Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 3 February 2010
Accepted 13 April 2011

Using molecular-dynamics simulations to understand and improve the treatment of anharmonic vibrations. II. Developing and assessing new Debye–Waller factors

Anthony M. Reilly,[a,b] Carole A. Morrison,[a]* David W. H. Rankin[a] and K. Robin McLean[c]

[a]School of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, Scotland, [b]Chair for Process Systems Engineering, Technische Universität München, D-85350 Freising, Germany, and [c]Department of Mathematical Sciences, Mathematical Sciences Building, University of Liverpool, Liverpool L69 7ZL, England. Correspondence e-mail: carole.morrison@ed.ac.uk

Two new anharmonic forms for the Debye–Waller factor, aimed at modelling curvilinear and asymmetric motion, have been introduced. These forms permit the refinement of structures with these types of anharmonic motion using a small number of additional parameters. Molecular-dynamics-derived numerical probability density functions (PDFs) have been used to assess the merit of these new functions in real space. The comparison is favourable particularly for the curvilinear PDF based on a parabolic coordinate system change of a trivariate Gaussian distribution. The initial results also suggest that high-order even terms from the Gram–Charlier series may be important for modelling methyl-group libration. The molecular-dynamics data sets provide useful insights into the nature of anharmonic thermal motion. Addressing the problem in real space allows intuitive PDFs to be developed but numerical methods may be necessary for these methods to be implemented in refinement programs as an analytical Debye–Waller factor cannot always be obtained.

1. Introduction

The Debye–Waller factor is one of the most important ingredients in a successful crystal structure refinement. It is formally defined as the Fourier transform or characteristic function of the three-dimensional probability density function (PDF), $P(\mathbf{u})$, that describes where an atom spends its time in space (Johnson, 1970; Castellano & Main, 1985; Kuhs, 1992). It can have contributions from disorder and systematic errors but our interest in the present work is its primary role: modelling the thermal motion of atoms. While there have been many advances in crystallographic technology, the majority of crystal structure refinements use a Debye–Waller factor devised over 50 years ago. This is despite the fact that numerous, more advanced forms have been developed in the intervening time. Traditionally, the development of these more advanced approaches has focused on the functional form of the Debye–Waller factor, giving generic and widely applicable methods. In this work we aim to illustrate how more specific, phenomenological models of thermal motion can be developed using the results of molecular-dynamics simulations. To achieve this we focus on modelling the intuitively more accessible form of the real-space PDF rather than the reciprocal-space Debye–Waller factor. In doing this the data could be fitted better, with fewer parameters than generic methods, leading to more accurate positional parameters.

The majority of small-molecule single-crystal studies determine six anisotropic displacement parameters (ADPs) which define the covariance matrix, $\mathbf{U}$, of an atom's motion about its mean position, $\mathbf{r}_a$, which is also directly refined as part of the structure-factor equation. In addition to the mean position, it is also possible to determine a most probable position for each atom, $\mathbf{r}_p$, which denotes the maximum (or mode) of its probability density function (Johnson, 1969). The six ADPs can be used to define a harmonic Debye–Waller factor (Kuhs, 1992; Johnson, 1969)

$$\hat{P}(\mathbf{Q})_{\rm harm} = \exp\left(-\frac{1}{2}\mathbf{Q}^T\mathbf{U}\mathbf{Q}\right), \qquad (1)$$

which is the Fourier transform of a trivariate Gaussian PDF:

$$P(\mathbf{u})_{\rm harm} = \frac{[\det(\mathbf{U}^{-1})]^{1/2}}{(2\pi)^{3/2}}\exp\left(-\frac{1}{2}\mathbf{u}^T\mathbf{U}^{-1}\mathbf{u}\right). \qquad (2)$$

The Gaussian or harmonic approximation is widespread in small-molecule crystallography and depictions of crystal

structures often use surfaces of constant probability, which for the Gaussian case are ellipsoidal in shape.

It has long been known that the harmonic approximation, while capturing most of the effect of thermal motion, can have its deficiencies. The harmonic modelling of librational motion of a molecule can lead to shorter intramolecular bond lengths than are physically reasonable (Cruickshank, 1956). This is because the mean atomic position moves towards the centre of libration, while the most probable position (the maximum of the probability density) remains at the more physically reasonable value. The mean and most probable positions for a harmonically modelled atom must coincide so the Debye–Waller factor is incapable of modelling the thermal motion properly. This leads not only to spurious geometries but also to incorrect intensities and poorer agreement factors. Similar issues arise from the asymmetric motion of atoms.

Numerous forms for the Debye–Waller factor have been devised to model librational and asymmetric motion accurately. The majority of these methods are based on series expansions of the harmonic Debye–Waller factor using polynomials or spherical harmonics. The Edgeworth (Johnson, 1969) and Gram–Charlier (Zucker & Schulz, 1982) methods focus on the statistical interpretation of $P(\mathbf{u})$. The one-particle potential (OPP) method (Willis, 1969) is based on the relationship between the PDF/Debye–Waller factor and the effective potential that an atom experiences, and has been formulated using polynomials (Willis, 1969; Tanaka & Marumo, 1983) and also spherical harmonics (Kurki-Suonio *et al.*, 1979). Other methods have been devised to model specific types of motion including librational motion upon an arc of a circle (Willis & Pawley, 1970; Pawley & Willis, 1970).

The most widely used methods are the statistical and OPP ones. The advantage of using series expansions is a high degree of flexibility. However, this flexibility comes at a cost. The harmonic Debye–Waller factor requires six parameters per atom whereas the anharmonic methods can require anywhere from 16 to more than 60 parameters *per atom*. The generic nature of these approximations and the resultant cost in terms of parameters stem from a lack of independent information on the nature and effect of anharmonic motion. Many of the anharmonic series expansions are not always convergent, leading to difficulties in assessing the convergence of the fit properly.

In situations where experiment alone cannot provide sufficient information, it is now common to turn to theoretical methods. Molecular dynamics (MD) is a powerful technique for exploring the dynamic behaviour of chemical systems. We have applied it previously to determine corrections to experimental crystal structures that convert the basic time-averaged atomic positions to equilibrium values (Reilly *et al.*, 2007). In addition, the MD simulation can be used to determine numerical PDFs, free of an approximation of their mathematical form.

In the preceding paper (Reilly *et al.*, 2011) we have looked at the different positional parameters and the effect of the PDF's functional form on the structures determined by diffraction experiments. In the present study we use the numerical PDFs to investigate two new forms of PDF (and Debye–Waller factor), comparing them to the traditional Gram–Charlier (GC) approach. The following section outlines the simulation and modelling approach used to determine and fit the numerical PDFs, while §3 outlines the GC method and §4 discusses orientating a PDF/Debye–Waller factor along particular directions to reduce the number of significant parameters. The two new PDFs for librational and asymmetric motion are discussed in §§5 and 6, respectively, and the application of these models to a selection of molecules is presented in §7.

2. Methodology

2.1. MD simulations

The results of three MD simulations on crystallographic models are used in the present work. The first system is the 1:1 adduct of urea and phosphoric acid (UPA), which features a short, strong hydrogen bond with the shared migratory proton possessing an asymmetric, egg-shaped PDF, at least at 150 K. Details of the MD simulations are given in the preceding paper (Reilly *et al.*, 2011). The second system is deuteronitromethane, chosen as a simple system possessing almost free rotation of a methyl group [discussed in Reilly *et al.*, 2010, 2011]. The third system is the 2:1 adduct of dimethylurea and oxalic acid (DMUOX), also chosen for methyl-group free-rotation behaviour. Details of the simulations performed for DMUOX are as follows. Gaussian plane-wave (GPW) density functional theory (DFT) (Lippert *et al.*, 1997) MD simulations of phase I of DMUOX (Pulham, 2009) were performed using the *CP2K* program (VandeVondele *et al.*, 2005). To simulate the periodic nature of the crystal structure correctly a $3 \times 1 \times 1$ supercell was modelled using periodic boundary conditions. The electronic wavefunction was represented in real space using double-Gaussian basis functions with polarization functions, while plane waves were used in reciprocal space to represent the electron density. A density cutoff energy of 4000 eV was used. Valence–core interactions were modelled using GPW-optimized analytical pseudopotentials (Goedecker *et al.*, 1996; Hartwigsen *et al.*, 1998; Krack, 2005). A Nosé–Hoover chain (Nosé, 1984; Hoover, 1985) was used to regulate the temperature at 350 K during a Born–Oppenheimer MD simulation with a time step of 0.55 fs. Data were collected for 21 ps. Initial analysis of the trajectory showed essentially free rotation of the methyl groups. For brevity, only the H atom marked by an asterisk in Fig. 1 is discussed here.

2.2. Analysis

The MD trajectories were analysed using Fortran code to calculate the time-averaged mean position and covariance matrix (the six unique U^{ij} values), as well as one-, two- and three-dimensional numerical PDFs for each atom. More details are given in Reilly *et al.* (2011).

The fitting of the numerical PDFs to specific analytical functions in direct space was carried out using the *Mathema-*

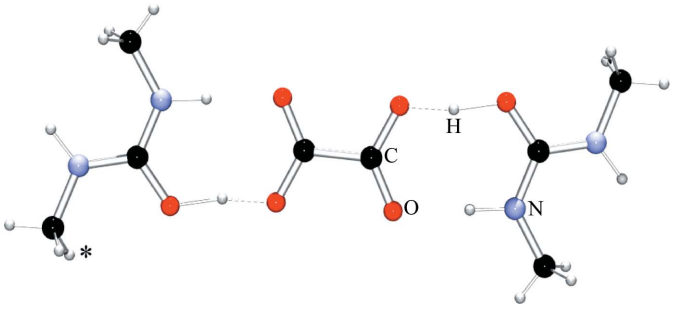

Figure 1
Molecular structure of the 2:1 adduct of dimethylurea and oxalic acid.

tica computer program (Wolfram Research Inc., 2007). It is convenient to have a measure of how well a particular function fits the numerical PDFs. This is defined in a similar manner to the R factor in crystallography:

$$R = \frac{\sum |P_{\rm n}(\mathbf{u}) - P_{\rm f}(\mathbf{u})|}{\sum P_{\rm n}(\mathbf{u})} \times 100, \qquad (3)$$

where $P_{\rm n}$ is the numerical (or observed) PDF and $P_{\rm f}$ is the model or fitted PDF. By definition $P_{\rm n}$ must be positive over all space and so there is no need to take its modulus. For those points at which $P_{\rm f}$ is less than zero $P_{\rm f}$ has been set equal to zero in equation (3). The summation is taken over data points on a cubic grid centred on the atom's mean position, but only those points where probability density is 'observed' (*i.e.* $P_{\rm n} \neq 0$) in the MD simulation have been included in the summation, as is normally the case for the crystallographic R factor. For nitromethane the grid consisted of 200^3 data points with a spacing of 0.015 Å. For the other molecules a smaller grid of 100^3 points (with spacings of 0.024 Å for UPA and 0.04 Å for DMUOX) was used.

In normal crystallographic usage, a good or acceptable value for the R factor is a value less than 10%. Here, however, the R factor should only be taken as a measure of the relative qualities of the fits of two or more models to a particular MD data set, since any high-frequency noise present in the numerical data set (which arises as a result of the binning process or a short trajectory) will adversely affect it. The fitting process, which focuses on the broad, low-frequency features of the distribution, should not be affected. For this reason, a low-pass Fourier filter has been used to aid the visualization of the numerical data sets; the reported R-factor values, however, were calculated using the unfiltered data sets, since the removal of the high-frequency noise would affect all the models used to the same degree.

3. Gram–Charlier series

The GC approach is one of the most widely used methods in anharmonic refinements and has been recommended by a number of authors (Trueblood *et al.*, 1996; Kuhs, 1992; Zucker & Schulz, 1982). It will therefore be used as the benchmark for the new PDFs detailed in §§5 and 6. The method involves an expansion of the harmonic PDF using high-order quasi-moments (Kuznetsov *et al.*, 1960), $c^{ijkl\ldots}$, and Hermite polynomials, $H_{ijkl\ldots}$:

$$P_{\rm GC}(\mathbf{u}) = P(\mathbf{u})_{\rm harm} \times \left[1 + \frac{1}{3!}c^{ijk}H_{ijk}(\mathbf{u}) + \frac{1}{4!}c^{ijkl}H_{ijkl}(\mathbf{u}) + \ldots\right], \qquad (4)$$

where $P(\mathbf{u})_{\rm harm}$ is a standard trivariate Gaussian PDF and indices that are repeated twice are implicitly summed over (from 1 to 3). The linear and quadratic polynomials are omitted as they can be correlated with the mean and the covariance matrix. The corresponding Debye–Waller factor is

$$\hat{P}_{\rm GC}(\mathbf{Q}) = \hat{P}_{\rm harm}(\mathbf{Q}) \times \left(1 - ic^{ijk}Q_iQ_jQ_k + c^{ijkl}Q_iQ_jQ_kQ_l + \ldots\right). \qquad (5)$$

The odd-order terms skew the distribution, while the even-order terms affect its 'peakedness' or physical extent. The GC series' flexibility has seen it used to model disorder and split-atom systems (Kuhs, 1983). Despite its repeated use, the GC series does have some drawbacks. There are ten third-order and 15 fourth-order parameters. In some cases, even higher-order terms might be necessary. In addition to requiring many more parameters than the harmonic approximation the GC series PDF is not strictly speaking a true PDF: the polynomial nature of the expansion means that there can be regions of negative (and therefore meaningless) probability.

4. Orientation of anharmonic PDFs

In our analysis of the numerical PDFs in Reilly *et al.* (2011) we employed GC series centred and orientated so that $U^{ij} = 0$ $(i \neq j)$ with the longest principal axis directed along the x axis and the shortest along the z axis. The required transformation matrix to diagonalize $\mathbf{U}$ is its eigenvector matrix, which can be easily determined numerically or analytically (Kronenburg, 2004). The fitting process is much faster when the covariances are zero as the GC and other PDF formulae are simpler. In addition, two-dimensional plots of the PDFs are easier to interpret when they are orientated along the principal axes. For normal crystallographic use the Hermite polynomials of the GC series are defined with the standard covariance matrix, $\mathbf{U}$, having non-zero off-diagonal elements. The skewing, bending *etc.* behaviour of the polynomials that results from using this covariance matrix is not directed along the crystallographic axes or any particularly significant directions unless symmetry or the covariances dictate that this be the case. This orientation and approach are suitable for generic anharmonicity or speculative refinements of higher-order terms.

For some situations it is possible to guess, based on chemical intuition, when anharmonicity may be important and even what form this anharmonicity might take. We might expect a significant degree of curvilinear motion perpendicular to the rotational axis of a methyl group and asymmetry in the PDF of a terminal atom or a migratory proton. Both effects have been seen in our MD simulations of nitromethane, UPA and DMUOX. For such systems it may be useful to replace the generic basis vectors of the GC PDF as usually defined by ones

Table 1
Third-order quasi-moments (in Å^3) for the general and orientated fits of the D1 atom of d_3-nitromethane at 15 K.

The pre-factor and multiplicities required for use in a crystallographic refinement have been omitted.

	General		Orientated	
	c^{ijk}	$\lvert c^{ijk}/\sigma_{c^{ijk}}\rvert$	c^{ijk}	$\lvert c^{ijk}/\sigma_{c^{ijk}}\rvert$
c^{111}	0.000114	100	−0.000593	48
c^{222}	−0.000381	277	0.000006	5
c^{333}	0.000025	113	−0.000001	2
c^{112}	−0.000155	52	0.000524	41
c^{122}	−0.000542	166	0.000059	9
c^{133}	−0.000184	213	0.000018	6
c^{113}	0.000550	376	0.001314	140
c^{123}	0.001457	530	0.000061	9
c^{223}	0.000896	550	0.000046	21
c^{233}	−0.000211	238	−0.000008	5

oriented along specific directions. This sort of approach has been used before for the OPP method (Tanaka & Marumo, 1983). The benefit of changing the orientation of the PDF is that fewer Hermite polynomials, and therefore parameters, may be required to model the anharmonicity. Such a re-orientation will not always be valid, especially when there is a possibility for coupling or mixing of different vibrational motions.

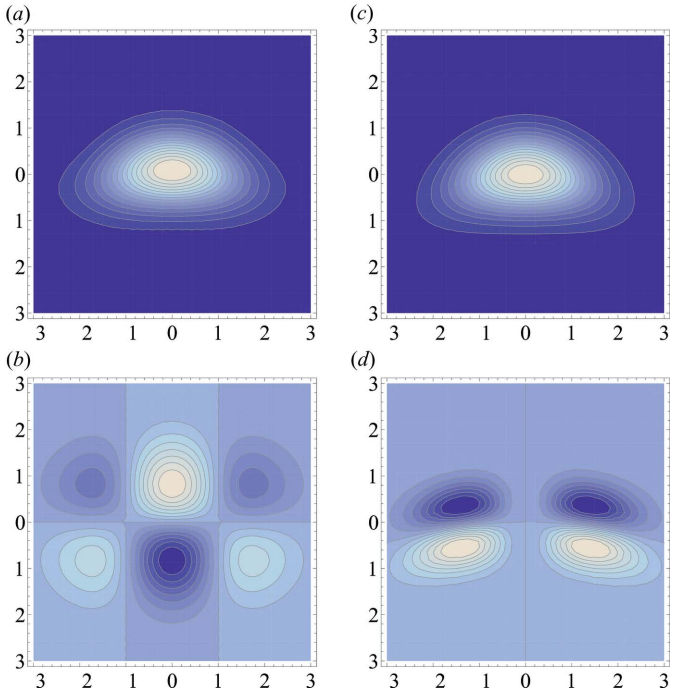

Figure 2
Probability density maps of (*a*) the anharmonic distribution $P(\mathbf{u})_{\rm harm}(1 + H_{122})$, (*b*) $[P(\mathbf{u})_{\rm anharm,GC} - P(\mathbf{u})_{\rm harm}]$, (*c*) an anharmonic distribution using the parabolic coordinate change $\hat{u}_2 = u_2 + 0.1u_1^2$ and (*d*) $[P(\mathbf{u})_{\rm anharm,para} - P(\mathbf{u})_{\rm harm}]$. The harmonic distribution, $P(\mathbf{u})_{\rm harm}$, is proportional to $\exp\{1/2[-(u_1^2/U^{11}) - (u_2^2/U^{22})]\}$, where $U^{11} = 1$ and $U^{22} = 0.2$. The u_1 axis is vertical; dark probability densities indicate negative contributions [or zero density in (*a*) and (*c*)] while lighter ones indicate positive contributions. The units on the axes are arbitrary.

In our simulations of d_3-nitromethane the anharmonicity in the D atoms is clearly along the largest principal axis with a PDF that is both curved and skewed. We have calculated the third-order GC fit to the PDF of the D1 atom of nitromethane at 15 K using a GC series defined along the crystallographic axes (the standard GC formulation) and along the principal axes of the harmonic approximation (as we have done previously for visualization and analysis purposes). For this comparison, the distributions have both been centred on the numerical mean as the standard GC PDF expressions are far more complex if the mean is non-zero. (In a crystallographic refinement the 'mean' is fitted separately from the Debye–Waller factor.) The resulting third-order quasi-moments for both models are given in Table 1. All ten of the standard GC series quasi-moments are large and significant as the $\lvert c^{ijk}/\sigma_{c^{ijk}}\rvert$ values show. In comparison the oriented distribution has only three large values, c^{111}, c^{112} and c^{113}. The remaining seven quasi-moments are at least an order of magnitude smaller and their corresponding contributions to the probability function are negligible compared to the other three.

As we have used an orientated frame of reference we can associate the quasi-moments with specific contributions. For example, without plotting any density maps, we know that H_{111} skews the PDF along the longest principal axis. At higher temperatures all of the c^{ijk} values become more significant but the c^{111}, c^{112} and c^{113} values still give the largest contributions, being more than an order of magnitude bigger than the other values.

In many cases the transformation of the GC PDF into a non-crystallographic basis may offer no benefits but it is clear from Table 1 that for some situations it may reduce the number of parameters. Performing the transformation so that the basis vectors coincide with the harmonic principal axes of thermal motion requires no extra parameters. A useful example of such a transformation is given by Tanaka & Marumo (1983), and rules and mathematics for transforming the coordinate system of the Hermite polynomials (and indeed other crystallographic quantities) are discussed elsewhere (Rowicka *et al.*, 2004; Giacovazzo *et al.*, 2002).

5. Curvilinear PDFs

For molecular crystals the most manifest form of anharmonic motion is librational or other curvilinear motion. The GC series can model curvature using its H_{iij} or H_{ijj} terms. In Table 1 the large and significant c^{112} and c^{113} terms indicate that the effect of the third-order parameters is to bend the longest axis (axis 1). The density contribution of a Hermite polynomial to the overall distribution is given by $P(\mathbf{u})_{\rm harm} \times c^{ijk\cdots} H(\mathbf{u})_{ijk\ldots}$. Fig. 2(*a*) shows an anharmonic GC PDF and Fig. 2(*b*) shows the corresponding difference density distribution.

5.1. Parabolic coordinate system

As well as perturbing a Cartesian Gaussian PDF with polynomials we might also consider a change of the coordinate system from a Cartesian one to a curved one. In particular,

librational motion could be represented as having a circular, parabolic or higher-order even polynomial. We consider a parabolic transformation of the coordinates, **u**, of a trivariate Gaussian:

$$\begin{aligned}\hat{u}_1 &= u_1\\ \hat{u}_2 &= u_2 + ku_1^2\\ \hat{u}_3 &= u_3 + k'u_1^2,\end{aligned} \tag{6}$$

where k and k' are suitable bending constants. The effect of this transformation is to move a point lying on the plane $u_2 = 0$ to one lying on the parabolic cylinder $u_2 = ku_1^2$ and likewise to move a point on the plane $u_3 = 0$ to one on $u_3 = k'u_1^2$. For simplicity we can formulate the PDF in a coordinate system where $U^{ij} = 0$ when $i \neq j$; the methods referenced in the previous section can be used to give a general PDF, which takes the form

$$P(\hat{\mathbf{u}}) = \frac{[\det(\mathbf{U}^{-1})]^{1/2}}{(2\pi)^{3/2}} \times \exp\left\{-\frac{1}{2}\left[\frac{u_1^2}{U^{11}} + \frac{(u_2 + ku_1^2)^2}{U^{22}} + \frac{(u_3 + k'u_1^2)^2}{U^{33}}\right]\right\}. \tag{7}$$

The resulting parabolic transformation takes the form shown in Fig. 2(*c*). Fig. 2(*d*) shows the corresponding density contribution. It gives the desired curving effect. A PDF of the form given in equation (7) might be a better approximation than the GC for a librating atom because the density contribution is a better model of motion along the arc of a circle, which is a reasonable approximation for the torsional motion in nitromethane for small amplitudes. The Hermite polynomial (Fig. 2*b*) changes sign twice in moving parallel to the longest axis (the u_2 axis), adding and subtracting density near the mean for $u_1 > 0$, while it removes density further away, whereas the parabolic PDF is a natural first-order approximation to the circular motion (Fig. 2*d*). To incorporate parameters representing the mean we can use a PDF of the form

$$P(\hat{\mathbf{u}}) = \frac{[\det(\mathbf{U}^{-1})]^{1/2}}{(2\pi)^{3/2}} \exp\left(-\frac{1}{2}\left\{\frac{(u_1 - x)^2}{U^{11}} + \frac{[u_2 + k(u_1 - x)^2 - y]^2}{U^{22}} + \frac{[u_3 + k'(u_1 - x)^2 - z]^2}{U^{33}}\right\}\right), \tag{8}$$

where $\mathbf{x} = (x, y, z)$ represents the mean when $k = k' = 0$.

5.2. Mean and variance

One of the strengths of the Gaussian, GC and Edgeworth series is that the mean and variances of the model distribution are directly determined as fitting parameters. One consequence of the parabolic PDF is that the mean and variance terms from a normal Gaussian distribution no longer correspond to the actual mean and variances of the distribution. Instead they must be found by using the equation for the non-centred moments of a PDF:

$$\langle u_i^n \rangle = \int_{-\infty}^{\infty} u_i^n P(\hat{\mathbf{u}})\, \mathrm{d}\mathbf{u}. \tag{9}$$

For $n = 1$ equation (9) gives the mean, so for a PDF of the form given by equation (8) it can be shown that

$$\begin{aligned}\overline{u}_1 &= x\\ \overline{u}_2 &= y - kU^{11}\\ \overline{u}_3 &= z - k'U^{11}.\end{aligned} \tag{10}$$

For $n = 2$ equation (9) gives the variance plus the square of the mean. Then, combining equations (9) and (10) yields variances of

$$\begin{aligned}\overline{u}_1^2 &= \sigma_1^2 = U^{11}\\ \overline{u}_2^2 &= \sigma_2^2 = U^{22} + 2k^2(U^{11})^2\\ \overline{u}_3^2 &= \sigma_3^2 = U^{33} + 2k'^2(U^{11})^2.\end{aligned} \tag{11}$$

Naturally, when k or k' is zero then **x** and **U** have their usual significance as the mean and covariance matrix of the distribution.

5.3. Structure factor

To use the parabolic PDF in a crystallographic refinement we have to find its Fourier transform or characteristic function. If we consider curvature only along a single direction then letting $k = 0$ in equation (8) gives a distribution of the form

$$P(\hat{\mathbf{u}}) \propto \exp\left(-\frac{1}{2}\left\{\frac{(u_1 - x)^2}{U^{11}} + \frac{(u_2 - y)^2}{U^{22}} + \frac{[u_3 + k'(u_1 - x)^2 - z]^2}{U^{33}}\right\}\right). \tag{12}$$

The characteristic function given by the expectation value of $\exp(i\mathbf{x}\cdot\mathbf{Q})$ [*i.e.* $\langle\exp(i\mathbf{x}\cdot\mathbf{Q})\rangle$] is

$$\hat{P}(\mathbf{Q}) = \exp[i(yQ_2 + zQ_3)] \exp\left[-\frac{1}{2}(Q_2^2U^{22} + Q_3^2U^{33})\right] \times \exp\left(ixQ_1 - \frac{iQ_1^2U^{11}}{-2i + 4k'Q_3U^{11}}\right)\left[(1 + 2iU^{11}k'Q_3)^{1/2}\right]^{-1}. \tag{13}$$

The first two exponential terms represent the standard harmonic parts of the Debye–Waller factor for the Q_2 and Q_3 directions. If $k' = 0$ then the usual Gaussian distribution is obtained for the Q_1 coordinate as well. For curvature in two directions [*i.e.* like equation (8)] only the third and fourth terms in equation (13) are affected:

$$\hat{P}(\mathbf{Q}) = \hat{P}(Q_2, Q_3)_{\mathrm{harm}} \exp\left[ixQ_1 - \frac{iQ_1^2U^{11}}{-2i + 4U^{11}(kQ_2 + k'Q_3)}\right] \times \left\{[1 + 2iU^{11}(kQ_2 + k'Q_3)]^{1/2}\right\}^{-1}. \tag{14}$$

Equations (13) and (14) are of a suitable form and length for use as analytical functions in refinement programs. Unfortunately, when the parabolic PDF is augmented with Hermite polynomials, as in equations (22)–(24), the analytical form for the structure factor becomes long and complicated, but this may not be an obstacle to their use in refinement programs.

An alternative to using an analytical Debye–Waller factor is to perform a numerical Fourier transform of the analytical PDF. Such an approach has been suggested before for methods such as the OPP approximation (Kuhs, 1992). It has also been applied by Hohlwein (1981) to calculate structure factors for orientationally disordered molecules. However, for a variety of reasons the numerical approach has not been widely tested. First, the analytical GC series is considered superior to the OPP approximation, the main anharmonic model that would benefit from a numerical approach. Secondly, the computational effort in terms of coding and more importantly the evaluation time would have been prohibitive when the majority of the investigations of such methods were carried out (in the 1970s and 1980s). Finally, it is preferable to have an analytical form to determine the various parameters directly in the refinement.

Neither of the second or third reasons is a substantial obstacle to implementing a PDF such as the parabolic one studied here. Harmonic refinements that took a significant amount of computational time and effort ten years ago can now be performed in minutes with modern codes and computer hardware. An analytical form for the PDF is preferable but the use of a numerical Debye–Waller factor would permit the application of a wider variety of anharmonic PDFs.

6. Skew-normal distribution

Curved PDFs only demonstrate one effect of anharmonic thermal motion. Skewed or asymmetric PDFs are found in MD simulations both for nitromethane (Reilly *et al.*, 2010) and for UPA. The GC series can effectively model asymmetry using the diagonal odd-order polynomials such as H_{111}. One disadvantage of using polynomial expansions is that for such odd-ordered polynomials there must exist a region of space where the probability is negative. In many cases this region will be far from the atom and also small in magnitude. However, at high temperatures negative regions can be a serious issue (Scheringer, 1988).

The skew-normal (SN) distribution was suggested by Azzalini (1985) for modelling asymmetric data sets. For a one-dimensional case it has the form

$$P(x)_{\rm SN} = 2P(x)\Phi(ax), \tag{15}$$

where $P(x)$ is a Gaussian with mean u and variance U, a is an asymmetry constant and $\Phi(ax)$ is a cumulative density function of the form

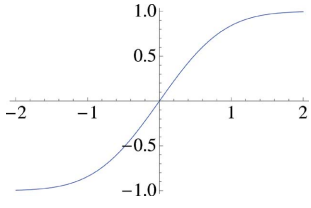

Figure 3
Graph of the error function erf(x) [see equation (17)].

$$\Phi(ax) = \frac{1}{2}\left\{1 + {\rm erf}\left[\frac{a(u-x)}{(2U)^{1/2}}\right]\right\}, \tag{16}$$

where the error function, erf(x), is given by

$${\rm erf}(x) = \frac{2}{\pi^{1/2}}\int_0^x \exp(-t^2)\,{\rm d}t. \tag{17}$$

The error function is limited to the range −1 to 1 and is shown in Fig. 3. Being limited to that range ensures that a PDF of the form of equation (15) cannot have any negative regions. The multivariate SN distribution is defined differently depending on the nature of the problem and whether or not correlation/orientation parameters are necessary. For our purposes it seems reasonable to start by skewing along the principal axes of the distribution. The distribution then takes the form (Gupta & Chen, 2004)

$$P(\mathbf{u})_{\rm SN} = P(\mathbf{u})_{\rm harm}\prod_{i=1}^{3}\left\{1 + {\rm erf}\left[\frac{a_i(x_i-u_i)}{(2U^{ii})^{1/2}}\right]\right\}. \tag{18}$$

The form of equation (18) skews each direction of the PDF in turn. This ensures that the distribution cannot have negative probability. When the skew coefficients a_i are zero we recover the normal Gaussian distribution. The mean of the distribution is

$$\overline{u}_i = x_i + (2/\pi)^{1/2}\delta_i(U^{ii})^{1/2}, \tag{19}$$

with $\delta_1 = a_i/(1+a_i^2)^{1/2}$. The variance is given by

$$\overline{u}_i^2 = \sigma_i^2 = U^{ii}[1-(2/\pi)\delta_i^2]. \tag{20}$$

The characteristic function of the PDF given by equation (18) is (Gupta & Chen, 2004; Azzalini, 2005)

$$\hat{P}(\mathbf{Q}) = \exp(i\mathbf{xQ})\exp\left(-\frac{1}{2}\mathbf{Q}^T\mathbf{UQ}\right)\prod_{i=1}^{3}\left\{1 + {\rm erf}\left[iQ_i\delta_i(U^{ii})^{1/2}\right]\right\}. \tag{21}$$

7. Application of anharmonic PDFs

Before applying the anharmonic PDFs introduced in the preceding sections, it is important to note that curvilinear or asymmetric motion will often not occur separately. Indeed, if we examine the values in Table 1 it is apparent that the D1 PDF is skewed as well as being bent. Therefore, to model anharmonic PDFs better, we could combine the different anharmonic approaches. For instance, to model the D atoms of nitromethane properly we can augment the parabolic PDF, $P(\hat{\mathbf{u}})$, with a third-order Hermite polynomial

$$P(\hat{\mathbf{u}})(1 - c^{111}H_{111}), \tag{22}$$

to produce a skewed parabolic PDF. It is quite likely that fourth-order contributions will be significant too (*cf.* Table 2 of Reilly *et al.*, 2011) so one might also supplement the above equation with the three main fourth-order Hermite polynomials

Table 2
Comparison of the various fits employed to model the numerical PDFs obtained from MD simulations at 15 and 228 K of some of the atoms of d_3-nitromethane.

See the text for details of the limited GC series used. (Note that the orientation of the model PDF is fixed to coincide with the harmonic principal axes, with the result being that there are three fewer parameters, namely the covariances, than would normally be required.)

Atom	Model	Parameters	*R* factor (%)
15 K			
D1	Harmonic	6	19.3
D1	Equation (22)	9	13.6
D1	Third-order GC series	16	14.2
D1	Limited third-order GC series	9	14.6
D1	Limited third-order GC series with error function	9	14.4
228 K			
D1	Harmonic	6	30.3
D1	Third-order GC series	16	22.5
D1	Limited third-order GC series	9	23.6
D1	Equation (22)	9	22.2
D1	Fourth-order GC series	19	17.9
D1	Equation (23)	12	18.3
D2	Fourth-order GC series	19	17.1
D2	Equation (23)	12	16.5
D2	Equation (24)	12	16.3
D2	Equation (25)	12	17.3
O1	Harmonic	6	14.9
O1	Third-order GC series	16	12.8
O1	Equation (18)	9	14.4
O1	Diagonal third-order GC series	9	14.4

$$P(\hat{\mathbf{u}})\left(1 - c^{111}H_{111} + \sum_{i=1}^{3} c^{iiii}H_{iiii}\right). \quad (23)$$

Two other combinations will be discussed in the following section. The first combines equation (22) with even-order Hermite polynomials

$$P(\mathbf{u}) = P(\hat{\mathbf{u}}) \times (1 - c^{111}H_{111} + c^{1111}H_{1111} + c^{111111}H_{111111} + c^{11111111}H_{11111111}), \quad (24)$$

while the second is the same equation but with the third-order Hermite polynomials replaced by an SN term:

$$P(\hat{\mathbf{u}})\left\{1 + \mathrm{erf}\left[\frac{a(x-u_1)}{(2U^{11})^{1/2}}\right] + c^{1111}H_{1111} + c^{111111}H_{111111} + c^{11111111}H_{11111111}\right\}. \quad (25)$$

7.1. Nitromethane

The D atoms of nitromethane represent a perfect system for studying modelling of the librational motion of methyl groups. In addition, at higher temperatures the O atoms display asymmetric motion, with egg-shaped PDFs obtained from the MD simulations (Reilly *et al.*, 2010). We begin by analysing the D1 PDF at 15 K.

7.1.1. D1 atom. When fitted to the numerical PDF the skewed parabolic PDF given by equation (22), with three anharmonic parameters, gave an agreement factor of 13.6%. The 16-parameter third-order GC fit (three means, three variances and ten quasi-moments) to the D1 atom of nitromethane at 15 K gave an *R* factor of 14.2%, while the trivariate harmonic model gave a fit of 19.3%. While the difference between the GC and parabolic fits may not be that large, it should be noted that the parabolic model has seven fewer parameters. It is also clear that they both provide a much better fit to the distribution than the harmonic model. The full GC model has seven more parameters than the parabolic model but if we restrict the GC series to the three most important terms (*cf.* Table 1) then they have the same number of parameters.

At 228 K the D1 atom shows significantly more anharmonicity than at 15 K and a number of different models have been employed to fit the numerical data; the results are presented in Table 2 and two-dimensional *xz* distributions are plotted in Fig. 4. Both the skewed parabolic and GC series give significantly lower *R* factors compared to the harmonic model. Both anharmonic distributions give similar values, which is encouraging from the point of view of using the parabolic PDF as it fits as well as the GC series with far fewer parameters. A limited third-order GC series, with only the three most important terms, performs slightly worse than the parabolic model. The k and k' values of the parabolic distribution are −0.170 (1) and −0.216 (1) Å^{-1}, respectively. While they have

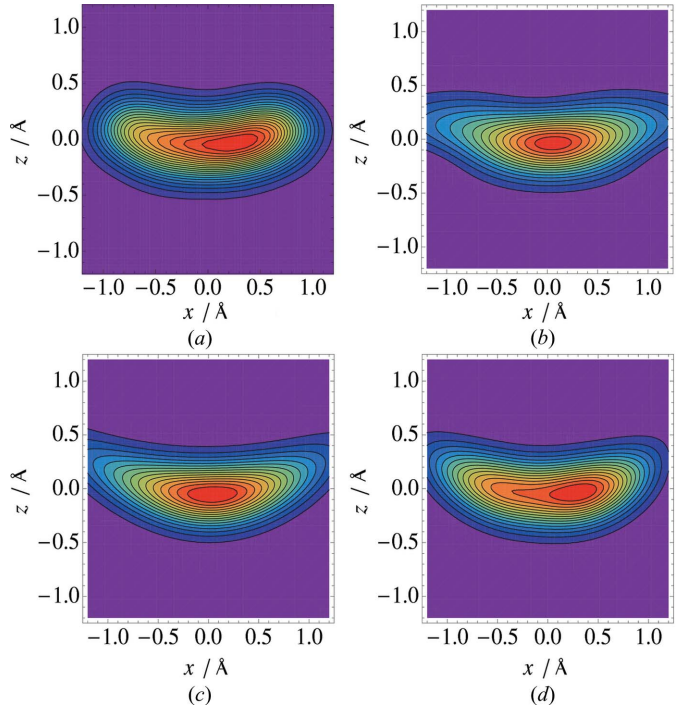

Figure 4
Two-dimensional *xz* PDFs of the D1 atom of nitromethane at 228 K: (*a*) Fourier-filtered numerical PDF, (*b*) third-order GC fit, (*c*) skewed parabolic PDF and (*d*) fourth-order skewed parabolic PDF. (The *x* axis represents the longest principal axis of thermal motion, while *z* is the shortest axis.)

dimensions of Å^{-1} it is important to remember the value of U^{ii} they are associated with. As U^{33} is much smaller than U^{22} the x axis will be bent mostly in the z direction.

However, as Figs. 4(b) and 4(c) show, while both of these anharmonic distributions are bent, neither matches the large asymmetry seen in the numerical PDF. They also extend much further out than the numerical distribution. In the preceding paper (Reilly *et al.*, 2011) it was shown how the fourth-order GC parameters were important for representing the size of the D1 PDF. By including the three diagonal fourth-order Hermite polynomials in the GC model and the parabolic model [equation (23)] we obtain significantly smaller R factors (see Table 2) and better visual agreement. The two-dimensional xz PDF of the fourth-order parabolic fit is presented in Fig. 4(d). Mathematically, the third- and fourth-order Hermite polynomials are not correlated and so in principle should work independently of each other. However, the high-probability region in the centre of the numerical distribution is so broad compared to a harmonic distribution that without fourth-order polynomials the third-order polynomial cannot skew the distribution without increasing the R factor. Visually, the agreement could still be better; the extent of the PDF remains too large. The parabolic PDF produces a more realistic curvature but evidently requires more even-order Hermite polynomial terms to produce a better fit. Of the three diagonal terms only the H_{1111} term seems important. Its value of $-6.64\,(1) \times 10^{-3}$ Å^4 is two orders of magnitude larger than the two other values. Examining the xy, xz and yz two-dimensional distributions it is clear that the vast majority of the anharmonicity is in the xz plane. In two dimensions there are only five unique fourth-order terms. Adding the H_{1111}, H_{1133}, H_{1333} and H_{1113} Hermite polynomials to the expansion in equation (22) reduces the R factor slightly but affects the visual appearance of the distribution only marginally, suggesting that the H_{1111} term is the most important in this type of system and that a full fourth-order refinement might not improve things greatly. Adding in the sixth-order diagonal terms also slightly reduces the parabolic PDF agreement R factor but again the general shape is still very similar to that in Fig. 4(d).

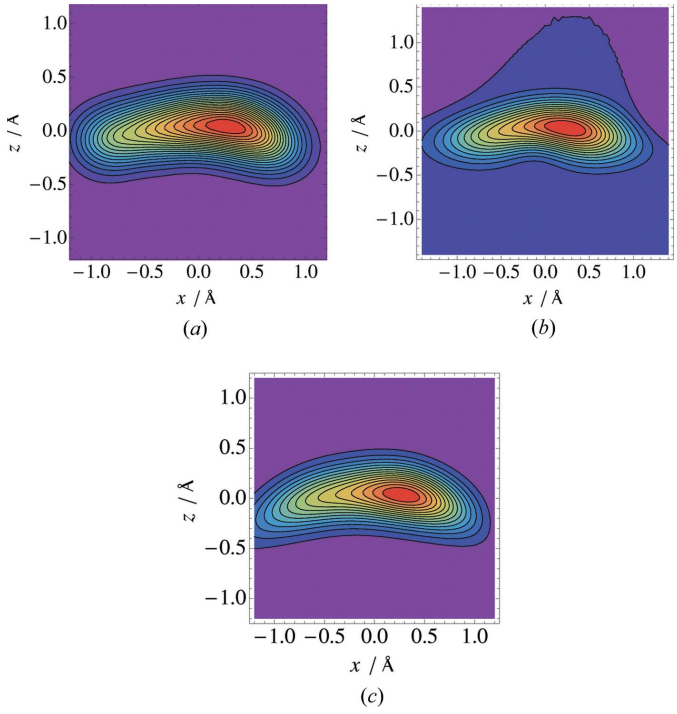

Figure 5
Two-dimensional xz PDFs of the D2 atom of nitromethane at 228 K: (a) Fourier-filtered numerical PDF, (b) fourth-order PDF GC fit and (c) fourth-order skewed parabolic PDF. The dark purple region indicates the lowest probability, which in the case of (b) is negative. (The x axis represents the longest principal axis of thermal motion, while z is the shortest axis.)

As well as considering the models in terms of R factors, we can compare the positions and bond lengths that result from them. Of particular importance is how well the anharmonic distributions correct for the bond-shortening effect of librational motion. The three-dimensional maximum of the fourth-order GC PDF, relative to the mean, is (0.1875, −0.0090, −0.0296), while the fourth-order parabolic PDF (Fig. 4d) has a three-dimensional maximum at (0.3078, −0.0009, −0.0343). The difference is quite large. Fitting a third-order GC series to the C atom and using the parabolic maximum for the D1 atom, a most probable C—D1 bond length of 1.084 Å is obtained, which compares well with the r_e value of 1.089 Å. The GC series fit for the D1 atom gives a bond length of 1.074 Å. The time-averaged bond length, which is what would be measured using a normal harmonic refinement, is 0.991 Å, showing that both methods correct for the vast majority of the librational shortening effect.

7.1.2. D2 atom. For the D1 atom of nitromethane the parabolic PDF of equation (23) consistently produced good agreement with the numerical data set. This is in part because the direction of curvature matches the principal axes of thermal motion quite well. The D2 atom of nitromethane has a much more skewed PDF (Fig. 5a), which should favour the GC series over the parabolic PDF. Again, a number of different models have been used to fit the numerical data set and these are presented in Table 2. Surprisingly, the parabolic distribution still outperforms the GC distribution despite the skewed nature of the numerical PDF. Figs. 5(b) and 5(c) show the GC and parabolic two-dimensional xz PDFs. A small region of negative probability is clearly visible for the fourth-order GC series. The potential for such regions of negative probability is one of the main disadvantages of PDFs based on polynomial expansions, although in some cases it is acceptable as long as the regions are small (Scheringer, 1988; Kuhs, 1992). Noting that the H_{1111} term was the only fourth-order term to be significant for the D1 atom, we have fitted a distribution with the corresponding sixth- and eighth-order terms [equation (24)], which gives a slightly smaller R factor but hardly any improvement in the visual fit. Refining a third-order GC series with the fourth, sixth and eighth-order x-axis terms also leads to little improvement over the fit shown in Fig. 5(b). The r_a value is 0.977 Å; the probable C—D2 bond length with the fourth-order parabolic model is 1.081 Å, while the eighth-order GC value is 1.068 Å. The parabolic distribution again is closer to the equilibrium value, which in this case is 1.089 Å.

Both anharmonic treatments capture the majority of the thermal-motion correction.

Given the problems with negative regions that result from using a polynomial expansion, we have also used the SN distribution, in the form of equation (25), to model the D2 atom. The R factor for this model is comparable to that for the fourth-order GC series. As the error function is contained within the GC series expansion it is still possible for the distribution to have negative regions but the limited range of the error function reduces its contributions to negative regions in the PDF. For systems with large amplitudes the error function may consequently be more useful than the H_{iii} functions.

7.1.3. O1 atom. At higher temperatures the simulated PDF of the O1 atom of nitromethane is egg or pear shaped. This should represent a good test case for the SN distribution. Three different models, harmonic, GC and SN, have been fitted to the numerical distribution at 228 K; the resulting R factors are presented in Table 2. The SN distribution is only slightly better than the harmonic model, whereas the GC fit is appreciably better. The SN model also gives a distribution that does not match the visual appearance of the numerical distribution very well. However, a GC series PDF with only the three H_{iii} terms produces an R factor identical to the SN fit. This suggests that the SN distribution is as effective as the GC series at skewing the distribution but that in some cases the off-diagonal GC polynomials are important as well. The biggest contributions from off-diagonal elements are from the H_{122}, H_{133} and H_{233} polynomials, showing that some 'bending' of the distribution occurs.

7.1.4. Positive-definite GC series. The error function may represent a useful way to define a PDF based on polynomial perturbations that cannot have negative regions. We may define a PDF of the form

$$P(\mathbf{u})_{\rm SN} = P(\mathbf{u})_{\rm harm}\{1 + {\rm erf}[f(\mathbf{u})]\}, \qquad (26)$$

where $f(\mathbf{u})$ is an appropriate polynomial function. If $f(-\mathbf{u}) = -f(\mathbf{u})$ then equation (26) represents a true PDF (Azzalini, 2005) and therefore must have a characteristic function. The third-order Hermite polynomials satisfy this condition when centred on the mean position. At 15 K the D1 atom of nitromethane required only three such polynomials to fit. If we let $f(\mathbf{u}) = c^{111}H_{111} + c^{112}H_{112} + c^{113}H_{113}$ then we obtain an R factor of 14.4%, which is slightly better than the 14.6% obtained by using them directly (*i.e.* without using the error function). The anharmonic parts of the resulting PDF obtained using the two types of perturbation are plotted in Fig. 6. This shows similar contour lines for both but with different scales.

Further investigation is obviously required but a PDF of the form of equation (26) may prove useful in crystallography, particularly at high temperatures, where the negative probability density inherent to the GC series can be an issue.

7.2. 1:2 Adduct of dimethylurea and oxalic acid

The methyl groups of DMUOX provide a useful data set to validate the results obtained by applying the parabolic PDFs [equations (22)–(24)] derived to model the PDFs of the D atoms of nitromethane. The PDF of a methyl-group hydrogen was fitted using the limited fourth-order GC series and equation (24) in the same way as for nitromethane. As with nitromethane the fourth-order parameters were essential to model the asymmetry of the distribution realistically. The numerical and parabolic distributions are shown in Fig. 7. The parabolic fit appears to be quite good but the underlying numerical distribution is particularly noisy because of the smaller data set collected in the DFT–MD simulation. This hinders detailed comparison of the parabolic and GC fits but it is clear that the parabolic model is capable of reproducing the general behaviour of the methyl group.

7.3. Urea–phosphoric acid

For the O atom of nitromethane the SN distribution was clearly inferior to the GC series in modelling the distribution's asymmetry. The UPA migratory proton exhibits similar asymmetry. At 150 K an agreement index of 27.5% is obtained

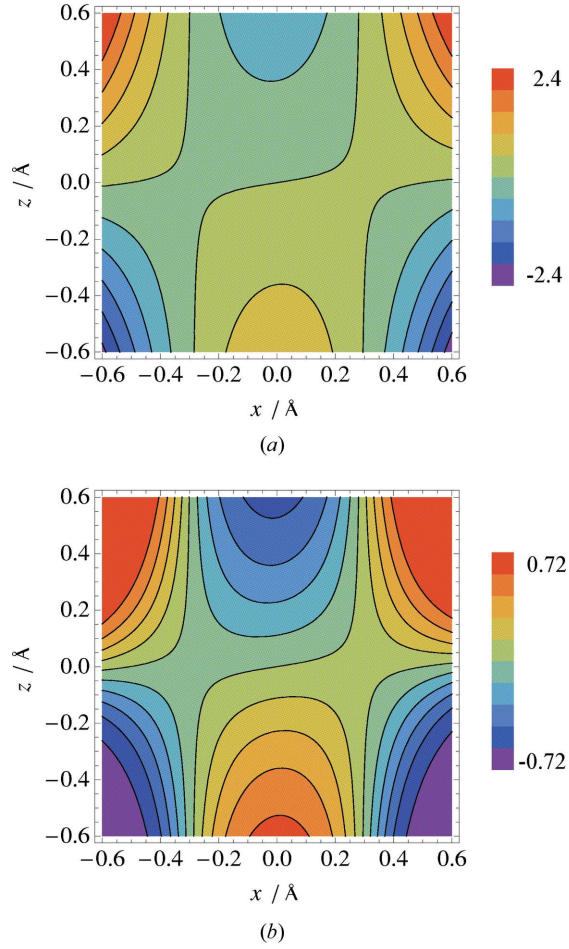

Figure 6
Two-dimensional xz difference maps of the fit to the D1 atom PDF of nitromethane at 15 K: (*a*) $c^{111}H_{111} + c^{112}H_{112} + c^{113}H_{113}$ and (*b*) ${\rm erf}(c^{111}H_{111} + c^{112}H_{112} + c^{113}H_{113})$. (Note the difference in scale.)

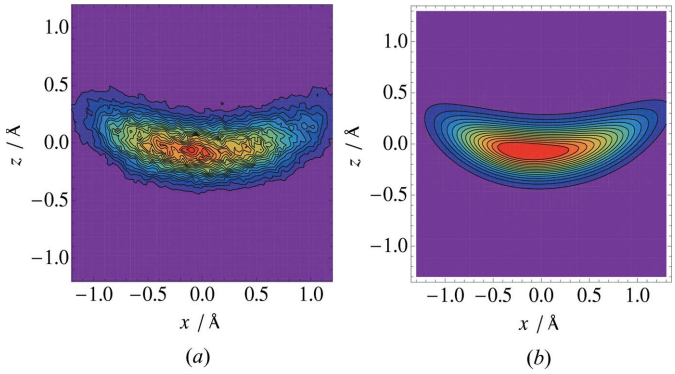

Figure 7
Two-dimensional xz PDFs of one of the H atoms of DMUOX: (*a*) Fourier-filtered numerical PDF, (*b*) eighth-order skewed parabolic PDF. (The x axis represents the longest principal axis of thermal motion, while z is the shortest axis.)

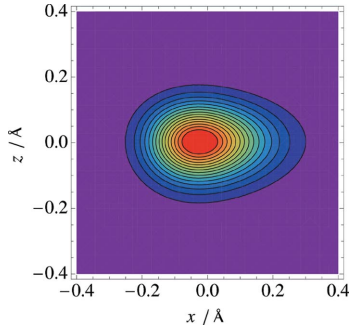

Figure 8
Two-dimensional xz PDF of the migratory proton of UPA at 150 K modelled with an SN distribution. See Fig. 4(*a*) of the preceding paper (Reilly *et al.*, 2011) for the numerical distribution.

by using the SN distribution. The large value is a result of the high-frequency noise present because the DFT–MD simulation collects far fewer data than the force-field MD used for nitromethane. A GC series with only the three H_{iii} polynomials gives essentially the same agreement factor. The full third-order GC series model (with ten polynomials) does give a slightly better fit of 27.1% but all three refinements yield similar values for the probable position. The SN distribution shown in Fig. 8 successfully reproduces the egg shape of the xz distribution.

Comparing the SN distribution's performance for the nitromethane O1 atom and the UPA H atom it is likely that the SN distribution will be as good at skewing a distribution as the third-order H_{iii} parameters. Whether the bending Hermite polynomials are important will probably depend on the profile of the distribution and the correlation between motion in different directions. Asymmetry that represents the complete skewing of the PDF from its centre may be better approximated using skew functions only, whereas distributions that are more harmonic near the maximum may require the flexibility of the full GC series.

8. Conclusion

The MD-derived numerical PDFs of a variety of atoms in different molecules have been modelled using a number of different analytical functions. In doing so the numerical data sets have shown their use in providing realistic models for developing new forms for the Debye–Waller factor. This offers the possibility not only to determine more physically meaningful structures but also to get more information on atomic displacements from diffraction data.

Two new forms of PDF have been fitted to the numerical PDFs. The first, based on a parabolic coordinate system, models the curvilinear motion of the hydrogen atoms in nitromethane and DMUOX as well as the GC series but with fewer parameters. In nitromethane, the parabolic distribution gave most probable C—D bond lengths that were closer to the equilibrium value than the GC model values. Irrespective of whether the third-order GC series or parabolic PDF was used to model the curvature, it was essential to include fourth-order GC terms (primarily the H_{iiii} terms) to model the asymmetry of the PDF. To cut down on the number of GC parameters it has been suggested that the Hermite polynomials employed be defined in the harmonic thermal motion coordinate system.

A skew-normal PDF based on the univariate form introduced by Azzalini (1985) has also been used to model anharmonic distributions. In this limited study it has been found to be as effective at skewing a PDF as the H_{iii} terms of the GC series but in some cases the flexibility of the full GC series permits one to model the data better. However, the SN distribution may provide a useful basis for a 'true' PDF that has no spurious negative regions – a potentially significant advantage over existing methods at high temperatures.

The EaStCHEM Research Computing Facility (http://www.eastchem.ac.uk/rcf) is acknowledged for its provision of computational resources. This work has also made use of the resources provided by the Edinburgh Compute and Data Facility (http://www.ecdf.ed.ac.uk/). Both of these facilities are partially supported by the eDIKT initiative (http://www.edikt.org.uk). AMR acknowledges the School of Chemistry, Edinburgh, for funding a studentship, and CAM the Royal Society for the award of a University Research Fellowship.

References

Azzalini, A. (1985). *Scand. J. Statist.* **12**, 171–178.
Azzalini, A. (2005). *Scand. J. Statist.* **32**, 159–188.
Castellano, E. E. & Main, P. (1985). *Acta Cryst.* A**41**, 156–157.
Cruickshank, D. W. J. (1956). *Acta Cryst.* **9**, 757–758.
Giacovazzo, C., Monaco, H. L., Artioli, G., Viterbo, D., Ferraris, G., Gilli, G., Zanotti, G. & Catti, M. (2002). *Fundamentals of Crystallography*. IUCr Texts on Crystallography, 2nd ed. Oxford University Press.
Goedecker, S., Teter, M. & Hutter, J. (1996). *Phys. Rev. B*, **54**, 1703–1710.
Gupta, A. K. & Chen, J. T. (2004). *Ann. Inst. Statist. Math.* **56**, 305–315.
Hartwigsen, C., Goedecker, S. & Hutter, J. (1998). *Phys. Rev. B*, **58**, 3641–3662.
Hohlwein, D. (1981). *Acta Cryst.* A**37**, 899–903.

Hoover, W. G. (1985). *Phys. Rev. A*, **31**, 1695–1697.
Johnson, C. K. (1969). *Acta Cryst.* A**25**, 187–194.
Johnson, C. K. (1970). *Thermal Neutron Diffraction*, p. 132. Oxford University Press.
Krack, M. (2005). *Theor. Chem. Acc.* **114**, 145–152.
Kronenburg, M. J. (2004). *Acta Cryst.* A**60**, 250–256.
Kuhs, W. F. (1983). *Acta Cryst.* A**39**, 148–158.
Kuhs, W. F. (1992). *Acta Cryst.* A**48**, 80–98.
Kurki-Suonio, K., Merisalo, M. & Peltonen, H. (1979). *Phys. Scr.* **19**, 57–63.
Kuznetsov, P. I., Stratonovich, R. L. & Tikhonov, V. I. (1960). *Theory Probab. Appl.* **5**, 80–97.
Lippert, G., Hutter, J. & Parrinello, M. (1997). *Mol. Phys.* **92**, 477–488.
Nosé, S. (1984). *J. Chem. Phys.* **81**, 511–519.
Pawley, G. S. & Willis, B. T. M. (1970). *Acta Cryst.* A**26**, 260–262.
Pulham, C. R. (2009). Personal communication.
Reilly, A. M., Habershon, S., Morrison, C. A. & Rankin, D. W. H. (2010). *J. Chem. Phys.* **132**, 094502.
Reilly, A. M., Morrison, C. A. & Rankin, D. W. H. (2011). *Acta Cryst.* A**67**, 336–345.
Reilly, A. M., Wann, D. A., Morrison, C. A. & Rankin, D. W. H. (2007). *Chem. Phys. Lett.* **448**, 61–64.
Rowicka, M., Kudlicki, A., Zelinka, J. & Otwinowski, Z. (2004). *Acta Cryst.* A**60**, 542–549.
Scheringer, C. (1988). *Acta Cryst.* A**44**, 343–349.
Tanaka, K. & Marumo, F. (1983). *Acta Cryst.* A**39**, 631–641.
Trueblood, K. N., Bürgi, H.-B., Burzlaff, H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., Shmueli, U. & Abrahams, S. C. (1996). *Acta Cryst.* A**52**, 770–781.
VandeVondele, J., Krack, M., Mohamed, F., Parrinello, M., Chassaing, T. & Hutter, J. (2005). *Comput. Phys. Commun.* **167**, 103–128.
Wolfram Research Inc. (2007). *Mathematica*. Version 6.0. Wolfram Research Inc., Champaign, Illinois.
Willis, B. T. M. (1969). *Acta Cryst.* A**25**, 277–300.
Willis, B. T. M. & Pawley, G. S. (1970). *Acta Cryst.* A**26**, 254–259.
Zucker, U. H. & Schulz, H. (1982). *Acta Cryst.* A**38**, 563–568.