JOURNAL OF COMPUTATIONAL PHYSICS **131,** 233–240 (1997)
ARTICLE NO. CP965612

Relaxation of Crystals with the Quasi-Newton Method

Bernd G. Pfrommer, Michel Côté, Steven G. Louie, and Marvin L. Cohen

Department of Physics, University of California at Berkeley, Berkeley, California 94720 and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Received June 13, 1996; revised November 4, 1996

A quasi-Newton method is used to simultaneously relax the internal coordinates and lattice parameters of crystals under pressure. The symmetry of the crystal structure is preserved during the relaxation. From the inverse of the Hessian matrix, elastic properties, and some optical phonon frequencies at the Brillouin zone center can be estimated. The efficiency of the method is demonstrated for silicon test systems. © 1997 Academic Press

1. INTRODUCTION

Ab-initio computations of the total energy within the framework of density functional theory (DFT) and the local density approximation (LDA) [1] have been successful in predicting the structural properties of materials [3]. At zero temperature and pressure, the structural parameters are determined by moving the constituting atoms to positions where the energy E is minimal. This can be done much more efficiently if the forces on the atoms can be computed [2]. If a pressure p is applied to the material, it is the enthalpy $\mathcal{H} = E + pV$ which has to be minimized with respect to all structural parameters, including the volume V.

Our focus will be on crystalline materials at zero temperature, where the unit cell shape and the coordinates of the atoms inside the unit cell are the parameters to be adjusted such that the enthalpy acquires a minimum. In this article, we report a symmetry-preserving algorithm to relax the unit cell shape and the atomic coordinates simultaneously by using the computed forces [2] and the stress [4]. This is a frequent task when structural phase transitions are studied, where one is interested in the properties of a phase with a given symmetry as a function of pressure. Although we demonstrate the efficiency of our method within the framework of DFT in LDA, it is of rather general use and can be applied to relax crystal structures whenever forces and stress are available.

Quantum molecular dynamics schemes of different flavors are commonly used to tackle this problem. In contrast to the Car–Parrinello method [5], we follow the more traditional approach [6] and relax the electronic degrees of freedom completely before moving the atoms and changing the shape of the unit cell. While the early molecular dynamics methods kept the unit cell shape of the crystal fixed and relaxed only the internal degrees of freedom, the more recent algorithms allow for a variation of both. The first molecular dynamics approach with a variable unit cell *shape* was proposed by Parrinello and Rahman (PR) [7], where a fictitious Lagrangian governed the time evolution of internal coordinates and the cell shape. However, the crystal symmetry is not preserved along the trajectories derived from their Lagrangian; in other words the symmetry of the crystal can be reduced during the relaxation process. This is undesirable when structural phase transitions are examined, where one would like to impose the crystal symmetry. Picking the strain ε as time-dependent variable instead of the lattice vectors, Wentzcovitch [8] modified the PR Lagrangian to generate symmetry-preserving trajectories.

We present here a relaxation scheme which preserves the symmetry and is *not* based on a molecular dynamics approach, but which uses a powerful quasi-Newton optimization scheme to search for the relaxed configuration. While this method has been applied to treat forces on atoms before [9], we report for the first time how to simultaneously relax the lattice parameters. The quasi-Newton method accumulates information about the enthalpy surface in the inverse of the Hessian matrix H, which renders it superior to the molecular dynamics algorithms proposed recently [6]. After the relaxation has been completed, H can be exploited to estimate elastic properties and the optical phonon energies at the center of the Brillouin zone. Using the H obtained from similar calculations further improves the performance. We demonstrate the efficiency of our scheme for silicon in the diamond and the R8 phases [10].

2. THE QUASI-NEWTON METHOD FOR CRYSTAL STRUCTURE RELAXATION

We start this section by establishing the notation and defining the configuration space coordinates. The crystal structure is determined by the matrix of lattice vectors $h = \{a, b, c\}$ and the coordinates s_i, $i = 1, \ldots N$, relative to

0021-9991/97 $25.00

h of the N atoms in the unit cell, which has a volume of $\Omega = \det(h)$. The energy E per unit cell is a function of h and the s_i's.

For convenience, we choose the finite strain tensor ε as a free variable, instead of the lattice vectors h. It has nine components and stretches a reference configuration h_0 into $h = (1 + \varepsilon)h_0$. Molecular dynamics schemes [8, 11] often constrain ε to be symmetric in order to avoid rotations of the unit cell. Since there is no notion of angular momentum in our scheme, we allow for an asymmetric ε to simplify the extraction of elastic properties and phonon modes in Section 3.

Relaxing a crystal structure with N atoms in the unit cell under the applied pressure p, thus, is an optimization problem for the enthalpy per unit cell $\mathcal{H} = E + p\Omega$ in a $(9 + 3N)$-dimensional space:

$$\mathcal{H} = \mathcal{H}(\varepsilon, s_1, ..., s_N). \tag{1}$$

Let us denote a point in configuration space by the column vector X. We define the first nine components of X to be the strain components ε, which are converted into a nine-element column vector by $X_{3(i-1)+j} = \varepsilon_{ij}$; $i, j = 1, 2, 3$. Then follow the coordinates of the atoms in the unit cell, $s_1, s_2, ..., s_N$. We will call the negative of the derivative of the enthalpy $\mathcal{H}$ with respect to X,

$$F = -\left.\frac{\partial \mathcal{H}}{\partial X}\right|_p, \tag{2}$$

the "force vector." The strain components of F are the derivatives of $\mathcal{H} = E + p\Omega$ with respect to ε,

$$f^{(\varepsilon)} = -(\sigma + p\Omega)(1 + \varepsilon^{\mathrm{T}})^{-1}, \tag{3}$$

where σ is the stress at a given configuration X:

$$\sigma = \left(\frac{\partial E((1 + \varepsilon')h)}{\partial \varepsilon'}\right)_{\varepsilon'=0}. \tag{4}$$

Notice that the right-hand side of (3) need not be symmetric and that we do not symmetrize it. Thus, ε can become asymmetric during the course of the relaxation.

The other $3N$ components of F are obtained by multiplying the forces on the atoms $f_1 ... f_N$ in lattice coordinates with the metric tensor $g = h^{\mathrm{T}}h$, such that the complete F can be written as

$$F = (f^{(\varepsilon)}, gf_1, ..., gf_N)^{\mathrm{T}}. \tag{5}$$

There exist a large number of algorithms for finding minima of multivariable functions if the first derivative is available. We favor the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton scheme [12] for its stability and efficiency [13]. A recent review of this method is given in Ref. [14]. Like all quasi-Newton schemes, BFGS accumulates information about the Hessian matrix and, therefore, about the shape of the enthalpy surface around the minimum. As we will show in Section 3, in many cases this allows us to estimate the frequencies of zone-center optical phonons, elastic stiffness coefficients, and the bulk modulus.

Sufficiently close to a minimum $X_{\min}$, the change in enthalpy $\delta\mathcal{H}$ can be approximated by

$$\delta\mathcal{H} = \tfrac{1}{2}(X - X_{\min}) \cdot A(X - X_{\min}). \tag{6}$$

In the vicinity of $X_{\min}$, complete knowledge of the Hessian matrix A would allow us to find the exact (local) minimum $X_{\min}$ from the force F with one relaxation step. However, A is unknown. The key idea of the quasi-Newton schemes is to start with an initial guess for A, and improve on A successively as the relaxation proceeds. Actually, it is not A, but the inverse $H = A^{-1}$ which is being developed. In relaxation step $i + 1$, the previous position X_i is updated according to

$$X_{i+1} = X_i + \lambda \Delta X_i, \tag{7}$$

$$\Delta X_i = H_i F_i, \tag{8}$$

where F_i is F evaluated at X_i and λ is the step length which is determined by an approximate line minimization along the step direction ΔX_i. The BFGS scheme takes as input an initial guess H_0 for H and updates it according to

$$\begin{aligned} H_i = H_{i-1} &- \frac{(X_i - X_{i-1}) \otimes (X_i - X_{i-1})}{(X_i - X_{i-1}) \cdot (F_i - F_{i-1})} \\ &- \frac{(H_{i-1}(F_i - F_{i-1})) \otimes (H_{i-1}(F_i - F_{i-1}))}{(F_i - F_{i-1}) \cdot H_{i-1}(F_i - F_{i-1})} \\ &+ [(F_i - F_{i-1}) \cdot H_{i-1}(F_i - F_{i-1})]U \otimes U, \end{aligned} \tag{9}$$

$$U = \frac{(X_i - X_{i-1})}{(X_i - X_{i-1}) \cdot (F_i - F_{i-1})} - \frac{H_{i-1}(F_i - F_{i-1})}{(F_i - F_{i-1}) \cdot H_{i-1}(F_i - F_{i-1})}.$$

If the enthalpy were perfectly quadratic in $X - X_{\min}$, H_i would converge to A^{-1} after the number of relaxation steps has reached the number of degrees of freedom in the system [15]. More precisely, H_i and A^{-1} would not be identical, but would have the same projection into the sampled subspace. The number of degrees of freedom equals the number of symmetry-compliant directions in configuration space, not counting rotations of the unit cell and an overall translation of the atoms. This amounts to the number of

symmetric optical phonon modes at the Brillouin zone center Γ plus the number of lattice parameters.

In molecular dynamics schemes, one has to choose suitable fictitious masses and proper time steps to get fast convergence. The masses are normally determined by optimizing the dynamical coupling to the internal degrees of freedom during test runs [8]. Analogously, in the present quasi-Newton method, H_0 has to be initialized properly to assure a reasonable step size during the first few relaxation steps. It is important that H_0 does not break the symmetry when it is applied to a force vector F. Evidently, the dependence of the enthalpy on ε is governed by the elastic stiffness coefficients B_{ijkl} [16], or in a coarser sense, by the bulk modulus. Similarly, the optical phonon frequencies at Γ should determine the increase of the enthalpy upon displacement of the internal coordinates from the equilibrium positions.

With this in mind, we suggest setting the strain part of H_0 to the (9×9) identity matrix multiplied by $(3\Omega B_0)^{-1}$, B_0 being an estimate for the bulk modulus. Thus, in the first step the strain components of the search direction ΔX_0 will be parallel to $f^{(\varepsilon)}$. For the internal coordinates, we propose to initialize H_0 as block diagonal with (3×3) matrices of the form $g_0^{-1}\bar{M}^{-1}\bar{\omega}_O^{-2}$, where $g_0 = h_0^{\mathrm{T}}h_0$ is the metric tensor of the initial configuration, $\bar{M}$ is the average mass of the atoms, and $\bar{\omega}_O$ is a guess for the average of the optical phonon frequencies at the center of the Brillouin zone:

$$H_0 = \begin{pmatrix} (3\Omega B_0)^{-1} & & & & & 0 \\ & \ddots & & & & \\ & & (3\Omega B_0)^{-1} & & & \\ & & & g_0^{-1}\bar{M}^{-1}\bar{\omega}_O^{-2} & & \\ & & & & \ddots & \\ 0 & & & & & g_0^{-1}\bar{M}^{-1}\bar{\omega}_O^{-2} \end{pmatrix}. \tag{10}$$

The motivation for Eq. (10) will become more transparent when we discuss the extraction of elastic properties and optical phonon modes in Section 3. We show in the Appendix that this initialization of H preserves the symmetry of the crystal during the relaxation. Note that (10) has only two "free" input parameters $\bar{\omega}_O$ and B_0, which makes it simple to use. There are certainly more sophisticated ways of choosing H_0, for instance with different values on the diagonal for the strain part, corresponding to strains for which the elastic stiffness coefficients are expected to differ substantially. However, this will only matter during the first few relaxation steps for a new structure. In the case that a similar relaxation has been done previously—for example the same structure, but at a different pressure—the fully builtup H of that calculation can be used with great benefit, as we will see in Section 4.3.

The line minimization required to find λ in Eq. (7) cannot be performed exactly, and one has to refrain to trial steps to find the minimum of $\mathcal{H}$ along the proposed search direction ΔX_i. It is often advantageous not to do any line minimization and to set $\lambda = 1$, because the increased number of relaxation steps required will be more than outweighed by the savings in force/stress evaluations for the line minimization. On the other hand, during the first few relaxation steps, the force F might be large due to a poor initialization of H, and an approximate line minimization is necessary to stabilize the algorithm. We find the following procedure a good compromise for several different systems. After a trial step with $\lambda = 1$, a linear fit to the forces F at $\lambda = 0$ and $\lambda = 1$ is performed. From the fit, we obtain λ_2 for which $\mathcal{H}$ should be minimal along ΔX_i. If λ_2 is smaller than 0.4 or larger than 1.6, we move by $\lambda_2 \Delta X_i$. Otherwise, we consider $\lambda = 1$ as sufficiently close to the minimum and move by ΔX_i, thereby saving a force/stress computation. Once H is built up, the trial step with $\lambda = 1$ will be close enough to the minimum to omit the additional step with $\lambda = \lambda_2$.

In general there are several local minima in configuration space, and the algorithm can get trapped in one of those. In that sense, we present a method to relax forces and stress, *not* to find the structure for which the enthalpy is globally minimal. The location of the global enthalpy minimum still requires the intuition of a good starting point. The very same fact allows us to apply negative pressures, where the system could lower its enthalpy $\mathcal{H} = E - |p|\Omega$ arbitrarily by increasing the cell volume Ω. This indeed happens for strong negative pressures if the initial crystal structure is weakly bound, e.g., by Van der Waals forces. For moderate negative pressures and crystals with ionic, metallic, or covalent bonds this is not a problem though.

3. OPTICAL PHONON MODES AND ELASTIC PROPERTIES

Given a perfectly quadratic form of the enthalpy around the minimum and assuming an exact line minimization, it can be shown [15] that H will converge to the inverse of the matrix A, and one would think that, conversely, a large amount of information about phonon frequencies and elastic properties can be extracted by analyzing H. However, the assumption of a quadratic form is only valid around the minimum, and we already mentioned that an accurate line minimization is expensive to do.

Also, the relaxation process might be converged long before the H-matrix is fully built up. This is especially true for a large number of degrees of freedom and if the initial

configuration is close to the relaxed one. Even though it is often possible to extract information from an incomplete H, for simplicity we will assume that all degrees of freedom have been sampled.

Moreover, depending on the crystal structure, only a subspace of the full configuration space might be sampled during the relaxation process, since the crystal symmetry is preserved. Thus, we can only get the elastic properties for strains which do not break the symmetry, and the symmetry-preserving optical phonon modes at Γ. On the other hand, since isotropic scaling leaves the symmetry unchanged, the cell volume Ω is always a free parameter and the bulk modulus is accessible irrespective of the crystal structure.

The extraction of information starts with manipulations of the H-matrix (cf. Eq. (9)) obtained from the last relaxation step. We first correct for the finite strain, which enters the force $f^{(\varepsilon)}$ in Eq. (3) in the form of a factor of $(1 + \varepsilon^{\mathrm{T}})^{-1}$. After defining the $(9 + 3N) \times (9 + 3N)$ matrix D as

$$D = \begin{pmatrix} (1+\varepsilon) & & & & & & 0 \\ & (1+\varepsilon) & & & & & \\ & & (1+\varepsilon) & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ 0 & & & & & & 1 \end{pmatrix}, \tag{11}$$

we arrive at the corrected $(9 + 3N) \times (9 + 3N)$ matrix

$$H' = D^{-1^{\mathrm{T}}} H D^{-1}. \tag{12}$$

In Eq. (12), H is transformed from the coordinate system of the *initial* configuration h_0 to the coordinate system of the relaxed configuration h, such that H' describes the changes in enthalpy around the relaxed configuration.

We then restrict the strain around the relaxed configuration to be symmetric by projecting H' from the full (9×9) strain space to the smaller (6×6) strain space in Voigt notation, which leads to the $(6 + 3N) \times (6 + 3N)$ matrix H''.

Next we find out which directions in configuration space have been sampled by examining the update $H''_{\mathrm{update}} = H'' - H''_0$, where H''_0 is obtained from H_0 the same way as H'' from H. The update H''_{update} has accumulated the step directions by means of the updating formula Eq. (9), and a singular value decomposition (SVD) of $H''_{\mathrm{update}} = U \cdot w \cdot V^{\mathrm{T}}$ (see [12]) yields a set of m orthonormal basis vectors of length $(6 + 3N)$ spanning the sampled subspace. We collect the basis vectors from the columns of U for which the corresponding diagonal elements of w are nonzero [12], into the columns of the $(6 + 3N) \times m$ matrix Y. Assuming that all directions in configuration space permitted by the symmetry constraints have been sampled, m is the number of degrees of freedom. In this case, we can choose the fist m_ε basis vectors in Y to have components in the six-dimensional symmetric strain space only, and the other m_s basis vectors to be pure displacements of the internal coordinates. Thus we have a symmetric strain tensor $\varepsilon^{(l)}$ for each of the strain-only basis vectors indexed by $l = 1, ..., m_\varepsilon$, and a set of atomic displacement vectors $s_i^{(k)}$, $i = 1, ..., N$, for each of the other basis vectors $k = 1, ..., m_s$. Such a decomposition is always possible, because the lattice degrees of freedom and those of the internal coordinates are independent of each other.

To determine the accessible elastic stiffness coefficients and phonon modes, it is convenient to write the distortion $\delta\varepsilon$ and the atomic displacement vectors u_i, $i = 1, ..., N$ (in Cartesian coordinates) as linear combinations of the basis vectors spanning the sampled subspace:

$$u_i = \sum_{j=1}^{m_s} \xi_j M_i^{-1/2} h s_i^{(j)}, \tag{13}$$

$$\delta\varepsilon = \sum_{j=1}^{m_\varepsilon} \zeta_j \varepsilon^{(j)}. \tag{14}$$

Equations (13) and (14) define the reduced coordinates ξ and ζ. The mass M_i of the atom i in (13) simplifies the notation when the phonon modes are calculated later. For a compact matrix notation, we further introduce the $(6 + 3N) \times (6 + 3N)$ matrix

$$\mu = \begin{pmatrix} I_6 & & & & & & 0 \\ & M_1 & & & & & \\ & & M_1 & & & & \\ & & & M_1 & & & \\ & & & & M_2 & & \\ & & & & & \ddots & \\ 0 & & & & & & M_N \end{pmatrix}, \tag{15}$$

which has the (6×6) identity I_6 and the masses M_i on the diagonal. With the definitions given by Eqs. (13), (14), and (15), the Hessian matrix $\bar{A}$ in the reduced coordinates is

$$\bar{A} = \begin{pmatrix} \bar{A}^{(\varepsilon)} & \bar{A}^{(\varepsilon,s)} \\ (\bar{A}^{(\varepsilon,s)})^{\mathrm{T}} & \bar{A}^{(s)} \end{pmatrix} = (Y^{\mathrm{T}} \mu^{1/2} H'' \mu^{1/2} Y)^{-1}, \tag{16}$$

and the change to the enthalpy $\delta\mathcal{H}$ around the relaxed configuration (cf. Eq. (6)) can be written as

$$\delta\mathcal{H} = \frac{1}{2}\left(\sum_{i=1}^{m_\varepsilon}\sum_{j=1}^{m_\varepsilon}\zeta_i\bar{A}^{(\varepsilon)}_{ij}\zeta_j + 2\sum_{k=1}^{m_\varepsilon}\sum_{l=1}^{m_s}\zeta_k\bar{A}^{(\varepsilon,s)}_{kl}\xi_l + \sum_{p=1}^{m_s}\sum_{q=1}^{m_s}\xi_p\bar{A}^{(s)}_{pq}\xi_q\right). \tag{17}$$

Equation (17) gives the change in enthalpy if the internal coordinates, the unit cell shape, or both together are varied along the reduced coordinates. This information is contained in the matrices $\bar{A}^{(s)}$, $\bar{A}^{(\varepsilon)}$, and $\bar{A}^{(\varepsilon,s)}$, respectively.

3.1. *Optical Phonon Modes*

It is straightforward to extract the normal modes and frequencies of the symmetry-preserving optical phonons at Γ from $\bar{A}^{(s)}$. With a fixed unit cell shape, the enthalpy can be expressed in terms of the displacement vectors u_i, $i = 1, ..., N$, of the N atoms

$$\delta\mathcal{H} = \frac{1}{2}\sum_{i,j=1}^{N} u_i\Phi^{(ij)}u_j, \tag{18}$$

where $\Phi^{(ij)}$ is the 3×3 matrix of force constants of second order [17] between atoms i and j. Its relation to the matrix $\bar{A}^{(s)}$ is

$$\bar{A}^{(s)}_{mn} = \sum_{i,j=1}^{N}(hs_i^{(m)})^{\mathrm{T}}\frac{\Phi^{(ij)}}{(M_iM_j)^{1/2}}(hs_j^{(n)}). \tag{19}$$

Thus $\bar{A}^{(s)}$ is the projection of the dynamical matrix at Γ into the subspace of symmetry-preserving displacement patterns.

To find the phonon modes and the frequencies, one substitutes (13) into the equations of motion for the phonons and arrives at the generalized eigenvalue problem

$$(\omega^2 S - \bar{A}^{(s)})\xi = 0, \tag{20}$$

where $S_{mn} = \sum_{i=1}^{N}(hs_i^{(m)})^{\mathrm{T}}(hs_i^{(n)})$ is a symmetric overlap matrix between the basis vectors. The angular frequencies ω and the vectors ξ that solve Eq. (20) contain all the accessible information about the phonon modes. To get the displacement patterns of the phonon modes in Cartesian coordinates, we substitute the vector ξ into Eq. (13).

3.2. *Elastic Stiffness Coefficients*

A fully builtup H-matrix allows the computation of some linear combinations of the elastic stiffness coefficients B_{ijkl} for finite pressure [18] from Eq. (17). The B_{ijkl}'s describe how the enthalpy changes upon lattice distortions around the equilibrium configuration, assuming that the lattice distortion is accompanied by a relaxation of the internal parameters. We will now relate B_{ijkl} to $\bar{A}$, which gives the change of enthalpy along the reduced coordinates. It is not sufficient to just use $\bar{A}^{(\varepsilon)}$, because the internal coordinates must be relaxed as the unit cell is deformed. The matrix $\bar{B}$ defined as

$$\bar{B}_{mn} = \bar{A}^{(\varepsilon)}_{mn} - (\bar{A}^{(\varepsilon,s)}\bar{A}^{(s)-1}\bar{A}^{(\varepsilon,s)\mathrm{T}})_{mn} \tag{21}$$

is the projection of B_{ijkl} into the subspace of symmetry-preserving, symmetric strains $\varepsilon^{(j)}$, $j = 1, ..., m_\varepsilon$:

$$\Omega\varepsilon^{(m)}_{ij}B_{ijkl}\varepsilon^{(n)}_{kl} = \bar{B}_{mn}. \tag{22}$$

If m_ε is smaller than six, the knowledge of $\bar{B}$ in Eq. (21) does not allow one to recover all elastic stiffness coefficients B_{ijkl}. However, the bulk modulus

$$B_0 = \left[\Omega\sum_{m,n=1}^{m_\varepsilon}\mathrm{Tr}(\varepsilon^{(m)})(\bar{B}^{-1})_{mn}\mathrm{Tr}(\varepsilon^{(n)})\right]^{-1} \tag{23}$$

can always be computed, since the cell volume Ω is a free parameter and guarantees $m_\varepsilon > 0$.

4. NUMERICAL TESTS

We show the efficiency of our method for three different silicon systems, treated within density functional theory (DFT) in the local density approximation (LDA) [19]. A norm-conserving pseudopotential [20] and a plane-wave basis set expansion up to an energy cutoff of 24 Rydbergs are used. Silicon is chosen because it is computationally simple, and because DFT in LDA is known to reproduce the experimentally observed structural parameters accurately [21, 22]. For the line minimization, we use the prescription outlined in Section 2.

4.1. *Bulk Modulus and Phonon Frequencies*

The accuracy of our formulae for the bulk modulus Eq. (23) and the phonon frequencies Eq. (20) are tested by stretching the silicon bond in the two-atom cell of the diamond lattice along the (111) direction by 14% or 0.34 a.u. (1 a.u. = 0.52918×10^{-10}m). This structure has three degrees of freedom: the angle between the lattice vectors, the bond length between the two atoms in the unit cell, and the cell volume. While for the starting configuration the angle is left at 60°, we decrease the volume per atom from 131 to 120 a.u.^{3}, and initialize H according to Eq. (10), assuming a bulk modulus of 500 GPa and an average phonon frequency of 8 THz. A Brillouin zone integration grid [23] of $8 \times 8 \times 8$ is used, and no external pressure is applied.

It takes 14 force/stress computations to reduce the stress and the forces to less than 10^{-3} GPa and 4×10^{-5} eV/a.u., respectively, thereby recovering the diamond structure. Using Eqs. (23) and (20), we compute the bulk modulus

to be 90 GPa and the optical phonon frequency at Γ to be 15 THz. This is close to the 96 GPa and 15.4 THz obtained from the Murnaghan equation of state [24] and a traditional frozen-phonon calculation. We cannot expect perfect agreement because of the imperfect line minimization and the anharmonicity of the enthalpy surface at the starting point. However, the accuracy is good enough to yield a reasonable estimate and provide insight into material properties. For example, when searching for hard materials [25], promising candidate structures can already be identified from the relaxation at ambient pressure, without going through the expensive computation of the equation of state.

4.2. *Relaxation of a 16-Atom Silicon Supercell*

To assess the performance of the quasi-Newton algorithm for a larger system, we apply it to a 16-atom supercell of silicon in the diamond structure, where the atoms are randomly displaced from the equilibrium positions with an amplitude of 0.05 a.u. in all three spatial directions. After displacing the atoms, the tips of the lattice vectors are distorted randomly with an amplitude of 5%. Finally, the volume is increased from 131.8 to 143.75 a.u.^{3} per atom. The resulting crystal has only primitive translations as symmetry operations, resulting in $16 \times 3 - 3 + 6 = 51$ degrees of freedom to be optimized. Brillouin zone integrations are performed on a $2 \times 2 \times 2$ grid [23]. No external pressure is applied. We initialize the H-matrix according to Eq. (10) with an estimated bulk modulus of 150 GPa and an average optical phonon frequency of 20 THz. Those differ intentionally from the 96 GPa and 15.4 THz of Section 4.1 in order to simulate realistic conditions, where the bulk modulus and the phonon frequencies are not known at the beginning of the calculation.

Figure 1 shows how the average force, the "stress" $(\sum_{ij}\sigma_{ij}^2)^{1/2}$, and the error in the enthalpy δH decrease as the relaxation proceeds. Since there are 51 degrees of freedom, it will take at least 51 relaxation steps to build up H completely. Thus in most cases, it takes two force/stress computations per relaxation step to perform the approximate line minimization. For the last three relaxation steps, and also the following three steps not shown on the graph, an approximate line minimization is not necessary, indicating that H has improved.

4.3. *Relaxation of the R8 Phase of Silicon*

The R8 phase of silicon under pressure has been observed experimentally [10] and has been studied with *ab-initio* calculations [26]. Its space group is $R\bar{3}$ with eight atoms in the rhombohedral unit cell. Two of the atoms are located at the Wyckoff positions 2(c) (u, u, u), the other six are on the 6(f) (x, y, z) sites [10]. Including the unit cell volume Ω and the angle α between the rhombohedral

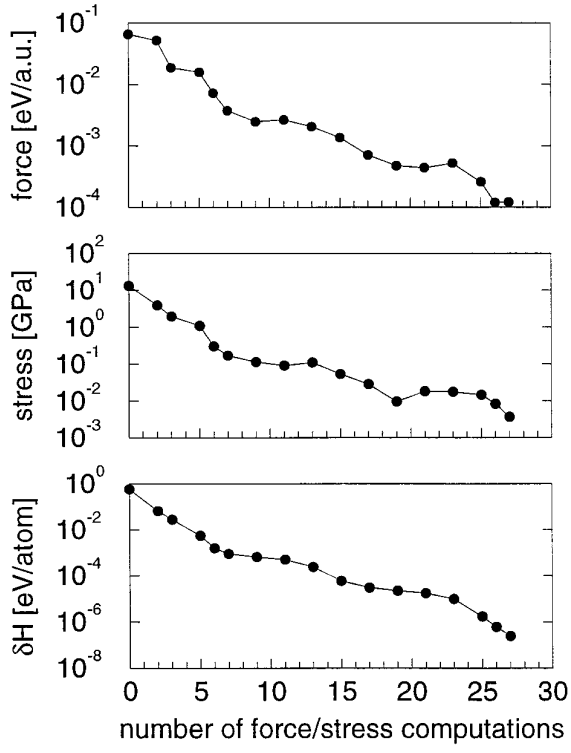

FIG. 1. Relaxation of the 16-atom silicon supercell. The atoms and the unit cell vectors have been randomly displaced from the perfect diamond structure, resulting in a system with 51 degrees of freedom. The average force, the "stress" $(\sum_{ij}\sigma_{ij}^2)^{1/2}$ and the error in the enthalpy δH are shown as a function of the number of force/stress computations. Each symbol represents a relaxation step, and most of the time it needs two force/stress computations to perform a relaxation step because of the line minimization.

lattice vectors, this structure has six degrees of freedom. It is thus a good example to show the performance of our method, because there are enough degrees of freedom to make a direct minimization impractical, and it is required that the symmetry of the crystal is preserved during the relaxation process. We use this test case also to demonstrate the benefits of a good starting guess for H.

We relax the R8 phase at several different pressures p, starting with $p = 8.2$ GPa, for which the experimentally observed parameters [26] are $\Omega = 902$ a.u.^{3}, $\alpha = 110.07^0$, $u = 0.2922$, $x = 0.4597$, $y = -0.0353$, and $z = 0.2641$. A $6 \times 6 \times 6$ grid [23] is used for the Brillouin zone integrations. The H-matrix is initialized assuming a bulk modulus of 100 GPa, and optical phonon frequencies of 15 THz. It takes 10 relaxation steps with a total of 16 force/stress computations to reduce the forces to less than 10^{-4} eV/a.u. and to converge the components of the stress tensor to better than 10^{-3} GPa. At the relaxed position, for $p = 8.2$ GPa, we find the computed structural parameters to be $\Omega = 861$ a.u.^{3}, $\alpha = 109.99^0$, $u = 0.280$, $x = 0.462$, $y = -0.034$, and $z = 0.269$.

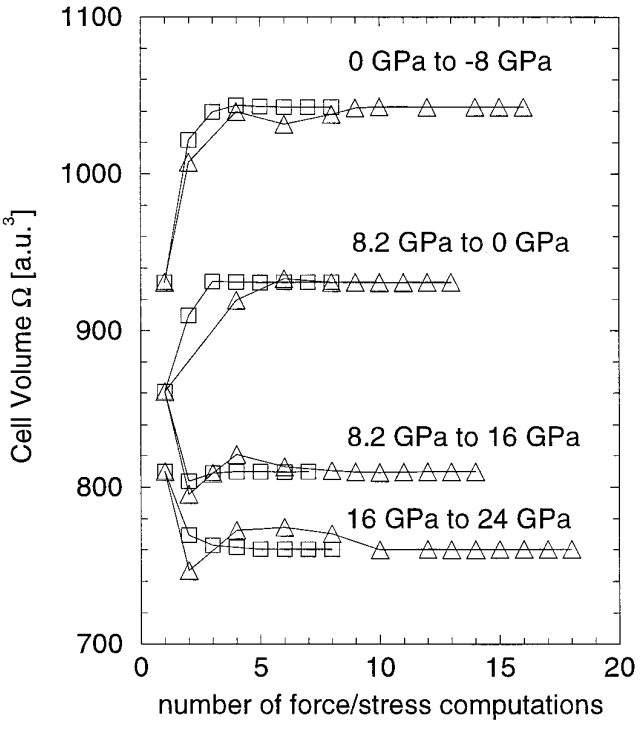

FIG. 2. Relaxation of the R8 phase of silicon, which has a total of six degrees of freedom. One of them—the unit cell volume—is shown as a function of the number of force/stress computations for relaxations at different pressure. Starting from the relaxed structural parameters obtained at 8.2 GPa, relaxations at 0 GPa and 16 GPa are performed. In the same way, relaxations from 0 GPa to −8 GPa, and from 16 GPa to 24 GPa are carried out. The triangles show how the volume develops if H is initialized with $B_0 = 100$ GPa and $\bar{\omega}_O/(2\pi) = 15$ THz. The relaxation is about twice as efficient if the H matrix of the starting point is used (squares). Each symbol represents a relaxation step, but with the inferior initialization of H (triangles), it often takes several force/stress computations per relaxation step, because the approximate line minimization has to be performed. This is especially the case during the first few relaxation steps, when H is not built up yet.

Starting with the computed parameters at 8.2 GPa, we increase the pressure to 16 GPa, and then to 24 GPa, relaxing after each increase in pressure. Analogously, we decrease the pressure from 8.2 GPa to 0 GPa, and from there to −8 GPa. Figure 2 shows how the cell volume changes during the relaxation processes. We perform the calculations with two different initializations of H: one intialized according to Eq. (10) with $B_0 = 100$ GPa, $\bar{\omega}_O/(2\pi) = 15$ THz (shown as triangles), and the other one (squares) taken from the fully builtup H of the previous pressure. For example to compute the square-marked curve from 8.2 GPa to 0 GPa, the final H obtained from the relaxation at 8.2 GPa is used. The convergence criterion is such that the forces are smaller than 10^{-4} eV/a.u., and the stress is accurate to better than 10^{-3} GPa.

Figure 2 shows the advantage of a superior initialization of H. Obviously, the H-matrix does not change too much with pressure, and carrying over H from a relaxation at a similar pressure cuts down the computational effort by half. Not only does a better H-matrix result in more efficient step directions, it also saves the line minimization, because the trial step of length $\lambda = 1$ in Eq. (7) is already sufficiently close to the minimum. With the superior starting guess for H, we find convergence after eight or fewer force/stress computations.

An analysis of H for $p = 0$, gives a bulk modulus of 95 GPa for the R8 structure, compared to 89 GPa from the Murnaghan equation of state [24]. Using Eq. (20), the frequencies for the symmetry-preserving A_g phonon modes at the Brillouin zone center are 4.0, 9.7, 10.7, and 14.1 THz.

5. SUMMARY

We propose an efficient quasi-Newton algorithm to simultaneously relax the internal coordinates and lattice parameters of a crystal while preserving its symmetry. As a by-product, elastic properties and some of the optical phonon modes can be estimated. We have demonstrated the efficiency of our method for silicon in a 16-atom diamond supercell and the R8 phase.

APPENDIX: SYMMETRY CONSERVATION

In this secton, we show that a relaxation with the BFGS scheme as outlined in Section 2 indeed preserves the symmetry of the crystal if H is properly initialized (10).

Let us first examine how a symmetry-compliant point X in configuration space must transform under a crystal symmetry operation $\{\alpha|\tau\}$, α being a unitary mirror-rotation matrix, and τ a nonprimitive translation vector. The first nine components of X are just the strain components ε, which have to remain invariant under the point symmetry operations:

$$\varepsilon' = \alpha^{-1}\varepsilon\alpha = \varepsilon \tag{24}$$

if the symmetry should be preserved. The other components of X are the positions s_i of the atoms in lattice coordinates, and they transform according to

$$s_i' = (h^{-1}\alpha h)\, s_i + h^{-1}\tau. \tag{25}$$

Since the symmetry operation $\{\alpha|\tau\}$ leaves the crystal invariant, it is always possible to find the permutation matrix P_{ij} which gives the index of the atom j into which atom i has been mapped:

$$s_i' = \sum_j P_{ij} s_j. \tag{26}$$

If the components of a configuration space point transform according to Eqs. (24) and (26), we shall call it *symmetric*.

The derivative F of the enthalpy with respect to X in Eq. (5) is *not* symmetric, but the step direction $\Delta X = HF$ is, as we will show below. The first nine components $f^{(\varepsilon)}$ of F are invariant under point symmetry operations according to Eqs. (3) and (5). The force f_i in lattice coordinates transform like the positions in Eq. (26),

$$f_i' = \sum_j P_{ij} f_j, \tag{27}$$

and is symmetric, but it is gf_i which forms the components of F, and this entity is not symmetric.

To show that ΔX is indeed symmetric at each step, we first notice that H can be decomposed into

$$H = H_0 + H^{(\text{update})}, \tag{28}$$

where the $H^{(\text{update})}$ is the sum of all updates (9). For HF to be symmetric, it is sufficient that the product of H_0 with the force vector F is symmetric. Given that, $H^{(\text{update})}F$ will be a linear combination of symmetric vectors by virtue of the updating formula (9), and so will be HF.

It remains to show that H_0F is always symmetric for our choice of H_0. According to Eq. (10), the stress components of F do not mix with the force components, nor do the forces on different atoms mix with each other when H_0 is applied. Since we initialize the strain components of H_0 to be a multiple of the identity matrix, the strain part of H_0F will trivially be symmetric. So are the atomic-position components Δs_i of H_0F, as we see from their transformation behavior under a symmetry operation $\{\alpha|\tau\}$. Because $h^{-1}\alpha h$ commutes with $g_0^{-1}g$, we have

$$\begin{aligned} \Delta s_i' &= h^{-1}\alpha h \bar{M}^{-1}\bar{\omega}_O^{-2}(g_0^{-1}gf_i) \\ &= \sum_j P_{ij}\bar{M}^{-1}\bar{\omega}_O^{-2}g_0^{-1}gf_j = \sum_j P_{ij}\Delta s_i, \end{aligned} \tag{29}$$

and, hence, H_0F is symmetric.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation Grant No. DMR-9520554 and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract DE-AC03-76SF-00098. S.G.L. acknowledges support from the Miller Institute for Research in Basic Science. We thank P. Delaney, Dr. F. Mauri, and O. Zakharov for helpful discussions.

REFERENCES

1. R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer Verlag, Berlin, 1990).
2. J. Ihm, A. Zunger, and M. L. Cohen, *J. Phys. C* **12,** 4409 (1979).
3. S. G. Louie, in *Electronic Structure, Dynamics, and Quantum Structural Properties of Condensed Matter*, edited by J. T. Devreese and P. V. Cassip (Plenum, New York, 1985), p. 335.
4. O. H. Nielsen and R. M. Martin, *Phys. Rev. B* **32,** 3792 (1985).
5. R. Car and M. Parrinello, *Phys. Rev. Lett.* **55,** 2471 (1985).
6. R. M. Wentzcovitch, J. L. Martins, and G. D. Price, *Phys. Rev. Lett.* **70,** 3947 (1993).
7. M. Parrinello and A. Rahman, *Phys. Rev. Lett.* **45,** 1196 (1980).
8. R. M. Wentzcovitch, *Phys. Rev. B* **44,** 2358 (1991).
9. J. Furthmüller, J. Hafner, and G. Kresse, *Phys. Rev. B* **50,** 15606 (1994).
10. J. Crain, G. J. Ackland, J. R. Maclean, R. O. Piltz, P. D. Hatton, and G. S. Pawley, *Phys. Rev. B* **50,** 13043 (1994).
11. S. Nosé and M. Klein, *Mol. Phys.* **50,** 1055 (1983).
12. W. H. Press *et al.*, *Numerical Recipes*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1992).
13. D. F. Shanno, *Math. Oper. Res.* **3,** 244 (1978).
14. V. Eyert, *J. Comput. Phys.* **124,** 271 (1996).
15. E. Polak, *Computational Methods in Optimization* (Academic Press, New York, 1971).
16. J. Wang, S. Yip, S. R. Phillpot, and D. Wolf, *Phys. Rev. Lett.* **71,** 4182 (1993).
17. A. A. Maradudin, in *Dynamical Properties of Solids*, edited by G. K. Horton and A. A. Maradudin (North-Holland, Amsterdam, 1974).
18. D. C. Wallace, *Thermodynamics of Crystals* (Wiley, New York, 1972).
19. D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45,** 566 (1980).
20. N. Troullier and J. L. Martins, *Phys. Rev. B* **43,** 1993 (1991).
21. M. T. Yin and M. L. Cohen, *Phys. Rev. Lett.* **50,** 1172 (1983).
22. J. Crain, S. J. Clark, G. J. Ackland, M. C. Payne, V. Milman, P. D. Hatton, and B. J. Reid, *Phys. Rev. B* **49,** 5329 (1994).
23. H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13,** 5188 (1976).
24. F. D. Murnaghan, *Proc. Nat. Acad. Sci. USA* **30,** 244 (1944).
25. M. Côté and M. L. Cohen, to be published.
26. R. O. Piltz, J. R. Maclean, S. J. Clark, G. J. Ackland, P. D. Hatton, and J. Crain, *Phys. Rev. B* **52,** 4072 (1995).