Free-energy derivatives and structure optimization within quasiharmonic lattice dynamics

M. B. Taylor, G. D. Barrera,* N. L. Allan,† and T. H. K. Barron
School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom
(Received 30 July 1997)

A method is presented for the calculation of the gradient of the free energy with respect to all the internal and external degrees of freedom of a periodic crystal. This gradient can be used in conjunction with a static-energy Hessian for efficient geometrical optimization of systems with large unit cells. The free energy is calculated using lattice statics and lattice dynamics in the quasiharmonic approximation, and its derivatives by means of first-order perturbation theory. In the present application of the method, particles are assumed to interact via arbitrary short-ranged spherically-symmetric pair potentials and long-ranged Coulomb forces, and polarizability effects are accounted for by use of the shell model. The method can be used directly as the basis for a computer program which makes efficient use of both storage and CPU time, especially for large unit cells. Detailed expressions for all the lattice sums are presented. [S0163-1829(97)01246-0]

I. INTRODUCTION

Theoretical calculation of the equilibrium thermodynamic properties of crystals, over a wide range of temperatures and pressures, has been approached in a number of ways. The most popular methods in recent times have been molecular dynamics (MD) and, to a lesser extent, classical Monte Carlo (MC) simulations. These approaches rely on generating a set of system states representative of the equilibrium configuration and forming an average over this set. The computationally expensive part is then generating new configurations, both to approach an equilibrium state from an initial configuration, and in sampling the configuration space near the equilibrium, more samples leading to greater accuracy. An alternative approach is lattice dynamics, in which the free energy of a given configuration, as well as dependent properties such as entropy and heat capacity, can be calculated directly.

In this paper we use lattice dynamics in the quasiharmonic approximation, which gives the equation of state to the first order in the anharmonicity of the potentials. The free energy for a given crystal geometry is taken to be that resulting from the harmonic approximation,[1] but the anharmonicity of the crystal potential causes the mode frequencies so obtained to be functions of the geometrical parameters, making the vibrational free energy a function of these parameters as well as of temperature. We take into account ionic polarization by using the shell model originally due to Dick and Overhauser,[2] in which each ion may consist of a massless ''shell'' and a massive core, the charge being distributed between the two and thereby, if the two are displaced relative to each other, giving rise to a dipole. The shell and core of a given ion are coupled typically by a springlike interaction.

Lattice dynamics has been somewhat neglected in recent years, possibly due to overcaution regarding the range of validity of the quasiharmonic approximation; however it can be remarkably robust even at elevated temperatures (examples of its successful use are Refs. 3–5; see also Ref. 6, and references therein) and has a number of characteristics that make it a powerful alternative or complementary technique to MD and MC. In particular, it takes into account zero-point energy and other quantum effects (although it can easily be adapted to generate results in the classical limit), and hence scores significantly over the inherently classical MD and MC methods below the Debye temperature, where quantum effects cannot be ignored. The investigation of normal vibrations aids interpretation of the results, and can reveal explicitly the mechanisms operating in thermal expansion and phase transitions; by identifying unstable vibrational modes it also provides a very sensitive test for interatomic potentials. Moreover, lattice dynamics does not rely on long runs for high precision, and is in fact relatively inexpensive, demanding typically an order of magnitude less computation time than MD or MC, and moreover avoiding the kinetic barriers and critical slowing down effects suffered by those techniques.[7]

Except at rather low temperatures, when using lattice dynamics the bulk of the computational effort is expended in the optimization problem of determining the equilibrium geometry of the crystal; given this, calculation of dependent properties is relatively straightforward. Efficient optimization is only possible if derivatives of the appropriate thermodynamic potential with respect to the geometrical coordinates (or strains) are available, and it is the generation and use of these derivatives which are addressed in this paper.

Kantorovich[8] has previously presented results for first and second free-energy derivatives within the model of deformable dipoles due to Tolpygo;[9,10] our choice of the related shell model[2] has been made because of the large number of shell parameters and interionic potentials available for a wide range of polar solids. If the elastic constants are required then the free-energy Hessian (second strain derivatives) as calculated directly by Kantorovich is necessary; however, considerable additional processing time and storage is required for the second derivatives, and there are also some unresolved problems regarding degeneracies in the associated second-order perturbation theory. As explained in Sec. III therefore we consider it efficient for the purposes of structure optimization to generate only the first derivatives.

The paper is organized as follows. Section II introduces our notation for the geometry of a crystal as defined in terms of a specific set of coordinates with respect to which free-energy derivatives can be taken. Section III discusses the

 © 1997 The American Physical Society

derivatives required for structural optimization of a crystal and how they may most efficiently be utilized. Section IV reviews the definition of the quasiharmonic dynamical matrix and its use in generating the appropriate thermodynamic potentials, and first-order perturbation theory is used in Sec. V to give expressions for their various strain derivatives. In Sec. VI detailed expressions for the required lattice sums for Coulombic and arbitrary short-ranged spherically symmetric pairwise interactions are presented, and Sec. VII concludes.

II. CRYSTAL STRUCTURE AND STRAIN COORDINATES

A crystal structure is periodic, defined by its lattice and the positions of all the atoms within a unit cell. In general, strain can alter both the shape of the lattice, which is specified by *external* coordinates $\mathcal{E}_l^{\rm ext}$, and the atomic positions, which are specified by *internal* coordinates $\mathcal{E}_k^{\rm int}$. The general method of optimization described in Sec. III requires partial derivatives of the free energy with respect to whatever coordinates are chosen to describe the state of strain. In the present application these are obtained by first deriving derivatives with respect to the strain rotation tensor of macroscopic elasticity theory and a related set of internal coordinates defined as follows.

The positions of all the particles (which in a shell model may be cores or shells) in a macroscopically strained crystal are given by Cartesian coordinates

$$r_{i\mathbf{x}}^{\alpha}=\sum_{\gamma}\ (\delta_{\alpha\gamma}+e^{\alpha\gamma})(x^{\gamma}+\rho_i^{\gamma}), \tag{1}$$

where Greek superscripts are Cartesian indices $1\ldots3$ (as elsewhere in this paper), $\mathbf{x}$ is a lattice vector of the unstrained crystal labeling a unit cell, $i=1\ldots n$ is an index that labels a specific particle within a unit cell, and $\delta_{\alpha\beta}$ is the Kronecker delta. Components of the tensor $e^{\alpha\beta}$ determine both the orientation and the macroscopic state of strain of the crystal.[11] The vector components ρ_i^{α} are internal coordinates determining the positions of particles within a unit cell.

The parameters $e^{\alpha\beta}$ and ρ_i^{α} have been chosen because it is convenient to obtain derivatives of the free energy with respect to them, and because they may be simply related to other parameters that are in common use. The Voigt macroscopic infinitesimal strain coordinates ϵ_λ are the components of the six-element vector

$$\boldsymbol{\epsilon}=\begin{pmatrix} e^{11}\\ e^{22}\\ e^{33}\\ e^{23}+e^{32}\\ e^{13}+e^{31}\\ e^{12}+e^{21}\end{pmatrix}. \tag{2}$$

Dimensionless ''basis internal coordinates'' u_i^t can be defined by

$$\rho_i^{\alpha}=\sum_t\ u_i^t a_t^{\alpha}=\sum_t\ u_i^t A^{t\alpha}, \tag{3}$$

where the matrix $A^{t\alpha}$ specifies a metric obtained from components of the lattice vectors $\mathbf{a}_t$ determining the unit cell in the unstrained lattice; $t=1\ldots3$ labels directions in unit-cell space.

In many applications symmetry reduces the number of independent coordinates. For example, when optimizing the geometry of a crystal it may be useful to characterize the state of external strain in terms of the crystallographic lattice parameters c_a, e.g., a and c for tetragonal or hexagonal symmetry. In the same way not all the u_i^t can be varied independently if the symmetry of the crystal is to be maintained. It is normally possible to define a set of ''symmetric internal coordinates'' w_m, which allows a description of the structure with a smaller number of variables. The w_m can be defined such that

$$u_i^t=g_i^t+\sum_m\ w_m g_{m,i}^t, \tag{4}$$

where g_i^t is constant for a given particle i, and the directions of the $g_{m,i}^t$ are determined by symmetry. Consider, for example, a rigid-ion model of rutile, TiO_2. The tetragonal unit cell contains six atoms, of which the two titanium positions are fixed by symmetry. The oxygen atoms lie at $[(0,0,0)\pm w_1(1,1,0)]$ and $[(0.5,0.5,0.5)\pm w_1(-1,1,0)]$. Thus for optimization only three parameters, a, c, and w_1 are needed. On the other hand, general elastic distortion destroys all symmetry except inversion; and so deriving elastic coefficients requires either the full set of the ρ_i^{α} or u_i^t, or else possibly a more extended set of symmetric coordinates w_m.

In general the geometry of a crystal can be characterized by, and the free energy differentiated with respect to, an $N_{\mathcal{E}}$-element vector of generalized coordinates $\mathcal{E}_A$, always including some external coordinates $\mathcal{E}_l^{\rm ext}$ and sometimes also some internal ones $\mathcal{E}_k^{\rm int}$.[12] In the present application the $\mathcal{E}_k^{\rm int}$ may be the ρ_i^{α}, u_i^t, or w_m, and the $\mathcal{E}_l^{\rm ext}$ may be the ϵ_λ or c_a.

III. STRUCTURE OPTIMIZATION

A. Derivatives required for structure optimization

In order to optimize a crystal structure, i.e., to find the most stable state under a given set of thermodynamic constraints, it is necessary to minimize the appropriate thermodynamic potential with respect to a set of $N_{\mathcal{E}}$ structure parameters $\mathcal{E}_A$. At finite temperature T under an applied hydrostatic pressure P_0 the appropriate potential is an availability $\widetilde{G}$, defined

$$\widetilde{G}=U-TS+P_0V=F+P_0V=\Phi_{\rm stat}+F_{\rm vib}+P_0V, \tag{5}$$

where U is the internal energy, S is the entropy, V is the volume, and F is the Helmholtz free energy which has static and vibrational parts $\Phi_{\rm stat}$ and $F_{\rm vib}$, respectively; all these quantities refer to the value per unit cell. Note this $\widetilde{G}$ differs from the availability $A=U-T_0S+P_0V$ defined by Pippard,[13] because here we optimize with respect only to varying strain, and not also to varying temperature.

The optimization condition is therefore to find that vector in parameter space at which, for all A,

$$\left(\frac{\partial \widetilde{G}}{\partial \mathcal{E}_A}\right)_{\mathcal{E}'} = 0, \tag{6}$$

where the notation $()_{\mathcal{E}'}$ indicates that all coordinates $\mathcal{E}_B$ for $B \neq A$ are held fixed. An appropriate set of coordinates $\mathcal{E}_A$ will be chosen as discussed in Sec. II.

To perform this minimization, it is effectively necessary to calculate at least the first, and possibly the second, derivatives of $\widetilde{G}$ with respect to the $\mathcal{E}_A$. All the derivatives can of course be calculated numerically using a finite difference formula. Care must be taken however in choosing the finite difference, since if it is too large then local nonlinearity of the function will result in a poor approximation, but if it is too small then precision may be lost. Furthermore if there are discontinuities in $\widetilde{G}(\mathcal{E})$, which may arise from sharp cutoffs often present in short-ranged interactions, the distortions introduced by the finite changes in the coordinates can lead to large inaccuracies in the numerical derivatives. These difficulties can usually be overcome, but the principal problem with numerical evaluation of derivatives is that it becomes very expensive for large $N_{\mathcal{E}}$.

Hitherto, because of the expense of calculating the derivatives numerically, only approximations to a full dynamic relaxation Eq. (6) have been made, as follows. The thermodynamic potential $\widetilde{G}$ can be defined as a sum of static and vibrational parts

$$\widetilde{G} = \widetilde{G}_{\text{stat}} + \widetilde{G}_{\text{vib}} \tag{7}$$

and the geometrical coordinates split into external ones $\mathcal{E}_l^{\text{ext}}$ and internal ones $\mathcal{E}_k^{\text{int}}$ as explained in Sec. II. The condition for full dynamical relaxation is

$$\left(\frac{\partial \widetilde{G}}{\partial \mathcal{E}_l^{\text{ext}}}\right)_{\mathcal{E}'} = \left(\frac{\partial \widetilde{G}}{\partial \mathcal{E}_k^{\text{int}}}\right)_{\mathcal{E}'} = 0. \tag{8}$$

However, although the derivatives of $\widetilde{G}_{\text{stat}}$ are quite straightforward to calculate analytically,[14] the theory for analytic calculation of the $\widetilde{G}_{\text{vib}}$ derivatives within the shell model has not previously been available. Accordingly, for large unit cells the zero static internal stress approximation (ZSISA) has often been used,[6] in which the equilibrium condition is taken to be

$$\left(\frac{\partial \widetilde{G}}{\partial \mathcal{E}_l^{\text{ext}}}\right)_{\mathcal{E}'} = \left(\frac{\partial \widetilde{G}_{\text{stat}}}{\partial \mathcal{E}_k^{\text{int}}}\right)_{\mathcal{E}'} = 0. \tag{9}$$

Since the number of the $\mathcal{E}_l^{\text{ext}}$ derivatives never exceeds 6, and is often 1 or 2, they can be calculated numerically without excessive expense (although the other considerations mentioned above may still present problems). An even more extreme, and less computationally expensive, approximation is the constant internal strain parameter approximation (CISPA), in which a set of internal coordinates is found by a full minimization of the static energy

$$\left(\frac{\partial \widetilde{G}_{\text{stat}}}{\partial \mathcal{E}_l^{\text{ext}}}\right)_{\mathcal{E}'} = \left(\frac{\partial \widetilde{G}_{\text{stat}}}{\partial \mathcal{E}_k^{\text{int}}}\right)_{\mathcal{E}'} = 0 \tag{10}$$

and subsequent dynamical optimizations with respect to external coordinates only are then performed while holding the internal ones fixed. With CISPA the results will be dependent on the set of internal coordinates chosen.

ZSISA can be shown[15] to yield the optimized external coordinates correctly to first order, but this is not so for the internal coordinates and hence for dependent thermodynamic quantities. Such quantities *may* be well approximated by ZSISA and CISPA, but it is hard to assess the errors thus introduced without a full dynamical free-energy optimization.

B. Optimization methods

Methods of unconstrained minimization of a function (e.g., availability $\widetilde{G}$) of many variables (e.g., a column vector of structure coordinates $\boldsymbol{\mathcal{E}}$) may be classified into two main groups:[16,17] those which require the gradient (first derivatives $\mathbf{y} = \boldsymbol{\nabla} \widetilde{G}$) and the Hessian (second derivatives $\mathbf{H} = \boldsymbol{\nabla} \mathbf{y}$), and those which require only the gradient.

If the Hessian is available then the Newton method is most straightforward, and usually very successful: a quadratic approximation to $\widetilde{G}$ in the region of its minimum is written

$$\widetilde{G}(\boldsymbol{\mathcal{E}} + \boldsymbol{\delta}) \approx \widetilde{G}(\boldsymbol{\mathcal{E}}) + \boldsymbol{\delta}^T \cdot \mathbf{y}(\boldsymbol{\mathcal{E}}) + \frac{1}{2} \boldsymbol{\delta}^T \cdot \mathbf{H}(\boldsymbol{\mathcal{E}}) \cdot \boldsymbol{\delta}, \tag{11}$$

where $\boldsymbol{\delta}^T$ is the transpose of $\boldsymbol{\delta}$, and $\delta = \boldsymbol{\mathcal{E}}' - \boldsymbol{\mathcal{E}}$, the step from $\boldsymbol{\mathcal{E}}$ to an improved value $\boldsymbol{\mathcal{E}}'$. This can then be differentiated with respect to $\boldsymbol{\delta}$ to give

$$\boldsymbol{\delta} = -\mathbf{H}(\boldsymbol{\mathcal{E}})^{-1} \cdot \mathbf{y}(\boldsymbol{\mathcal{E}}). \tag{12}$$

If $\widetilde{G}$ is quadratic then solving for $\boldsymbol{\mathcal{E}}'$ will give its minimizing value; otherwise $\boldsymbol{\mathcal{E}}'$ should be an improved estimate and repeated iterations of Eq. (12) should approach the minimum. Some choices of sets of variables can result in a singular $\mathbf{H}$ and therefore failure of Eq. (12); in particular the Hessian with respect to all $3n$ basis coordinates u_i^α or ρ_i^α will be singular because of the translational invariance of the lattice. This problem can be avoided by optimizing with respect to the coordinates of $n-1$ rather than n particles, equivalent to fixing the position of one particle.

If only the gradient is available then a quasi-Newton method can be used, the essence of which is to maintain an approximation $\mathbf{J}$ to the inverse Hessian $\mathbf{H}^{-1}$ and at each step to use a gradient evaluation to produce an improved estimate $\mathbf{J}'$. Several prescriptions exist for this update, but the most efficient is generally considered to be the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update[17]

$$\mathbf{J}' = \mathbf{J} + \frac{\boldsymbol{\delta} \cdot \boldsymbol{\delta}^T}{\boldsymbol{\delta}^T \cdot \boldsymbol{\gamma}} - \frac{(\mathbf{J} \cdot \boldsymbol{\gamma}) \cdot (\mathbf{J} \cdot \boldsymbol{\gamma})^T}{\boldsymbol{\gamma}^T \cdot \mathbf{J} \cdot \boldsymbol{\gamma}} + (\boldsymbol{\gamma}^T \cdot \mathbf{J} \cdot \boldsymbol{\gamma}) \mathbf{s} \cdot \mathbf{s}^T, \tag{13}$$

where

$$\boldsymbol{\gamma} = \mathbf{y}(\boldsymbol{\mathcal{E}} + \boldsymbol{\delta}) - \mathbf{y}(\boldsymbol{\mathcal{E}}), \tag{14}$$

$$\mathbf{s} = \frac{\boldsymbol{\delta}}{\boldsymbol{\delta}^T \cdot \boldsymbol{\gamma}} - \frac{\mathbf{J} \cdot \boldsymbol{\gamma}}{\boldsymbol{\gamma}^T \cdot \mathbf{J} \cdot \boldsymbol{\gamma}}. \tag{15}$$

y and **J** are then used as per Eq. (12) to update $\boldsymbol{\mathcal{E}}$. This new $\boldsymbol{\mathcal{E}}$ may then be refined by performing a linear search (using repeated evaluations of $\widetilde{G}$ and **y**) for the function minimum in the direction indicated; if **J** is a poor approximation to $\mathbf{H}^{-1}$ the linear search will be indispensable, but if they are close it may be omitted.

Our experience has been that if the quasi-Newton method is started reasonably close to the optimal structure, with a reasonable approximation to the Hessian, it provides very satisfactory optimization without the necessity for linear searches. Using the static part $\partial^2\widetilde{G}_{\rm stat}/\partial\mathcal{E}_A\partial\mathcal{E}_B$ of the Hessian as the initial approximation to the full dynamic value turns out to work very well in the cases we have tried, since the dynamical contributions to this are usually quite small. For the initial configuration an optimized structure at a nearby temperature, or if that is unavailable the statically optimized structure, is usually adequate. If the best starting configuration is far from its optimized state however, or the dynamic contributions to the Hessian are large, it is sometimes necessary to use another method, such as a conjugate gradient method, or a quasi-Newton method including linear minimizations, to approach the free-energy minimum.

Our usual algorithm for structural optimization is therefore:

(1) Set $\boldsymbol{\mathcal{E}}$ to optimized values at nearby temperature (or static),

(2) Calculate static Hessian $\mathbf{H}_{\rm stat}(\boldsymbol{\mathcal{E}})$,

(3) Set approximate inverse Hessian $\mathbf{J}=\mathbf{H}_{\rm stat}^{-1}$,

(4) Calculate dynamic gradient $\mathbf{y}(\boldsymbol{\mathcal{E}})$,

(5) Stop if gradient magnitude $(\mathbf{y}^T\cdot\mathbf{y})^{1/2}$ is small,

(6) Update **J** using BFGS formula (13) with $\boldsymbol{\delta}=-\mathbf{J}\cdot\mathbf{y}$ (except the first time round),

(7) Update $\boldsymbol{\mathcal{E}}$ using $\boldsymbol{\mathcal{E}}'=\boldsymbol{\mathcal{E}}-\mathbf{J}\cdot\mathbf{y}$,

(8) Return to step 4.

Progress of the optimization of course varies according to the details of the problem under consideration, but typically the magnitude of the gradient decreases by about an order of magnitude per iteration, and an optimization completes to high accuracy in fewer than 10 iterations.

An optimization therefore requires one static Hessian calculation, and a few dynamic gradient calculations. In the subsequent sections we present a unified approach to calculating the static Hessian and the dynamic gradient, as well as the free energy itself. The method outlined here is attractive because these quantities can be calculated analytically, and relatively rapidly. In particular, except for very small $N_{\mathcal{E}}$, this method is a great deal faster than those which require repeated exact calculation of the dynamic Hessian, or numerical evaluation of the derivatives, or frequent linear minimizations.

We conclude this section by stressing that there are many methods for multivariate optimization not discussed here, as well as several subtleties and tricks which are sometimes required to ensure rapid convergence, depending on the details of the problem under consideration; the interested reader is referred to the discussion in Ref. 16. Our approach has been pragmatic, and is based on the observation that the method outlined here is seen to be highly effective, rather than on formal proofs of efficacy.

IV. FREE-ENERGY EVALUATION

Calculation of the static energy $\Phi_{\rm stat}$ is straightforward, and an expression for it under the assumption of spherically symmetric pairwise potentials appears in Sec. VI B. We give a brief outline in this section of the method for calculating the vibrational contribution $F_{\rm vib}$ to $\widetilde{G}$ within the quasiharmonic approximation and using the shell model.

Following standard lattice-dynamical practice[1,10] we obtain the squares of the vibrational frequencies $\omega_{\mathbf{q}s}^2$ as eigenvalues of the dynamical matrix $D_{\kappa\kappa'}^{\alpha\beta}(\mathbf{q})$, where the indices κ, κ' refer to atoms within the unit cell with masses m_κ, $m_{\kappa'}$. Taking plane-wave solutions to the equations of motion leads to the following form for the atomic displacement $u_{\kappa\mathbf{x}}^{\alpha}$ of an atom κ in the unit cell displaced **x** from the origin:

$$u_{\kappa\mathbf{x}}^{\alpha}=\frac{1}{\sqrt{m_\kappa}}\sum_{\mathbf{q}s} A_{\mathbf{q}s}f_{\mathbf{q}s\kappa}^{\alpha}e^{i(\mathbf{q}\cdot\mathbf{x}+\omega_{\mathbf{q}s}t)}, \tag{16}$$

where $f_{\mathbf{q}s\kappa}^{\alpha}$ are the normalized eigenvectors of $D_{\kappa\kappa'}^{\alpha\beta}(\mathbf{q})$ and $A_{\mathbf{q}s}$ are the corresponding vibrational amplitudes. To obtain $D_{\kappa\kappa'}^{\alpha\beta}(\mathbf{q})$ we first define the matrices

$$B_{ij}^{\alpha\beta}(\mathbf{q})=\sum_{\mathbf{x}} e^{i\mathbf{q}\cdot\mathbf{x}}\frac{\partial^2\Psi}{\partial u_{i\mathbf{0}}^{\alpha}\partial u_{j\mathbf{x}}^{\beta}}, \tag{17}$$

where Ψ is the static energy of the whole crystal, the indices i, j refer to all cores and shells within the unit cell, and the summation is over all lattice vectors, $\mathbf{x}=l_1\mathbf{a}_1+l_2\mathbf{a}_2+l_3\mathbf{a}_3$ for all integer l_1, l_2, l_3. The B matrix is therefore composed of the submatrices B_{CC}, B_{CS}, B_{SC}, and B_{SS} where, for example, in $[B_{CS}]_{ij}^{\alpha\beta}$ i ranges only over the cores and j only over the shells. The dynamical matrix is then given by[10,18]

$$D_{\kappa\kappa'}^{\alpha\beta}(\mathbf{q})=\frac{1}{\sqrt{m_\kappa m_{\kappa'}}}[B_{CC}-B_{CS}B_{SS}^{-1}B_{SC}]_{\kappa\kappa'}^{\alpha\beta}(\mathbf{q}). \tag{18}$$

Thus for models in which there are no shells, the particle indices i,j can be identified with the atom indices κ,κ' to give simply

$$D_{\kappa\kappa'}^{\alpha\beta}(\mathbf{q})=\frac{1}{\sqrt{m_\kappa m_{\kappa'}}}B_{\kappa\kappa'}^{\alpha\beta}(\mathbf{q}). \tag{19}$$

Given the mode frequencies $\omega_{\mathbf{q}s}$ the vibrational contributions to the quasiharmonic Helmholtz free energy can be written

$$F_{\rm vib}=\sum_{\mathbf{q}s} F_{\mathbf{q}s}, \tag{20}$$

where

$$F_{\mathbf{q}s}=\frac{1}{2}\hbar\omega_{\mathbf{q}s}+k_BT\ \ln(1-e^{-\beta\hbar\omega_{\mathbf{q}s}}), \tag{21}$$

and expressions for other thermodynamic quantities such as entropy, internal energy, and constant strain heat capacity can similarly be derived.[6,10] The **q** summation is over all

wave vectors in the first Brillouin zone, and may be approximated using a set of vectors chosen using one of the standard schemes.[19,20]

V. STRAIN DERIVATIVES OF THE FREE ENERGY

As discussed in Sec. III A the availability $\widetilde{G}$ is the quantity to be differentiated. From Eq. (5) it follows for a general coordinate $\mathcal{E}$

$$\frac{\partial \widetilde{G}}{\partial \mathcal{E}} = \frac{\partial \Phi_{\rm stat}}{\partial \mathcal{E}} + \frac{\partial F_{\rm vib}}{\partial \mathcal{E}} + P_0 \frac{\partial V}{\partial \mathcal{E}}. \tag{22}$$

The dimensions of the unit cell depend only on the external coordinates, so that for any internal coordinate $\mathcal{E}^{\rm int}$

$$\frac{\partial V}{\partial \mathcal{E}^{\rm int}} = 0, \tag{23}$$

and it can be shown that, using the notation of Eq. (2)

$$\left(\frac{\partial V}{\partial e^{\alpha\beta}} \right)_0 = V \delta_{\alpha\beta}, \tag{24}$$

where the $(\,)_0$ notation indicates that the derivatives are evaluated at zero external strain, i.e., when all $e^{\mu\nu}=0$. We therefore have

$$\frac{\partial \widetilde{G}}{\partial \mathcal{E}^{\rm int}} = \frac{\partial \Phi_{\rm stat}}{\partial \mathcal{E}^{\rm int}} + \frac{\partial F_{\rm vib}}{\partial \mathcal{E}^{\rm int}}, \tag{25}$$

$$\left(\frac{\partial \widetilde{G}}{\partial \epsilon_\lambda} \right)_0 = \left(\frac{\partial \Phi_{\rm stat}}{\partial \epsilon_\lambda} \right)_0 + \left(\frac{\partial F_{\rm vib}}{\partial \epsilon_\lambda} \right)_0 - V \sigma_\lambda, \tag{26}$$

where $\boldsymbol{\sigma}$ is the external stress vector, which for hydrostatic pressure P_0 has the value

$$\boldsymbol{\sigma} = (-P_0, -P_0, -P_0, 0, 0, 0)^T. \tag{27}$$

If anisotropic external stress is to be considered these expressions will remain the same, but the form of $\boldsymbol{\sigma}$ will be different.

Using Eqs. (25) and (26) therefore we can generate the derivatives of $\widetilde{G}$ from derivatives of $\Phi_{\rm stat}$ and of $F_{\rm vib}$. The first and second $\Phi_{\rm stat}$ derivatives are obtained by differentiation of the appropriate lattice sum with respect to the strain coordinates in question, and expressions for these under the assumption of pairwise potentials appear in Sec. VI. The general theory of differentiating $F_{\rm vib}$ is given in Sec. V A and conversion between different geometrical derivatives is discussed in Sec. V B.

A. Vibrational free-energy derivatives

Since there is no explicit expression for $F_{\rm vib}$ in terms of the geometry of the lattice, it is not possible to determine its strain derivatives by direct differentiation of a lattice sum as with $\Phi_{\rm stat}$. Instead from Eqs. (20) and (21), $F_{\rm vib}$ is known in terms of the eigenvalues of the dynamical matrix, so the eigenvalue strain derivatives are required. Specifically, differentiating Eq. (21) with respect to an arbitrary strain $\mathcal{E}$ gives

$$F'_{\mathbf{q}s} \equiv \frac{\partial F_{\mathbf{q}s}}{\partial \mathcal{E}} = \frac{\hbar}{2\omega_{\mathbf{q}s}} \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega_{\mathbf{q}s}} - 1} \right) \frac{\partial \omega^2_{\mathbf{q}s}}{\partial \mathcal{E}}, \tag{28}$$

this can alternatively be written in terms of the mode-Grüneisen function $\gamma_{\mathbf{q}s} = -\partial \ln \omega_{\mathbf{q}s} / \partial \mathcal{E}$ and mode contribution to the internal energy $U_{\mathbf{q}s}$,

$$F'_{\mathbf{q}s} = -\gamma_{\mathbf{q}s} U_{\mathbf{q}s}. \tag{29}$$

Note that the wave vector $\mathbf{q}$, like $\mathbf{x}$, is a label defined to remain invariant under the strain $\partial\mathcal{E}$. Writing for convenience

$$\lambda_s = \omega^2_{\mathbf{q}s}, \tag{30}$$

and using the notation $X' = \partial X / \partial \mathcal{E}$, then if all eigenvalues are distinct, standard first-order perturbation theory[21] applied to the dynamical matrix gives the following result for the eigenvalue derivatives:

$$\lambda'_s = \Delta_{ss}, \tag{31}$$

where

$$\Delta_{uv} = \sum_{\kappa\kappa'\alpha\beta} f^{\alpha}_{\mathbf{q}u\kappa} {}^{*} [D^{\alpha\beta}_{\kappa\kappa'}(\mathbf{q})]' f^{\beta}_{\mathbf{q}v\kappa'}. \tag{32}$$

The derivatives of $D^{\alpha\beta}_{\kappa\kappa'}(\mathbf{q})$ can be written in terms of the derivatives of $B^{\alpha\beta}_{ij}(\mathbf{q})$. We can therefore differentiate Eq. (19) for rigid-ion models, giving

$$(D^{\alpha\beta}_{\kappa\kappa'})' = \frac{1}{\sqrt{m_\kappa m_{\kappa'}}} (B^{\alpha\beta}_{\kappa\kappa'})' \tag{33}$$

and Eq. (18) for shell models, giving

$$\begin{aligned} (D^{\alpha\beta}_{\kappa\kappa'})' = & \frac{1}{\sqrt{m_\kappa m_{\kappa'}}} [B'_{CC} - B'_{CS} B^{-1}_{SS} B_{SC} - B_{CS} (B^{-1}_{SS})' B_{SC} \\ & - B_{CS} B^{-1}_{SS} B'_{SC}]^{\alpha\beta}_{\kappa\kappa'} \\ = & \frac{1}{\sqrt{m_\kappa m_{\kappa'}}} [B'_{CC} - B'_{CS} B^{-1}_{SS} B_{SC} \\ & + B_{CS} B^{-1}_{SS} B'_{SS} B^{-1}_{SS} B_{SC} - B_{CS} B^{-1}_{SS} B'_{SC}]^{\alpha\beta}_{\kappa\kappa'}. \end{aligned} \tag{34}$$

Here we have used the matrix equivalence $(M^{-1})' = -M^{-1} M' M^{-1}$, which can readily be obtained by differentiation of $(M^{-1}M)$.

Note that since the theory is being used to construct a derivative, the perturbation of the matrix is infinitesimal, so that these expressions are not an approximation but are exact.

Special consideration must be given if some of the vibrational states are degenerate. To discuss this case we relabel state s as state (r,t), where r labels a particular energy and $t = 1 \ldots n_r$ labels states with that energy. In order for Eq. (31) to be correct, the eigenvectors of the infinitesimally perturbed state must differ only infinitesimally from the eigenvectors of the unperturbed state. If several states share the same eigenvalue (i.e., $n_r > 1$) then diagonalizing the dynamical matrix does not give a unique set of eigenvectors, but rather some arbitrary linear combination of the eigenvectors

with the same r. If a perturbation lifts this degeneracy then the eigenvectors for the perturbed states will be uniquely determined, and unless care is taken to choose the eigenvectors of the degenerate states to be limiting cases of those of the perturbed states then discontinuities will result and the value for $\lambda'_{(r,t)}$ will be incorrect.

This correct choice can be ensured by diagonalizing $\Delta_{(r)(r)}$, being the $n_r \times n_r$ submatrix of $\Delta_{(r,t)(r,t)}$ spanning the degenerate subspace of states all labeled r. This having been done, the result of Eq. (31)

$$\lambda'_{(r,t)} = \Delta_{(r,t)(r,t)} \quad [\text{if } \Delta_{(r)(r)} \text{ is diagonalized}] \tag{35}$$

is recovered. With the Δ_{ss} formed by thus diagonalizing the degenerate submatrices therefore, all the quantities dependent on the frequency strain derivatives $\omega'_{\mathbf{q}s}$ can be obtained.

However, often only quantities like $F'_{\mathbf{q}} = \Sigma_s F'_{\mathbf{q}s}$ which are summed over all modes s, and not the mode contributions like $F'_{\mathbf{q}s}$, are required. In this case, summing the mode contributions (28) and labeling the degenerate states explicitly we have

$$F'_{\mathbf{q}} = \sum_{rt} F'_{\mathbf{q}(r,t)} = \sum_{rt} \frac{\hbar}{2\omega_{\mathbf{q}(r,t)}} \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega_{\mathbf{q}(r,t)}} - 1} \right) \lambda'_{\mathbf{q}(r,t)}, \tag{36}$$

then since by the definition of degeneracy $\omega_{(r,u)} \equiv \omega_{(r,v)} \equiv \omega_{(r)}$, this can be written

$$F'_{\mathbf{q}} = \sum_{r} \frac{\hbar}{2\omega_{\mathbf{q}(r)}} \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega_{\mathbf{q}(r)}} - 1} \right) \sum_{t} \lambda'_{\mathbf{q}(r,t)}. \tag{37}$$

If the degenerate submatrix $\Delta_{(r)(r)}$ has been diagonalized as described above then at a given $\mathbf{q}$ the sum over t in Eq. (37) can be written

$$\sum_{t} \lambda'_{(r,t)} = \sum_{t} (\Delta_{(r)(r)})_{tt} = \mathrm{Tr}[\Delta_{(r)(r)}], \tag{38}$$

and since the trace of such a matrix is invariant with diagonalization, then the diagonalization has no effect.

It can therefore be seen that no special treatment of degenerate states is required within the first-order perturbation theory to evaluate the free-energy derivatives. This argument applies equally to any thermodynamic quantity which requires a sum over all modes. However, for quantities specific to particular modes, such as the mode-Grüneisen functions $\gamma_{\mathbf{q}s}$, diagonalization of degenerate submatrices $\Delta_{(r)(r)}$ would be required.

B. Conversion between geometrical derivatives

Using the terms defined in Sec. II, if the derivatives of free energy with respect to ρ_i^α and $e^{\mu\nu}$ are known, then those with respect to all the other geometrical parameters (u_i^α, w_m, ϵ_λ, and c_a as defined in Sec. II) can be generated straightforwardly as follows.

Internal strain derivatives can be obtained from ρ_i^α derivatives using Eqs. (3) and (4)

$$\frac{\partial}{\partial u_i^t} = \sum_\alpha A^{t\alpha} \frac{\partial}{\partial \rho_i^\alpha}, \tag{39}$$

$$\frac{\partial}{\partial w_m} = \sum_{i\alpha} h^\alpha_{m,i} \frac{\partial}{\partial \rho_i^\alpha}, \tag{40}$$

where

$$h^\alpha_{m,i} = \sum_t g^t_{m,i} A^{t\alpha}. \tag{41}$$

First derivatives with respect to the symmetrized external strains ϵ_λ can be calculated from $e^{\mu\nu}$ derivatives using the equivalence which follows from Eq. (2):

$$\left(\frac{\partial}{\partial \epsilon_\lambda} \right)_0 = \frac{1}{2} \left[\left(\frac{\partial}{\partial e^{\mu\nu}} \right)_0 + \left(\frac{\partial}{\partial e^{\nu\mu}} \right)_0 \right], \tag{42}$$

where μ, ν for a given λ are determined as in Eq. (2) by the Voigt mapping.

Derivatives with respect to each of the lattice parameters c_a are obtained from those with respect to ϵ_λ using the formula

$$\left(\frac{\partial}{\partial c_a} \right)_0 = \sum_\lambda J_{\lambda a} \left(\frac{\partial}{\partial \epsilon_\lambda} \right)_0, \tag{43}$$

where $J_{\lambda a}$ is the transformation matrix

$$J_{\lambda a} = \left(\frac{\partial \epsilon_\lambda}{\partial c_a} \right)_0. \tag{44}$$

The actual expressions for $J_{\lambda a}$ for a given Bravais lattice will depend on the arbitrary choice of lattice vectors used; a typical choice for a tetragonal unit cell would result in $J_{11} = J_{21} = 1/c_1$ and $J_{32} = 1/c_2$ (other elements being zero), where c_1 and c_2 are the tetragonal lattice parameters a and c, respectively.

VI. LATTICE SUMS FOR PAIR POTENTIALS

Sections IV and V A contained the general theory of evaluation of the Helmholtz free energy and of its derivatives. We present in this section the lattice sums required to implement these evaluations when the particle interactions have the form of spherically symmetric pair potentials. These will be a sum of Coulombic and short-ranged parts, so that the potential between a particle of type i and a particle of type j separated by a vector $\mathbf{r}$ has the form

$$\phi_{ij}(r) = \phi_{ij[\mathrm{Coul}]}(r) + \phi_{ij[\mathrm{short}]}(r). \tag{45}$$

$\phi_{ij[\mathrm{Coul}]}$ is the usual Coulombic potential acting between all particles except the shell and core of the same ion. The short-ranged part $\phi_{ij[\mathrm{short}]}$ may be any smooth function which operates over a finite range. [Specifically, $V^0_{ij}(r)$, $V^1_{ij}(r)$, $V^2_{ij}(r)$, and $V^3_{ij}(r)$ as defined in Eqs. (46) and (47) must exist for values of r at which they will be evaluated.]

In practice within the framework of the shell model, $\phi_{ij[\mathrm{short}]}(r)$ will be a sum of two parts: the first, typically a sum of r^{-n} and e^{-r} terms, operates between different ions

(where an ion consists of a paired core and shell) and must have a cutoff beyond which the potential is small enough to be considered zero (typically around 10 Å). The second part, typically an r^2 ''spring'' term, operates only between shell and core of the same ion. By defining the cutoff of this second part at some distance greater than the maximum shell-core separation of a single ion, and less than the minimum separation of two different ions (say 0.2 Å) no special arrangements need to be made for identifying shell and core of the same ion, which simplifies the formalism somewhat. For rigid ions, only the former part will be present.

A. Strategy for lattice sum evaluation

In calculating the lattice sums for large unit cells, both computational efficiency and economy of storage must be considered. A serious conflict between these requirements arises in construction of the dynamical matrix derivatives $D_{ij}^{\alpha\beta}(\mathbf{q})'$. A naive algorithm would be to calculate each $D_{ij}^{\alpha\beta}(\mathbf{q})'$ sequentially, which requires order n^2 words of storage but repeats much of the calculation work. The naive alternative would be to construct all $N_{\mathcal{E}}$ of them simultaneously, since much of the work to be done in calculation of one can be used for the others. Unless the unit cell is highly symmetrical however this will require storage of order n^3 words, which dramatically limits the size of the unit cell which can reasonably be considered.

Fortunately a compromise is possible, by performing the computationally intensive part of the lattice sums once and storing the results in a compact (i.e., n^2) form. We give an implementation of this strategy in the following section: for each wave vector $\mathbf{q}$, the arrays defined in Eqs. (53)–(61) are calculated first, and the $D_{ij}^{\alpha\beta}(\mathbf{q})'$ are constructed (without much additional work per derivative) from these and used one at a time. Given that the $D_{ij}^{\alpha\beta}(\mathbf{q})'$ are constructed in this way, it is convenient and efficient to calculate the quantities $D_{ij}^{\alpha\beta}(\mathbf{q})$, Φ_{stat}, Φ'_{stat}, and Φ''_{stat} in a similar fashion, since many of the required intermediate steps are common to the two sets of quantities. Alternative expressions for the last four are available elsewhere.[10,14]

B. Direct space expressions for the lattice sums

It is convenient to define a set of functions related to the derivatives of the potentials, represented $V_{ij}^n(r)$, which can be defined recursively:

$$V_{ij}^0(r)=\phi_{ij}(r), \tag{46}$$

$$V_{ij}^n(r)=\frac{1}{r}\frac{d}{dr}V_{ij}^{n-1}(r). \tag{47}$$

Then writing for compactness in this and subsequent sections

$$\rho^{\alpha}=\rho_{ij}^{\alpha}=\rho_j^{\alpha}-\rho_i^{\alpha} \tag{48}$$

$$\rho_{\mathbf{x}}^{\alpha}=\rho_{ij}^{\alpha}+x^{\alpha}, \tag{49}$$

defining $\mathbf{r}$ and $\mathbf{r}_{\mathbf{x}}$ from $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_{\mathbf{x}}$,

$$r^{\alpha}=r_{ij}^{\alpha}=\sum_{\gamma}\,(\delta_{\alpha\gamma}+e^{\alpha\gamma})\rho^{\gamma}, \tag{50}$$

$$r_{\mathbf{x}}^{\alpha}=\sum_{\gamma}\,(\delta_{\alpha\gamma}+e^{\alpha\gamma})\rho_{\mathbf{x}}^{\gamma}, \tag{51}$$

and finally writing

$$V_{\mathbf{x}}^n=V_{ij}^n(r_{\mathbf{x}}), \tag{52}$$

we define the following quantities:

$$S_{ij}=\sum_{\mathbf{x}}{}'\;V_{\mathbf{x}}^0, \tag{53}$$

$$S_{ij}^{\alpha}=\sum_{\mathbf{x}}{}'\;r_{\mathbf{x}}^{\alpha}V_{\mathbf{x}}^1, \tag{54}$$

$$S_{ij}^{\alpha,\mu}=\sum_{\mathbf{x}}{}'\;r_{\mathbf{x}}^{\alpha}r_{\mathbf{x}}^{\mu}V_{\mathbf{x}}^1, \tag{55}$$

$$S_{ij}^{\alpha\beta}=\sum_{\mathbf{x}}{}'\;[\delta_{\alpha\beta}V_{\mathbf{x}}^1+r_{\mathbf{x}}^{\alpha}r_{\mathbf{x}}^{\beta}V_{\mathbf{x}}^2], \tag{56}$$

$$S_{ij}^{\alpha\beta,\mu}=\sum_{\mathbf{x}}{}'\;[\delta_{\alpha\beta}V_{\mathbf{x}}^1+r_{\mathbf{x}}^{\alpha}r_{\mathbf{x}}^{\beta}V_{\mathbf{x}}^2]r_{\mathbf{x}}^{\mu}, \tag{57}$$

$$S_{ij}^{\alpha\beta,\mu\nu}=\sum_{\mathbf{x}}{}'\;[\delta_{\alpha\beta}V_{\mathbf{x}}^1+r_{\mathbf{x}}^{\alpha}r_{\mathbf{x}}^{\beta}V_{\mathbf{x}}^2]r_{\mathbf{x}}^{\mu}r_{\mathbf{x}}^{\nu}, \tag{58}$$

$$T_{ij}^{\alpha\beta}(\mathbf{q})=\sum_{\mathbf{x}}{}'\;e^{i\mathbf{q}\cdot\mathbf{x}}[\delta_{\alpha\beta}V_{\mathbf{x}}^1+r_{\mathbf{x}}^{\alpha}r_{\mathbf{x}}^{\beta}V_{\mathbf{x}}^2], \tag{59}$$

$$T_{ij}^{\alpha\beta\gamma}(\mathbf{q})=\sum_{\mathbf{x}}{}'\;e^{i\mathbf{q}\cdot\mathbf{x}}[(\delta_{\beta\gamma}r_{\mathbf{x}}^{\alpha}+\delta_{\alpha\gamma}r_{\mathbf{x}}^{\beta}+\delta_{\alpha\beta}r_{\mathbf{x}}^{\gamma})V_{\mathbf{x}}^2 + r_{\mathbf{x}}^{\alpha}r_{\mathbf{x}}^{\beta}r_{\mathbf{x}}^{\gamma}V_{\mathbf{x}}^3], \tag{60}$$

$$T_{ij}^{\alpha\beta\gamma,\mu}(\mathbf{q})=\sum_{\mathbf{x}}{}'\;e^{i\mathbf{q}\cdot\mathbf{x}}[(\delta_{\beta\gamma}r_{\mathbf{x}}^{\alpha}+\delta_{\alpha\gamma}r_{\mathbf{x}}^{\beta}+\delta_{\alpha\beta}r_{\mathbf{x}}^{\gamma})V_{\mathbf{x}}^2 + r_{\mathbf{x}}^{\alpha}r_{\mathbf{x}}^{\beta}r_{\mathbf{x}}^{\gamma}V_{\mathbf{x}}^3]r_{\mathbf{x}}^{\mu}. \tag{61}$$

Here and subsequently i and j take values $1\ldots n$, referring to all particles in the unit cell. The summations are over all lattice vectors $\mathbf{x}$, the prime indicating that the term $i=j,\mathbf{x}=\mathbf{0}$ is omitted. There are several symmetry properties of these values which allow a reduction in the number of lattice sums which have to be performed. In the first place, it is obvious that any superscripts not separated by a comma (and additionally those of $S_{ij}^{\alpha,\mu}$) can be freely permuted, e.g.,

$$S_{ij}^{\alpha,\mu}=S_{ij}^{\mu,\alpha}, \tag{62}$$

$$T_{ij}^{\alpha\beta}(\mathbf{q})=T_{ij}^{\beta\alpha}(\mathbf{q}), \tag{63}$$

$$T_{ij}^{\alpha\beta\gamma,\mu}(\mathbf{q})=T_{ij}^{\alpha\gamma\beta,\mu}(\mathbf{q})=T_{ij}^{\beta\gamma\alpha,\mu}(\mathbf{q})=\cdots, \tag{64}$$

and so on. Additionally, all these matrices have some symmetry with respect to exchange of i and j; $S^{\alpha,\mu}$, $S^{\alpha\beta}$, $S^{\alpha\beta,\mu\nu}$, $T^{\alpha\beta}$, and $T^{\alpha\beta\gamma,\mu}$ are Hermitian, while S^{α}, $S^{\alpha\beta,\mu}$, and $T^{\alpha\beta\gamma}$ are anti-Hermitian.

Using Eqs. (1), (47), (49), and (51) the following useful relations can be derived:

$$\frac{\partial \rho_{\mathbf{x}}^{\alpha}}{\partial \rho_{k}^{\alpha}} = \delta_{kj} - \delta_{ik}, \tag{65}$$

$$\frac{\partial r_{\mathbf{x}}^{\alpha}}{\partial e^{\mu\nu}} = \delta_{\alpha\mu}\rho_{\mathbf{x}}^{\nu}, \tag{66}$$

$$\frac{\partial r_{\mathbf{x}}^{\alpha}}{\partial \rho_{\mathbf{x}}^{\beta}} = \delta_{\alpha\beta} + e^{\alpha\beta}, \tag{67}$$

$$\frac{\partial V_{\mathbf{x}}^{n}}{\partial e^{\mu\nu}} = V_{\mathbf{x}}^{n+1} r_{\mathbf{x}}^{\mu} \rho_{\mathbf{x}}^{\nu}, \tag{68}$$

$$\frac{\partial V_{\mathbf{x}}^{n}}{\partial \rho_{\mathbf{x}}^{\alpha}} = V_{\mathbf{x}}^{n+1} \sum_{\gamma} r_{\mathbf{x}}^{\gamma} (\delta_{\gamma\alpha} + e^{\gamma\alpha}). \tag{69}$$

By using these in conjunction with the definition of the static energy Φ_{stat} and the unscaled dynamical matrix $B_{ij}^{\alpha\beta}(\mathbf{q})$ from Sec. IV it is possible to construct the energy, dynamical matrix, and the required internal and external coordinate derivatives in terms of the S and T matrices:

$$\Phi_{\text{stat}} = \frac{1}{2} \sum_{ij} S_{ij}, \tag{70}$$

$$\left(\frac{\partial \Phi_{\text{stat}}}{\partial \rho_{k}^{\alpha}} \right)_{0} = \sum_{i} (1 - \delta_{ik}) S_{ik}^{\alpha}, \tag{71}$$

$$\left(\frac{\partial \Phi_{\text{stat}}}{\partial e^{\mu\nu}} \right)_{0} = \frac{1}{2} \sum_{ij} S_{ij}^{\mu,\nu}, \tag{72}$$

$$\left(\frac{\partial^2 \Phi_{\text{stat}}}{\partial \rho_{k}^{\alpha} \partial \rho_{l}^{\beta}} \right)_{0} = \delta_{kl} \sum_{i} (1 - \delta_{ik}) S_{ik}^{\alpha\beta} - (1 - \delta_{kl}) S_{kl}^{\alpha\beta}, \tag{73}$$

$$\left(\frac{\partial^2 \Phi_{\text{stat}}}{\partial \rho_{k}^{\alpha} \partial e^{\mu\nu}} \right)_{0} = \sum_{i} (1 - \delta_{ik}) [S_{ik}^{\alpha\nu,\mu} + \delta_{\alpha\mu} S_{ik}^{\nu}], \tag{74}$$

$$\left(\frac{\partial^2 \Phi_{\text{stat}}}{\partial e^{\mu\nu} \partial e^{\gamma\lambda}} \right)_{0} = \frac{1}{2} \sum_{ij} S_{ij}^{\mu\gamma,\nu\lambda}, \tag{75}$$

$$B_{ij}^{\alpha\beta}(\mathbf{q}) = \delta_{ij} \sum_{k} T_{ik}^{\alpha\beta}(0) - T_{ij}^{\alpha\beta}(\mathbf{q}), \tag{76}$$

$$\left(\frac{\partial B_{ij}^{\alpha\beta}(\mathbf{q})}{\partial \rho_{k}^{\gamma}} \right)_{0} = \delta_{ik}\delta_{jk} \sum_{l} T_{lk}^{\alpha\beta\gamma}(0) + \delta_{ij} T_{ik}^{\alpha\beta\gamma}(0) + (\delta_{ik} - \delta_{jk}) T_{ij}^{\alpha\beta\gamma}(\mathbf{q}), \tag{77}$$

$$\left(\frac{\partial B_{ij}^{\alpha\beta}(\mathbf{q})}{\partial e^{\mu\nu}} \right)_{0} = \delta_{ij} \sum_{k} T_{ik}^{\alpha\beta\mu,\nu}(0) - T_{ij}^{\alpha\beta\mu,\nu}(\mathbf{q}), \tag{78}$$

where the expressions are given at zero external strain ($e^{\mu\nu} = 0$) as indicated by the $()_0$ notation. Derivatives with respect to other strains can be derived from these using the expressions in Sec. V B; in particular the w_m derivatives are

$$\left(\frac{\partial \Phi_{\text{stat}}}{\partial w_{m}} \right)_{0} = \frac{1}{2} \sum_{ij} (1 - \delta_{ij}) \sum_{\alpha} h_{m,ij}^{\alpha} S_{ij}^{\alpha}, \tag{79}$$

$$\left(\frac{\partial^2 \Phi_{\text{stat}}}{\partial w_{m} \partial w_{n}} \right)_{0} = \frac{1}{2} \sum_{ij} (1 - \delta_{ij}) \sum_{\alpha\beta} h_{m,ij}^{\alpha} h_{n,ij}^{\beta} S_{ij}^{\alpha\beta}, \tag{80}$$

$$\left(\frac{\partial^2 \Phi_{\text{stat}}}{\partial w_{m} \partial e^{\mu\nu}} \right)_{0} = \frac{1}{2} \sum_{ij} (1 - \delta_{ij}) \sum_{\alpha} h_{m,ij}^{\alpha} [S_{ij}^{\alpha\nu,\mu} + \delta_{\alpha\mu} S_{ij}^{\nu}], \tag{81}$$

$$\left(\frac{\partial B_{ij}^{\alpha\beta}(\mathbf{q})}{\partial w_{m}} \right)_{0} = \sum_{\gamma} \left[\delta_{ij} \sum_{l} h_{m,lj}^{\gamma} T_{lj}^{\alpha\beta\gamma}(0) - h_{m,ij}^{\gamma} T_{ij}^{\alpha\beta\gamma}(\mathbf{q}) \right], \tag{82}$$

where $h_{m,ij}^{\alpha} = h_{m,j}^{\alpha} - h_{m,i}^{\alpha}$, analogously to the definition (48) of ρ_{ij}^{α}.

Applying the symmetry relations, already noted, to these expressions, and observing from Eqs. (59)–(61) that the T matrices are real for $\mathbf{q} = 0$, it can be seen that $B_{ij}^{\alpha\beta}(\mathbf{q})$ and its derivatives are Hermitian.

We make one additional practical point here: for very large n calculation time can be dominated by the matrix multiplication (32) since the number of operations required for that step in calculating the gradient scales as $N_{\mathcal{E}} n^3$, while no other part of the code has an execution time which scales faster than n^3 or $N_{\mathcal{E}} n^2$. In one particular case this can be improved: if the $N_{\mathcal{E}} = 3n$ basis internal strain derivatives of a rigid-ion system are being calculated, then the dynamical matrix derivatives required, $\partial D_{ij}^{\alpha\beta} / \partial \rho_{k}^{\gamma}$, are very sparse (any element for which i, j, and k are all different is zero). Explicit calculation and manipulation of the $3n$ dynamical matrix derivatives ($\sim n^4$ operations) can then be avoided in favor of generating the eigenvalue derivatives directly ($\sim n^3$ operations) by substituting Eq. (77) with Eq. (33) in Eqs. (31) and (32) to give

$$\frac{\partial \lambda_{s}}{\partial \rho_{k}^{\gamma}} = \sum_{l} \sum_{\alpha\beta} \left[\left(\frac{f_{\mathbf{q}sk}^{\alpha} * f_{\mathbf{q}sk}^{\beta}}{m_{k}} + \frac{f_{\mathbf{q}sl}^{\alpha} * f_{\mathbf{q}sl}^{\beta}}{m_{l}} \right) T_{lk}^{\alpha\beta\gamma}(0) - \frac{2}{\sqrt{m_{l} m_{k}}} \operatorname{Re}(f_{\mathbf{q}sl}^{\alpha} * f_{\mathbf{q}sk}^{\beta} T_{lk}^{\alpha\beta\gamma}(\mathbf{q})) \right]. \tag{83}$$

This method cannot be applied if the cell contains shells however, since then D' is constructed using Eq. (34) and is no longer sparse.

C. Ewald lattice sums for the Coulomb interactions

As required in Sec. VI $\phi_{ij[\text{short}]}(r)$ acts over a finite range so that the sums (53)–(61) can be taken directly, however $\phi_{ij[\text{Coul}]}(r)$ remains significant for large values of r and special measures must be taken to form good approximations to the sums in a reasonable number of terms. An Ewald summation is therefore used for the Coulombic contribution to the T and S arrays, and this section gives detailed expressions for this part of the calculation.

Using the definitions of $\mathbf{r}$ and $\mathbf{r}_\mathbf{x}$ in Eqs. (50) and (51) and additionally

$$\mathbf{q}_\mathbf{k}=\mathbf{q}+\mathbf{k}, \tag{84}$$

we first define the following standard Ewald expression,[22] $C_{ij}(\mathbf{q})$:

$$C_{ij}(\mathbf{q})=\sum_\mathbf{x}\frac{e^{i\mathbf{q}\cdot\mathbf{x}}}{r_\mathbf{x}}(1-\delta'_{ij\mathbf{x}}) \tag{85}$$

$$=\sum_\mathbf{k}{}'\frac{4\pi}{V}\frac{e^{-i\mathbf{q}_\mathbf{k}\cdot\mathbf{r}-q_\mathbf{k}^2/4\eta^2}}{q_\mathbf{k}^2}+\sum_\mathbf{x}e^{i\mathbf{q}\cdot\mathbf{x}}\frac{\mathrm{erfc}(\eta r_\mathbf{x})-\delta'_{ij\mathbf{x}}}{r_\mathbf{x}}. \tag{86}$$

Here the symbol $\delta'_{ij\mathbf{x}}$ is zero normally, but equal to unity if either (a) $\mathbf{x}=0$ and $i=j$ or (b) $\mathbf{x}=0$ and i,j are shell and core of the same ion. In a rigid-ion system, only case (a) arises. The factor $(1-\delta'_{ij\mathbf{x}})$ therefore properly eliminates the Coulombic interaction of an ion with itself. Additionally the $\mathbf{x}$ summation is over all lattice vectors, the $\mathbf{k}$ summation is over all reciprocal-lattice vectors omitting $\mathbf{q}=\mathbf{k}=0$ (as indicated by the prime), V is the unit-cell volume, and η is an arbitrary parameter, which can be chosen so that good accuracy is achieved when both sums are truncated after a small number of terms going out from their respective origins. $\eta=\sqrt{\pi}V^{-1/3}$ is a reasonable choice.

Comparing Eq. (85) with the r^{-1} form of the Coulombic pair potential $V^0_{ij[\mathrm{Coul}]}(r)\equiv\phi_{ij[\mathrm{Coul}]}(r)$ it can be seen that

$$\frac{z_iz_j}{4\pi\varepsilon_0}C_{ij}(\mathbf{q})=\sum_\mathbf{x}e^{i\mathbf{q}\cdot\mathbf{x}}V^0_{ij[\mathrm{Coul}]}(r_\mathbf{x}), \tag{87}$$

where z_i is the charge on particle i.By manipulating Eq. (87) with aid of the easily obtained relations

$$\frac{\partial r}{\partial r^\alpha}=\frac{r^\alpha}{r}, \tag{88}$$

$$\frac{\partial r^\beta}{\partial r^\alpha}=\delta_{\alpha\beta}, \tag{89}$$

$$\frac{\partial V^n_{ij}(r)}{\partial r^\alpha}=r^\alpha V^{n+1}_{ij}(r), \tag{90}$$

the Coulombic parts of expressions (53)–(61) can be expressed as derivatives of $C_{ij}(\mathbf{q})$ with respect to appropriately chosen variables, as follows:

$$S_{ij[\mathrm{Coul}]}=Z_{ij}C_{ij}(0), \tag{91}$$

$$S^\alpha_{ij[\mathrm{Coul}]}=Z_{ij}\frac{\partial}{\partial r^\alpha}C_{ij}(0), \tag{92}$$

$$S^{\alpha,\mu}_{ij[\mathrm{Coul}]}=Z_{ij}\left[\frac{\partial}{\partial p^\mu}\left(-ie^{i\mathbf{p}\cdot\mathbf{r}}\frac{\partial}{\partial r^\alpha}C_{ij}(\mathbf{p})\right)\right]_{\mathbf{p}=0}, \tag{93}$$

$$S^{\alpha\beta}_{ij[\mathrm{Coul}]}=Z_{ij}\frac{\partial^2}{\partial r^\alpha\partial r^\beta}C_{ij}(0), \tag{94}$$

$$S^{\alpha\beta,\mu}_{ij[\mathrm{Coul}]}=Z_{ij}\left[\frac{\partial}{\partial p^\mu}\left(-ie^{i\mathbf{p}\cdot\mathbf{r}}\frac{\partial^2}{\partial r^\alpha\partial r^\beta}C_{ij}(\mathbf{p})\right)\right]_{\mathbf{p}=0}, \tag{95}$$

$$S^{\alpha\beta,\mu\nu}_{ij[\mathrm{Coul}]}=Z_{ij}\left[\frac{\partial^2}{\partial p^\mu\partial p^\nu}\left(-e^{i\mathbf{p}\cdot\mathbf{r}}\frac{\partial^2}{\partial r^\alpha\partial r^\beta}C_{ij}(\mathbf{p})\right)\right]_{\mathbf{p}=0}, \tag{96}$$

$$T^{\alpha\beta}_{ij[\mathrm{Coul}]}(\mathbf{q})=Z_{ij}\frac{\partial^2}{\partial r^\alpha\partial r^\beta}C_{ij}(\mathbf{q}), \tag{97}$$

$$T^{\alpha\beta\gamma}_{ij[\mathrm{Coul}]}(\mathbf{q})=Z_{ij}\frac{\partial^3}{\partial r^\alpha\partial r^\beta\partial r^\gamma}C_{ij}(\mathbf{q}), \tag{98}$$

$$T^{\alpha\beta\gamma,\mu}_{ij[\mathrm{Coul}]}(\mathbf{q})=Z_{ij}\left[\frac{\partial}{\partial p^\mu}\left(-ie^{i\mathbf{p}\cdot\mathbf{r}}\frac{\partial^3}{\partial r^\alpha\partial r^\beta\partial r^\gamma}C_{ij}(\mathbf{q}+\mathbf{p})\right)\right]_{\mathbf{p}=0}, \tag{99}$$

where Z_{ij} is the Coulomb factor

$$Z_{ij}=\frac{z_iz_j}{4\pi\varepsilon_0}. \tag{100}$$

These differentiations can then be performed term by term on the Ewald sum expression (86) to provide computationally tractable formulas for the required quantities.

The manipulations can be made easier by rewriting Eq. (86) as

$$C_{ij}(\mathbf{q})=\sum_\mathbf{k}{}'\frac{4\pi}{V}e^{-i\mathbf{q}_\mathbf{k}\cdot\mathbf{r}}K^0(q_\mathbf{k})+\sum_\mathbf{x}e^{i\mathbf{q}\cdot\mathbf{x}}X^0(r_\mathbf{x}), \tag{101}$$

where

$$K^0(q_\mathbf{k})=\frac{1}{q_\mathbf{k}^2}e^{-q_\mathbf{k}^2/4\eta^2}, \tag{102}$$

$$X^0(r_\mathbf{x})=\frac{\mathrm{erfc}(\eta r_\mathbf{x})-\delta'_{ij\mathbf{x}}}{r_\mathbf{x}}. \tag{103}$$

We also define recursively in terms of K^0 and X^0 two sets of functions K^n and X^n

$$K^{n+1}(q_\mathbf{k})=\frac{1}{q_\mathbf{k}}\frac{\partial}{\partial q_\mathbf{k}}K^n(q_\mathbf{k}), \tag{104}$$

$$X^{n+1}(r_\mathbf{x})=\frac{1}{r_\mathbf{x}}\frac{\partial}{\partial r_\mathbf{x}}X^n(r_\mathbf{x}), \tag{105}$$

which definitions lead to the following expressions for K^n and X^n:

$$K^1(q_\mathbf{k})=\left(-\frac{1}{2\eta^2}-\frac{2}{q_\mathbf{k}^2}\right)K^0(q_\mathbf{k}), \tag{106}$$

$$K^2(q_{\mathbf{k}})=\left(-\frac{1}{2\eta^2}-\frac{2}{q_{\mathbf{k}}^2}\right)K^1(q_{\mathbf{k}})+\frac{4}{q_{\mathbf{k}}^4}K^0(q_{\mathbf{k}}), \quad (107)$$

$$X^n(r_{\mathbf{x}})=\frac{1}{r_{\mathbf{x}}^2}\left[\frac{(-2)^n\eta^{2n-1}}{\sqrt{\pi}}e^{-\eta^2r^2}-(2n-1)X^{n-1}(r_{\mathbf{x}})\right], \quad (108)$$

and additionally imply the relations

$$\frac{\partial}{\partial q^\mu}K^n(q_{\mathbf{k}})=q_{\mathbf{k}}^\mu K^{n+1}(q_{\mathbf{k}}), \quad (109)$$

$$\frac{\partial}{\partial r^\alpha}X^n(r_{\mathbf{x}})=r_{\mathbf{x}}^\alpha X^{n+1}(r_{\mathbf{x}}). \quad (110)$$

Substituting Eq. (101) in Eqs. (91)–(99) and making use of Eqs. (109) and (110), the expressions for the Ewald sums are therefore given by

$$\frac{S_{ij[\mathrm{Coul}]}}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}\cos(\mathbf{k}\cdot\mathbf{r})K^0(k)+\sum_{\mathbf{x}}\ X^0(r_{\mathbf{x}}), \quad (111)$$

$$\frac{S^\alpha_{ij[\mathrm{Coul}]}}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}[-\sin(\mathbf{k}\cdot\mathbf{r})]k^\alpha K^0(k)+\sum_{\mathbf{x}}\ r_{\mathbf{x}}^\alpha X^1(r_{\mathbf{x}}), \quad (112)$$

$$\frac{S^{\alpha,\mu}_{ij[\mathrm{Coul}]}}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}[-\cos(\mathbf{k}\cdot\mathbf{r})][\delta_{\alpha\mu}K^0(k)+k^\alpha k^\mu K^1(k)]+\sum_{\mathbf{x}}\ r_{\mathbf{x}}^\alpha r_{\mathbf{x}}^\mu X^1(r_{\mathbf{x}}), \quad (113)$$

$$\frac{S^{\alpha\beta}_{ij[\mathrm{Coul}]}}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}[-\cos(\mathbf{k}\cdot\mathbf{r})]k^\alpha k^\beta K^0(k)+\sum_{\mathbf{x}}\ [\delta_{\alpha\beta}X^1(r_{\mathbf{x}})+r_{\mathbf{x}}^\alpha r_{\mathbf{x}}^\beta X^2(r_{\mathbf{x}})], \quad (114)$$

$$\frac{S^{\alpha\beta,\mu}_{ij[\mathrm{Coul}]}}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}\sin(\mathbf{k}\cdot\mathbf{r})[(\delta_{\alpha\mu}k^\beta+\delta_{\beta\mu}k^\alpha)K^0(k)+k^\alpha k^\beta k^\mu K^1(k)]+\sum_{\mathbf{x}}\ [\delta_{\alpha\beta}X^1(r_{\mathbf{x}})+r_{\mathbf{x}}^\alpha r_{\mathbf{x}}^\beta X^2(r_{\mathbf{x}})]r_{\mathbf{x}}^\mu, \quad (115)$$

$$\frac{S^{\alpha\beta,\mu\nu}_{ij[\mathrm{Coul}]}}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}\cos(\mathbf{k}\cdot\mathbf{r})[(\delta_{\alpha\mu}\delta_{\beta\nu}+\delta_{\beta\mu}\delta_{\alpha\nu})K^0(k)+(\delta_{\alpha\mu}k^\beta k^\nu+\delta_{\beta\mu}k^\alpha k^\nu+\delta_{\alpha\nu}k^\beta k^\mu+\delta_{\beta\nu}k^\alpha k^\mu+\delta_{\mu\nu}k^\alpha k^\beta)K^1(k)$$
$$+k^\alpha k^\beta k^\mu k^\nu K^2(k)]+\sum_{\mathbf{x}}\ [\delta_{\alpha\beta}X^1(r_{\mathbf{x}})+r_{\mathbf{x}}^\alpha r_{\mathbf{x}}^\beta X^2(r_{\mathbf{x}})]r_{\mathbf{x}}^\mu r_{\mathbf{x}}^\nu, \quad (116)$$

$$\frac{T^{\alpha\beta}_{ij[\mathrm{Coul}]}(\mathbf{q})}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}[-e^{-i\mathbf{q}_{\mathbf{k}}\cdot\mathbf{r}}]q_{\mathbf{k}}^\alpha q_{\mathbf{k}}^\beta K^0(q_{\mathbf{k}})+\sum_{\mathbf{x}}\ e^{i\mathbf{q}\cdot\mathbf{x}}[\delta_{\alpha\beta}X^1(r_{\mathbf{x}})+r_{\mathbf{x}}^\alpha r_{\mathbf{x}}^\beta X^2(r_{\mathbf{x}})], \quad (117)$$

$$\frac{T^{\alpha\beta\gamma}_{ij[\mathrm{Coul}]}(\mathbf{q})}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}ie^{-i\mathbf{q}_{\mathbf{k}}\cdot\mathbf{r}}q_{\mathbf{k}}^\alpha q_{\mathbf{k}}^\beta q_{\mathbf{k}}^\gamma K^0(q_{\mathbf{k}})+\sum_{\mathbf{x}}\ e^{i\mathbf{q}\cdot\mathbf{x}}[(\delta_{\beta\gamma}r_{\mathbf{x}}^\alpha+\delta_{\alpha\gamma}r_{\mathbf{x}}^\beta+\delta_{\alpha\beta}r_{\mathbf{x}}^\gamma)X^2(r_{\mathbf{x}})+r_{\mathbf{x}}^\alpha r_{\mathbf{x}}^\beta r_{\mathbf{x}}^\gamma X^3(r_{\mathbf{x}})], \quad (118)$$

$$\frac{T^{\alpha\beta\gamma,\mu}_{ij[\mathrm{Coul}]}(\mathbf{q})}{Z_{ij}}=\sum_{\mathbf{k}}{}'\ \frac{4\pi}{V}e^{-i\mathbf{q}_{\mathbf{k}}\cdot\mathbf{r}}[(\delta_{\alpha\mu}q_{\mathbf{k}}^\beta q_{\mathbf{k}}^\gamma+\delta_{\beta\mu}q_{\mathbf{k}}^\alpha q_{\mathbf{k}}^\gamma+\delta_{\gamma\mu}q_{\mathbf{k}}^\alpha q_{\mathbf{k}}^\beta)K^0(q_{\mathbf{k}})+q_{\mathbf{k}}^\alpha q_{\mathbf{k}}^\beta q_{\mathbf{k}}^\gamma q_{\mathbf{k}}^\mu K^1(q_{\mathbf{k}})]$$
$$+\sum_{\mathbf{x}}\ e^{i\mathbf{q}\cdot\mathbf{x}}[(\delta_{\beta\gamma}r_{\mathbf{x}}^\alpha+\delta_{\alpha\gamma}r_{\mathbf{x}}^\beta+\delta_{\alpha\beta}r_{\mathbf{x}}^\gamma)X^2(r_{\mathbf{x}})+r_{\mathbf{x}}^\alpha r_{\mathbf{x}}^\beta r_{\mathbf{x}}^\gamma X^3(r_{\mathbf{x}})]r_{\mathbf{x}}^\mu. \quad (119)$$

Care must be taken in evaluating the functions $X^n(r)$ when $\delta'_{ij\mathbf{x}}$ is unity, in which case $r_{\mathbf{x}}=r$ will be small. When r is small but nonzero it is important in practice to ensure that the evaluation is carried out without loss of precision, for instance by using a few terms of a polynomial series expansion of the $X^n(r)$ instead of direct evaluation of Eq. (108). In the limit $r=0$,

$$\lim_{r\to 0}X^n(r)=\frac{(-2)^{n+1}\eta^{2n+1}}{(2n+1)\sqrt{\pi}}. \quad (120)$$

VII. CONCLUSION

A method has been presented for the analytic calculation of the fully quasiharmonic dynamical free-energy first derivatives, and static energy second derivatives, for a periodic crystal with respect to various parameters of its geometry, and it is shown how these derivatives can be used for efficient structural optimization of such a crystal. Given the optimized geometry, equilibrium thermodynamic quantities such as free energy, heat capacity, and thermal expansion can accurately and rapidly be generated. This procedure can

be used over a wide range of temperatures and pressures.

Detailed expressions for the lattice sums required have been given, where particles are taken to interact via a sum of Coulombic and arbitrary short-ranged spherically symmetric pairwise potentials. Care has been taken that the method as presented can be directly implemented in a computer program making efficient use of processing time and available storage, and such a program has indeed been written[23] and used successfully for fully dynamical optimization of unit cells containing up to ~700 ions.[24]

Future work will include the treatment of three- and four-body short-ranged interactions, and a consideration of further anharmonic corrections. A wide range of applications including complex oxides, defective lattices, high-pressure phase transitions, and surface properties is currently in progress and will be reported separately.

ACKNOWLEDGMENTS

This work was supported by EPSRC Grant No. GR/K05979. G.D.B. gratefully acknowledges financial support from la Universidad de Buenos Aires. His contribution to this work was made possible by means of a grant from el Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina. The authors wish to acknowledge useful discussions with W. C. Mackrodt and with L. Kantorovich over a long period.

*Permanent address: Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Química Inorgánica, Analítica y Química Física, Pabellón 2, Ciudad Universitaria, 1428 Buenos Aires, Argentina.

†Electronic address: n.l.allan@bristol.ac.uk

[1] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, Oxford, 1954).

[2] B. G. Dick and A. W. Overhauser, Phys. Rev. **112**, 90 (1958).

[3] G. D. Barrera, M. B. Taylor, N. L. Allan, T. H. K. Barron, L. N. Kantorovich, and W. C. Mackrodt, J. Chem. Phys. **107**, 4337 (1997).

[4] N. L. Allan, M. Braithwaite, D. L. Cooper, W. C. Mackrodt, and S. C. Wright, J. Chem. Phys. **95**, 6792 (1991).

[5] N. L. Allan, M. Braithwaite, D. L. Cooper, W. C. Mackrodt, and B. Petch, J. Chem. Soc. Faraday Trans. **89**, 4369 (1993).

[6] G. W. Watson, P. Tschaufeser, A. Wall, R. A. Jackson, and S. C. Parker, in *Computer Modeling in Inorganic Crystallography*, edited by C. R. A. Catlow (Academic, San Diego, 1997), p. 55.

[7] K. Binder and D. Stauffer, in *Application of the Monte Carlo Method in Statistical Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1984), p. 1.

[8] L. N. Kantorovich, Phys. Rev. B **51**, 3520 (1995).

[9] K. B. Tolpygo, Zh. Eksp. Teor. Fiz. **20**, 497 (1950).

[10] A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic, New York, 1971).

[11] J. F. Nye, *Physical Properties of Crystals*, 2nd ed. (Clarendon, Oxford, 1985).

[12] T. H. K. Barron, T. G. Gibbons, and R. M. Munn, J. Phys. C **4**, 2805 (1971).

[13] A. B. Pippard, *The Elements of Classical Thermodynamics* (Cambridge University Press, Cambridge, 1957).

[14] C. R. A. Catlow and W. C. Mackrodt, in *Computer Simulation of Solids*, edited by C. R. A. Catlow and W. C. Mackrodt, Lecture Notes in Physics Vol. 166 (Springer-Verlag, Berlin, 1982), p. 7.

[15] N. L. Allan, T. H. K. Barron, and J. A. O. Bruno, J. Chem. Phys. **105**, 8300 (1996).

[16] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, London, 1981).

[17] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1989).

[18] G. Venkataraman, L. A. Feldkamp, and V. C. Sahni, *Dynamics of Perfect Crystals* (MIT, Cambridge, 1975).

[19] D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[20] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[21] See, e.g., P. C. W. Davies and D. S. Betts, *Quantum Mechanics* (Chapman and Hall, London, 1994).

[22] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1964). Note however that the erfc function is incorrectly written as erf in some of the expressions given in this reference.

[23] M. B. Taylor, G. D. Barrera, N. L. Allan, T. H. K. Barron, and W. C. Mackrodt (unpublished).

[24] M. B. Taylor, G. D. Barrera, N. L. Allan, T. H. K. Barron, and W. C. Mackrodt, Farad. Discuss. (to be published).