Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 18 March 2004
Accepted 12 August 2004

Molecular crystal global phase diagrams. I. Method of construction

Jonathan A. Mettes, J. Brandon Keith and Richard B. McClurg*

Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA. Correspondence e-mail: mcclurg@cems.umn.edu

A method is described to produce global phase diagrams for single-component molecular crystals with separable internal and external modes. The phase diagrams present the equilibrium crystalline phase as a function of the coefficients of a general intermolecular potential based on rotational symmetry-adapted basis functions. It is assumed that phase transitions are driven by orientational ordering of molecules with a fixed time-averaged shape. The mean-field approximation is utilized and the process begins in a high-temperature disordered reference state, then spontaneous symmetry-breaking phase transitions and phase structure information at lower temperature are sought. The information is mapped onto phase diagrams using the intermolecular expansion coefficients as independent variables. This is illustrated by global phase diagrams for molecules having tetrahedral symmetry (*e.g.* carbon tetrachloride, adamantane and white phosphorus). Uses of global phase diagrams include crystal structure data mining, guidance for crystal design and enumeration of likely or missing polymorphic structures.

1. Introduction

Molecular crystals are lattices of discrete molecular groups held together in the solid state by non-covalent interactions such as van der Waals forces or hydrogen bonding. They have been the subject of interest due to the potential for controlling solid-state properties with appropriate variations to the molecular components. Indeed, many interesting properties once thought to be unique to inorganics have been discovered in molecular crystals, including electrical conductivity (Wudl, 1984), nonlinear optical behavior (Chemla & Zyss, 1987) and ferromagnetism (Miller *et al.*, 1988). Since the structure of the constituent molecules can be modified *via* traditional organic synthetic chemistry, organic molecular crystals seem to hold an advantage for crystal design. Unlike covalent bonds within molecules, however, intermolecular interactions are less specific with regard to relative orientation and are much weaker. These weak non-specific bonds make crystal structure design and control difficult, though there have been a number of advances both in crystal engineering (Desiraju, 1989; Braga *et al.*, 1999; Holman *et al.*, 2001; Moulton & Zaworotko, 2001) and in crystal structure prediction (Gavezzotti, 1994; Beyer *et al.*, 2001; Leusen, 2003). Ultimately, crystal structure prediction remains difficult because physical and chemical properties of crystalline solids are dependent not only on the properties of the constituent molecules but also on the spatial arrangement of the molecules within the crystal lattice.

The rigorous study of the relationship between molecular structure and crystal structure has been in development since at least the early 1960s, when Kitaigorodskii (1961) and Williams, (1966) reported crystal structure predictions by rigid-body molecular-packing analysis. X-ray diffraction data were used to verify and refine the initial proposed packing. It was not until the last decade, with recent progress of computing power, that energy-minimization approaches became possible (Gdanitz, 1992; Holden *et al.*, 1993; Karfunkel *et al.*, 1993; Williams, 1996). In these approaches, the energy of the crystal is represented by a force field, as implemented in molecular mechanics (Dinur & Hagler, 1991). To describe fully the intermolecular interactions, the potential is represented as a sum of terms for electrostatic and van der Waals interactions, and for bond bending and stretching (Coombes *et al.*, 1996; Price, 1996; Gavezzotti, 2002). Once a potential function is derived or assumed, a search methodology is used to locate the global minimum and obtain the energetically favored structure. There are several computational programs that have been developed for the prediction of stable crystal structures (Gavezzotti, 1991; Gdanitz, 1992; Holden *et al.*, 1993; Karfunkel *et al.*, 1993; van Eijck *et al.*, 1995; Tajima *et al.*, 1995; Chaka *et al.*, 1996; Schmidt & Englert, 1996; Williams, 1996; Pillardy *et al.*, 2001), reviews of which have been given in the literature (Gdanitz, 1997, 1998; Verwer & Leusen, 1998).

Recently, the Cambridge Crystallographic Data Centre conducted two blind tests of some of the currently available crystal-structure prediction methods. Participants were given only the chemical structures for six organic compounds with unpublished but accurately known crystal structures (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002). Each participating group was allowed to submit a list of three crystal

structures for each compound, and the submission was considered correct if the observed structure was on the list. The study found that, with 15 groups participating, the overall success rate for predicting the crystal structure was approximately 17%. No method had a significantly higher success rate.

A post-analysis of the results from the blind tests revealed that the programs placed the experimental structure among a long list of potential polymorphs, but often it was removed by a few kJ mol^{-1} from the global minimum given by their model potential. Thus, the observed structure was located further down the energetic ranking and not selected, despite being only slightly higher in energy. Another study involving acetic acid resulted in approximately 100 different structural candidates, all within only 5 kJ mol^{-1} of each other (Mooij *et al.*, 1998). Apparently, the success of a given method depends less on the search method than accurately and precisely ranking many possible structures that differ little in energy (Chaka *et al.*, 1996).

The major difficulty in crystal-structure prediction lies in the energy surface itself, which is often quite flat with many local minima of similar energy. Small details in the parametrization of the intermolecular potential (IP), such as position of interaction sites, heavily influence the energy ranking on this surface. Given results computed for a particular IP, small perturbations to the parameters can reshuffle the energetic order of the local minima. Owing to this extreme parameter sensitivity, we have chosen to consider the inverse problem. We are developing methods that use known crystal structures to place strong limits on interaction potentials consistent with the observed structures. The set of such limits can be represented on a *global phase diagram*, which summarizes the crystal phase behavior as a function of intermolecular potential parameters. This type of tool is well established for classifying and rationalizing liquid–vapor phase behavior in terms of molecular parameters (van Konynenburg & Scott, 1980). By constructing global phase diagrams for solid phases, we seek to highlight the trends and relationships among molecular crystals.

To generate global phase diagrams, we require an intermolecular potential with a minimum number of adjustable parameters, but still general enough to be broadly applicable. This motivates the selection of a potential for systems dominated by rotational coupling as detailed in §2.2 and discussed in §5. The potential that we have used is appropriate for fixed-shape molecules interacting through van der Waals and/or hydrogen-bonding interactions. Since the potential is based on nearest-neighbor pairwise additive interactions, it is not appropriate for ionic, covalent or metallic systems. These restrictions are discussed further in §2.2 and §5. For molecules with non-trivial point-group symmetry, it is convenient to reduce the dimensionality of the phase diagrams through adaptation of the potential to molecular symmetry. This basic approach of expanding the angle-dependent intermolecular interactions into symmetry-adapted rotation functions was originally developed for phase transitions in heavy methane by James & Keenan (1959) and has since been applied to crystals such as solid C_{60} (Michel *et al.*, 1992; Lamoen & Michel, 1999). Rather than investigating structures for a specific molecule with a given IP, however, we construct global phase diagrams for all molecules with a selected point-group symmetry.

Global phase diagrams have several uses. Firstly, there is a wealth of structural data available today, such as in the Cambridge Structural Database (Allen & Motherwell, 2002), so we gain a data mining tool for insight into the intermolecular forces that form crystals. Secondly, by locating neighboring structures that may be more desirable than a given crystal, global phase diagrams provide feedback for crystal engineering and molecular synthesis to achieve a desired crystal. Finally, global phase diagrams can provide rationalization and/or prediction of crystal polymorphs for a given molecule. These applications are discussed in greater detail in §5.

The goal of this paper is to describe methods for producing global phase diagrams like those illustrated in Fig. 1, which has been constructed for tetrahedral molecules. Global phase diagrams give the thermodynamically stable phase as a function of temperature and parameters $\boldsymbol{v}$ in the intermolecular potential, as defined in §2.3. The remainder of the paper is organized as follows. §2 details the development of the statistical mechanical and group-theoretical framework used to describe the orientational state of the crystal. There we indicate the assumptions made and describe the intermolecular potential used. A brief synopsis of the model has been given elsewhere (Keith *et al.*, 2004). In §3, we present the elements of the solution method, including specification of the reference state and the computational approach used to locate and identify phase-transition surfaces in IP parameter space. §4 provides an example of our method applied to single-component crystals composed of molecules with tetrahedral symmetry. §5 summarizes the uses of these diagrams and contains some discussion of their limitations. Finally, the Appendices present a few of the more technical details of the calculations.[1]

2. Model development

§2.1 introduces our approach to the free energy. §2.2 discusses the separation of the partition function into rotational, translational and intramolecular components. Based on close-packing and mode separation arguments, we suggest molecular orientation dominantly influences phase-diagram construction. A general orientational potential is presented in §2.3 and the mean-field approximation and crystallographic group theory invoked to calculate the free energy in §2.4 and assign the phase in §2.5.

[1] Appendices *A*–*F* have been deposited as supplementary material. These data are available from the IUCr electronic archives (Reference: PZ5009). Services for accessing these data are described at the back of the journal.

2.1. Free energy

Beginning with a crystal of N identical molecules each with M atoms, the classical canonical partition function Z is given by

$$Z = \frac{1}{h^{3NM}} \iint \exp(-\mathcal{H}/kT)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}, \qquad (1)$$

where h is Planck's constant, k is Boltzmann's constant, T is the absolute temperature, $\mathcal{H}$ is the crystal Hamiltonian, and the vectors $\mathbf{x}$ and $\mathbf{p}$ are atomic positions and linear momenta, respectively. Use of this partition function requires molecules with large enough mass and moments of inertia such that the degrees of freedom can be treated classically at the temperature of interest. It should not be used for quantum rotors such as H_2 and NH_3 at low temperature. Thermodynamic properties of the crystal can be calculated based on the relationship between the partition function and the Helmholtz free energy A,

$$A = -kT \ln Z. \qquad (2)$$

Since $A(T, V)$ is a fundamental equation of state, the other thermodynamic state variables (*e.g.* internal energy, entropy *etc.*) may be expressed as functions of Z and its derivatives (McQuarrie, 2000). Although equation (1) is exact, it is impractical to use since it is a $6NM$-dimensional integral where N is a macroscopic quantity. To simplify calculations, we make several assumptions that are described and justified in the following paragraphs.

2.2. Separation of modes

The intramolecular vibrations and distortions (internal modes) of molecules generally occur on a time scale that is much faster than the corresponding rotations and translations of the molecule as a whole (external modes). The characteristic frequencies for internal vibrations are often greater than 1200 cm^{-1} (36 THz), compared to less than 100 cm^{-1} (3 THz) for rotations (Sherwood, 1979). This separation of time scales causes the vibrational coordinates to be largely independent of the external coordinates. Therefore, to an excellent approximation, the partition function is separable into products of internal and external mode contributions,

$$Z = (z_{\mathrm{int}})^N Z_{\mathrm{ext}}, \qquad (3)$$

where

$$z_{\mathrm{int}} = \frac{1}{h^{3M-d-3-t}} \iint \exp(-\mathcal{H}_{\mathrm{int}}/kT)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p} \qquad (4)$$

$$Z_{\mathrm{ext}} = \frac{1}{h^{N(d+3+t)}} \iint \exp(-\mathcal{H}_{\mathrm{ext}}/kT)\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{p}. \qquad (5)$$

The molecular dimensionality d has a value of 2 for linear molecules and 3 for nonlinear molecules, while t is the number of shape-changing torsions, discussed below. The internal degrees of freedom in a typical molecule are contained in z_{int}, while Z_{ext} contains the external degrees of freedom, including the center-of-mass translations, rigid rotations and their conjugate momenta.

When considering the separation of internal and external modes, molecular torsions can fall into either category. Some torsions have large moments of inertia and dramatically change molecular shape, such as the C_2-C_3 torsions in butane. These often have frequencies comparable to those of rigid rotations, and should be contained in the external partition function. The number of such torsions is denoted by t in equations (4) and (5). However, for other torsions the

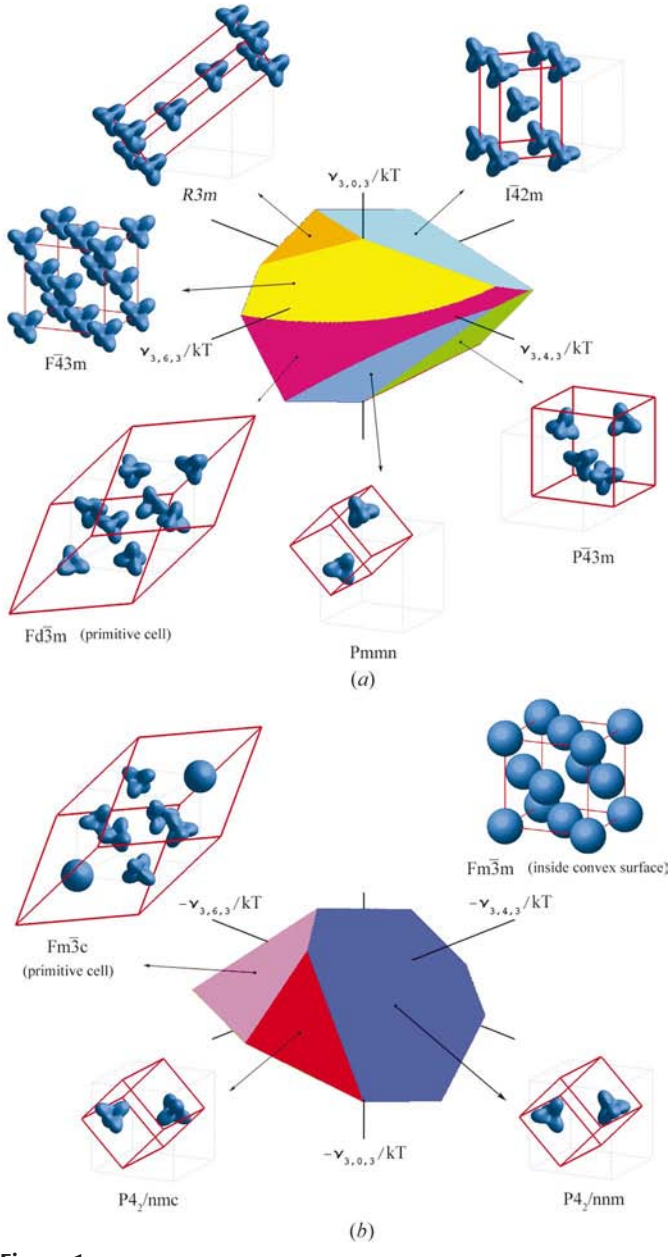

Figure 1
(*a*) [111] view of T_d global phase diagram as a function of intermolecular potential coefficients $\boldsymbol{\nu}$ scaled by kT. The origin of the axes is the infinite-temperature reference state, discussed in §3. Rays protruding from the surfaces pointing at distinct phases represent ratios of potential coefficients that form tetrahedral molecular crystals of indicated symmetry. Colored surfaces are boundaries between daughter phases and the reference f.c.c. phase. (*b*) [$\bar{1}\bar{1}\bar{1}$] view of the T_d global phase diagram. Some molecules such as those in the $Fm\bar{3}m$ structure appear 'spherical' because of rapid molecular reorientation.

moment of inertia is relatively small and the torsions have little effect on molecular symmetry, such as the terminal CH_3 groups in butane. These appropriately belong in $z_{\rm int}$ since they are not expected to influence strongly or be influenced by crystal structure. This work does not consider molecules having slow molecular shape-changing torsions, and therefore does not include torsional degrees of freedom in the external modes ($t = 0$). This assumption is discussed further in §5.

Since $z_{\rm int}$ is independent of the relative positions and orientations of the molecules, it is a function only of T and does not influence phase equilibrium. Therefore, we focus on the external partition function. The interaction between center-of-mass displacements and large-amplitude orientational motion, termed translation–rotation coupling, is most pronounced in ionic compounds (*e.g.* NH_4Br, $NaNO_2$), while less important for phase transitions of neutral molecular crystals (Lynden-Bell & Michel, 1994). The exceptions to this are for slightly prolate molecules such as C_{70} (Lamoen & Michel, 1999) and possibly N_2 (Löding *et al.*, 1997). Since this coupling is weak in the majority of the organic molecular crystals that are the focus of this work, we separate $Z_{\rm ext}$ into translational and orientational contributions

$$Z_{\rm ext} = Z_{\rm tr} Z_{\rm or}, \tag{6}$$

where

$$Z_{\rm tr} = \frac{1}{h^{3N}} \iint \exp(-\mathcal{H}_{\rm tr}/kT)\, {\rm d}\mathbf{x}\, {\rm d}\mathbf{p}_x \tag{7}$$

$$Z_{\rm or} = \frac{1}{h^{Nd}} \iint \exp(-\mathcal{H}_{\rm or}/kT)\, {\rm d}\boldsymbol{\omega}\, {\rm d}\mathbf{p}_\omega. \tag{8}$$

The vector $\boldsymbol{\omega}$ contains the Euler angles (Varshalovich *et al.*, 1988) describing the orientation of each molecule and $\mathbf{p}_\omega$ is the vector of conjugate angular momenta.

Combining equations (2), (3) and (6) yields

$$A = -kT(\ln z_{\rm int}^N + \ln Z_{\rm tr} + \ln Z_{\rm or}) \tag{9}$$

for the crystal free energy, with the internal, translational and orientational contributions separated. When considering solid/solid phase transitions, we are interested in the free-energy difference ΔA between two crystal structures at the same temperature. Since we have assumed that the internal modes are largely decoupled from relative motions of molecules, the changes in internal modes between structures are insignificant and there is a minor impact on phase behavior from this contribution. Likewise, the molecular translations accompanying most solid/solid phase transitions are generally a small fraction of the lattice parameters. We hypothesize that the small translations accompanying solid/solid phase equilibria are of secondary importance and it is orientational ordering that drives solid/solid phase change in molecular crystals. Therefore, we choose to focus on the orientational Helmholtz free energy $A_{\rm or} = -kT \ln Z_{\rm or}$ as the main arbiter of crystal structure. The consequences of neglecting translational relaxation of the molecules are revisited in §5.

2.3. Orientational crystal Hamiltonian

To calculate the partition function $Z_{\rm or}$ using equation (8), we must specify an orientational Hamiltonian $\mathcal{H}_{\rm or} = \mathcal{K} + V$. The kinetic energy contribution is that of a rotor with principal moments of inertia $\{I_a, I_b, I_c\}$ (McQuarrie, 2000). Integrating over angular momenta yields

$$Z_{\rm or} = z_{\rm rot}^N \int \exp(-V/kT)\, {\rm d}(\boldsymbol{\omega}), \tag{10}$$

where

$$z_{\rm rot} = 8\pi^2 \left(\frac{2\pi (I_a I_b I_c)^{1/3} kT}{h^2} \right)^{3/2} \tag{11}$$

$$ {\rm d}(\boldsymbol{\omega}) = \prod_i {\rm d}(\boldsymbol{\omega}_i) = \prod_i \frac{\sin\beta_i\, {\rm d}\boldsymbol{\omega}_i}{8\pi^2} \tag{12}$$

are the partition function for the rotor and measure on the rotation group $SO(3)$, respectively (Vilenkin & Klimyk, 1991; Chirikjian & Kyatkin, 2000). Equations (11) and (12) are for nonlinear molecules ($d = 3$). $z_{\rm rot}$ and the integral measure are different for linear molecules ($d = 2$). Since $z_{\rm rot}$ is independent of the relative position and orientation of molecules, it does not determine crystal structure.

Crystal structure is determined by V and T through the integration over orientations. For van der Waals solids in which electrostatic, exchange and dispersion forces dominate, the interaction potential V in $\mathcal{H}_{\rm or}$ can be written as a sum over two-molecule terms (Stone, 1996). Crystals of strongly dipolar molecules such as ice or ionic crystals in which induction forces produce significant three-body terms (Briels *et al.*, 1986) are not considered here.

We choose to expand the potential using a complete set of two-molecule basis functions that span the rotational space $SO(3)$ parametrized by Euler angles $\boldsymbol{\omega}$ of molecules i, j and their intervening orientational space S^2 parametrized by solid angle $\boldsymbol{\Omega}_{ij} = (\theta_{ij}, \phi_{ij})$. Coupling two $SO(3)$ irreducible representations (IR's) $D^{\ell_i}_{m_i n_i}(\boldsymbol{\omega}_i)$ and $D^{\ell_j}_{m_j n_j}(\boldsymbol{\omega}_j)$ and a spherical harmonic $C^{\ell}_{m}(\boldsymbol{\Omega}_{ij})$ gives

$$W^{n_i n_j}_{\ell_i \ell \ell_j} = \sum_{m_i m m_j} \begin{pmatrix} \ell_i & \ell & \ell_j \\ m_i & m & m_j \end{pmatrix} D^{\ell_i}_{m_i n_i} C^{\ell}_{m} D^{\ell_j}_{m_j n_j}, \tag{13}$$

where the explicit functional dependence of C^{ℓ}_{m} and $D^{\ell_j}_{m_j n_j}$ has been dropped. Definitions of basis functions and a derivation of the coupling is provided in Appendix *A*. The functions D^{ℓ}_{mn} are called Wigner functions. Wigner functions and spherical harmonics are orthogonal and complete, making their irreducible decomposition $W^{n_i n_j}_{\ell_i \ell \ell_j}$ a complete set of orthogonal IP basis functions over the eight-dimensional space $SO(3) \times S^2 \times SO(3)$. Their properties have been studied previously (van der Avoird *et al.*, 1980; Briels, 1980; Stone & Tough, 1984; van der Avoird *et al.*, 1994).

While the basis functions $W^{n_i n_j}_{\ell_i \ell \ell_j}$ are general, it is of great advantage computationally to project out the point-group symmetry of the molecule and that of the Wyckoff point. Appendix *B* reviews our use of projection operators which amount to matrix multiplication by a sparse unitary matrix $S^{\ell_i}_{n_i n_\sigma}$, where σ refers to a point-group IR and n_σ is a particular

component of the IR. For example, using projection operators for the molecular point group yields symmetry-adapted matrix elements

$$T^{\ell_i}_{m_i n_\sigma}(\boldsymbol{\omega}_i) = \sum_{n_i} D^{\ell_i}_{m_i n_i} S^{\ell_i}_{n_i n_\sigma}, \tag{14}$$

which may be coupled to produce symmetry-adapted basis functions

$$F^{n_\sigma n_\mu}_{\ell_i \ell \ell_j} = \sum_{m_i m m_j} \begin{pmatrix} \ell_i & \ell & \ell_j \\ m_i & m & m_j \end{pmatrix} T^{\ell_i}_{m_i n_\sigma} C^{\ell}_{m} T^{\ell_j}_{m_j n_\mu}. \tag{15}$$

With these basis functions, the potential is

$$V = \tfrac{1}{2} \sum_{ij} \sum_{\ell_i \ell \ell_j n_\sigma n_\mu} v^{n_\sigma n_\mu}_{\ell_i \ell \ell_j}(r_{ij}) F^{n_\sigma n_\mu}_{\ell_i \ell \ell_j}(\boldsymbol{\omega}_i, \boldsymbol{\Omega}_{ij}, \boldsymbol{\omega}_j), \tag{16}$$

where the half avoids overcounting, $\ell_i, \ell_j \in \mathbb{N}$, $|\ell_i - \ell_j| \leq \ell \leq \ell_i + \ell_j$ and $v^{n_\sigma, n_\mu}_{\ell_i, \ell, \ell_j}(r_{ij})$ are coefficients that are functions of the distance r_{ij} between molecular centers. The full set of these in the potential is termed $\boldsymbol{v}$. Subscripts σ, μ are a compound index referring to multiple copies of the molecular point-group unit IR subduced in the ℓ_i, ℓ_jth manifold of $SO(3)$ and n_σ, n_μ is its dimension, which is always $n_\sigma, n_\mu = 1$. Point-group IR subduction frequencies in spherical harmonics are discussed elsewhere (Bradley & Cracknell, 1972). All other point-group IR's do not have the full molecular symmetry and so are zero to first order. Thus electronic structure is averaged over short time vibrations and excitations so the molecule is invariant under the full molecular point group (Califano *et al.*, 1981). These coefficients can be calculated efficiently from *ab initio* or empirical potentials using a fast Fourier transform for $SO(3)$ (Maslen & Rockmore, 1997; Chirikjian & Kyatkin, 2000).

Projecting out Wyckoff point symmetry from the functions $T^{\ell_i}_{m_i n_\sigma}$ gives rotator functions

$$U^{\ell_i}_{m_\tau n_\sigma}(\boldsymbol{\omega}_i) = \sum_{m_i n_i} S^{\ell_i *}_{m_i m_\tau} D^{\ell_i}_{m_i n_i} S^{\ell_i}_{n_i n_\sigma}. \tag{17}$$

Expressing the potential using coupled rotator functions gives

$$V/kT = \tfrac{1}{2} \sum_{ij} \sum_{\ell_i \ell_j m_\tau m_\rho n_\sigma n_\mu} U^{\ell_i}_{m_\tau n_\sigma} K^{\ell_i \ell_j}_{m_\tau n_\sigma m_\rho n_\mu} U^{\ell_j}_{m_\rho n_\mu}, \tag{18}$$

where

$$K^{\ell_i \ell_j}_{m_\tau n_\sigma m_\rho n_\mu}(\boldsymbol{\Omega}_{ij}) = \sum_{\ell m_i m m_j} \frac{v^{n_\sigma, n_\mu}_{\ell_i, \ell, \ell_j}(r_{ij})}{kT} \begin{pmatrix} \ell_i & \ell & \ell_j \\ m_i & m & m_j \end{pmatrix} S^{\ell_i}_{m_\tau m_i} C^{\ell}_{m} S^{\ell_j}_{m_\rho m_j} \tag{19}$$

is a dimensionless coupling function. Subscripts τ, ρ are a compound index referring to multiple copies of the Wyckoff point-group IR's subduced in the ℓ_i, ℓ_jth manifold of $SO(3)$ and m_τ, m_ρ goes over the dimensions of each IR. It is sometimes convenient to write equation (18) using a shorthand notation

$$V/kT = \tfrac{1}{2} \sum_{ij} \mathbf{U}_i \cdot \mathbf{K}_{ij} \cdot \mathbf{U}_j, \tag{20}$$

in which $\mathbf{U}_i$ is a vector covering all elements of $U^{\ell_i}_{m_\tau n_\sigma}$ and $\mathbf{K}_{ij}$ is a coupling matrix. The generalized dot product matches multiple repeated indices in the potential.

By symmetry adapting in two steps, we are able to define the expansion coefficients in equation (16) and expand the potential using rotator functions in equation (18). Thus the expansion coefficients $v^{n_\sigma, n_\mu}_{\ell_i, \ell, \ell_j}$ are functions of molecular parameters and intermolecular distance, but are independent of crystal structure. They serve as independent variables in our method. The rotator functions $\mathbf{U}_i$ are functions of both molecular and crystal symmetries. These functions lead to order parameters for phase transitions as discussed in §2.4.

2.4. Space-group order parameters

Using the potential outlined in §2.3, the partition function in equation (10) must be calculated efficiently to construct the diagrams. The method we have chosen is a variational mean-field theory in which molecules interact with the time-averaged field of their neighbors (Chandler, 1987). Equation (20) thus becomes

$$V/kT \approx V^{\rm mf}/kT = \sum_i V^{\rm mf}_i/kT = \sum_i \mathbf{U}_i \cdot \mathbf{h}_i, \tag{21}$$

where the vector $\mathbf{h}_i$ is the mean field coupling to molecule i, and $V^{\rm mf}_i$ is the corresponding mean-field IP. This formulation is analogous to standard mean-field theory for systems of continuous spins in the field of their neighbors. This case is somewhat different because the rotator functions $\mathbf{U}_i$ have multiple components $U^{\ell_i}_{m_\tau n_\sigma}$ for each molecule i, and there is a corresponding field for each. Therefore, the dot product in (21) is to be interpreted as a summation over multiple indices as discussed in §2.3.

Following the treatment in Appendix *C*, this leads to the potential

$$V^{\rm mf}/kT = \sum_{ij} \mathbf{U}_i \cdot \mathbf{K}_{ij} \cdot \langle \mathbf{U}_j \rangle. \tag{22}$$

The absence of a factor of 1/2 compared to equation (20) is a result of the variational principle. The thermodynamic averages $\langle \mathbf{U}_i \rangle$ for the rotator functions are

$$\langle \mathbf{U}_i \rangle = \frac{1}{z_i} \int \mathbf{U}_i \exp(-V^{\rm mf}_i/kT)\, {\rm d}(\boldsymbol{\omega}_i), \tag{23}$$

where

$$z_i = \int \exp(-V^{\rm mf}_i/kT)\, {\rm d}(\boldsymbol{\omega}_i). \tag{24}$$

$\langle \mathbf{U}_i \rangle$ appears on both sides of the equation and must be solved self-consistently. Also from Appendix *C*, the variational Helmholtz free energy is

$$A^{\rm var}_{\rm or} = -kT \ln\left(z^N_{\rm rot} \prod_i z_i \right) - \frac{kT}{2} \sum_{ij} \langle \mathbf{U}_i \rangle \cdot \mathbf{K}_{ij} \cdot \langle \mathbf{U}_j \rangle. \tag{25}$$

The second term in (25), which is a general feature of the variational free energy, compensates for the lack of the 1/2 factor in equation (22). The resulting free energy is more accurate than if a 1/2 had been introduced in equation (22) and this term eliminated.

With d_U as the dimension of $\langle \mathbf{U}_i \rangle$, equation (23) for all i is a set of $d_U N$ nonlinear coupled integral equations with a very large set of solutions. To find the lowest free-energy phase, information correlating the roots and possible space groups is desirable. Thus we symmetry adapt the potential using space-group IR's and define associated order parameters.

To keep the notation simple, we take the special case of one molecule i per primitive unit cell. The generalization to more is straightforward. Appendix D reviews the necessary background of space groups. To symmetry adapt the potential to G, we first use a discrete Fourier transform over the translational subgroup $\mathcal{T}$ which can also be viewed as a projection operator over elements $t_i \in \mathcal{T}$

$$V/kT = \tfrac{1}{2}\sum_{\mathbf{k}} \mathbf{U}^{\mathbf{k}*} \cdot \mathbf{K}^{\mathbf{k}} \cdot \mathbf{U}^{\mathbf{k}}, \tag{26}$$

where

$$\mathbf{U}^{\mathbf{k}} = \frac{1}{|\mathcal{T}|^{1/2}} \sum_{t_i \in \mathcal{T}} \exp(-i\mathbf{k} \cdot \mathbf{t}_i)\mathbf{U}_i \tag{27}$$

$$\mathbf{K}^{\mathbf{k}} = \frac{1}{|\mathcal{T}|} \sum_{t_i, t_j \in \mathcal{T}} \exp[-i\mathbf{k} \cdot (\mathbf{t}_i - \mathbf{t}_j)]\mathbf{K}_{ij}. \tag{28}$$

The summation over $\mathbf{k}$ goes over reciprocal-lattice vectors for all unit cells of the crystal.

A diagonalization of $\mathbf{K}^{\mathbf{k}}$ (Michel & Copley, 1997; Birman, 1984) can be used to express the potential in terms of eigenfrequencies $\mathbf{\Theta}^{\mathbf{k}} = (\mathbf{E}^{\mathbf{k}})^{-1} \cdot \mathbf{K}^{\mathbf{k}} \cdot \mathbf{E}^{\mathbf{k}}$ and normal coordinates $\mathbf{Q}^{\mathbf{k}} = (\mathbf{E}^{\mathbf{k}})^{-1} \cdot \mathbf{U}^{\mathbf{k}}$. In terms of the normal coordinates, the potential is

$$V/kT = \tfrac{1}{2}\sum_{\mathbf{k}} \mathbf{Q}^{\mathbf{k}*} \cdot \mathbf{\Theta}^{\mathbf{k}} \cdot \mathbf{Q}^{\mathbf{k}}. \tag{29}$$

As discussed in Appendix D, each $\mathbf{k}$ vector belongs to a star $\mathbf{s}$. Creating a vector of $\mathbf{Q}^{\mathbf{k}}$ belonging to a star and grouping components from each $\mathbf{Q}^{\mathbf{k}}$ by IR gives an ordered form of equation (29)

$$V/kT = \tfrac{1}{2}\sum_{\mathbf{s}} \mathbf{Q}^{\mathbf{s}*} \cdot \mathbf{\Theta}^{\mathbf{s}} \cdot \mathbf{Q}^{\mathbf{s}}. \tag{30}$$

This form has the IR components of each $\mathbf{Q}^{\mathbf{s}}$ in consecutive order. Appendix D discusses the matrices $\{\mathbf{D}^{\mathbf{s}}\}$ onto which a space-group IR maps elements of G. Since the action of G on $\mathbf{Q}^{\mathbf{s}}$ induces this representation

$$g \circ Q^{\mathbf{s}}_{p_\tau} = \sum_{r_\tau} Q^{\mathbf{s}}_{r_\tau} D^{\mathbf{s}}_{r_\tau p_\tau}(g), \tag{31}$$

the $\mathbf{Q}^{\mathbf{s}}$ are basis functions of a crystallographic IR space.

Although equation (30) gives the proper value of the potential, a Hamiltonian constructed with this potential leads to equations of motion that are not invariant under time reversal $\hat{T}$ that maps the complex field onto itself. To remedy this, $\mathbf{Q}^{\mathbf{s}}$ must be made to transform like physically irreducible representations of the space–time symmetry group $\mathcal{G} = G + \hat{T}G$. The mappings $\mathbf{Q}^{\mathbf{s}} \rightarrow \mathbf{q}^{\mathbf{s}}$, $\mathbf{\Theta}^{\mathbf{s}} \rightarrow \mathbf{\Lambda}^{\mathbf{s}}$ accomplishing this are discussed in Appendix E. The potential symmetric under time reversal is

$$V/kT = \tfrac{1}{2}\sum_{\mathbf{s}} \mathbf{q}^{\mathbf{s}} \cdot \mathbf{\Lambda}^{\mathbf{s}} \cdot \mathbf{q}^{\mathbf{s}}, \tag{32}$$

where IR's from this point on are implicitly real physically irreducible representations. Whereas the $\mathbf{Q}^{\mathbf{s}}$ are frequently complex, the $\mathbf{q}^{\mathbf{s}}$ are real-valued and the associated Hamiltonian yields equations of motion that are symmetric under time reversal.

The mean-field version of equation (32) is constructed in analogy to equation (22), giving

$$V^{\rm mf}/kT = \sum_{\mathbf{s}} \mathbf{q}^{\mathbf{s}} \cdot \mathbf{\Lambda}^{\mathbf{s}} \cdot \langle \mathbf{q}^{\mathbf{s}} \rangle \tag{33}$$

with orientational probability density

$$\rho = \exp(-V^{\rm mf}/kT) \Big/ \int \exp(-V^{\rm mf}/kT)\, \mathrm{d}(\boldsymbol{\omega}) \tag{34}$$

and free energy

$$A^{\rm var}_{\rm or} = -kT \ln Z^{\rm mf}_{\rm or} - \frac{kT}{2} \sum_{\mathbf{s}} \langle \mathbf{q}^{\mathbf{s}} \rangle \cdot \mathbf{\Lambda}^{\mathbf{s}} \cdot \langle \mathbf{q}^{\mathbf{s}} \rangle, \tag{35}$$

where

$$\langle \mathbf{q}^{\mathbf{s}} \rangle = \int \mathbf{q}^{\mathbf{s}} \rho \, \mathrm{d}(\boldsymbol{\omega}). \tag{36}$$

The IR components of $\langle \mathbf{q}^{\mathbf{s}} \rangle$ are order parameters (OP's) with amplitude [1, −1]. Whereas the number of $\langle \mathbf{U}_i \rangle$ is proportional to N, the number of $\langle \mathbf{q}^{\mathbf{s}} \rangle$ important to phase transitions is small as discussed in the following section.

2.5. Phase-transition information

With basis functions symmetry-adapted to a space group IR, we relate IR distortions to crystal-structure changes. To do this, we expand ρ in the basis functions $\mathbf{q}^{\mathbf{s}}$, which are a complete set over configurational space $SO(3) \times \mathbb{Z}^3$

$$\rho = \sum_{\mathbf{s}} \mathbf{a}^{\mathbf{s}} \cdot \mathbf{q}^{\mathbf{s}} \tag{37}$$

and by orthogonality determine the coefficients

$$\mathbf{a}^{\mathbf{s}} = \mathbf{w}^{\mathbf{s}} \cdot \int \mathbf{q}^{\mathbf{s}} \rho \, \mathrm{d}(\boldsymbol{\omega}) = \langle \mathbf{q}^{\mathbf{s}} \rangle \cdot \mathbf{w}^{\mathbf{s}}, \tag{38}$$

where $\mathbf{w}^{\mathbf{s}}$ is a diagonal matrix of weights $2l_i + 1$ arising from the $SO(3)$ manifold dependence of the components of $\mathbf{q}^{\mathbf{s}}$. This gives the Ginzburg–Landau form of ρ (Chaikin & Lubensky, 1995)

$$\rho = \sum_{\mathbf{s}} \langle \mathbf{q}^{\mathbf{s}} \rangle \cdot \mathbf{w}^{\mathbf{s}} \cdot \mathbf{q}^{\mathbf{s}}. \tag{39}$$

The order parameters have slightly different transformation rules to the basis functions, which can be seen by the action of G on ρ

$$g \circ \rho = \sum_{\mathbf{s}} \sum_{p_\tau} \langle q^{\mathbf{s}}_{p_\tau} \rangle w^{\mathbf{s}}_{p_\tau p_\tau} \sum_{r_\tau} q^{\mathbf{s}}_{r_\tau} D^{\mathbf{s}}_{r_\tau p_\tau}(g) \tag{40}$$

from which the action of G on the OP's is

$$g \circ \langle \mathbf{q}^{\mathbf{s}} \rangle = \sum_{p_\tau} D^{\mathbf{s}}_{r_\tau p_\tau}(g) \langle q^{\mathbf{s}}_{p_\tau} \rangle. \tag{41}$$

At a temperature higher than a phase transition, ρ is invariant under the parent phase group G_0 ($g \circ \rho = \rho = \rho_0$). Below the phase-transition temperature, certain order parameters become nonzero giving the probability density change

$$\delta\rho = \rho - \rho_0 = \sum_{\mathbf{s}} \langle \mathbf{q}^{\mathbf{s}} \rangle \cdot \mathbf{w}^{\mathbf{s}} \cdot \mathbf{q}^{\mathbf{s}}, \tag{42}$$

where **s** goes over the stars of those OP's that are nonzero. These OP's can be primary, coupled primary or secondary OP's.

The space-group operators leaving all $\langle \mathbf{q}^{\mathbf{s}} \rangle$ invariant in equation (41) determine the phase. As space groups are discrete infinite groups, this requires some care. Appendix *D* discusses how each IR maps all the elements g in the parent phase G_0 to matrices $\mathbf{D}^{\mathbf{s}}(g)$. If any of these matrices are identical, this mapping is a homomorphism. The set of *distinct* matrices $\{\mathbf{D}^{\mathbf{s}}\}$ forms a group called the image I_0 of G_0 whose kernel K is the subset of elements g such that the $\{\mathbf{D}^{\mathbf{s}}\}$ coincide with the identity. As the quotient group G_0/K is isomorphic to I_0, we write $G_0 = I_0 \times K$. Although K is an infinite group, I_0 is finite so it is useful to use the image instead of the full group when considering the invariance of $\delta\rho$. The operators of I_0 that leave $\langle \mathbf{q}^{\mathbf{s}} \rangle$ invariant form a group I giving the *isotropy group* $G = I \times K$. Algorithms exist (Hatch & Stokes, 1985) to tabulate possible primary OP directions and the images that leave them invariant (Stokes & Hatch, 1988, 2002*a*). This gives (i) a list of all possible solutions to equation (36) for a group–subgroup transition and (ii) a method to identify nonsubgroup transitions based on common isotropy groups (Hatch *et al.*, 2001; Stokes & Hatch, 2002*b*).

Being able to correlate an OP direction with a specific space group allows one to reduce the system size and make physically meaningful initial estimates for the OP's. For example, most phase transitions occur at high-symmetry points in the Brillouin zone. Thus the number of stars in equation (33) may be limited to those corresponding to high-symmetry points. This also means the number of molecules N in equations (23), (77) (Appendix *C*) and (25) may be limited to those within the kernel of a high-symmetry IR. Specifically, these are the molecules within the lattice vectors $\mathbf{t}_i$ in G_0 for which

$$\exp(i\mathbf{k} \cdot \mathbf{t}_i) = 1 \tag{43}$$

for all arms **k** of the star. OP directions leading to distinct space groups may be used to seed efficient numerical solutions of equation (36).

To summarize, phase transitions are branching points of solutions of equation (36) and the branch that minimizes $A_{\rm or}^{\rm var}$ is different on either side of the transition. Typically, the high-temperature phase has fewer nonzero OP's. The new OP's and their relative magnitudes determine the phase. Algorithms exist to enumerate possible solutions and their corresponding space groups. Each possibility can be used in turn to find the lowest free-energy solution of equation (36). A discontinuous jump of the OP's indicates a first-order phase transition while a higher-order transition has a continuous bifurcation.

3. Numerical solution methods

This section discusses some of the issues relating to calculation of the free energy of the new phase such as quadrature and choice of approximate packing of the molecules. An algorithm is also presented for constructing phase diagrams.

3.1. Numerical integration

The consistency relations [equation (23) or (36)] form a set of homogeneous nonlinear integral equations for the OP's, which must be solved numerically. As derived, the rotator functions in the integrands are in terms of trigonometric functions of Euler angles. We have chosen to convert these into algebraic functions by parametrizing the equations using quaternions (Kuipers, 1999), which can be accomplished using Wigner matrices in terms of quaternions (Lynden-Bell & Stone, 1989). This change of variables involves a 2-to-1 mapping of the unit quaternion space S^3 onto the Euler angle space $SO(3)$.

The quaternion integration region is the three-dimensional surface of a four-dimensional sphere. Taking advantage of this well studied region, we use spherical product Gauss formulae (Stroud & Secrest, 1966), whereby the three-dimensional integrals in equation (23) are transformed into the product of three integrals which are approximated using Gauss–Chebyshev (first kind), Gauss–Legendre, and Gauss–Chebyshev (second kind) one-dimensional quadrature formulas. For this publication, the integrals were approximated using 16 nodes in each quadrature for a total of 4096 function evaluations per integration. The 2-to-1 mapping permits using only half of these points.

3.2. Reference state

The resulting set of nonlinear algebraic equations and analytic Jacobian are solved using Newton–Raphson iteration (Press *et al.*, 1992). This requires an initial guess sufficiently close to the global free-energy minimum [*via* equation (25)]. Thus it is helpful to have a reference state in a particular lattice that minimizes the free energy. For this purpose, we define a high-temperature ($T \to \infty$) reference state of the crystal, in which the substance remains crystalline and the molecules maintain their identity. In this limit, equation (23) reduces to

$$\lim_{T\to\infty} \langle \mathbf{U}_i \rangle = \int U_{m_\tau n_\sigma}^{\ell_i} \, \mathrm{d}(\boldsymbol{\omega}_i) = \delta_{\ell_i,0}, \tag{44}$$

where the second equality follows from the properties of Wigner functions (Varshalovich *et al.*, 1988). Therefore, the only nonzero order parameter in this high-temperature limit is the one resulting from the angle-independent rotator function, $\langle U_{m_\tau n_\sigma}^0 \rangle = 1$. The orientational distribution of the molecules thus has the highest possible symmetry, that of a sphere. This corresponds to a fully disordered plastic crystalline (FDPC) phase, with translational but not orientational order within the crystal (Timmermans, 1961; Sherwood, 1979). The FDPC structure is observed for molecules such as methane (Neumann *et al.*, 2003) and C_{60} (Michel & Copley, 1997). Most molecular crystals melt, sublime or decompose prior to reaching this state. Even if the FDPC state is not observed for a particular molecular crystal, it is a useful reference state analogous to the ideal-gas reference state for vapors.

The plastic crystalline state is a solution of the consistency relations [equation (23) or (36)] for *any* lattice of molecular centers, but the proper lattice for the reference state is the one

Table 1
Candidate high-temperature lattices, their number of nearest neighbors, the nearest-neighbor separation as a function of specific volume v and presence of inversion center at molecular sites in the FDPC phase.

Lattice	N_{nn}	r_{nn}		Inversion
f.c.c./h.c.p.	12	$(2^{1/2}v)^{1/3}$	$\approx 1.1225v^{1/3}$	Yes/No
b.c.c.	8	$(3\times 3^{1/2}v/4)^{1/3}$	$\approx 1.0911v^{1/3}$	Yes
s.c.	6	$v^{1/3}$	$= 1.0000v^{1/3}$	Yes
Diamond	4	$(3\times 3^{1/2}v/8)^{1/3}$	$\approx 0.86603v^{1/3}$	No

that minimizes A_{or}^{var}. Evaluating the free energy [equation (25)] in the high-temperature limit and neglecting terms of order T^{-2} gives

$$\lim_{T\to\infty} A_{or}^{var} = -N\left[kT\ln z_{rot} - \frac{N_{nn}}{2}\nu_{0,0,0}(r_{nn})\right], \qquad (45)$$

where N_{nn} is the number of nearest neighbors and r_{nn} is the nearest-neighbor distance at fixed molecular volume v. For clarity, the superscripts on $\nu_{0,0,0}$ in (45) have been dropped. Since z_{rot} is independent of the lattice, the second term in (45) determines the high-temperature lattice. The number of nearest neighbors and their relative distances are summarized in Table 1 for several candidate lattices. For attractive potentials ($\nu_{0,0,0} < 0$), the free energy is minimized for a close-packed lattice (f.c.c. or h.c.p.) at the density that minimizes $\nu_{0,0,0}$. For long-range repulsive potentials ($\nu_{0,0,0} > 0$), the lattice that minimizes A_{or}^{var} depends upon the form of the potential. Stiff potentials favor more neighbors at a larger nearest-neighbor distance (*i.e.* f.c.c./h.c.p.) whereas soft potentials favor fewer neighbors at a shorter nearest-neighbor distance (*i.e.* diamond). Hard spheres are a special case. They form f.c.c. lattices due to translational entropy, an effect that has been neglected along with translational/rotational coupling in our model.

Once the high-temperature lattice has been selected, $\mathbf{\Omega}_{ij}$ is determined for each nearest-neighbor pair. The coupling matrix $\mathbf{K}_{ij}$ and thus the order parameters ($\langle\mathbf{U_i}\rangle$ or $\langle\mathbf{q^s}\rangle$) are then functions only of the scaled Fourier coefficients $\boldsymbol{\nu}/kT$. Since the order parameters provide a fingerprint for the crystal structure, we are able to construct global phase diagrams using the scaled Fourier coefficients at finite temperature as independent parameters.

3.3. Global phase diagram construction

Our algorithm for constructing global phase diagrams is illustrated in Fig. 2. Any of the molecular point groups may be chosen. Reference lattices are discussed in §3.2 above. In this section, we describe the remaining steps from Fig. 2.

To find the first transition temperature for a particular set of $\nu_{\ell_i\ell\ell_j}^{n_\sigma n_\mu}$, we align the set of $\langle\mathbf{q^s}\rangle$ in a vector $\langle\mathbf{q}\rangle$ and generalize equations (33) and (36)

$$V^{mf}/kT = \mathbf{q}\cdot\mathbf{\Lambda}\cdot\langle\mathbf{q}\rangle \qquad (46)$$

$$\langle\mathbf{q}\rangle = \int \mathbf{q}\rho\,\mathrm{d}(\omega) \qquad (47)$$

and observe at the new phase

$$\langle\mathbf{q}\rangle = \langle\mathbf{q}_{cf}\rangle + \mathbf{\Delta}, \qquad (48)$$

where $\langle\mathbf{q}_{cf}\rangle$ is the crystal field and is nonzero only for the unit IR's in the Γ point OP of the parent phase $\mathbf{q}^\Gamma$. The second term $\mathbf{\Delta}$ is a vanishingly small OP for the primary IR of the new phase that rises continuously from zero at the second-order phase-transition temperature T_2 (Heid, 1993). Inserting this solution in equation (47) and expanding the exponential around $\mathbf{\Delta}$ gives an eigenvalue problem for T_2

$$-(\langle\mathbf{q}\times\mathbf{q}\rangle - \langle\mathbf{q}_{cf}\rangle\times\langle\mathbf{q}_{cf}\rangle)\cdot\tilde{\mathbf{\Lambda}}\cdot\mathbf{\Delta} = kT_2\mathbf{\Delta} + \mathcal{O}(\mathbf{\Delta}^2), \qquad (49)$$

where $\times$ is a direct product, $\tilde{\mathbf{\Lambda}} = kT\mathbf{\Lambda}$ is independent of temperature and all thermal averages are in the crystal field potential

$$V_{cf}/kT = \mathbf{q}\cdot\mathbf{\Lambda}\cdot\langle\mathbf{q}_{cf}\rangle. \qquad (50)$$

The crystal field OP's $\langle\mathbf{q}_{cf}\rangle$ are obtained by solving their consistency relations. As this is also temperature dependent, it must be solved simultaneously with equation (49) for T_2 for each possible IR of $\mathbf{\Delta}$. The highest $T_2(\mathbf{k})$ gives the IR of the symmetry-breaking distortion.

For a f.c.c. reference state, each IR can be identified by the nonzero components of the corresponding eigenvector which are the point-group IR's τ_w of the Wyckoff point group w appearing in the representation of w subduced by the space-

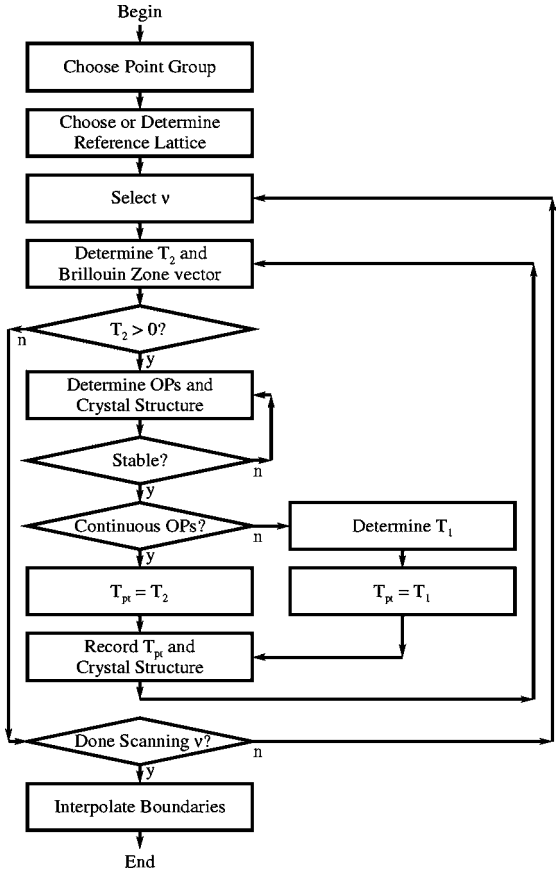

Figure 2
Flow chart illustrating the algorithm for constructing global phase diagrams.

group IR τ_G, $\tau_G \downarrow w$ (Bradley & Cracknell, 1972). This nonzero subduction frequency is

$$n = \frac{1}{|w|} \sum_{g \in w} \chi^{\tau_G}(g)^* \chi^{\tau_w}(g). \tag{51}$$

Special rules apply for physically irreducible representations and are explained by Stokes *et al.* (1991). The star corresponding to this IR gives the number of independent molecules in the system, see equation (43). While this method worked well for the f.c.c. lattice, we have since found that it is inadequate for other lattices. A more general approach to identify IR's, appropriate for other reference lattices, has been outlined by Maradudin & Vosko (1968).

The dimension of the IR is equal to the number of degenerate zero eigenvalues. If this dimension is greater than one, any linear combination of the $\mathbf{\Delta}$ is a solution of equation (49) up to $\mathcal{O}(\mathbf{\Delta}^2)$. This gives a limited domain in which to search for $\langle \mathbf{q}_\tau \rangle$. As discussed in §2.4, the OP directions corresponding to distinct isotropy subgroups of τ provide a list of all possible $\langle \mathbf{q}_\tau \rangle$. We find that initializing a Newton–Raphson solver with an OP direction with a magnitude that minimizes the free energy improves convergence to a self-consistent solution minimizing $A_{\rm or}^{\rm var}$. Thus we evaluate $A_{\rm or}^{\rm var}$ for multiple random values for each OP direction, without regard for the consistency relations, and keep track of the trial OP that minimizes $A_{\rm or}^{\rm var}$. This provides an initial estimate that rapidly converges towards a self-consistent solution for the OP's using the consistency relations. The OP directions, images, isotropy subgroups, new Wyckoff points, and **k**-stars can be calculated using *ISOTROPY* (Stokes & Hatch, 2002*a*).

Once the phase-transition temperature is located and the free energy minimized, the stability of the solution is tested. To do this, we use linear response theory (Chaikin & Lubensky, 1995) to determine how the system responds to an infinitesimal perturbation away from equilibrium at fixed temperature. Defining the response of the order parameters to an applied disturbance in the fields (see Appendix *F*),

$$\boldsymbol{\chi}^o \equiv \left. \frac{\partial \langle \mathbf{U} \rangle}{\partial \mathbf{h}} \right|_o, \tag{52}$$

the solution is stable if the Hessian

$$\frac{\partial^2 A_{\rm or}^{\rm var}}{\partial \langle \mathbf{U} \rangle \partial \langle \mathbf{U} \rangle} = -(\boldsymbol{\chi}^o)^{-1} + \mathbf{K} \tag{53}$$

is positive definite. The eigenvectors of the Hessian evaluated at the phase-transition temperature provide equivalent information to $\mathbf{\Delta}$ regarding the subspace of OP's that must be searched to minimize the free energy.

The order parameters for the low-temperature solution $\langle \mathbf{q} \rangle$ fall into two categories

$$(a) \quad \lim_{T \to T_2} \langle \mathbf{q} \rangle - \langle \mathbf{q}_{\rm cf} \rangle = 0$$

$$(b) \quad \lim_{T \to T_2} \langle \mathbf{q} \rangle - \langle \mathbf{q}_{\rm cf} \rangle \neq 0.$$

In case (*a*), the phase transition is second or higher order and $T_{\rm pt} = T_2$. In case (*b*), the phase transition is first order and $T_{\rm pt} > T_2$. To determine $T_{\rm pt}$, the free energies of the two solutions are set equal,

$$A_{\rm or}^{\rm var}(T_{\rm pt}, \langle \mathbf{q} \rangle) = A_{\rm or}^{\rm var}(T_{\rm pt}, \langle \mathbf{q}_{\rm cf} \rangle), \tag{54}$$

and solved for $T_{\rm pt}$, where free energies and consistency relations were discussed in §2.4.

To construct a global phase diagram, the above procedure is repeated for multiple sets of potential coefficients $\boldsymbol{v}$. After a sampling of points from the hypersphere of potential coefficients has been completed, we interpolate between corresponding phase transitions to locate the boundaries of the phases on the diagram. Fig. 1 is typical of global phase diagrams displaying the highest-temperature transition parametrized by three potential coefficients. Equilibrium loci for two particular phases are assigned a unique color. A global phase diagram presenting multiple transitions, where possible, as a function of two potential coefficients has been presented elsewhere (Keith *et al.*, 2004).

4. Example: tetrahedral molecules

To illustrate the method, we consider the class of molecules invariant under the tetrahedral point group (*e.g.* methane, adamantane, white phosphorous). This is because of its historical significance (Nagamiya, 1951; James & Keenan, 1959; Yamamoto *et al.*, 1977), applications to homologous series such as diamondoid lattice synthons (Zaworotko, 1994) and the high symmetry which leads to a reduction in the number of potential coefficients.

In a study of the seven most populated space groups from the Cambridge Structural Database, Motherwell (1997) concluded that there was a tendency for even nonspherical molecules to pack as if they were spheres, suggesting a way of looking at complex packing structures as approximations to the packing of spheres. Based on this, we choose to illustrate our method using a cubic close-packed (f.c.c.) high-temperature reference lattice $Fm\bar{3}m$, although a complete analysis requires consideration of alternative reference lattices.

As there are an infinite number of potential coefficients and we need a small number of independent variables, in this example we take equation (16) with $\ell_{\rm trunc} = 3$

$$V = \tfrac{1}{2} \sum_{ij} \sum_{\ell_i, \ell_j \in \{0,3\}} \sum_{\ell=|\ell_i-\ell_j|}^{\ell_i+\ell_j} \sum_{n_\sigma, n_\mu \in 1_{A_1}} v_{\ell_i \ell \ell_j}^{n_\sigma n_\mu}(r_{ij}) F_{\ell_i \ell \ell_j}^{n_\sigma n_\mu}(\boldsymbol{\omega}_i, \mathbf{\Omega}_{ij}, \boldsymbol{\omega}_j). \tag{55}$$

This truncation is discussed further in §5. From subduction frequencies in $SO(3)$, the unit IR appears once in the zeroth and third manifolds. Therefore, all $v_{\ell_i \ell \ell_j}^{n_i n_j}$ with $\ell_i, \ell_j = 1$ or 2 are zero and $n_\sigma, n_\mu = 1$ for $\ell_i, \ell_j = 0$ or 3 and $\{\sigma, \mu\} = A_1$. Also, the behavior of $F_{\ell_i \ell \ell_j}^{n_i n_j}$ under inversion allows one to show that (van der Avoird *et al.*, 1994)

$$v_{\ell_i \ell \ell_j}^{n_\sigma n_\mu} = 0 \qquad \text{if } \ell_i + \ell + \ell_j = \text{odd} \tag{56}$$

and, if molecules i and j are identical,

$$v_{\ell_i \ell \ell_j}^{n_\sigma n_\mu} = (-1)^{\ell_i + \ell_j} v_{\ell_j \ell \ell_i}^{n_\mu n_\sigma}. \tag{57}$$

Table 2
High-symmetry IR basis functions and the number of independent molecules N in the kernel of $G_0 \rightarrow I_0$.

$\mathbf{q}^s$	N
$\mathbf{q}^\Gamma = \{\mathbf{q}_{\Gamma_1^+}, \mathbf{q}_{\Gamma_2^-}, \mathbf{q}_{\Gamma_4^-}, \mathbf{q}_{\Gamma_5^-}\}$	1
$\mathbf{q}^X = \{\mathbf{q}_{X_1^+}, \mathbf{q}_{X_2^-}, \mathbf{q}_{X_3^-}, \mathbf{q}_{X_4^-}, \mathbf{q}_{X_5^-}, \mathbf{q}_{X_5^-}\}$	4
$\mathbf{q}^L = \{\mathbf{q}_{L_1^+}, \mathbf{q}_{L_1^-}, \mathbf{q}_{L_2^-}, \mathbf{q}_{L_2^-}, \mathbf{q}_{L_3^-}, \mathbf{q}_{L_3^-}\}$	8
$\mathbf{q}^W = \{\mathbf{q}_{W_1}, \mathbf{q}_{W_1}, \mathbf{q}_{W_2}, \mathbf{q}_{W_3}, \mathbf{q}_{W_5}, \mathbf{q}_{W_5}\}$	32

Therefore, for T_d molecules the only nonzero coefficients in the truncated potential are

$$\{\nu_{0,0,0}, \nu_{0,3,3}, \nu_{3,3,0}, \nu_{3,0,3}, \nu_{3,2,3}, \nu_{3,4,3}, \nu_{3,6,3}\}, \qquad (58)$$

where the redundant superscripts have been dropped. The first coefficient has a corresponding basis function that is isotropic. It influences the reference lattice through its radial dependence but does not favor any particular rotationally ordered phase. Therefore it is not included in the global phase diagram. By equation (57), the second and third coefficients are equal in magnitude so only one is independent. Furthermore, for molecules in a reference lattice with an inversion center at the occupied Wyckoff points such as f.c.c., b.c.c. or s.c., neither $\nu_{0,3,3}$ nor $\nu_{3,3,0}$ appear in the potential due to pairwise cancellation. This reduces the number of potential coefficients to four. We take the $\nu_{3,2,3} = 0$ projection since $\nu_{3,2,3}$ is small for a representative tetrahedral molecule (Missaghi *et al.*, 2004) and the remaining three coefficients

$$\boldsymbol{\nu} = \{\nu_{3,0,3}, \nu_{3,4,3}, \nu_{3,6,3}\} \qquad (59)$$

serve as independent parameters.

Next we symmetry adapt the rotator functions to O_h, the point group of Wyckoff point *a* in $Fm\bar{3}m$, and expand the sum in equation (33) over stars belonging to high-symmetry points of the Brillouin zone. This gives $\mathbf{q} = \{\mathbf{q}^\Gamma, \mathbf{q}^X, \mathbf{q}^L, \mathbf{q}^W\}$ as shown in Table 2 where the star superscript has been dropped when repeated in the IR notation of Miller & Love (1967). Apart from $\mathbf{q}_{\Gamma_1^+} = U^0_{1,1} = 1$, the trivial case, there is no $\mathbf{q}_{\Gamma_1^+}$ basis function in the Γ star so there is no crystal field and the variance in equation (49) is the identity. As a result, there is no crystal field and equation (49) can be solved directly for T_2 as a function of $\boldsymbol{\nu}$ in this case. Generally, T_2 must be solved for iteratively due to the consistency relations for the crystal field.

The eigenvalues of equation (49) provide the temperatures leading to the first phase-transition surfaces in Fig. 1. The origin is the reference state and the axes are scaled potential parameters $\boldsymbol{\nu}/kT$. Each ray from the origin is a particular molecular crystal as a function of temperature. Some of the phase surfaces are planar while others are curved. Each differently colored surface corresponds to a different lower-temperature phase and each volume to a phase.

To identify each phase, equation (36) is solved for a system of N independent molecules as shown in Table 2 with the resulting OP directions in Table 3. Phase boundaries beyond the first phase transition are discussed elsewhere (Keith *et al.*, 2004). The transition temperatures T_2 are also shown as a function of wavevector in Fig. 3 for three example sets of potential coefficients discussed in the following paragraphs.

Table 3
Space-group IR OP directions with corresponding space group and Wyckoff points of the molecules.

Cases with two IR's and OP's give the primary IR first and the secondary IR's second. Letters *a*, *b* in the 'OP directions' column are temperature-dependent constants. Each solution is only one of several energetically equivalent domains.

IR	OP directions	Space group		Wyckoff point	Z
Γ_1^+	{1}	225	$Fm\bar{3}m$	*a*	4
Γ_2^-	{*a*}	216	$F\bar{4}3m$	*a*	4
Γ_4^-, Γ_2^-	{*a*, *a*, *a*}, {*b*}	160	$R3m$	$a\,(z = 0)$	3
Γ_5^-	{*a*, 0, 0}	121	$I\bar{4}2m$	*a*	2
X_2^-	{*a*, 0, 0}	137	$P4_2/nmc$	*b*	2
X_3^-, Γ_2^-	{*a*, *a*, *a*}, {*b*}	215	$P\bar{4}3m$	$e\,(x = 3/4)$	4
X_4^-	{*a*, 0, 0}	134	$P4_2/nnm$	*b*	2
X_5^-, X_5^-	{*a*, 0, 0, 0, 0, 0}, {*b*, 0, 0, 0, 0, 0}	59	$Pmmn$	$a\,(z = 1/4)$	2
L_1^-	{*a*, *a*, *a*, *a*}	226	$Fm\bar{3}c$	*a*, *c*	32
L_2^-, L_2^-	{*a*, *a*, *a*, −*a*}, {*b*, *b*, *b*, −*b*}	227	$Fd\bar{3}m$	$e\,(x = 3/4)$	32

For $\boldsymbol{\nu}/kT = f\{1, 0, -5\}$, where f is a temperature-dependent scaling factor, the phase is f.c.c. at high temperature ($f < 0.82$) and, as shown in Fig. 4, the free energy has a single minimum. Fig. 3(*a*) shows that a Γ-point mode has the highest T_2. The corresponding eigenvector has nonzero entries for T_{2u} components of $\mathbf{U}_i$. The nonzero subduction frequency of T_{2u} in $\Gamma_5^- \downarrow O_h$ identifies this mode as Γ_5^-. Possible isotropy subgroups and their associated OP directions calculated using *ISOTROPY* are listed in Table 4. These OP directions include the six possible daughter phases. As the temperature is

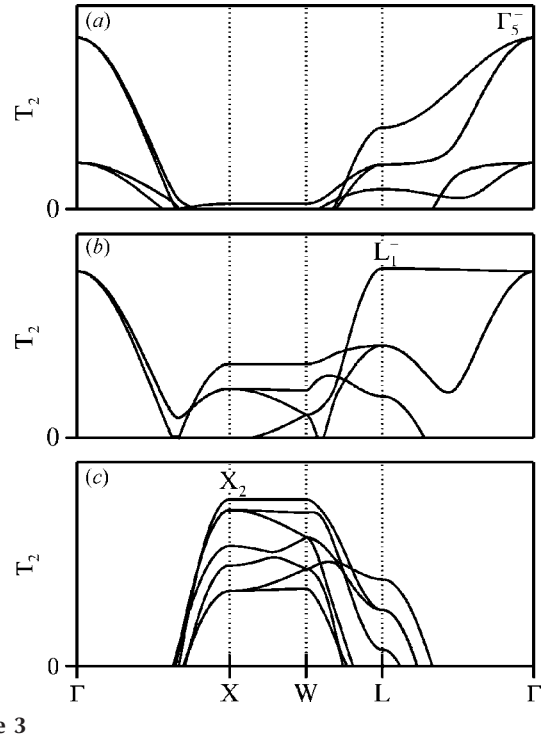

Figure 3
Second-order transition temperatures in arbitrary units along some high-symmetry points and lines for three sets of expansion coefficients (*a*) {1, 0, −5}, (*b*) {0, 0, −1} and (*c*) {35, −11, −5}. The highest temperature in each case corresponds to the symmetry-breaking IR.

Table 4
Isotropy subgroups and OP directions for the Γ_5^- IR.

Subgroups		OP directions	Lattice vectors of new unit cell
121	$I\bar{4}2m$	$\{a, 0, 0\}$	$\{1/2, 1/2, 0\}, \{-1/2, 1/2, 0\}, \{0, 0, 1\}$
44	$Imm2$	$\{a, a, 0\}$	$\{1/2, 0, 1/2\}, \{0, 1, 0\}, \{-1/2, 1/2, 0\}$
155	$R32$	$\{a, a, a\}$	$\{-1/2, 1/2, 0\}, \{0, -1/2, 1/2\}, \{1, 1, 1\}$
8	Cm	$\{a, b, 0\}$	$\{0, 0, -1\}, \{0, 1, 0\}, \{1/2, 0, 1/2\}$
5	$C2$	$\{a, a, b\}$	$\{-1/2, 1, -1/2\}, \{-1/2, 0, 1/2\}, \{1/2, 0, 1/2\}$
1	$P1$	$\{a, b, c\}$	$\{0, 1/2, 1/2\}, \{1/2, 0, 1/2\}, \{1/2, 1/2, 0\}$

lowered ($f > 0.82$) and the system passes the phase boundary, equation (36) gives a stable OP direction $\{a, 0, 0\}$ for $\langle \mathbf{q}_{\Gamma_5^-} \rangle$ as the free energy displays new minima. See the insets to Fig. 4. As stated previously, the phase transition is continuous so no search for T_1 is necessary. Other solutions such as $\{0, a, 0\}$ and $\{0, 0, a\}$ with the same free energy correspond to distinct domains of this phase with different orientations. The new phase is $I\bar{4}2m$, a symmorphic phase as expected from a Γ-point IR. The molecules are at Wyckoff point a. This is a body-centered tetragonal phase with all molecules in equivalent orientations and lattice constants $a = b = c/2^{1/2}$. This is one point on the light blue surface in Fig. 1(a).

For $\boldsymbol{v}/kT = f\{0, 0, -1\}$ at high temperature ($f < 8.60$) in a f.c.c. phase, Fig. 3(b) shows the highest T_2 at the L point. Its eigenvector is nonzero for the T_{2u} point-group IR, giving the four-dimensional L_1^- space group IR. Just past $f = 8.60$, equation (36) converges to the lowest free-energy OP $\langle \mathbf{q}_{L_1^-} \rangle = \{a, a, a, a\}$ giving the new phase $Fm\bar{3}c$ with molecules at Wyckoff points a and c. Other equivalent directions are $\{-a, -a, a, a\}$, $\{-a, a, a, -a\}$ and $\{a, -a, a, -a\}$. The unusual feature of this phase is that two of the eight molecules at Wyckoff point a in the primitive unit cell are disordered while the other six at Wyckoff point c assume a fixed equilibrium position. See the inset labeled $Fm\bar{3}c$ in Fig. 1(b). This structure is observed in methane and deuteromethane (Press, 1972; Yamamoto *et al.*, 1977).

For $\boldsymbol{v}/kT = f\{35, -11, -5\}$, Fig. 3($c$) shows a phase transition of the same temperature for three places in the Brillouin zone to lowest order. A subduction frequency calculation shows these modes are the 12-dimensional V_2 IR in the V line star, the three-dimensional X_2^- IR and the six-dimensional W_3 IR. As equation (49) truncated at second order gives all three with the same transition temperature, a calculation must be done for all three IR's to find the lowest free-energy phase. The V line involves a system size of $N = 256$, so specialized techniques are employed as described in Appendix G. Comparing the five possibilities, we find the X_2^- IR with direction $\{a, 0, 0\}$ has the lowest free energy and the space group is $P4_2/nmc$ with molecules at Wyckoff point b.

To fully explore the three-dimensional surface depicted in Fig. 1, nine different sets of potential coefficients were considered in the same manner as the three discussed above. Other structures listed in Table 3 were identified in the same manner giving three symmorphic phases at the Γ point and six nonsymmorphic ones from the X and L stars. Only $P4_2/nmc$ required consideration of the V line and W point. The unit-cell structures for each phase in Fig. 1 are drawn with tetrahedral subunits meant to indicate the orientation of the threefold axes and to suggest a molecular component. The size and precise shape of the subunits was chosen arbitrarily. Dotted lines indicate the relationship to the reference lattice in cases where the Bravais lattice is different to the reference lattice.

5. Discussion

Since the purpose of this publication is to present a method for producing molecular crystal global phase diagrams, it is appropriate to comment at some length on the limitations, features and applications of our work. There are several assumptions used in the derivation of our methods that limit their applicability. Although they are discussed separately within §2, it seems appropriate to summarize them in §5.1 and §5.2. There are also several noteworthy features of our model that may not be apparent from the derivation. They are discussed in §5.3. Finally, we discuss some of the ways in which global phase diagrams may be used in §5.4.

5.1. Assumptions and their justifications

The partition function in equation (1) is for a classical crystal. It should not be used at or below the translational and rotational excitation temperatures based on the molecular mass and moments of inertia. In practice, this means that it should not be applied to crystals of quantum gases such as H_2 or He, nor should it be applied to crystals of quantum rotors such as CH_4 (James & Keenan, 1959; Yamamoto *et al.*, 1977). Molecules in which the heavy atoms are collinear but H atoms are off-axis, such as ethylene, have two classical and one quantum rotation. Replacing hydrogen with deuterium is generally sufficient to make the classical partition function applicable in the relevant temperature range. Crystals of molecules with three or more noncollinear heavy atoms are well modeled using equation (1) at temperatures relevant for their phase transitions. Our model is appropriate for molecular crystals composed of molecules whose dimensionality

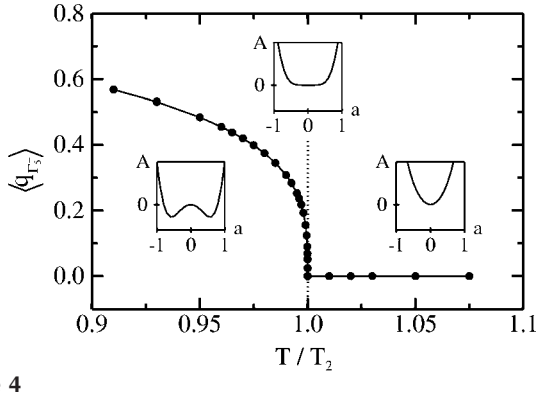

Figure 4
Orientational OP *versus* temperature for $\boldsymbol{v} \propto \{1, 0, -5\}$. At high T, $\langle \mathbf{q}_{\Gamma_5^-} \rangle = \{0, 0, 0\}$ giving a f.c.c. phase. At low T, $\langle \mathbf{q}_{\Gamma_5^-} \rangle = \{a, 0, 0\}$, where a is temperature-dependent, giving the phase $I\bar{4}2m$. Insets show $A_{\rm or}^{\rm var}$ *versus* a at three temperatures showing a single minimum at the origin at high temperature and multiple equivalent minima at low temperature.

with all atoms considered is the same as when H atoms are disregarded.

Equation (1) is also written for a single-component crystal of N identical molecules. It is clearly applicable to crystals that exclude their solvent, but it can reasonably be applied in other cases as well. Crystals of multiple components that are tightly and rigidly bound may be treated as if they are composed of a single supermolecular component. At the other extreme, proteins often crystallize with a large and variable volume fraction of water. Such crystals can be modeled using our methods as a single component whose interactions are mediated by implicit residual solvent. Both of these cases may occur together. Our model as written is not suitable for nonrigid stoichiometric solvates and other multicomponent crystals. In this case, the relative positions and orientations of each component should be treated separately. There are many more high-temperature reference phases to be considered for multicomponent crystals, particularly if the stoichiometry is not fixed. Also, there are many more potential coefficients $\boldsymbol{v}$ to be considered if the components interact both with themselves and with each other. Both effects greatly complicate our analysis, but may be included in principle.

There is a detailed discussion of our assumptions regarding separation of modes in §2.2 so we only summarize it briefly here. We have invoked the common assumption that internal vibrational modes decouple from the external modes. Torsions are the exception. There a distinction is made between torsions that do not significantly change the molecular shape (such as for methyl groups) and shape-changing torsions. The latter type leads to the possibilities of conformational polymorphs and of crystals in which the same molecule adopts multiple conformations. Also, such torsions can lead to translation–rotation coupling preventing the separation of these modes in equations (7) and (8). These complexities have been avoided in the current work. Thus, as presented here, our model is applicable to molecules without torsions that couple strongly to molecular rotational and/or translational modes, $t = 0$.

We have further assumed that the rotational modes drive phase changes between crystalline phases in molecular solids. This is not always the case for multicomponent ionic crystals such as alkali metal cyanides/superoxides, sodium nitrate/nitrite and ammonium halides (Lynden-Bell & Michel, 1994). In these cases, the rotations of the anions are strongly coupled to cation translations while the cations screen the rotational couplings between anions. Since we have eliminated multicomponent systems from consideration, we do not discuss these further here. Order/disorder phase transitions in crystals of neutral molecules are generally dominated by rotational coupling (Lynden-Bell & Michel, 1994). We do not mean to imply that molecules do not translate during phase transitions, only that the translations are secondary to the orientational structure changes.

An example may serve to illuminate the role of translations. Heavy methane (CD_4) crystallizes at atmospheric pressure near 90 K forming a fully disordered f.c.c. plastic crystal called phase I. At 27.0 K, a new phase with $Fm\bar{3}c$ symmetry forms as shown in Fig. 1. In this phase, called phase II, the molecules remain on a f.c.c. lattice but adopt orientational preferences that break some of the high-temperature phase symmetries. At 22.1 K, the crystal phase changes to *Cmca* in which the molecules not only adopt new orientational preferences but also move a few percent of a lattice parameter into new lattice sites (Neumann *et al.*, 2003). This phase is called phase III. This phase does not appear in Fig. 1 due to our choice of potential coefficients as discussed below. We contend that the molecular translations accompanying the transition from phase II to phase III are of secondary importance. To test this hypothesis, we took the *Cmca* structure and translated the molecules back to their phase II locations while preserving their phase III orientations. This structure also has space group *Cmca*. Thus the translations did not break any additional crystallographic symmetries but merely relaxed by a small fraction of the lattice parameter in directions permitted by the rotationally determined space group. Thus the translations are tied to secondary order parameters while the primary order parameters have rotational origins. Small density changes associated with solid/solid phase changes are also tied to these secondary order parameters, at least at low pressures. The quantitative effects caused by omitting translations are expected to be similar to those related to omission of other secondary order parameters (Michel & Copley, 1997). The work of Motherwell (1997) indicates that the molecular centers in the majority of the Cambridge Structural Database share similar close-packing tendencies. We conclude that the database contains many structures determined primarily by rotational coupling. Our method focuses on these structures.

In writing the interaction potential in equation (16), we assumed that it is pairwise additive over molecules. It is well known that there are important three-body effects when considering interactions among atoms (Stone, 1996). Such effects are included implicitly in the values of the expansion coefficients $\boldsymbol{v}$ if the atoms are contained in two molecules. Not included explicitly in the model is the influence of an entire molecule on the interactions between two other molecules. Since most molecules are larger than their characteristic interaction distance, this is a reasonable approximation. Also, since we consider phase change at constant density, the multimolecular effects are included implicitly in the values of the potential coefficients as a mean-field contribution. As a brief aside, this contribution to the intermolecular potential is not easily calculated using *ab initio* methods. In fact, there is no available method to calculate the dispersion interaction in an infinite crystal from *ab initio* calculations (Wilson, 2003). This inability to calculate accurately the dominant binding mechanism for van der Waals crystals currently limits the effectiveness of *ab initio* methods for molecular crystal structure prediction. It is much less of a limitation for metallic, covalent and some strongly hydrogen-bonded structures. *Ab initio* methods have been quite successful in describing these structures (Wilson, 2003). Our model is not appropriate for metallic or covalent structures because it neglects many-body interactions. Likewise any limitation to nearest-neighbor interactions precludes ionic or strongly dipolar molecules

which require Ewald sums over distant neighbors. However, the pairwise additive potential over molecules with an implicit multibody contribution is a good approximation for the neutral molecules considered here.

The mean-field approximation does not include the effects of correlation in molecular motions. This is not believed to affect the properties of phases far from phase transitions or near first-order phase transitions, but does have significant effects on properties near continuous phase transitions. In principle, classical density functional methods can fully account for correlations, but the unknown free-energy functional introduces other uncertainties (Oxtoby, 2002). Renormalization group theory contains methods for incorporating correlation effects in mean field models, but we find it difficult to automate (Chandler, 1987). We anticipate that correlation effects may change the magnitudes of phase-transition temperatures, but are unlikely to change the identities of the phases involved. Since global phase diagrams are used primarily to indicate relations among phases rather than to produce high-precision predictions, we choose to ignore such correlation effects for simplicity.

The ability to construct low-dimensional global phase diagrams depends upon the representation of the IP using as few parameters as possible. Potential-energy expansions similar to equation (16) have been used by previous authors who have concluded that the convergence of the summations is too slow for practical applications (Briels, 1980). However, the same authors acknowledge that this is surprising considering the early success in representing phase behavior with only one parameter (James & Keenan, 1959). We contend that the fault lies in the methods used to determine the expansion coefficients $\boldsymbol{v}$ rather than the expansion itself. Since the basis set in the expansion is orthogonal, the standard technique is to multiply both sides of equation (16) by one of the functions $F(\boldsymbol{\omega}_i, \boldsymbol{\Omega}_{ij}, \boldsymbol{\omega}_j)$ and integrate over all orientations. This method is known to provide optimal estimates for $\boldsymbol{v}$ in the least-squares sense (Davis, 1989). The drawback of this method is that it equally weights all possible orientations whereas crystal structures typically contain only low-energy molecular interactions. What is needed is a potential that is very accurate in the vicinity of the energy minima while sacrificing accuracy for the less-important high-energy interactions. To this end, we are refining a process for evaluating $\boldsymbol{v}$ that equates the multidimensional Taylor-series expansion of the potential and its approximation in equation (16). The result is a potential that accurately represents the low-energy interactions but underestimates the strongly repulsive portions of the potential. This type of potential is ideal for crystallization modeling, but would be less suitable for molecular-beam scattering studies, for instance. Another benefit of this potential parametrization is that it naturally produces a coarse-grained approximation to very complicated potentials. This is crucial for protein modeling for which a fully atomistic potential would require far too many parameters. We have in mind a coarse-grained potential more in keeping with the patch model of Lenhoff and co-workers (Hloucha *et al.*, 2001). For the purposes of this paper, it is sufficient to acknowledge that the expansion in equation (16) can give a suitable potential with a small number of expansion coefficients. Details of this method will be published elsewhere (Missaghi *et al.*, 2004).

5.2. Additional assumptions in the example

For the purposes of our example in §4, we have made additional assumptions which simplified the example but are not absolutely necessary given our methods. These are discussed in this subsection.

We have assumed that molecular distortions due to the crystal field are small. Therefore only basis functions consistent with the unit representation of the molecular point-group symmetry were included in equation (16). Strong crystal fields can induce a distortion of the molecule that changes the magnitudes of the expansion coefficients and can introduce additional coefficients if the molecular symmetry is reduced. For instance, cubic molecules cannot have a dipole moment in vacuum but a dipole can be induced if the molecule is placed in a suitable crystal. We believe that these induced potential coefficients are likely to be small and therefore minor contributors to the crystal phase behavior. This is consistent with the common belief that molecules in crystals have structures similar to those in other phases, except that torsion angles may be fixed in the crystal. Again, we have not considered external torsional modes in this work.

We have also assumed that the potential is dominated by nearest-neighbor interactions. As a result, the potential coefficients which are defined as functions of intermolecular distance have been evaluated at the nearest-neighbor distance in the crystal and treated as scalars. This is appropriate if the molecule is large relative to its range of interaction and the molecule is globular. The first condition is met for all but the smallest of molecules. The second is met for many of the compounds in the Cambridge Structural Database, but does not apply to molecules that form interpenetrating networks (Zaworotko, 1994). Crystals composed of interpenetrating networks form when nonconvex molecules interact favorably with second or more distant neighbors, but do not strongly repel their nearest neighbors. This case may be handled by defining potential coefficients for multiple sets of neighbors. The effect is to multiply the number of potential coefficients in the potential by the number of significant sets of neighbors. Since interpenetrating network crystals are an interesting but relatively uncommon occurrence, we have retained only first neighbors in our example.

When choosing which terms to retain in equation (16), we simply truncated the potential for $\ell_i, \ell_j \leq 3$. As noted earlier, this prevented the low-temperature phase of heavy methane *Cmca* from appearing on the slices of the global phase diagrams in Fig. 1. We expect that it is possible to observe the new phase by a less severe truncation of V and plotting a different projection of the potential coefficients (Keith & McClurg, 2004).

Based on the work of Motherwell (1997) and the high-temperature phase for methane, we chose the f.c.c. reference phase. A complete global phase diagram for a particular

molecular point group is the union of figures like Fig. 1 with other high-temperature reference phases such as h.c.p., b.c.c. and s.c. This opens up the possibility of martensitic transformations in which the phase transition is between phases with different reference phases. We are currently producing a more complete set of global phase diagrams for tetrahedral molecules (Keith & McClurg, 2004).

5.3. Features

One of the important features of our model is that it includes temperature-dependent entropic effects by minimizing the free energy (A) rather than simply the potential (V). Though it has been noted that the free energy is a better criterion for structure prediction than is the lattice energy (van Eijck *et al.*, 2001), many of the currently available software packages are based on the assumption that molecules adopt single low-energy orientations at low temperature (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002). This becomes inadequate at finite temperatures and for crystals with disordered phases (*i.e.* plastic crystals). Our model includes contributions from the kinetic energy and from the entire distribution of molecular orientations, both of which are important to represent adequately the physical behavior of the molecular crystals at finite temperature.

Our method is a systematic procedure to search for low-temperature crystalline phases starting from a small number of high-temperature reference lattices. By starting with a highly symmetric structure and seeking symmetry-breaking transitions, we avoid the need to scan a large number of trial structures. Other search algorithms choose a series of space groups and a variable number of molecules per primitive cell to propose suitable structures (Verwer & Leusen, 1998). The most common space groups and small-integer numbers of molecules per cell are considered first, but this procedure is not exhaustive. Therefore the possibility of a less-common space group and/or a larger unit cell cannot be discounted. Our method eliminates this uncertainty by starting in one of a small number of fully disordered plastic crystalline structures. Subsequent symmetry-breaking phase transitions lead to less-symmetric space groups and/or larger unit cells without having to specify these *a priori*.

By summarizing the phase behavior of an entire class of molecules with the same point-group symmetry, rather than one molecule at a time, we are able to answer several important questions.

- Why is a given phase observed and not an alternative phase?
- What phase is to be expected for a member of a homologous series given the structures of the other members?
- How will a perturbation to the intermolecular potential affect the equilibrium phase?
- What perturbation to the intermolecular potential is required to cause a desired change in phase?
- Which phases have only slightly higher free energy than the equilibrium phase?

The answers to these questions are the basis for several uses of molecular crystal global phase diagrams.

5.4. Uses

Global phase diagrams like the ones shown in Fig. 1 are maps between intermolecular potentials and crystalline phase behavior. The maps could be used to predict crystal structure if the potential coefficients were known with sufficient accuracy and precision. Unfortunately, the large number of nearly iso-energetic structures common in molecular crystals make it likely that the global minimum free energy changes abruptly as a result of small changes in the interaction potential. This is an instance of extreme parameter sensitivity. The requisite accuracy and precision for crystal-structure prediction is difficult to achieve given the difficulty in accurately estimating dispersion interactions for infinite crystals using *ab initio* methods, as discussed above. Therefore, we recommend using our global phase diagrams for the inverse problem. The global phase diagrams serve as tools for delineation of the family of interaction potentials that is consistent with an observed crystal structure. Owing to the extreme parameter sensitivity that hinders crystal structure prediction, the inverse problem yields a small family of consistent potential parameters. Of course the solution is not unique. This process, which we call reverse engineering of the crystal, has several uses.

Reverse engineering of a crystal is a tool for data mining in the extensive Cambridge Structural Database and Protein Data Bank. Examining molecules of the same symmetry with similar chemical structures (*e.g.* homologous series) allows us to compare the interaction potential parameters that led to each structure in an attempt to understand the relationship between interaction potential and crystal structure.

We are able to provide guidance in crystal design to produce desired crystal structures. For some applications, including nonlinear optical and organic electronic materials, a particular space group or packing motif is desired. Currently, a candidate molecule with desirable properties is synthesized and crystallized, but too often the resulting crystal structure is incompatible with the desired bulk properties. The key is to direct changes to the candidate molecule to influence crystal phase behavior in a desired manner. This search is largely unguided, requiring extensive synthesis and crystallization of candidate molecular libraries. Using molecular crystal global phase diagrams, the IP parameters can be estimated for both the observed and the desired crystal structures. The direction in phase space indicated by the differences between the two parameter sets is an indication of the perturbation to the intermolecular potential needed to achieve the desired crystal. This perturbation is in a form that can be used to screen molecules using *ab initio* methods. We have previously noted that such methods may not have the requisite accuracy for reliable crystal-structure prediction, but they may be useful for indicating the perturbation to the intermolecular potential resulting from a perturbation of the molecular structure. There is reason for optimism because differences are often more easily estimated than absolute magnitudes. Thus the

global phase diagrams provide a feedback mechanism from the synthetic chemist to the *ab initio* property simulator.

The methods used to produce the global phase diagrams can also provide information about metastable phases. The range of metastability of each phase is not shown in Fig. 1 for clarity, but in most cases the low-energy metastable phases are the neighboring phases in the global phase diagram. Each of these is a possible crystal polymorph. If information about one or more polymorphs is known, the family of consistent intermolecular potential parameters can be further refined. Alternatively, the structures and relative energies for various polymorphs can be computed for a particular choice of intermolecular potential parameters. Certainly kinetics plays an important role in determining which polymorphs are observed. However, high-energy structures are unlikely to be observed and the observed polymorphs must be at least metastable. A set of observed polymorphs that surrounds, but does not include, a particular phase in the global phase diagram would be a strong indication that the missing phase is a yet to be observed polymorph which is likely to be the thermodynamically stable structure. The inference regarding the stable structure derives from its supposed central location. On the other hand, if the observed polymorphs form a convex set within the global phase diagram, then the most stable of the existing polymorphs is likely to be the thermodynamically stable one.

6. Conclusions

We have described a method for construction of molecular crystal global phase diagrams. The phase diagrams summarize the phase behavior of molecules with a given molecular point-group symmetry. The method is illustrated for the tetrahedral point group T_d. Molecular crystal global phase diagrams are useful tools for data mining, crystal design and polymorph enumeration.

RBM received support through a 3M Non-Tenured Faculty Grant. Additional financial support was provided by the Graduate School of the University of Minnesota. Computational resources maintained by the University of Minnesota Supercomputer Institute were used for portions of this research. We had helpful discussions with Professor David Morse regarding mean-field theory.

References

Allen, F. H. & Motherwell, W. D. S. (2002). *Acta Cryst.* B**58**, 407–422.
Avoird, A. van der, Wormer, P. E. S. & Moszynski, R. (1994). *Chem. Rev.* **94**, 1931–1974.
Avoird, A. van der, Wormer, P. E. S., Mulder, F. & Berns, R. M. (1980). *Topics Curr. Chem.* **93**, 1–50.
Beyer, T., Lewis, T. & Price, S. L. (2001). *CrystEngComm*, **3**, 178–212.
Birman, J. L. (1984). *Theory of Crystal Space Groups and Lattice Dynamics.* New York: Springer-Verlag.
Bradley, C. J. & Cracknell, A. P. (1972). *The Mathematical Theory of Symmetry in Solids.* New York: Oxford University Press.
Braga, D., Grepioni, F. & Orpen, A. G. (1999). Editors. *Crystal Engineering: from Molecules and Crystals to Materials.* Boston: Kluwer Academic Publishers.
Briels, W. J. (1980). *J. Chem. Phys.* **73**, 1850–1861.
Briels, W. J., Jansen, A. P. J. & van der Avoird, A. (1986). *Adv. Quantum Chem.* **18**, 131–206.
Califano, S., Schettino, V. & Neto, N. (1981). *Lattice Dynamics of Molecular Crystals.* New York: Springer-Verlag.
Chaikin, P. M. & Lubensky, T. C. (1995). *Principles of Condensed Matter Physics.* New York: Cambridge University Press.
Chaka, A. M., Zaneiwski, R., Youngs, W., Tessier, C. & Klopman, G. (1996). *Acta Cryst.* B**52**, 165–p183.
Chandler, D. (1987). *Introduction to Modern Statistical Mechanics.* New York: Oxford University Press.
Chemla, D. S. & Zyss, J. (1987). In *Nonlinear Optical Properties of Organic Molecules and Crystals*, edited by D. S. Chemla & J. Zyss, Vol. 1, pp. 23–191. Orlando: Academic Press.
Chirikjian, G. S. & Kyatkin, A. B. (2000). Editors. *Engineering Applications of Noncommutative Harmonic Analysis.* New York: CRC Press.
Coombes, D. S., Price, S. L., Willock, D. J. & Leslie, M. (1996). *J. Phys. Chem.* **100**, 7352–7360.
Davis, H. F. (1989). *Fourier Series and Orthogonal Functions.* New York: Dover.
Desiraju, G. R. (1989). *Crystal Engineering: the Design of Organic Solids.* New York: Elsevier.
Dinur, U. & Hagler, A. T. (1991). *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz & D. B. Boyd, Vol. 2, pp. 99–164. New York: Wiley-VCH.
Eijck, B. P. van, Mooij, W. T. M. & Kroon, J. (1995). *Acta Cryst.* B**51**, 99–103.
Eijck, B. P. van, Mooij, W. T. M. & Kroon, J. (2001). *J. Phys. Chem. B*, **105**, 10573–10578.
Gavezzotti, A. (1991). *J. Am. Chem. Soc.* **113**, 4622–4629.
Gavezzotti, A. (1994). *Acc. Chem. Res.* **27**, 309–314.
Gavezzotti, A. (2002). *Model. Simul. Mater. Sci. Eng.* **10**, R1–R29.
Gdanitz, R. J. (1992). *Chem. Phys. Lett.* **190**, 391–396.
Gdanitz, R. J. (1997). In *Theoretical Aspects and Computer Modeling of the Molecular Solid State*, edited by A. Gavezzotti, pp. 185–201. New York: Wiley.
Gdanitz, R. J. (1998). *Curr. Opin. Solid State Mater. Sci.* **3**, 414–418.
Hatch, D. M., Lookman, T., Saxena, A. & Stokes, H. T. (2001). *Phys. Rev. B*, **64**, 060104(R)-1–060104(R)-4.
Hatch, D. M. & Stokes, H. T. (1985). *Phys. Rev. B*, **31**, 2908–2912.
Heid, R. (1993). *Phys. Rev. B*, **47**, 15912–15922.
Hloucha, M., Lodge, J. F. M., Lenhoff, A. M. & Sandler, S. I. (2001). *J. Cryst. Growth*, **232**, 195–203.
Holden, J. R., Du, Z. & Ammon, H. (1993). *J. Comput. Chem.* **14**, 422–437.
Holman, K. T., Pivovar, A. M., Swift, J. A. & Ward, M. D. (2001). *Acc. Chem. Res.* **34**, 107–118.
James, H. M. & Keenan, T. A. (1959). *J. Chem. Phys.* **31**, 12–41.
Karfunkel, H. R., Leusen, F. J. J. & Gdanitz, R. J. (1993). *J. Comput. Aided Mater. Des.* **1**, 177–185.
Keith, J. B. & McClurg, R. B. (2004). In preparation.
Keith, J. B., Mettes, J. A. & McClurg, R. B. (2004). *Cryst. Growth Des.* **4**, 1009–1012.
Kitaigorodskii, A. I. (1961). *Organic Chemical Crystallography.* New York: Consultants Bureau.
Konynenburg, P. H. van & Scott, R. L. (1980). *Philos. Trans. R. Soc. London Ser. A*, **298**, 495–540.
Kuipers, J. B. (1999). Editor. *Quaternions and Rotation Sequences.* Princeton University Press.
Lamoen, D. & Michel, K. H. (1999). *Phase Transit.* **67**, 789–807.
Leusen, F. J. J. (2003). *Cryst. Growth Des.* **3**, 189–192.
Löding, D., Müser, M. H. & Nielaba, P. (1997). *Z. Phys. B*, **102**, 505–511.

Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst.* B**56**, 697–714.
Lynden-Bell, R. M. & Michel, K. H. (1994). *Rev. Mod. Phys.* **66**, 721–762.
Lynden-Bell, R. M. & Stone, A. J. (1989). *Mol. Simulat.* **3**, 271–281.
McQuarrie, D. A. (2000). *Statistical Mechanics.* Sausalito, CA: University Science Books.
Maradudin, A. A. & Vosko, S. H. (1968). *Rev. Mod. Phys.* **40**, 1–37.
Maslen, D. K. & Rockmore, D. N. (1997). In *DIMACS Series in Discrete Mathematics and Theoretical Compuing Science*, pp. 183–237. Providence, RI: American Mathematical Society.
Michel, K. H. & Copley, J. R. D. (1997). *Z. Phys. B*, **103**, 369–376.
Michel, K. H., Copley, J. R. D. & Neumann, D. A. (1992). *Phys. Rev. Lett.* **68**, 2929–2932.
Miller, J. S., Epstein, A. J. & Reiff, W. M. (1988). *Acc. Chem. Res.* **21**, 114–120.
Miller, S. C. & Love, W. F. (1967). *Tables of Irreducible Representations of Space Groups and Co-Representations of Magnetic Space Groups.* Boulder, CO: Pruett.
Missaghi, M. N., Mettes, J. A. & McClurg, R. B. (2004). Submitted.
Mooij, W. T. M., van Eijck, B. P., Price, S. L., Verwer, P. & Kroon, J. (1998). *J. Comput. Chem.* **19**, 459–474.
Motherwell, W. D. S. (1997). *Acta Cryst.* B**53**, 726–736.
Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Dzyabchenko, A., Erk, P., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Lommerse, J. P. M., Mooij, W. T. M., Price, S. L., Scheraga, H., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2002). *Acta Cryst.* B**58**, 647–661.
Moulton, B. & Zaworotko, M. J. (2001). *Chem. Rev.* **101**, 1629–1658.
Nagamiya, T. (1951). *Prog. Theor. Phys. (Kyoto)*, **6**, 702–713.
Neumann, M. A., Press, W., Noldeke, C., Asmussen, B., Prager, M. & Ibberson, R. M. (2003). *J. Chem. Phys.* **119**, 1586–1589.
Oxtoby, D. W. (2002). *Annu. Rev. Mater. Res.* **32**, 39–52.
Pillardy, J., Arnautova, Y. A., Czaplewski, C., Gibson, K. D. & Scheraga, H. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 12351–12356.
Press, W. (1972). *J. Chem. Phys.* **56**, 2597–2609.
Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1992). *Numerical Recipes in FORTRAN: the Art of Scientific Computing.* New York: Cambridge University Press.
Price, S. L. (1996). *J. Chem. Soc. Faraday Trans.* **92**, 2997–3008.
Schmidt, M. U. & Englert, U. (1996). *J. Chem. Soc. Dalton Trans.* pp. 2077–2082.
Sherwood, J. N. (1979). *The Plastically Crystalline State.* New York: Wiley.
Stokes, H. T. & Hatch, D. M. (1988). *Isotropy Subgroups of the 230 Crystallographic Space Groups.* Teaneck, NJ: World Scientific.
Stokes, H. T. & Hatch, D. M. (2002*a*). *ISOTROPY*. http://stokes.byu.edu/isotropy.html.
Stokes, H. T. & Hatch, D. M. (2002*b*). *Phys. Rev. B*, **65**, 144114-1–144114-12.
Stokes, H. T., Hatch, D. M. & Wells, J. D. (1991). *Phys. Rev. B*, **43**, 11010–11018.
Stone, A. J. (1996). *The Theory of Intermolecular Forces.* New York: Clarendon Press.
Stone, A. J. & Tough, R. J. A. (1984). *Chem. Phys. Lett.* **110**, 123–129.
Stroud, A. H. & Secrest, D. (1966). *Gaussian Quadrature Formulas.* Englewood Cliffs, NJ: Prentice-Hall.
Tajima, N., Tanaka, T., Arikawa, T., Sukarai, T., Teramae, S. & Hirano, T. (1995). *Bull. Chem. Soc. Jpn*, **68**, 519–527.
Timmermans, P. J. (1961). *J. Phys. Chem. Solids*, **18**, 1–8.
Varshalovich, D. A., Moskalev, A. N. & Khersonskii, V. K. (1988). *Quantum Theory of Angular Momentum.* Teaneck, NJ: World Scientific.
Verwer, P. & Leusen, F. J. J. (1998). In *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz & D. B. Boyd, Vol. 12, pp. 327–366. New York: Wiley-VCH.
Vilenkin, N. I. & Klimyk, A. U. (1991). *Representations of Lie Groups and Special Functions.* Boston: Kluwer Academic Publishers.
Williams, D. E. (1966). *Acta Cryst.* **21**, 340–349.
Williams, D. E. (1996). *Acta Cryst.* A**52**, 326–328.
Wilson, S. (2003). Editor. *Handbook of Molecular Physics and Quantum Chemistry*, Vol. 3. Chichester: Wiley.
Wudl, F. (1984). *Acc. Chem. Res.* **17**, 227–232.
Yamamoto, T., Kataoka, Y. & Okada, K. (1977). *J. Chem. Phys.* **66**, 2701–2730.
Zaworotko, M. J. (1994). *Chem. Soc. Rev.* pp. 283–288.