior if we multiply $f(r)$ by r. For a solid with α atomic species we get

$$\begin{aligned} \int_a^b f(r) r \, {\rm d}r &= \int_a^b [1/(N\langle f \rangle^2)] \sum_{i \neq j} f_j^* f_i \delta(r - r_{ij}) \, {\rm d}r \\ &= [1/(N\langle f \rangle^2)] \sum_i \sum_{j \in S} f_j^* f_i \\ &= [1/(\langle f \rangle^2)] \sum_\alpha c_\alpha f_\alpha \sum_{j \in S} f_j^*, \qquad (19) \end{aligned}$$

where S is the set of atoms with distance from atom i greater than a and less than b. In the case of just one atomic species, this reduces to

$$\int_a^b f(r) r \, \mathrm{d}r = (f^2/f^2) \sum_{j \in S} 1$$
$$= N_{ab}, \qquad (20)$$

as required. Thus,

$$f(r) = R(r)/r$$
$$= 4\pi r \rho(r). \qquad (21)$$

The second expression in equation (21) comes from the relationship between the RDF and the pair density. The pair density is defined such that $\int \mathrm{d}r \, \mathrm{d}\varphi \, \mathrm{d}\theta \, r^2 \sin(\theta) \rho(r) = \int R(r) \, \mathrm{d}r$, so that $4\pi r^2 \rho(r) = R(r)$.

We now see the slightly unexpected result that the commonly used definition, equation (16), and our rederivation of it, equation (21), are not equivalent. The definition in equation (16) does *not* yield $G(r)$ [equation (2)] and strictly $(2/\pi) \int_0^\infty F(Q) \sin(Qr) \, \mathrm{d}Q = R(r)/r \neq G(r)$. However, we see below that in practice it is $G(r)$ and not $R(r)$ that is indeed obtained experimentally in most cases.

Finally, rearranging equation (21) we find

$$\rho(r) = f(r)/4\pi r$$
$$= [1/(4\pi r^2 N \langle f \rangle^2)] \sum_{i \neq j} f_j^* f_i \delta(r - r_{ij}). \qquad (22)$$

In reality, $I_c(Q)$ is measured down to a minimum Q due to the experimental setup. This means that in general the forward-scattering contributions are lost. We will consider the impact of this on the measured real-space function $f(r)$. We rewrite the expression for the experimental $f(r)$ as

$$f(r; Q_{\min}) = (2/\pi) \int_{Q_{\min}}^{\infty} F(Q) \sin(Qr) \, \mathrm{d}Q$$
$$= 4\pi r \rho(r) - (2/\pi) \int_0^{Q_{\min}} F(Q) \sin(Qr) \, \mathrm{d}Q. \qquad (23)$$

Of course, we have a finite $Q_{\max}$ as well, but this will be disregarded during the following discussion, as the effects are well understood (Toby & Egami, 1992).

3. Low-angle-scattering intensity

We will now consider a number of explicit examples to understand how the missing forward scattering affects the measured $f(r; Q_{\min})$.

It is instructive to first consider the RDF of an infinite scatterer of uniform density, ρ_0. The intensity of such a system is typically expressed as a delta function at $Q = 0$. We note that this gives $f(r) = 0$ when used in equations (12) and (16), so we must use a more rigorous description of the scattered intensity. Since the scattering length is defined per atom, the scattering amplitude of a volume element $\mathrm{d}\mathbf{r}$ at position $\mathbf{r}$ is $\psi = \rho_0 \langle f \rangle \exp(i\mathbf{Q} \cdot \mathbf{r}) \, \mathrm{d}\mathbf{r}$. Then, for an infinite uniform scatterer, the intensity is given by

$$I_c(\mathbf{Q}) = \rho_0^2 \langle f \rangle^2 \int \int \exp[i\mathbf{Q} \cdot (\mathbf{r} - \mathbf{r}')] \, \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}', \qquad (24)$$

where the integrals are over all space. Using this in equation (12) and integrating over Q first in equation (16) gives $f(r) = 4\pi r \rho_0$, as expected. As a side note, using $I_c(Q) = 2\pi^2 N \langle f \rangle^2 \rho_0 \delta(Q)/Q$ gives the correct result, but we forgo this derivation for the one below, where we consider in detail the general case of an arbitrary scatterer.

To evaluate the scattered intensity for an arbitrary scatterer of uniform density we define a shape function $s(\mathbf{r})$ such that inside the shape $s = 1$ and outside the shape $s = 0$. For such a material

$$I_c(\mathbf{Q}) = \rho_0^2 \langle f \rangle^2 \int \int s(\mathbf{r}) s(\mathbf{r}') \exp[i\mathbf{Q} \cdot (\mathbf{r} - \mathbf{r}')] \, \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}'. \qquad (25)$$

We redefine variables so that $\mathbf{r}'' = \mathbf{r} - \mathbf{r}'$, and $\mathrm{d}\mathbf{r}'' = \mathrm{d}\mathbf{r}$, in which case we have

$$I_c(\mathbf{Q}) = \rho_0^2 \langle f \rangle^2 \int \int s(\mathbf{r}') s(\mathbf{r}' + \mathbf{r}'') \exp(i\mathbf{Q} \cdot \mathbf{r}'') \, \mathrm{d}\mathbf{r}' \, \mathrm{d}\mathbf{r}''$$
$$= \rho_0^2 \langle f \rangle^2 \int \mathrm{d}\mathbf{r}'' \exp(i\mathbf{Q} \cdot \mathbf{r}'') \int s(\mathbf{r}') s(\mathbf{r}' + \mathbf{r}'') \, \mathrm{d}\mathbf{r}'. \qquad (26)$$

The second integral is a self convolution, or autocorrelation, of the shape function. Let us define

$$\gamma_0(\mathbf{r}) = (1/V) \int s(\mathbf{r}') s(\mathbf{r}' + \mathbf{r}) \, \mathrm{d}\mathbf{r}', \qquad (27)$$

where $V = \int s(\mathbf{r}) \, \mathrm{d}\mathbf{r}$ is the volume defined by the shape function. This $\gamma_0(\mathbf{r})$ is the characteristic function of the shape (Guinier *et al.*, 1955), and has been called the nanoparticle form factor in the PDF literature (Qiu *et al.*, 2005; Kodama *et al.*, 2006; Masadeh *et al.*, 2007). Defined as such, $\gamma_0(\mathbf{r})$ has the following properties:

$$\gamma_0(0) = 1,$$
$$\int \gamma_0(\mathbf{r}) \, \mathrm{d}\mathbf{r} = V. \qquad (28)$$

This definition, and a convenient dropping of the double primes, gives

$$I_c(\mathbf{Q}) = \rho_0^2 \langle f \rangle^2 V \int \gamma_0(\mathbf{r}) \exp(i\mathbf{Q} \cdot \mathbf{r}) \, \mathrm{d}\mathbf{r}. \qquad (29)$$

In analogy with the discrete case we want to convert this to $S(\mathbf{Q}) - 1 = [I_c/(N \langle f \rangle^2)] - [\langle f^2 \rangle / \langle f \rangle^2]$, and using the fact that $N = \rho_0 V$ we get

$$S(\mathbf{Q}) - 1 = [1/(N \langle f \rangle^2)] \rho_0^2 \langle f \rangle^2 V \int \gamma_0(\mathbf{r}) \exp(i\mathbf{Q} \cdot \mathbf{r}) \, \mathrm{d}\mathbf{r}$$
$$- \langle f^2 \rangle / \langle f \rangle^2$$
$$= \rho_0 \int \gamma_0(\mathbf{r}) \exp(i\mathbf{Q} \cdot \mathbf{r}) \, \mathrm{d}\mathbf{r} - \langle f^2 \rangle / \langle f \rangle^2. \qquad (30)$$

The second term, $\langle f^2 \rangle / \langle f \rangle^2$, is very small compared to the first term. It is order unity, where the first term scales as $N = \rho_0 V$, and can safely be ignored in most cases.

Now we want to take the orientational average. This must be done with care as, in general, the orientation of the nanoparticle shape and the underlying structure are correlated. For example, the morphology of the particles (plates or needles) depends on easy growth directions of the underlying structure. As discussed by Gilbert (2008), this means that the shape function and the internal structure of the particle are not, in general, separable and it is not correct to get the scattered intensity by convolving the reciprocal-space intensity with the Fourier transform of the characteristic function. Things are greatly simplified in the case where the underlying structure, or the nanoparticle shape, or both, are isotropic, or approximately so. Then we can denote the angle-averaged characteristic function as

$$\overline{\gamma_0(\mathbf{r})} = \gamma_0(r) = \frac{\int \mathrm{d}\varphi \int \mathrm{d}\theta\, \sin(\theta) r^2 \gamma_0(\mathbf{r})}{\int \mathrm{d}\varphi \int \mathrm{d}\theta\, r^2 \sin(\theta)}. \tag{31}$$

Using this and equation (13) we get

$$\begin{aligned} S(Q) - 1 &= \rho_0 \textstyle\int_0^\infty \mathrm{d}r \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\theta\; \overline{\gamma_0(\mathbf{r})}\, \overline{\exp(i\mathbf{Q}\cdot\mathbf{r})} r^2 \sin\theta \\ &= \rho_0 \textstyle\int_0^\infty \mathrm{d}r \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\theta\; \overline{\gamma_0(\mathbf{r})}\, \overline{\exp(i\mathbf{Q}\cdot\mathbf{r})} r^2 \sin\theta \\ &= \rho_0 \textstyle\int_0^\infty \mathrm{d}r \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\theta\, \gamma_0(r)[\sin(Qr)/(Qr)] r^2 \sin\theta \\ &= \rho_0 \textstyle\int_0^\infty \gamma_0(r)[\sin(Qr)/(Qr)] 4\pi r^2\, \mathrm{d}r. \end{aligned} \tag{32}$$

Since the particles have no preferred orientation in space, we have broken the average of the product in the first line into the product of the averages. This gives

$$F(Q) = \int\limits_0^\infty 4\pi \rho_0 r \gamma_0(r) \sin(Qr)\, \mathrm{d}r. \tag{33}$$

Noting that this is the direct sine-Fourier transform, we take the inverse transform to get

$$\begin{aligned} f_u(r) &= (2/\pi) \int\limits_0^\infty F(Q) \sin(Qr)\, \mathrm{d}Q \\ &= 4\pi \rho_0 r \gamma_0(r), \end{aligned} \tag{34}$$

where the subscript u indicates that this result is for a solid of uniform density distribution, ρ_0.

Next we consider a macroscopic crystal. The difference in $F(Q)$ is at higher Q where, instead of complete cancellation of all the discrete intensity it appears at distinct reciprocal-lattice points as sharp Bragg peaks. Importantly, in the region of Q below the first Bragg peak, the distinct scattering is zero except at very low Q, where the small-angle-scattering region is reached. The *small-angle*-scattering intensity, $I^{\rm sas}$, from the crystal is identical to that from the solid with uniform density: $I_u^{\rm sas} = I_{\rm crystal}^{\rm sas}$. The small- and wide-angle-scattering regions are well separated in Q and $I_{\rm sas}$ decays to zero before $Q_{\min}$ is reached in the crystal. Thus,

$$\begin{aligned} f_{\rm sas}(r) &= (2/\pi) \int\limits_0^{Q_{\min}} F(Q) \sin(Qr)\, \mathrm{d}Q \\ &= f_u(r) \end{aligned} \tag{35}$$

and therefore

$$(2/\pi) \int\limits_0^{Q_{\min}} F(Q) \sin(Qr)\, \mathrm{d}Q = 4\pi \rho_0 r \gamma_0(r). \tag{36}$$

We are now in a position to understand in detail the nature of the measured PDF $f(r; Q_{\min})$. Substituting equation (36) into equation (23) we get

$$f(r; Q_{\min}) = 4\pi r \rho(r) - 4\pi r \rho_0 \gamma_0(r). \tag{37}$$

This is similar to the definition of the PDF from equation (2), except that $\gamma_0(r)$ appears in the sloping baseline term.

4. Calculating *f*(*r*; $Q_{\min}$) from models

We can now consider the calculation of measured PDFs using equation (37) at a number of interesting limits.

4.1. Calculating in real space for bulk crystals

In the case of bulk crystals, the region of interest in the PDF is usually $r \ll D$, D being the smallest dimension of the crystal. In this region, $\gamma_0(r) \simeq 1$. Thus,

$$\begin{aligned} f(r; Q_{\min}) &= G(r) \\ &= 4\pi r[\rho_{\rm bulk}(r) - \rho_0], \end{aligned} \tag{38}$$

which is the familiar definition of $G(r)$ in equation (2). The pair density function, $\rho_{\rm bulk}(r)$, is calculated from a model with periodic boundary conditions (Billinge, 1998; Proffen & Billinge, 1999), or from a box of atoms that is much larger in extent than the range of r of interest (McGreevy & Pusztai, 1988), using equation (22). The average number density ρ_0 is given by the number of atoms per unit volume, which in the case of crystals is the number of atoms in the unit cell divided by the unit-cell volume.

Equation (22) can be modified to account for peak broadening due to atomic vibrations and finite maximum Q. These modifications are discussed elsewhere (Egami & Billinge, 2003).

4.2. Calculating in real space for nanoparticles modeled as attenuated bulk crystals

In this case $\rho_{\rm bulk}(r)$ is determined using a model of a bulk structure as described in §4.1. The pair density, $\rho(r)$, in equation (37) is the function for the nanoparticle, which is approximated as $\gamma_0(r)\rho_{\rm bulk}(r)$ (Guinier, 1963). Thus

$$f(r; Q_{\min}) = 4\pi r \gamma_0(r)[\rho_{\rm bulk}(r) - \rho_0]. \tag{39}$$

This approach has been implemented in the *PDFgui* modeling software (Farrow *et al.*, 2007) and used successfully on rather well ordered CdSe nanocrystals (Masadeh *et al.*, 2007). The main shortcoming is that effects that cannot be incorporated in the average structure, such as surface relaxations or core–shell inhomogeneities, cannot be modeled.

If there is a distribution of nanoparticle sizes and shapes, the characteristic function, $\gamma_0(r)$, can be replaced with an appropriately averaged characteristic function,

$$\gamma(r) = \int \gamma_0(r; R_1, R_2, \ldots) p(R_1, R_2, \ldots)\, \mathrm{d}R_1\, \mathrm{d}R_2 \ldots . \tag{40}$$

Here $p(R_1, R_2, \ldots)$ is the normalized distribution of nanoparticle shapes parameterized by R_1, $R_2, \ldots$. For example, for spherical nanoparticles of radius R, $p(R_1, R_2, \ldots) = p(R)$, the distribution of nanoparticle radii. Finally, we replace equation (39) with

$$f(r; Q_{\min}) = 4\pi r \gamma(r)[\rho_{\rm bulk}(r) - \rho_0]. \tag{41}$$

Great care should be taken to ensure that the result is unique when refining a number of nanoparticle morphology parameters beyond one or two.

4.3. Calculating as the Fourier transform of the properly normalized Debye function

This approach has been successfully used by a number of authors (Zhang *et al.*, 2003; Cervellino *et al.*, 2006). The $F(Q)$ function is evaluated using equation (15) and then Fourier transformed to obtain the desired real-space function. To account for thermal and zero-point motion in reciprocal-space calculations, equation (15) is replaced with a version that includes Debye–Waller effects,

$$F(Q) = \frac{1}{N\langle f\rangle^2}\sum_{i\neq j} f_j^* f_i \exp(-\tfrac{1}{2}\sigma_{ij}^2 Q^2)\frac{\sin(Qr_{ij})}{r_{ij}}. \qquad (42)$$

Here, σ_{ij}^2 is the correlated broadening factor for the atom pair (Proffen & Billinge, 1999; Thorpe *et al.*, 2002; Jeong *et al.*, 2003).

As we show here, for a quantitative comparison with measured data care must be taken with the Fourier transform so that it is carried out over the same range of Q as the experiment. The main drawback of the reciprocal-space approach is that it can be very slow compared to direct real-space calculation due to the long-range extent of the signal from each pair. It is generally preferred for smaller systems. However, due to recent algorithmic advances, the calculation of the Debye equation for larger systems can be greatly accelerated under certain circumstances (Cervellino *et al.*, 2006).

Using this method, the termination effects coming from the finite Q range are implicity included provided the Fourier transforms to obtain the model and data PDFs are terminated with the same $Q_{\min}$ and $Q_{\max}$ values.

4.4. Calculating in real space from discrete nanoparticle models

In this case, equation (37) is used directly, where $\rho(r)$ is calculated from a finite model of the discrete nanoparticle using equation (22). The difficulty arises in determining a correct form for the baseline $-4\pi r\rho_0\gamma_0(r)$. Up until now, the shape of the baseline has been approximated using expansions of *ad hoc* mathematical functions (Korsunskiy & Neder, 2005; Neder & Korsunskii, 2005; Korsunskiy *et al.*, 2007; Neder *et al.*, 2007). This is successful at approximating the behavior of the baseline. However, in this work we derive the explicit form of the baseline shape in terms of the characteristic function of the nanoparticle, the autocorrelation function of the nanoparticle shape. This suggests a number of approaches to calculating the PDF baseline in a more physical way.

If we have accurate small-angle-scattering data from the samples, from equation (37) we see that we can compute the PDF baseline from the measured SAS *via* a Fourier transform. However, care must be exercised as the derivation assumes that the sample is made up of discrete nanoparticles. In general, clusters and aggregates of nanoparticles will form and small-angle-scattering signals from these structures on different length scales will be present and must be separated. Also, scattering-density fluctuations of any sort in the sample will affect the SAS signal, as discussed elsewhere (Cargill, 1971). None of these effects need be explicitly considered if the SAS signal is not used in the PDF definition, as in equation (37), although an alternative method is then required to determine the baseline.

The inclusion of SAS data has the potential to add significant value to any refinement of nanoparticle models from the PDF. Both small- and large-angle scattering contain information about the shape and size of nanoparticles, but this information is decoupled from the internal nanoparticle structure in the small-angle scattering. This same information is in the PDF, but it can be obscured by structural features, such as when the PDF prematurely attenuates due to a complex or amorphous surface structure. In this case, the nanoparticle size obtained from SAS will be larger than the apparent nanoparticle size obtained from the PDF, which reflects the size of the coherent core structure. Even without the inclusion of SAS data into PDF refinements, we can learn much from SAS analysis techniques. PDF nanoparticle refinements usually start with a simple model that includes atomic positions restricted by a shape. Therefore, PDF analysis can benefit from the various *ab initio* methods (Svergun & Stuhrmann, 1991; Chacón *et al.*, 1998) for determining the shape of a scatterer from the SAS.

Without the use of the small-angle intensity for determining $\gamma_0(r)$, we can consider approaches to determining it self-consistently from the model, since it is the autocorrelation of the particle shape, which is directly available from the model itself by determining a ‘shrink-wrapping’ of the atomistic model. On the contrary, when the internal structure is well known, but the size and shape distribution of nanoparticles is not, then the characteristic function can be parameterized and refined to obtain the approximate nanoparticle dimensions as done by Masadeh *et al.* (2007).

5. The extent of small-angle scattering

We have considered the two asymptotic situations here of including or excluding all the SAS. If the SAS is retained in the intensity that is Fourier transformed, the resulting real-space function obtained is $4\pi r\rho(r) = R(r)/r$ [equation (18)]. Excluding it all results in $G(r)$ [equation (37)]. We now consider the possibility that some, but not all, of the small-angle scattering is included in the Fourier transform. This might occur in the case of very small nanoparticles, for example, when the SAS extends to wider angles. Here we estimate the circumstances under which a significant amount of small-angle intensity will appear in a wide-angle PDF experiment for the case of a sphere of uniform density. The scattering intensity is given by Rayleigh (1914),

$$I(Q) \propto [9/(QR)^6][\sin(QR) - QR\cos(QR)]^2, \qquad (43)$$

where R is the radius of the sphere. By integrating this equation from 0 to $Q_{\min}$, and dividing by the total integrated intensity, we get an expression for the proportion of small-angle intensity above $Q_{\min}$ for a given nanoparticle diameter (Rayleigh, 1914):

$$i(x) = 1 - [1/(2\pi r^5)][(2x^4 - x^2 + 3)\cos(2x) + x(x^2 + 6)\sin(2x) + 4x^5\mathrm{Si}(2x) - (5x^2 + 3)]. \quad (44)$$

Here $x = Q_{\min}R$ and $\mathrm{Si}(x)$ represents the sine integral, $\mathrm{Si}(x) = \int_0^x [\sin(x')/x']\,\mathrm{d}x'$. For a very small nanoparticle of radius 5 Å, we see that $i(5) < 0.01$, corresponding to $Q_{\min} \simeq 1$ Å^{-1}, which is typical for a rapid-acquisition pair distribution function (RAPDF) (Chupas *et al.*, 2003) experiment. Thus, for even quite small nanoparticles, practically all small-angle intensity is below $Q_{\min} = 1$ Å^{-1} and equation (37) is appropriate. However, care should be taken not to extend $Q_{\min}$ too low in Q in a measurement of a nanoparticulate system.

In the few-atom limit, such as the case of discrete small molecules, the small- and wide-angle scattering are not cleanly separated. To produce a complete real-space signal, one can approximate the small-angle scattering from a candidate structure model. This approach is commonly used in the study of small molecules in the gas phase (Hargittai & Hargittai, 1988). The Fourier transform of the estimated scattering approximates $R(r)/r$, the nominal 'experimental' or 'modified' RDF. An equivalent method for obtaining the modified RDF from wide-angle scattering alone is to add a baseline estimated from a model structure to $f(r, Q_{\min})$ (Ruan *et al.*, 2007). This approach has been successful for calculating $R(r)/r$ for many-atom nanoparticles.

6. Summary

The PDF is a valuable tool for identifying the form and the interior composition of nanoscale materials. Whereas the oscillating component of the PDF gives information about the interatomic distances within the material, the PDF baseline is a function of the characteristic function, a measure of nanoparticle shape which has its origin in the SAS that is usually disregarded in a powder-diffraction experiment. This characteristic function goes unnoticed in macroscopic particles, where the PDF is observed at distances that are much smaller than the particle diameter. For nanoparticles, the PDF baseline, and therefore the characteristic function, cannot be disregarded. We have presented a full derivation of the PDF equation taking into account the missing SAS and have reviewed different methods for calculating the PDF for nanoparticles. Given the relationship between the PDF and SAS equations, there is potential benefit in incorporating SAS data and analysis methods into PDF studies.

The authors would like to thank Phil Duxbury and Pavol Juhás for careful proofreading and discussion of this manuscript. The authors acknowledge enlightening conversations with Matteo Leoni, Reinhard Neder, Thomas Proffen, Chong-Yu Ruan, Paolo Scardi and Mike Thorpe, and thank an anonymous referee for insightful and conductive comments. This work was supported by the US National Science Foundation through grant DMR-0703940.

References

Bedford, N., Dablemont, C., Viau, G., Chupas, P. & Petkov, V. (2007). *J. Phys. Chem. C*, **111**, 18214–18219.

Billinge, S. J. L. (1998). *Local Structure from Diffraction*, edited by S. J. L. Billinge & M. F. Thorpe, p. 137. New York: Plenum.

Billinge, S. J. L. (2008). *J. Solid State Chem.* **181**, 1698–1703.

Billinge, S. J. L. & Kanatzidis, M. G. (2004). *Chem. Commun.* pp. 749–760.

Billinge, S. J. L., Kwei, G. H. & Takagi, H. (1994). *Phys. Rev. Lett.* **72**, 2282.

Billinge, S. J. L. & Levin, I. (2007). *Science*, **316**, 561–565.

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

Cervellino, A., Giannini, C. & Guagliardi, A. (2006). *J. Comput. Chem.* **27**, 995–1008.

Chacón, P., Morán, F., Díaz, J. F., Pantos, E. & Andreu, J. M. (1998). *Biophys. J.* **74**, 2760–2775.

Chupas, P. J., Qiu, X., Hanson, J. C., Lee, P. L., Grey, C. P. & Billinge, S. J. L. (2003). *J. Appl. Cryst.* **36**, 1342–1347.

Debye, P. (1915). *Ann. Phys. (Berlin)*, **46**, 809–823.

Debye, P. & Menke, H. (1930). *Phys. Z.* **31**, 797–798.

Dinnebier, R. E. & Billinge, S. J. L. (2008). Editors. *Powder Diffraction: Theory and Practice.* London: Royal Society of Chemistry.

Egami, T. & Billinge, S. J. L. (2003). *Underneath the Bragg Peaks: Structural Analysis of Complex Materials.* Oxford: Pergamon Press/Elsevier.

Egami, T., Rosenfeld, H. D., Toby, B. H. & Bhalla, A. (1991). *Ferroelectrics*, **120**, 11.

Ehm, L., Antao, S. M., Chen, J., Locke, D. R., Michel, F. M., Martin, C. D., Yu, T., Parize, J. B., Antao, S. M., Lee, P. L., Chupas, P. J., Shastri, S. D. & Guo, Q. (2007). *Powder Diffr.* **22**, 108–112.

Farrow, C. L., Juhás, P., Liu, J. W., Bryndin, D., Božin, E. S., Bloch, J., Proffen, T. & Billinge, S. J. L. (2007). *J. Phys. Condens. Matter*, **19**, 335219.

Franklin, R. E. (1950). *Acta Cryst.* **3**, 107–121.

Franklin, R. E. (1951). *Proc. R. Soc. London Ser. A*, **209**, 196.

Gateshki, M., Hwang, S.-J., Park, D., Ren, Y. & Petkov, V. (2004). *J. Phys. Chem. B*, **108**, 14956–14963.

Gilbert, B. (2008). *J. Appl. Cryst.* **41**, 554–562.

Gilbert, B., Huang, F., Zhang, H., Waychunas, G. A. & Banfield, J. F. (2004). *Science*, **305**, 651–654.

Glatter, O. & Kratky, O. (1982). Editors. *Small Angle X-ray Scattering*, 1st ed. London: Academic Press Inc.

Guinier, A. (1963). *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies.* San Francisco: W. H. Freeman.

Guinier, A., Fournet, G., Walker, C. & Yudowitch, K. (1955). *Small-Angle Scattering of X-rays.* New York: John Wiley and Sons, Inc.

Hargittai, I. & Hargittai, M. (1988). Editors. *Stereochemical Applications of Gas-Phase Electron Diffraction*, 1st ed. New York: VHC Publishers.

Jeong, I. K., Heffner, R. H., Graf, M. J. & Billinge, S. J. L. (2003). *Phys. Rev. B*, **67**, 104301.

Johnson, P. A. V., Wright, A. C. & Sinclair, R. N. (1982). *J. Non-Cryst. Solids*, **50**, 281–311.

Kaplow, R., Rowe, T. A. & Averbach, B. L. (1968). *Phys. Rev.* **168**, 1068.

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

Keen, D. A. (2001). *J. Appl. Cryst.* **34**, 172–177.

Klug, H. P. & Alexander, L. E. (1974). *X-ray Diffraction Procedures for Polycrystalline and Amorphous Materials*, 2nd ed. New York: Wiley.

Kodama, K., Iikubo, S., Taguchi, T. & Shamoto, S. (2006). *Acta Cryst.* A**62**, 444–453.

Korsunskii, V. I. (1985). *J. Struct. Chem.* **26**, 208–216.

Korsunskiy, V. I. & Neder, R. B. (2005). *J. Appl. Cryst.* **38**, 1020–1027.

Korsunskiy, V. I., Neder, R. B., Hofmann, A., Dembski, S., Graf, C. & Rühl, E. (2007). *J. Appl. Cryst.* **40**, 975–985.

Levashov, V. A., Billinge, S. J. L. & Thorpe, M. F. (2007). *J. Comput. Chem.* **28**, 1865–1882.

Liz-Marzán, L. M., Giersig, M. & Mulvaney, P. (1996). *Langmuir*, **12**, 4329–4335.

McGreevy, R. L. & Pusztai, L. (1988). *Mol. Simul.* **1**, 359–367.

McKenzie, D. R., Davis, C. A., Cockayne, D. J. H., Muller, D. A. & Vassallo, A. M. (1992). *Nature (London)*, **355**, 622–624.

Masadeh, A. S., Božin, E. S., Farrow, C. L., Paglia, G., Juhás, P., Karkamkar, A., Kanatzidis, M. G. & Billinge, S. J. L. (2007). *Phys. Rev. B*, **76**, 115413.

Nanao, S., Dmowski, W., Egami, T., Richardson, J. W. & Jorgensen, J. D. (1987). *Phys. Rev. B*, **35**, 435–440.

Neder, R. B. & Korsunskii, V. I. (2005). *J. Phys. Condens. Matter*, **17**, S125–S134.

Neder, R. B., Korsunskii, V. I., Chory, C., Müller, G., Hofmann, A., Dembski, S., Graf, C. & Rühl, E. (2007). *Phys. Status Solidi C*, **4**, 1610–1634.

Page, K., Proffen, T., Terrones, H., Terrones, M., Lee, L., Yang, Y., Stemmer, S., Seshadri, R. & Cheetham, A. K. (2004). *Chem. Phys. Lett.* **393**, 385–388.

Petkov, V. & Danev, R. (1998). *J. Appl. Cryst.* **31**, 609–619.

Petkov, V., Gateshki, M., Choi, J., Gillan, E. G. & Ren, Y. (2005). *J. Mater. Chem.* **15**, 4654.

Petkov, V., Gerber, T. & Himmel, B. (1998). *Phys. Rev. B*, **58**, 11982–11989.

Pradhan, S. K., Mao, Y., Wong, S. S., Chupas, P. & Petkov, V. (2007). *Chem. Mater.* **19**, 6180–6186.

Proffen, T. & Billinge, S. J. L. (1999). *J. Appl. Cryst.* **32**, 572–575.

Qiu, X., Proffen, T., Mitchell, J. F. & Billinge, S. J. L. (2005). *Phys. Rev. Lett.* **94**, 177203.

Rayleigh, L. (1914). *Proc. R. Soc. London Ser. A*, **90**, 219.

Ruan, C.-Y., Murooka, Y., Raman, R. K. & Murdick, R. A. (2007). *Nano Lett.* **7**, 1290–1296.

Soper, A. K. (2007). *J. Phys. Condens. Matter*, **19**, 335206.

Soper, A. K. & Silver, R. N. (1982). *Phys. Rev. Lett.* **49**, 471–474.

Svergun, D. I. & Stuhrmann, H. B. (1991). *Acta Cryst.* A**47**, 736–744.

Thorpe, M. F. & Lei, M. (2007). *Nanoparticle Shape Factors for Spheroids*. Unpublished.

Thorpe, M. F., Levashov, V. A., Lei, M. & Billinge, S. J. L. (2002). *From Semiconductors to Proteins: Beyond the Average Structure*, edited by S. J. L. Billinge & M. F. Thorpe, pp. 105–128. New York: Kluwer/Plenum.

Toby, B. H. & Egami, T. (1992). *Acta Cryst.* A**48**, 336–346.

Tucker, M. G., Dove, M. T. & Keen, D. A. (2001). *J. Appl. Cryst.* **34**, 630–638.

Warren, B. E. (1934). *J. Phys. Chem.* **2**, 551.

Warren, B. E. (1990). *X-ray Diffraction*. New York: Dover.

Warren, B. E., Krutter, H. & Morningstar, O. (1936). *J. Am. Ceram. Soc.* **19**, 202–206.

Wright, A. C. (1985). *J. Non-Cryst. Solids*, **76**, 187.

Wright, A. C. (1998). *Glass Phys. Chem.* **24**, 148–179.

Zhang, H. Z., Gilbert, B., Huang, F. & Banfield, J. F. (2003). *Nature (London)*, **424**, 1025–1029.