Geometry optimization of periodic systems using internal coordinates

Tomáš Bučko\textsuperscript{b}) and Jürgen Hafner

\textit{Computational Materials Science, Institut für Materialphysik, Universität Wien, Sensengasse 8/12, A-1090 Wien, Austria}

János G. Ángyán\textsuperscript{b})

\textit{Laboratoire de Cristallographie et de Modélisation des Matériaux Minéraux et Biologiques, UMR 7036, CNRS-Université Henri Poincaré, Boîte Postale 239, F-54506 Vandœuvre-lès-Nancy, France}

(Received 16 November 2004; accepted 11 January 2005; published online 30 March 2005)

An algorithm is proposed for the structural optimization of periodic systems in internal (chemical) coordinates. Internal coordinates may include in addition to the usual bond lengths, bond angles, out-of-plane and dihedral angles, various “lattice internal coordinates” such as cell edge lengths, cell angles, cell volume, etc. The coordinate transformations between Cartesian (or fractional) and internal coordinates are performed by a generalized Wilson B-matrix, which in contrast to the previous formulation by Kudin et al. [J. Chem. Phys. 114, 2919 (2001)] includes the explicit dependence of the lattice parameters on the positions of all unit cell atoms. The performance of the method, including constrained optimizations, is demonstrated on several examples, such as layered and microporous materials (gibbsite and chabazite) as well as the urea molecular crystal. The calculations used energies and forces from the \textit{ab initio} density functional theory plane wave method in the projector-augmented wave formalism. © 2005 American Institute of Physics. DOI: 10.1063/1.1864932

I. INTRODUCTION

Recently, there has been an increasing interest in electronic structure calculations of three-dimensional periodic systems. Several high-performance density functional theory (Refs. 1–5) and Hartree–Fock\textsuperscript{5,6} codes are available, permitting the study of structural features by gradient optimization. Until very recently, geometry relaxations in solids were done exclusively in terms of Cartesian and/or fractional coordinates, and lattice vector components.\textsuperscript{1,4,7–9} Although the Cartesian/fractional coordinates are simple and universally applicable, the use of internal (atomic and lattice) coordinates as control parameters in geometry optimizations offers several advantages. As it has been extensively demonstrated for molecular examples,\textsuperscript{10–12} in internal coordinates (i) it is easy to have a good initial Hessian matrix guess, (ii) the coupling of different modes is reduced in comparison with the Cartesian coordinates, and (iii) the handling of constraints is simple and straightforward.

Inspired by the experience in the domain of finite molecular systems, several groups have attempted to improve the convergence of geometry optimization of solids by adopting alternative coordinate systems. Approximate normal mode coordinates derived from a simple model of the dynamical matrix were proposed by Fernández-Serra et al.\textsuperscript{13} A scaling of these coordinates by estimated force constants considerably improves the condition number (ratio of highest and lowest eigenvalue) of the Hessian and accelerates convergence. Another approach is based on the construction of a redundant set of primitive internal coordinates, such as bond lengths, valence angles, torsional angles, etc., that are translationally unique. Solid optimizations can be either performed in these redundant coordinates directly\textsuperscript{14} or in a non-redundant linear combination of them, in delocalized internal coordinates.\textsuperscript{15} To our knowledge, the problem of lattice vector optimizations in the context of internal coordinates has been discussed only in Ref. 14.

In the present work we describe an approach that uses delocalized internal coordinates for the optimization of atomic positions, similar to Ref. 15. The lattice parameter optimization follows the principles of the Kudin’s method,\textsuperscript{14} with one significant improvement: while in their procedure cell parameter variations are described by a small subset of primitive internal coordinates, our method treats all atoms of the unit cell on an equal footing letting them contribute to cell parameter variations.

In Sec. II, after a short reminder of the Wilson B-matrix formalism\textsuperscript{16} for geometry optimizations in internal coordinates, the necessary extensions for the periodic case are discussed along with the handling of constraints. Some results are presented in Sec. III, demonstrating that our procedure allows one to optimize efficiently atomic coordinates and cell parameters simultaneously. The gain in the number of iterations for structures relaxed to a comparable degree of precision with respect to Cartesian relaxations with a unit Hessian matrix guess ranges from a factor of 2 to 10. In addition to the better convergence properties, our approach offers a significantly increased versatility in performing constrained lattice optimizations.

\textsuperscript{a}Electronic mail: tomas.bucko@univie.ac.at

\textsuperscript{b}Electronic mail: janos.angyan@lcm3b.uhp-nancy.fr
II. METHOD

A. Cartesian and fractional coordinates

The structure of a periodic system is characterized by three lattice vectors, arranged in the matrix \( \mathbf{h} = [a_1, a_2, a_3] \) and by the 3N fractional coordinates, \( s_i \). The Cartesian positions of atom \( a \) in the \( \mathbf{L} = (l_1, l_2, l_3) \) unit cell of the solid is given by the linear transformation

\[
s_a = \sum_{\beta=1}^{3} h_{a\beta} s_\beta^a + r_a^0.
\]

The three lattice vectors form, in the general case, an unnormalized and nonorthogonal basis. The fractional coordinates can be obtained from the Cartesian ones by the inverse transformation in terms of the reciprocal lattice vectors, \( (\mathbf{h}^i)^{-1} = [a_1, a_2, a_3] \), defined by \( a_i^* = \delta_{ij} \) as

\[
s_a^i = \text{mod} \left( \sum_{\beta} h_{a\beta} r_{a\beta}^i, s_a^i \right).
\]

The structural deformations in periodic systems arise from the change of atomic positions at constant lattice vectors on the one hand, and from the deformation of the lattice vectors at constant fractional coordinates. The variation of lattice vectors is conveniently described by the strain tensor \( \varepsilon_{a\beta} \) as

\[
h_{a\beta}' = \sum_{\gamma} (\delta_{a\gamma} + \varepsilon_{a\gamma}) h_{\gamma\beta}.
\]

Any change in \( h_{a\beta} \) implies a variation of atomic Cartesian coordinates by Eq. (1) which obey an analogous relationship

\[
r_a' = \sum_{\beta} (\delta_{a\beta} + \varepsilon_{a\beta}) r_\beta.
\]

Out of the total number of 3N+9 deformation variables (3N atomic coordinates and 9 strain tensor elements) the energy is invariant with respect to six degrees of freedom: the position of the origin and the orientation of the lattice vectors. Although the origin of the lattice is arbitrary, it is usually chosen on the basis of symmetry considerations. Overall rotations of the lattice can be avoided by taking the symmetric part of the strain tensor, \( \varepsilon_{a\beta} = \frac{1}{2} (\varepsilon_{a\beta} + \varepsilon_{\beta a}) \), or the metric matrix formed from the six unique scalar products of lattice vectors, \( \mathbf{F} \) as independent lattice variables.

Most of the geometry optimization algorithms are based on a harmonic expansion of the total energy around an initial structure. Using exact first and approximate second derivatives, a convenient step is predicted that takes the system closer to the desired critical point (minimum or saddle point) with vanishing gradients. In periodic systems the simplest coordinate system is constructed from the 3N fractional atomic positions and nine lattice vector component variations, \( \mathbf{\delta x} = (\mathbf{\delta s}, \mathbf{\delta h}) \) leading to the following expansion:

\[
E(x + \delta x) - E(x) = - f^0 \cdot \delta x + \frac{1}{2} \delta x^T \mathbf{F} \delta x + \cdots + .
\]

The elements of the column vector \( f \) are either the forces, \( f_a^0 = -dE/ds_a^0 \), or the lattice vector derivatives of the total energy, \( f_a^\beta = -dE/dh_{a\beta} \). The latter are related to the stress tensor elements, i.e., the negative volume-normalized strain derivatives of the energy:

\[
\sigma_{a\beta} = - \frac{1}{\Omega} \frac{dE}{dh_{a\beta}}
\]

by the relationship (cf. the Appendix)

\[
\frac{dE}{dh_{a\beta}} = - \Omega \sum_{\gamma} \sigma_{a\gamma} (h^i)^{-1}_{\gamma\beta}.
\]

Electronic structure codes usually provide Cartesian force components, which can be transformed to fractional coordinate forces as

\[
\frac{dE}{ds_a^i} = \sum_{\beta} \frac{\partial F_{a\beta}}{\partial s_a^i} \frac{dE}{dh_{a\beta}^i} = \sum_{\beta} \frac{\partial F_{a\beta}}{\partial s_a^i} \sum_{\gamma} \sigma_{a\gamma} (h^i)^{-1}_{\gamma\beta} + \sum_{\gamma} \frac{\partial F_{a\beta}}{\partial s_a^i} \sigma_{a\gamma} (h^i)^{-1}_{\gamma\beta}.
\]

It should be noted that in addition to the explicit lattice vector dependence of the total energy, atomic forces contribute also to the strain-derivative tensor, according to the following expression:

\[
\frac{dE}{dh_{a\beta}} = \frac{\partial E}{\partial h_{a\beta}} + \sum_{\alpha} \frac{\partial E}{\partial s_{\alpha}} \frac{dE}{dh_{a\beta}} = \frac{\partial E}{\partial h_{a\beta}} + \sum_{\alpha} \frac{\partial E}{\partial s_{\alpha}} h_{\alpha\beta}.
\]

As we shall see later, in Sec. II D, the same relationships can be obtained directly from the B-matrix formalism, to be developed below.

The extended Hessian matrix \( \mathbf{F} \) in Eq. (5) is built up from several blocks: that of the second derivatives of the energy with respect to atomic positions (related to the dynamical matrix), that of the stress-strain derivatives, as well as cross terms.

B. Curvilinear internal coordinates for periodic systems

In contrast to the set of external coordinates, \( \mathbf{delta x} = (\mathbf{delta s}, \mathbf{delta h}) \), including overall rotations and translations of the system, one can define a set of internal coordinates, \( \delta \xi = (\delta q, \delta \tilde{q}) \), which involve only internal degrees of freedom. Internal atomic coordinates, such as bond lengths, bond angles, torsion angles, etc., or their linear combinations are, in general, nonlinear functions of Cartesian coordinates, \( q_i = f((r_a^{1L})) \), and by the virtue of the linear relationship Eq. (1), also of the fractional coordinates and the lattice vectors. Similarly, internal lattice coordinates, such as cell edge length, cell angle, volume, etc., are determined by the three lattice vectors, \( \delta \tilde{q}_i = f((h_{a\beta})) \).

Internal coordinate \textit{deformations} are related to external coordinate deformations by a nonlinear (curvilinear) transformation, usually approximated by a truncated Taylor expansion

\[
\delta \xi_i = \sum_{s} \frac{\delta \xi_i}{\delta x_s} \delta x_s + \frac{1}{2} \sum_{s, t} \frac{\delta^2 \xi_i}{\delta x_s \delta x_t} \delta x_s \delta x_t + \cdots
\]

\[
= (\mathbf{B} \delta \mathbf{x})_i + \frac{1}{2} \delta \mathbf{x}^T \mathbf{C} \delta \mathbf{x} + \cdots
\]

Inserting expansion (10) into the harmonic expansion of the total energy with respect to internal coordinates

\[
E(x + \delta x) - E(x) = - f^0 \cdot \delta x + \frac{1}{2} \delta x^T \mathbf{F} \delta x + \cdots
\]
\[
E(\xi + \delta\xi) - E(\xi) = -\varphi^T \delta\xi + \frac{1}{2} \delta\xi^T \mathbf{H} \delta\xi + \cdots
\]

and equating terms of the same order, we obtain a relationship between external and internal force/stress components

\[
\mathbf{B}^T \varphi = f,
\]

while the external coordinate deformations can be obtained from a set of internal distortions by the first-order relationship

\[
\mathbf{B} \delta x = \delta \xi.
\]

The matrix \( \mathbf{B} \) is a generalization of Wilson’s B-matrix\(^\text{16} \) for periodic systems. As in the molecular case, the number of internal coordinates that can be constructed for an \( N \)-atom periodic structure is usually much more than the \( 3N+3 \) geometrical degrees of freedom. Since \( \delta x, f \in \mathbb{R}^{3N+9} \) and \( \delta\xi, \varphi \in \mathbb{R}^M, B \in \mathbb{R}^{M \times 3N+9}, \) the B-matrix (unless \( M=3N+3 \) and the set of internal coordinates is nonredundant) and the solutions of the above equations are given by the Moore–Penrose pseudoinverses as

\[
\varphi = (\mathbf{B}^T)^+ f, \quad \delta\xi = \mathbf{B}^+ \delta\xi.
\]

The same B-matrix and its pseudoinverse appear in the transformation of the fractional (\( \mathbf{F} \)) and internal (\( \mathbf{H} \)) second derivative matrices

\[
\mathbf{F} = \mathbf{B}^T \mathbf{H}, \quad \mathbf{H} = (\mathbf{B}^T)^+ \mathbf{F} \mathbf{B}^+,
\]

where the correction term involving the second derivative of the internal coordinates with respect to the fractional ones has been neglected.\(^\text{12} \)

In the case of really large systems (>1000 atoms), the coordinate and force transformation steps may become the bottleneck of the optimization procedure. Various methods have been proposed to solve efficiently the above equations for extended systems.\(^\text{19–23} \) In the present implementation delocalized internal coordinates\(^\text{15,21,24,25} \) are used, which are particularly appropriate for medium-sized problems.

### C. Delocalized internal coordinates

Even for small systems, the number of generated internal coordinates is usually much larger than the number of ionic degrees of freedom (i.e., \( 3N-6 \) for molecules and \( 3N+3 \) for systems with three-dimensional periodicity). The handling of this set of redundant coordinates may lead to a significant increase of the computational time and may cause convergence problems in the coordinate back-transformation step. To avoid these problems, Baker et al. proposed the use of nonredundant linear combinations of primitive internal coordinates.\(^\text{24} \) Taking the singular value decomposition of the B-matrix, as \( \mathbf{B} = \mathbf{U} \mathbf{S} \mathbf{B} \mathbf{V} \), where \( \mathbf{U} \) and \( \mathbf{V} \) are unitary matrices and \( \mathbf{B} \) is a diagonal matrix formed from the eigenvectors associated with the nonzero eigenvalues of \( \mathbf{B} \mathbf{B}^T \). After multiplication of the B-matrix equation from left by \( \mathbf{U} \), one obtains

\[
\mathbf{B}_V \mathbf{V} x = \mathbf{U} \mathbf{q}.
\]

which can be considered as the defining relationship of the delocalized internal coordinate transformation:

\[
\tilde{\mathbf{B}} x = \tilde{\mathbf{q}}
\]

with \( \tilde{\mathbf{B}} = \mathbf{B} \mathbf{V} \) and \( \tilde{\mathbf{q}} = \mathbf{U} \mathbf{q} \). Note that \( \mathbf{B}^+ \), the pseudoinverse of \( \tilde{\mathbf{B}} \), is simply \( \mathbf{V} \mathbf{B}^{-1} \mathbf{V}^T \). The transformation relations Eqs. (14)–(17), for the primitive internal coordinates remain valid for the delocalized internal coordinates after switching from \( \mathbf{B} \) to \( \tilde{\mathbf{B}} \).

### D. B-matrix for periodic systems

The construction of the B-matrix for periodic systems, defined by atomic position and lattice vector components, needs some special consideration as compared to the molecular case. Kudin et al. remarked\(^\text{14} \) that primitive internal coordinates involving atoms that belong to different unit cells are explicitly dependent on the lattice vector components \( h_{\alpha\beta} \). They have written expression (1) in the following form:

\[
r_{\alpha} = r_{\alpha0} + \sum_{\beta} h_{\alpha\beta} l_{\beta},
\]

and generated the corresponding B-matrix elements by the application of the chain rule. In their formulation only intercell coordinates (\( \mathbf{L} \neq 0 \)) are allowed to contribute to the lattice B-matrix elements. The optimization of lattice parameters is treated in an indirect manner, through intercell distances between atoms and their periodic images as well as through angles between three replicates of the same atom lying in different unit cells. This procedure seems to be well adapted to simple molecular crystals, but it is much less convenient for ionic systems, oxides or atomic lattices, where the distinction between intracell and intercell coordinates is less obvious. In the following we propose a “democratic” approach taking into consideration that any unique internal coordinate deformation may influence the lattice parameters.

Let us consider the augmented B-matrix equation

\[
\begin{pmatrix} \delta\mathbf{q} \\ \delta\xi \end{pmatrix} = \begin{pmatrix} \mathbf{B}^{\alpha\xi} & \mathbf{B}^{\alpha\ell} \\ \mathbf{B}^{\ell\xi} & \mathbf{B}^{\ell\ell} \end{pmatrix} \begin{pmatrix} \delta\mathbf{q} \\ \delta\xi \end{pmatrix},
\]

where the individual blocks \( \mathbf{B}^{\alpha\xi} \) and \( \mathbf{B}^{\ell\xi} \) describe the linear transformation of the atomic positions \( \mathbf{s}_i^{\alpha} \), while the blocks \( \mathbf{B}^{\alpha\ell} \) and \( \mathbf{B}^{\ell\ell} \) describe transformations involving the lattice vector components \( h_{\alpha\beta} \) to atomic \( \tilde{\mathbf{q}}_i \) and lattice internal \( \tilde{\mathbf{q}}_j \) coordinates. The blocks of the augmented B-matrix can be calculated from the relationships (1) and (2) using the chain rule.

An internal coordinate may depend on the Cartesian coordinates of atoms in different unit cells. Therefore the B-matrix elements between fractional atomic distortions and unique internal coordinates should be calculated as
Disregarding the transformation from Cartesian to fractional coordinates, the B-matrix elements for internal coordinates that “remain” entirely in the unit cell are identical to those of a nonperiodical system. If several translated copies of the same atom participate in a given internal coordinate, the contributions from different cells should be summed up. This leads to somewhat unexpected consequences. For instance, in the case of a monoatomic lattice, a zero B-matrix element leads to somewhat unexpected consequences. For instance, the following relationship holds for \( q \hat{h} \):

\[
\frac{\partial q}{\partial h_{\alpha\beta}} = \sum_{a} \sum_{\gamma} \frac{\partial q_{a}}{\partial h_{\alpha\beta}} (s_{a}^{\gamma} + f_{a}^{\gamma})
\]

where the superscript \( qh \) refers to the type of B-matrix element.

The “lattice internal” coordinates \( q \), such as cell edge lengths, \( a/b \) ratio, cell angles, volume, etc., do not depend explicitly on the fractional atomic positions, therefore the block of the extended B-matrix is zero, \( B^{qh} = 0 \). The transformation between the lattice vector changes and various “internal” lattice parameters is described by the matrix elements

\[
B^{qh}_{\alpha\beta} = \frac{\partial q}{\partial h_{\alpha\beta}}.
\]

For instance, the following relationship holds for \( q_{i} = \Omega \), the unit cell volume:

\[
B^{qh}_{\Omega,\alpha\beta} = \Omega (h')^{-1}.
\]

Other specific lattice B-matrix elements can also be derived from the definition of lattice vector lengths and lattice angles.

In order to appreciate the role of the extended B-matrix in lattice optimizations, let us consider the special case of Cartesian coordinates as internals, i.e., \( \delta \xi = \{ \delta \mathbf{r}, \delta \mathbf{h} \} \). According to the coordinate transformation relationship, Eq. (13),

\[
\delta \mathbf{r} = B^{\gamma} \delta \mathbf{s} + B^{qh} \delta \mathbf{h}
\]

and using the matrix elements given by Eqs. (22) and (23), we retrieve the expected result that the variation of a Cartesian coordinate can be decomposed into a variation of the fractional coordinate, \( \delta s_{a} \), and a variation of the lattice vector components, \( \delta h_{\alpha\beta} \).

The transformation equations for the forces and stresses, Eq. (12), lead to the following relationships [cf. Eqs. (8) and (9)] between the energy derivatives with respect to fractional coordinates, cell parameters, and Cartesian coordinates:

\[
\frac{dE}{ds_{a}} = \sum_{\beta} (h_{a\beta})^{\gamma} \frac{dE}{d\beta},
\]

\[
\frac{dE}{dh_{a\beta}} = \sum_{a} (s_{a}^{\alpha} + f_{a}^{\alpha}) \frac{dE}{ds_{a}} + \frac{\partial E}{d\alpha\beta}.
\]

The first of these equations is a simple linear transformation of the force components from a Cartesian to a lattice vector reference system. The second equation tells that the total derivatives with respect to the lattice parameters have a contribution from the partial derivative of the energy with respect to the lattice parameters (explicit dependence) and a “virial contribution” proportional to the atomic force components. The latter term, which obviously vanishes if the atomic forces are zero, is analogous to the expression of the virial pressure discussed by several authors in a somewhat different context.

E. Constraints

Constrained geometry optimizations are helpful and even necessary in most of the applications of \( ab \) \( initio \) calculations to chemical reactions, phase transitions, etc. The general strategy is to make vanish, exactly or at least approximately, forces along the constrained coordinates to avoid deformations involving the variation of the constrained coordinates. The principal advantage of using internal coordinates is that one can impose exact internal coordinate constraints during the optimization. Various algorithms, such as the use of projection operator,\(^{23}\) orthogonalization,\(^{24}\) and Lagrange multiplier\(^{31}\) techniques have been proposed in the past. Our implementation follows essentially the orthogonalization algorithm of Ref. 24.

In the first step, the B-matrix is modified in such a way that first derivatives of the active coordinates (those coordinates which are allowed to be relaxed) are orthogonalized with respect to each constrained coordinate \( q_{c} \). The rows \( B_{j} \) are modified according to the formula

\[
\bar{B}_{j} = B_{j} - \sum_{c} \left( B_{c}^{j} \right) \left( B_{c}^{\gamma} \right)^{-1} (B_{c}^{\gamma} B_{c}^{\gamma})^{-1} B_{c}^{\gamma}.
\]

where the summation is over the constrained coordinates. If one of the rows of the original B-matrix is identical to a constrained coordinate, it is exactly annihilated in this step.

Delocalized internal coordinates and corresponding gradients are generated from the modified matrix \( \bar{B} \), as described in previous sections. The delocalized coordinates are
Farkas and Schlegel suggested an improved DIIS algorithm, which allows one to adjust automatically the dimension of the iterative subspace. The idea is that the ionic step produced by DIIS $\Delta \vec{q}_i = \vec{q}_{i+k} - \vec{q}_k$ is compared to a simple quasi-Newton step $\Delta \vec{q}_{QN} = \bar{H}^{-1} \bar{\vec{q}}$, and should meet the following criteria.

(i) The direction of the DIIS step $\Delta \vec{q}$ deviates from $\Delta \vec{q}_{QN}$ by an angle $\phi$. The step is accepted if $\cos(\phi)$ is larger than 0.97, 0.84, 0.71, 0.67, 0.62, 0.56, 0.49, and 0.41 for two to nine recent relaxation steps used in DIIS. For a dimension of ten or higher this criterion is not taken into account.

(ii) The norm of DIIS step is limited to be not more than ten times larger than that of the reference step.

(iii) The sum of all positive coefficients $c_i$ should not exceed the value of 15.

(iv) The magnitude of $c_i |r_i^2|$ should not exceed $10^8$.

If one of these criteria is not fulfilled, the step is not accepted and the most remote vector is removed from the iterative subspace. This procedure is repeated until all criteria are fulfilled.

One of the major advantages of the use of internal instead of Cartesian coordinates is that a reasonable guess for a Hessian matrix can be constructed. Even a very simple model Hessian which is just a diagonal matrix with elements 0.5, 0.2, and 0.1 a.u. for bonds, angles, and torsions, respectively, usually works very well. Lindh et al. proposed a model Hessian, constructed from force constants that are simple functions of nuclear positions

$$ k_{ij} = k_{ij}^0, $$

$$ k_{ij} = k_{ij}^\phi p_{ij}^\phi, $$

$$ k_{ijkl} = k_{ij} p_{ij}^0 p_{kl}, $$

with

$$ p_{ij}^\phi = \exp[\alpha_{ij}(r_{ij}^2 - r_{ij}^0)]. $$

For the first three rows of the periodic table one needs altogether 15 independent parameters for the quantities $k_r$, $k_{\phi}$, $k_{\tau}$, $\alpha_{ij}$, and $r_{0,ij}$ that are collected in Table I. The model Hessian can be easily transformed to delocalized internal coordinates using the formula

$$ \bar{H} = U^T H U. $$

In the course of the relaxation, the Hessian matrix is updated using the Broyden–Fletcher–Goldfarb–Shanno algorithm:

$$ \bar{H}_k = \bar{H}_{k-1} - \frac{\Delta \bar{q}_k \Delta \bar{q}_k^T}{\Delta \bar{q}_k^T \Delta \bar{q}_k} - \frac{\bar{H}_{k-1} \Delta \bar{q}_{k-1} \Delta \bar{q}_{k-1}^T}{\Delta \bar{q}_{k-1}^T \Delta \bar{q}_{k-1}}, $$

where $\Delta \bar{q}_k = \bar{q}_k - \bar{q}_{k-1}$ is the change of gradients associated with the relaxation step $\Delta \bar{q}_k = \bar{q}_k - \bar{q}_{k-1}$.

The performance of our optimization engine GADGET is checked against the "native" Vienna ab initio simulation package (VASP) optimizer using the conjugate gradient algorithm and performed in cartesian coordinates. The conju-
the plane wave basis-set incompleteness error.\(^{37}\) The applied cutoff energy of 800 eV is pretty high and it is usually assumed that in this case the Pulay stress is negligible. Somehow unexpectedly, we have found relatively small but nevertheless significant residual stresses at the equilibrium volume obtained by fitting a Murnaghan equation of state to a series of fixed-volume relaxed structures. This residual stress (Pulay stress) can be added as a constant in the automatic relaxation procedure, in order to correct the finite-basis error in the direct lattice optimizations.

Different kinds of optimizations were performed, all starting from the same initial geometry. (i) Only the atomic positions are optimized at constant lattice parameters, using delocalized internal coordinates (GADGET). (ii) Atomic position relaxation with the native VASP optimizer. (iii) Simultaneous atomic parameter and lattice parameter optimization in delocalized internal coordinates with GADGET. (iv) Full atomic position and lattice parameter relaxation with VASP. (v) Multistep geometry relaxation with GADGET using a series of fixed-volume relaxations to determine a Murnaghan-type equation of state (EOS) and reoptimizing the structure at the volume of the interpolated minimum. (vi) Full delocalized internal coordinate relaxation using the Pulay stress deduced from the EOS optimization.

### III. APPLICATIONS

The electronic structure calculations were done with the VASP,\(^{1}\) using the density functional theory in the generalized gradient approximation. The PW91 exchange-correlation functional was used, using the projector-augmented wave formalism.\(^{35,36}\) The calculations were done with high precision, i.e., the wave function was developed on a plane wave basis with a cutoff of 800 eV, and the support grid for the representation of the charge density was sufficiently precise to avoid any wraparound errors. The electronic wave functions were converged in each step to \(2.7 \times 10^{-8}\) hartree. It is expected that these computational parameters allow us to obtain quite reliable forces.

The geometry optimizations were done with the external optimizer GADGET, written in Python. GADGET reads the geometry, energy, and gradients from the VASP output, sets up internal coordinates, estimates an optimal move, calculates the new set of lattice parameters and Cartesian coordinates, and starts a new VASP calculation, until convergence. As we shall see, the additional overhead related to the repeated restarts of VASP is largely compensated by the gain in the number of iterations. Our optimizer can be easily interfaced with other packages that calculate total energies and forces/stresses, and such an interface is already operational with the GAUSSIAN 03 package.\(^{5}\) Optimizations were carried out using the geometrical DIIS method with an iterative subspace of dimension 5. Convergence criteria involve the simultaneous fulfillment of an energy change less than \(1 \times 10^{-6}\) a.u., a maximal gradient less than \(2 \times 10^{-4}\) a.u., and a volume change (in lattice relaxations) that is smaller than 0.05%.

In lattice parameter optimizations, involving volume changes, we are faced with the problem of “Pulay stress,” i.e., the gate gradient algorithm tries to improve a simple steepest descent step (i.e., in the direction of gradient) by conjugating the search direction from the most recent step. It starts with the steepest descent direction, but in all the following steps the search direction \(s_{k+1}\) is given by

\[
g = \frac{f_k + f_{k-1} \cdot f_g}{f_{k-1} \cdot f_{k-1}},
\]

\[
s_{k+1} = f_k + \gamma s_k.
\]

The conjugate gradient algorithm requires the line minimization along search direction.

### A. Microporous material: SiO\(_2\) chabazite

Chabazite (Fig. 1) is microporous aluminosilicate mineral (zeolite) in which every Si (Al) atom is coordinated by four O atoms. The space groups of the purely siliceous cha-

---

**TABLE I.** Parameters of the model Hessian proposed by Lindh \textit{et al.} (Ref. 34). Indices \(i\) and \(j\) designate the rows of the periodic table of elements to which the atoms correspond. The universal stretch, bend, and torsion coefficients are \(k_{ij} = 0.45\), \(k_{ij} = 0.15\), and \(k_{ij} = 0.005\). All quantities are given in atomic units.

<table>
<thead>
<tr>
<th></th>
<th>First period</th>
<th>Second period</th>
<th>Third period</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_{ij})</td>
<td>1.0000</td>
<td>0.3949</td>
<td>0.3949</td>
</tr>
<tr>
<td>(r_{0ij})</td>
<td>1.35</td>
<td>2.10</td>
<td>2.53</td>
</tr>
</tbody>
</table>

---

**FIG. 1.** Rhombohedral unit cell of purely siliceous chabazite, Si\(_{12}\)O\(_{24}\).
TABLE II. Optimization of the chabazite structure: the initial configuration (Initial); relaxation of the ionic degrees of freedom using delocalized internal coordinates \([\langle s \rangle]\) and with the VASP native optimizer \([\langle s, h \rangle]\); relaxation of the lattice parameters and ionic positions using delocalized internal coordinates \([\langle q, \dot{q} \rangle]\) and with the VASP native optimizer \([\langle s, h \rangle]\); minimum of the fitted Murnaghan equation of state (EOS); and the full relaxation using delocalized internal coordinates with stress corrected for a Pulay stress \([\langle q, \dot{q} ; \sigma \rangle]\). Listed are cell parameters \([a, b, c \, (\text{Å}); \alpha, \beta, \gamma \, (\text{deg.})]\), energy (hartree), gradient (a.u.), stress (kB), cell volume (Å³), and the number of relaxation steps.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial</th>
<th>([\langle q \rangle])</th>
<th>([\langle s \rangle])</th>
<th>([\langle q, \dot{q} \rangle])</th>
<th>([\langle s, h \rangle])</th>
<th>EOS</th>
<th>([\langle q, \dot{q} ; \sigma \rangle])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>9.301(9)</td>
<td>9.349(7)</td>
<td>9.354(1)</td>
<td>9.349(6)</td>
<td>9.351(6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>9.307(2)</td>
<td>9.344(5)</td>
<td>9.380(9)</td>
<td>9.349(6)</td>
<td>9.347(6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>9.303(9)</td>
<td>9.347(7)</td>
<td>9.347(6)</td>
<td>9.349(9)</td>
<td>9.350(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\alpha)</td>
<td>94.122(4)</td>
<td>94.263(5)</td>
<td>94.113(4)</td>
<td>94.179(4)</td>
<td>94.158(3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\beta)</td>
<td>94.058(9)</td>
<td>94.286(7)</td>
<td>94.228(1)</td>
<td>94.186(0)</td>
<td>94.179(9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\gamma)</td>
<td>94.104(5)</td>
<td>94.280(5)</td>
<td>94.166(0)</td>
<td>94.187(8)</td>
<td>94.176(0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(E)</td>
<td>−10.378557</td>
<td>−10.535669</td>
<td>−10.53566</td>
<td>−10.53709</td>
<td>−10.53710</td>
<td></td>
<td>−10.53710</td>
</tr>
<tr>
<td>(g_{max}^s)</td>
<td>0.13218</td>
<td>0.00015</td>
<td>0.00015</td>
<td>0.00011</td>
<td>0.00005</td>
<td>0.00014</td>
<td></td>
</tr>
<tr>
<td>(\sigma_{xx})</td>
<td>36.30</td>
<td>36.00</td>
<td>36.00</td>
<td>36.00</td>
<td>36.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sigma_{yy})</td>
<td>35.74</td>
<td>35.74</td>
<td>35.74</td>
<td>35.74</td>
<td>35.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sigma_{zz})</td>
<td>17.15</td>
<td>17.15</td>
<td>17.15</td>
<td>17.15</td>
<td>17.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sigma_{xy})</td>
<td>2.17</td>
<td>2.17</td>
<td>2.17</td>
<td>2.17</td>
<td>2.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sigma_{yz})</td>
<td>24.34</td>
<td>24.34</td>
<td>24.34</td>
<td>24.34</td>
<td>24.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sigma_{zx})</td>
<td>4.17</td>
<td>4.17</td>
<td>4.17</td>
<td>4.17</td>
<td>4.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Omega)</td>
<td>799.00</td>
<td>799.00</td>
<td>799.00</td>
<td>799.00</td>
<td>799.00</td>
<td>809.50</td>
<td>809.69</td>
</tr>
<tr>
<td>(N_{rel})</td>
<td>22</td>
<td>25</td>
<td>25</td>
<td>62</td>
<td>62</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Layered oxide: Gibbsite

Unlike chabazite, gibbsite \([\text{Al(OH)}_3]\) (see Fig. 2) is a layered mineral. The individual layers are charge balanced which results into very weak interlayer interactions. The primitive cell of gibbsite is monoclinic with space group symmetry \(P12_1/n1\) (No. 14). There has been several recent \(ab\) \(initio\) structural studies on gibbsite and other \(\text{Al(OH)}_3\) polymorphs and we are currently working on the theoretical vibrational spectrum of gibbsite and bayerite in light of some new experimental results.

In this system the conventional primitive internal coordinates (bonds, angles, and torsions) are insufficient to describe all atomic degrees of freedom. In order to describe correctly the mutual position of subsequent layers we have used inverse power distance coordinates \((5/r)\) coordinates) proposed by Baker and Pulay. The \(5/r\) coordinates were generated only for pairs of atoms lying in different structural segments (layers) and satisfying the condition that the corresponding interatomic distance is smaller than 1.6 times the sum of the covalent radii. In this way we have generated 664 conventional primitive and 179 inverse-power distance coordinates. Data for the initial configuration and for the relaxed structures are collected in Table III. The performance of our method is by a factor of 10 better than the conjugate gradient method.

![FIG. 2. Crystal structure of gibbsite, Al(OH)₃.](http://jcp.aip.org/jcp/copyright.jsp)
algorithm implemented in VASP. The full relaxation with the compensation for a Pulay stress of \(\{-0.64, -1.81, -0.50\}\) leads to the configuration close to the minimum obtained from the Murnaghan equation of state.

### C. Molecular crystal: Urea

Urea \([\text{NH}_2\text{CO}]\) (Fig. 3) is an example of a molecular crystal (space group \(P\bar{4}2_1m\), No. 113), with strong electrostatic and hydrogen bond interactions. In the crystal, the molecules occupy a special position compatible with the full molecular symmetry \((mm\bar{2})\). High-resolution x-ray\(^{45}\) and neutron\(^{45,46}\) diffraction studies show small differences in the cell parameters, but some significant changes in the atomic positions.

The conventional primitive internal coordinates are generated only for intramolecular degrees of freedom, the mutual positions of individual urea molecules are described using the \(5/r\) coordinates generated as described in the preceding section. Total number of 152 conventional primitive and 160 inverse-power distance coordinates have been used to generate the delocalized internal coordinates. The data concerning this example are collected in Table IV. The initial configuration has already the correct space group symmetry \((P\bar{4}2_1m)\). As pointed out by Baker \(^{25}\) delocalized internal coordinates are automatically symmetry adapted. Therefore the symmetry is preserved if the relaxation is started with correct symmetry. As in the previous examples, Pulay stress causes the underestimation of the equilibrium cell volume. The structure relaxed with stress compensated for this effect is very close to the minimum of the Murnaghan fit. The performance of our algorithm is by a factor of 5 better than the native optimizer in VASP.

### D. Constrained relaxation: Urea

To illustrate the ability of our algorithm to impose virtually any geometrical constraint on the structure, we have chosen example of urea with fixed intramolecular angles and torsions. The relaxations have been started from the same configuration as in the previous example. The relaxation criteria are the same as in the unconstrained relaxations. Note, however, that the contributions to the cartesian gradients corresponding to the constrained coordinates had to be projected out. Results of relaxations of atomic positions, atomic positions, and lattice parameters, data for the minimum of the Murnaghan equation of state and for the relaxation with the correction for the Pulay stress (estimated in the unconstrained relaxations) are shown in Table V.

### IV. CONCLUSIONS

We have shown that the geometry optimization of periodic systems in internal coordinates offers a highly advantageous alternative to perform structural relaxations of solids...
and surfaces. It has been demonstrated that the present implementation is able to handle not only atomic position but also lattice parameter relaxations. The final geometries obtained by the native optimizer of VASP and the internal coordinate optimizer GADGET are the same within the numerical precision of the electronic code. For the selected model systems GADGET outperforms the native optimizer by a factor of 2–10. One has to be aware of the fact that a considerable portion of this better performance can be ascribed to the quality of the initial Hessian guess and to the use of more efficient optimization algorithms.

The most significant advantage of the internal coordinate optimization lies undeniably in the handling of constraints, as demonstrated by the simple examples of rigid molecule optimization of the urea molecular crystal. In any case, the possibility of constraining virtually any physically/chemically motivated combination of internal coordinates opens a broad field of application of periodic electronic structure codes to complicated problems, such as phase transitions, solid state chemical reactions, etc.

The internal coordinate optimizer GADGET is an independent software, easy to interface with various molecular and solid state total energy codes. It allows the user to optimize not only periodic but also finite (molecular) systems in delocalized internal coordinates by ignoring the lattice parameter block of the extended B-matrix, i.e., by the standard method widely documented in the literature (see, e.g., Ref. 12). In addition to the innovative treatment of periodic systems, the GADGET program includes a large selection of algorithms permitting the efficient handling of Hessian update, transition state optimization, etc. An exhaustive overview of these features will be the subject of a forthcoming paper that describes the technical aspects of our implementation of these algorithms in GADGET.

### Table IV

Optimization of the urea structure: the initial configuration (Initial): relaxation of the ionic degrees of freedom using delocalized internal coordinates ([q]), and using the VASP native optimizer ([s]); relaxation of the lattice parameters and ionic positions using delocalized internal coordinates ([q, q]) and with the VASP native optimizer ([s, h]); minimum of the fitted Murnaghan equation of state (EOS); and the full relaxation using delocalized internal coordinates with stress corrected for a Pulay stress ([q, q; σ]). Listed are the cell parameters [a, b, c (Å); α, β, γ (deg.)], energy (hartree), gradient (a.u.), stress (kB), cell volume (Å³), and the number of relaxation steps.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial</th>
<th>[q]</th>
<th>[s]</th>
<th>[q, q]</th>
<th>[s, h]</th>
<th>EOS</th>
<th>[q, q; σ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>5.565(0)</td>
<td></td>
<td></td>
<td>5.758(2)</td>
<td>5.744(0)</td>
<td>5.790(6)</td>
<td>5.788(6)</td>
</tr>
<tr>
<td>b</td>
<td>5.565(0)</td>
<td></td>
<td></td>
<td>5.758(2)</td>
<td>5.744(0)</td>
<td>5.790(6)</td>
<td>5.788(6)</td>
</tr>
<tr>
<td>c</td>
<td>4.684(0)</td>
<td></td>
<td></td>
<td>4.696(9)</td>
<td>4.695(5)</td>
<td>4.705(9)</td>
<td>4.703(6)</td>
</tr>
<tr>
<td>g&lt;sub&gt;max&lt;/sub&gt;</td>
<td>0.075 04</td>
<td>0.000 3</td>
<td>0.000 18</td>
<td>0.000 4</td>
<td>0.000 04</td>
<td>0.000 04</td>
<td>0.000 17</td>
</tr>
<tr>
<td>σ&lt;sub&gt;xx&lt;/sub&gt;</td>
<td>79.14</td>
<td>9.38</td>
<td>9.84</td>
<td>−0.02</td>
<td>0.83</td>
<td>−1.17</td>
<td>−1.18</td>
</tr>
<tr>
<td>σ&lt;sub&gt;xy&lt;/sub&gt;</td>
<td>79.14</td>
<td>9.38</td>
<td>9.84</td>
<td>−0.02</td>
<td>0.83</td>
<td>−1.17</td>
<td>−1.18</td>
</tr>
<tr>
<td>σ&lt;sub&gt;xz&lt;/sub&gt;</td>
<td>145.60</td>
<td>8.18</td>
<td>8.63</td>
<td>0.04</td>
<td>−0.02</td>
<td>−1.75</td>
<td>−1.42</td>
</tr>
<tr>
<td>σ&lt;sub&gt;yx&lt;/sub&gt;</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>σ&lt;sub&gt;yz&lt;/sub&gt;</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Volume</td>
<td>145.06</td>
<td>145.06</td>
<td>145.06</td>
<td>155.73</td>
<td>154.92</td>
<td>157.80</td>
<td>157.61</td>
</tr>
<tr>
<td>N&lt;sub&gt;max&lt;/sub&gt;</td>
<td>6</td>
<td>27</td>
<td>9</td>
<td>49</td>
<td></td>
<td></td>
<td>9</td>
</tr>
</tbody>
</table>

### Table V

Relaxation of urea with constrained intramolecular angles and torsions: the initial configuration (Initial): relaxation of the ionic degrees of freedom using delocalized internal coordinates ([q]); relaxation of the lattice parameters and ionic positions using delocalized internal coordinates ([q, q]); minimum of the fitted Murnaghan equation of state (EOS); and the relaxation of the lattice parameters and ionic positions using delocalized internal coordinates with stress corrected for a Pulay stress ([q, q; σ]). Listed are the cell parameters [a, b, c (Å); α, β, γ (deg.)], energy (hartree), gradient (a.u.), stress (kB), cell volume (Å³), and the number of relaxation steps.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial</th>
<th>[q]</th>
<th>[q, q]</th>
<th>EOS</th>
<th>[q, q; σ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>5.565(0)</td>
<td></td>
<td>5.768(0)</td>
<td>5.802(4)</td>
<td>5.802(7)</td>
</tr>
<tr>
<td>b</td>
<td>5.565(0)</td>
<td></td>
<td>5.768(0)</td>
<td>5.802(4)</td>
<td>5.802(7)</td>
</tr>
<tr>
<td>c</td>
<td>4.684(0)</td>
<td></td>
<td>4.710(7)</td>
<td>4.717(3)</td>
<td>4.716(9)</td>
</tr>
<tr>
<td>g&lt;sub&gt;max&lt;/sub&gt;</td>
<td>0.075 04</td>
<td>0.000 03</td>
<td>0.000 08</td>
<td>0.000 08</td>
<td>0.000 05</td>
</tr>
<tr>
<td>σ&lt;sub&gt;xx&lt;/sub&gt;</td>
<td>79.14</td>
<td>15.17</td>
<td>3.53</td>
<td>2.18</td>
<td>2.12</td>
</tr>
<tr>
<td>σ&lt;sub&gt;xy&lt;/sub&gt;</td>
<td>79.14</td>
<td>15.17</td>
<td>3.53</td>
<td>2.18</td>
<td>2.12</td>
</tr>
<tr>
<td>σ&lt;sub&gt;xz&lt;/sub&gt;</td>
<td>145.60</td>
<td>1.20</td>
<td>−7.23</td>
<td>−8.23</td>
<td>−8.43</td>
</tr>
<tr>
<td>σ&lt;sub&gt;yx&lt;/sub&gt;</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>σ&lt;sub&gt;yz&lt;/sub&gt;</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Volume</td>
<td>145.06</td>
<td>145.06</td>
<td>156.72</td>
<td>158.82</td>
<td>158.82</td>
</tr>
<tr>
<td>N&lt;sub&gt;max&lt;/sub&gt;</td>
<td>6</td>
<td>19</td>
<td></td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENT

One of the authors (J.G.A.) is indebted to the Computational Materials Science College for having supported his stay in Vienna.

APPENDIX: THE STRESS TENSOR

During a deformation of the lattice, all atomic positions and virtual grid points, including the lattice vector endpoints, change according to the relationship

\[ u'_\alpha = \sum_\beta \left( \delta_{\alpha\beta} + e_{\alpha\beta} \right) u_\beta. \tag{A1} \]

where \( e_{\alpha\beta} \) are the elements of the strain tensor. The stress tensor elements \( \sigma_{\alpha\beta} \) are usually defined as the volume-normalized negative strain derivatives of the energy:

\[ \sigma_{\alpha\beta} = -\frac{1}{\Omega} \frac{\partial E}{\partial e_{\alpha\beta}}. \tag{A2} \]

The Cartesian to fractional transformation matrix \( h_{\alpha\beta} \) is formed from the three column vectors \( d_\beta \) of the lattice as \( h_{\alpha\beta} = (d_\beta)^\alpha \). Using Eq. (A1) it follows that the variation of the lattice vectors can be expressed as

\[ h'_{\alpha\beta} = \sum_\gamma \left( \delta_{\alpha\gamma} + e_{\alpha\gamma} \right) h_{\gamma\beta} \tag{A3} \]

and the strain derivative of the lattice vectors is

\[ \frac{\partial h'_{\alpha\beta}}{\partial \varepsilon_{\mu\nu}} = \delta_{\mu\alpha} h_{\gamma\beta}. \tag{A4} \]

The strain derivative of the energy can be written in terms of the lattice vector component derivatives, using Eq. (A4), as

\[ \frac{\partial E}{\partial e_{\alpha\beta}} = \sum_\mu \frac{\partial E}{\partial h_{\mu\nu}} \frac{\partial h_{\mu\nu}}{\partial e_{\alpha\beta}} = \sum_\nu \frac{\partial E}{\partial h_{\alpha\nu}} h_{\nu\beta} \tag{A5} \]

leading to the alternative form of the stress tensor

\[ \sigma_{\alpha\beta} = -\frac{1}{\Omega} \sum_\gamma \frac{\partial E}{\partial h_{\alpha\gamma}} h'_{\gamma\beta}. \tag{A6} \]

The lattice-vector derivatives of the energy can be obtained from the stress tensor, since

\[ -\Omega \sum_\beta \sigma_{\alpha\beta}(h')^{-1}_{\beta\gamma} = \sum_\gamma \frac{\partial E}{\partial h_{\alpha\gamma}} h'_{\gamma\beta}. \tag{A7} \]

so

\[ \frac{\partial E}{\partial h_{\alpha\beta}} = -\Omega \sum_\gamma \sigma_{\alpha\gamma}(h')^{-1}_{\gamma\beta}. \tag{A8} \]

5M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., GAUSSIAN 03, 2004.