PHYSICAL REVIEW B VOLUME 48, NUMBER 16 15 OCTOBER 1993-II

Electronic-structure calculations in adaptive coordinates

François Gygi*
IBM Research Division, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland

(Received 8 January 1993)

The plane-wave method for electronic-structure calculations is reformulated in generalized curvilinear coordinates. This introduces a new set of basis functions that depend continuously on a coordinate transformation, and can adapt themselves to represent optimally the solutions of the Schrödinger equation. As a consequence, the effective plane-wave energy cutoff is allowed to vary in the unit cell in an unbiased way. The efficiency of this method is demonstrated in the calculation of the equilibrium structures of the CO and H_2O molecules using the local-density approximation of density-functional theory, and norm-conserving, nonlocal pseudopotentials. The easy evaluation of forces on all degrees of freedom makes the method suitable for *ab initio* molecular-dynamics applications.

I. INTRODUCTION

With the advent of *ab initio* molecular dynamics (MD) based on density-functional theory (DFT), tremendous progress in the determination of electronic and structural properties of materials has been made.[1] Its application to a wide range of molecules, clusters, solids, and surfaces has demonstrated that MD calculations can successfully predict structural properties of most ionic and covalent systems. The success of such an approach relies largely on an accurate and numerically stable representation of the forces acting on atoms. This was first achieved using plane-wave basis functions and pseudopotentials, which are still used in most *ab initio* MD calculations. Many properties of the plane-wave basis have contributed in an essential way to the success of *ab initio* MD simulations. Among these, we note (i) their orthonormality, (ii) the use of fast Fourier transform algorithms to obtain a real space representation of wave functions, (iii) the smooth control of convergence provided by the plane-wave energy cutoff, (iv) the independence of the basis set on atomic positions, and (v) the unbiased way in which orbitals are described throughout the unit cell.

These advantages, however, are obtained at the cost of efficiency, since a large number of plane waves is needed in order to reach convergence in large systems, or for atoms having a rapidly varying potential. This lack of efficiency, even though partially remedied by the use of pseudopotentials, makes the plane-wave approach unwieldy for the calculation of electronic properties of compounds involving first-row elements. Furthermore, the large number of plane waves needed in such calculations also affects adversely the rate of convergence of the iterative algorithms used in the energy minimization.

A way around these difficulties could be to turn to localized basis functions (e.g., Gaussians, Slater-type orbitals, atomic orbitals), which are better suited to the description of solutions of the Schrödinger equation, and usually give a good description of physical properties with a reasonably small number of basis functions. However, this increase in efficiency usually implies the loss of property (v), which is acceptable, but also of properties (i)–(iv), which leads to additional complications in the calculations. First, lack of orthonormality of the basis functions implies the use of an overlap matrix in the eigenvalue problem, which is known to become numerically untractable as the number of localized basis functions is increased in order to attain convergence. Furthermore, choosing localized basis functions which lead to a rapid convergence of the results can be a delicate task. A systematic procedure to increase the quality of the basis set is usually lacking. Lastly, localized basis functions are usually centered on atoms, so that the basis set and the overlap matrix depend *explicitly* on the ionic coordinates. This introduces additional terms (Pulay forces) in the calculation of the atomic forces.

In this paper, we present a formalism which generalizes the plane-wave formalism and dramatically improves its efficiency without losing its most important properties. This is achieved by reformulating the plane-wave method in arbitrary curvilinear coordinates.[2] The resulting basis set is similar to plane waves, but depends continuously on a given three-dimensional coordinate transformation. Treating this coordinate transformation as a variational parameter then allows the unbiased optimization of the basis set, while keeping the number of basis functions constant. An immediate consequence of this approach is that the effective energy cutoff of the plane-wave expansion can vary *locally* so as to adapt optimally to the physics of the problem. In typical examples, the increase of efficiency of the basis set is manifested in a fourfold to ninefold *local* increase of the effective plane-wave cutoff in the vicinity of the atoms, i.e., where the wave functions vary rapidly. This implies that a calculation that requires a cutoff of 160 Ry in a small part of the unit cell can be carried out with a number of plane waves corresponding to a cutoff of 20 Ry throughout the unit cell. An application of the method to the calculation of structural properties of CO and H_2O shows that a high degree of accuracy can be attained with only a limited number of

0163-1829/93/48(16)/11692(9)/$06.00

basis functions. A reduction in the number of basis functions by a factor of 10–25 can be achieved with respect to a conventional plane-wave calculation, with equivalent accuracy.

The rest of this paper is organized as follows. In Sec. II, the basic tools needed for the construction of the new basis set are developed. Section III describes the application of the adaptive coordinates to the problem of electronic structure within the local-density approximation of the DFT, and using norm-conserving, nonlocal pseudopotentials. An application to the calculation of the equilibrium geometries of the CO and H_2O molecules is then given in Sec. IV.

II. BASIC FORMALISM

A. Properties of the basis functions

In this section, we derive the properties of plane waves in arbitrary curvilinear coordinates. The solutions of the one-particle Schrödinger equation are then described using this formalism. We consider a set of arbitrary curvilinear coordinates for a periodic system with a unit cell of volume Ω, spanned by the Cartesian coordinates $\xi^i, i = 1,\dots,3$. We then define an invertible map $\xi \mapsto \mathbf{x}(\xi)$ on the unit cell Ω (which, because of the periodic boundary conditions, can be identified with the 3-torus T^3), and require that this map be twice continuously differentiable everywhere in T^3. The Riemannian metric tensor can be expressed as[3,4]

$$g_{ij} = \frac{\partial x^k}{\partial \xi^i}\frac{\partial x^k}{\partial \xi^j} \tag{1}$$

(summation over repeated indices is used throughout this paper). This particularly simple expression of g_{ij} is a consequence of the flatness of T^3 or, in other words, of the fact that there always exists a coordinate system on T^3 for which $g_{ij} \equiv \delta_{ij}$ [i.e., Euclidean coordinates $\mathbf{x}(\xi) \equiv \xi$].

Plane waves in curvilinear coordinates are then defined by

$$\begin{aligned}\chi_{\mathbf{k}}(\mathbf{x}) &= \frac{1}{\sqrt{\Omega}}\, g^{-1/4} \exp[i\mathbf{k}\xi(\mathbf{x})] \\ &= \frac{1}{\sqrt{\Omega}} \left|\frac{\partial \xi}{\partial x}\right|^{\frac{1}{2}} \exp[i\mathbf{k}\xi(\mathbf{x})],\end{aligned} \tag{2}$$

where $g = \det g_{ij}$, $|\partial\xi/\partial x|$ is the Jacobian of the transformation $\mathbf{x} \mapsto \xi(\mathbf{x})$, and $\mathbf{k}$ is a reciprocal lattice vector

$$\mathbf{k} = 2\pi\left(\frac{n_1}{a_1}, \frac{n_2}{a_2}, \frac{n_3}{a_3}\right), \quad n_1, n_2, n_3 \in Z, \tag{3}$$

where a_1, a_2, and a_3 are the dimensions of the unit cell, which we choose to be orthorhombic for simplicity. The functions $\chi_{\mathbf{k}}(\mathbf{x})$ reduce to ordinary plane waves in the case of Euclidean coordinates. We now show that for any given map $\mathbf{x}(\xi)$, the functions $\chi_{\mathbf{k}}$ form a complete orthonormal basis of $L^2(T^3, d^3x)$. The orthonormality follows immediately from the change of variables

$$\begin{aligned}\langle \mathbf{k}|\mathbf{k}'\rangle &= \frac{1}{\Omega}\int \exp[i(\mathbf{k}'-\mathbf{k})\xi(\mathbf{x})]\left|\frac{\partial\xi}{\partial x}\right| d^3x \\ &= \frac{1}{\Omega}\int \exp[i(\mathbf{k}'-\mathbf{k})\xi]\, d^3\xi \\ &= \delta_{\mathbf{k}-\mathbf{k}'}.\end{aligned} \tag{4}$$

To prove completeness in $L^2(T^3, d^3x)$, we note that the functions $\chi_{\mathbf{k}}$ can be seen as the characters of the unitary representations of the compact Lie group T^3. The theorem of Peter and Weyl[5,6] states that these characters form a complete basis, independent of the coordinates chosen.

Given a smooth map $\mathbf{x}(\xi)$, the functions $\chi_{\mathbf{k}}$ can therefore be used to represent any function $\psi \in L^2(T^3, d^3x)$,

$$\psi(\mathbf{x}) = \sum_{\mathbf{k}} c_{\mathbf{k}}\chi_{\mathbf{k}}(\mathbf{x}). \tag{5}$$

B. One-electron Schrödinger equation

In the following example we consider the use of this basis set to represent the solutions of a one-particle Schrödinger equation. The corresponding Hamiltonian is

$$H = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}) \tag{6}$$

and its matrix elements in the $\chi_{\mathbf{k}}$ basis are

$$\begin{aligned}\langle \mathbf{k}|H|\mathbf{k}'\rangle = \frac{1}{\Omega}\int \Bigg[&\frac{\hbar^2}{2m}(k_i + iA_i)g^{ij}(k'_j - iA_j) \\ &+V(\mathbf{x}(\xi))\Bigg] \exp[i(\mathbf{k}'-\mathbf{k})\xi]\, d^3\xi,\end{aligned} \tag{7}$$

where g^{ij} denotes the inverse of the Riemannian metric tensor g_{ij}, and by analogy with the notation of a gauge potential, A_i is defined by

$$A_i = \frac{1}{2}\frac{\partial}{\partial \xi^i}\ln\left|\frac{\partial x}{\partial \xi}\right|. \tag{8}$$

Note the close connection of A_i with the contracted Christoffel symbol Γ^j_{ji}. All the quantities appearing in Eq. (7) can be evaluated once a coordinate system $\mathbf{x}(\xi)$ has been specified. Using the expansion of a solution $\psi(\mathbf{x})$ in the $\chi_{\mathbf{k}}$ basis,

$$\psi(\mathbf{x}) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}} c_{\mathbf{k}}\left|\frac{\partial\xi}{\partial x}\right|^{\frac{1}{2}} \exp[i\mathbf{k}\xi(\mathbf{x})], \tag{9}$$

we can write the energy as

$$\begin{aligned}E[\{c_{\mathbf{k}}\}, \{x^i(\xi)\}] = \sum_{\mathbf{k}\mathbf{k}'} c^*_{\mathbf{k}}c_{\mathbf{k}'}\frac{1}{\Omega}\int \Bigg[&\frac{\hbar^2}{2m}(k_i + iA_i)g^{ij}(k'_j - iA_j) \\ &+V(\mathbf{x}(\xi))\Bigg] \exp[i(\mathbf{k}'-\mathbf{k})\xi]\, d^3\xi.\end{aligned} \tag{10}$$

Note that expressions (7) and (10) are nothing but a coordinate-free representation of the Hamiltonian matrix elements and of the energy. By definition, the coordinate-free energy (10) is independent of the choice of coordinates $\mathbf{x}(\xi)$ *if the solution* ψ *is represented exactly in* $L^2(T^3, d^3x)$, i.e., if ψ is represented by an infinite sum in Eq. (9). Any truncation of the expansion (9) to a finite number of terms makes the energy coordinate-dependent. Clearly this is the case in all practical applications, in which only a finite number of functions $\chi_{\mathbf{k}}$ can be used to represent the solution. Thus, in practice, some coordinate systems will be better suited than others to the representation of the solutions. We will consider below strategies to obtain optimal coordinates, in the sense that the energy (10) is minimized.

In order to measure locally the efficiency of the basis set, we define a *local* effective energy cutoff by

$$E_{\mathrm{cut}}(\mathbf{x}) = g(\mathbf{x})^{-1/3} E^0_{\mathrm{cut}}, \tag{11}$$

where $g(\mathbf{x})$ is the determinant of the metric tensor at position $\mathbf{x}$, and E^0_{cut} is the usual plane-wave energy cutoff defined in terms of the largest reciprocal lattice vector $\mathbf{k}_{\mathrm{max}}$ used in the expansion (9),

$$E^0_{\mathrm{cut}} = \frac{\hbar^2 \mathbf{k}^2_{\mathrm{max}}}{2m}. \tag{12}$$

In the limit of slowly varying coordinates, the largest value ($E^{\mathrm{max}}_{\mathrm{cut}}$) of the local effective cutoff in the unit cell gives an estimate of the energy cutoff needed to reproduce the same results using conventional plane waves. However, in practical calculations, $E^{\mathrm{max}}_{\mathrm{cut}}$ is usually *smaller* than the energy cutoff of an equivalent calculation in conventional plane waves. This can be understood by considering the extreme case of a function represented with only one ($\mathbf{k} = 0$) basis function. In this case, even though $E^{\mathrm{max}}_{\mathrm{cut}} = E^0_{\mathrm{cut}} = 0$, any positive function can be represented exactly by this single basis function by choosing appropriate coordinates.

Intuitively, it appears advantageous to use coordinate systems such that the effective energy cutoff is large in the regions where solutions vary rapidly (e.g., in the vicinity of atoms). As is clear from Eq. (2), this also makes the basis functions $\chi_{\mathbf{k}}$ oscillate more rapidly and have a larger amplitude in these regions.

C. Determination of optimal coordinates

We now focus on the problem of choosing an optimal coordinate system $\mathbf{x}(\xi)$, which, as will be seen below, can increase dramatically the efficiency of the basis set. A possible approach would consist of choosing arbitrarily a local effective cutoff function $E_{\mathrm{cut}}(\mathbf{x})$, although it is not always clear *a priori* where this function should be peaked (e.g., on atoms or on bond centers in covalent systems). In the next step, one would have to find a set of coordinates $\mathbf{x}(\xi)$ such that Eq. (11) holds. This involves the additional work of solving a set of coupled nonlinear partial differential equations for the coordinates $\mathbf{x}(\xi)$. We also note that a definition of the effective cutoff function in terms of atomic positions would introduce an explicit dependence of the basis set on atomic positions, and therefore complicate the calculation of ionic forces.

Instead of making uncontrollable assumptions about the optimal choice of $\mathbf{x}(\xi)$, we consider the coordinate system as a variational parameter, and look for coordinates that minimize the energy. For practical reasons, the coordinate system must not only be represented by a finite number of parameters, it must be done in a way that facilitates the calculation of the derivatives $\partial x^i / \partial \xi^j$ needed for the evaluation of g_{ij} and A_i. We chose to represent $\mathbf{x}(\xi)$ as

$$\mathbf{x}(\xi) = \xi + \sum_{\mathbf{Q}} \mathbf{x}_{\mathbf{Q}} \exp[i\mathbf{Q}\xi], \tag{13}$$

where the $\mathbf{Q}$'s are reciprocal lattice vectors, and the Fourier expansion is limited to $N_{\mathbf{Q}}$ plane waves. Although $N_{\mathbf{Q}}$ can be arbitrarily large, care must be taken to insure that no fast Fourier transform (FFT) aliasing errors appear in the representation of the metric tensor and of its determinant.

The energy, therefore, has the general form $E[\{c_{\mathbf{k}}\}, \{\mathbf{x}_{\mathbf{Q}}\}, \{\mathbf{R}_n\}]$, where the dependence on atomic positions $\mathbf{R}_n$ has been included. Its minimization can now be carried out using, e.g., a simulated annealing approach, whereby the coefficients $c_{\mathbf{k}}$ and $\mathbf{x}_{\mathbf{Q}}$ and the atomic positions $\mathbf{R}_n$ are varied simultaneously until the energy is stationary. The Lagrangian describing such a dynamical system is

$$\begin{aligned} L[\{c_{\mathbf{k}}\}, \{\dot{c}_{\mathbf{k}}\}, \{\mathbf{x}_{\mathbf{Q}}\}, \{\dot{\mathbf{x}}_{\mathbf{Q}}\}, \{\mathbf{R}_n\}, \{\dot{\mathbf{R}}_n\}] \\ = K_{\mathrm{el}}(\{\dot{c}_{\mathbf{k}}\}) + K_{\mathrm{coord}}(\{\dot{\mathbf{x}}_{\mathbf{Q}}\}) + K_{\mathrm{ion}}(\{\dot{\mathbf{R}}_n\}) \\ - E[\{c_{\mathbf{k}}\}, \{\mathbf{x}_{\mathbf{Q}}\}, \{\mathbf{R}_n\}], \end{aligned} \tag{14}$$

where $K_{\mathrm{el}}(\{\dot{c}_{\mathbf{k}}\})$ and $K_{\mathrm{coord}}(\{\dot{\mathbf{x}}_{\mathbf{Q}}\})$ are fictitious kinetic energies associated with the electronic and coordinate degrees of freedom, respectively, and $K_{\mathrm{ion}}(\{\dot{\mathbf{R}}_n\})$ is the ionic kinetic energy. Fictitious masses are assigned to the electronic degrees of freedom as well as to the coordinates. Lagrange multipliers can be added to this expression in the usual way[1] in order to enforce additional constraints, such as orthonormality of the eigenfunctions.

We can now derive the generalized forces on all degrees of freedom from the Lagrangian (14). Functional derivatives $\delta L / \delta \psi$ are obtained in the usual way from the Hamiltonian

$$\frac{\delta L}{\delta \psi} = -H\psi = \left(\frac{\hbar^2}{2m} \Delta - V \right) \psi. \tag{15}$$

Functional derivatives with respect to the coordinates $\mathbf{x}(\xi)$ are obtained using the general expression

$$\frac{\delta L}{\delta x^p} = \frac{\partial L}{\partial x^p} - \frac{\partial}{\partial \xi^q} \frac{\partial L}{\partial \frac{\partial x^p}{\partial \xi^q}} + \frac{\partial^2}{\partial \xi^q \xi^r} \frac{\partial L}{\partial \frac{\partial^2 x^p}{\partial \xi^q \partial \xi^r}}. \tag{16}$$

Contributions from the electronic kinetic energy involve the functional derivatives $\delta g^{ij} / \delta x^p$ and $\delta A_i / \delta x^p$ for which detailed expressions are given in Appendix A. Terms arising from the ionic potentials are simpler, since

$$\delta\langle V\rangle = \sum_n \int \frac{\delta}{\delta x^p}[V_n(\mathbf{x}-\mathbf{R}_n)\rho(\mathbf{x})\, d^3x]\delta x^p, \qquad (17)$$

where $V_n(\mathbf{x})$ is the ionic potential associated with atom n, and $\rho(\mathbf{x}) = |\psi(\mathbf{x})|^2$ is the electronic density. This expression can be simplified by noting that the measure

$$\begin{aligned} d\rho &= \rho(\mathbf{x})\, d^3x \\ &= \frac{1}{\Omega}\sum_{\mathbf{kk}'} c^*_{\mathbf{k}} c_{\mathbf{k}'} \exp[i(\mathbf{k}'-\mathbf{k})\xi(\mathbf{x})] \left|\frac{\partial \xi}{\partial x}\right| d^3x \\ &= \frac{1}{\Omega}\sum_{\mathbf{kk}'} c^*_{\mathbf{k}} c_{\mathbf{k}'} \exp[i(\mathbf{k}'-\mathbf{k})\xi]\, d^3\xi \end{aligned} \qquad (18)$$

is invariant under any change of the coordinates. Therefore,

$$\delta\langle V\rangle = \sum_n \int \frac{\partial V_n(\mathbf{x}-\mathbf{R}_n)}{\partial x^p}\delta x^p\, \rho(\mathbf{x})\, d^3x. \qquad (19)$$

More detailed expressions for the forces on the Fourier coefficients of the coordinates can be found in Appendix A. An example of application of these formulas to the solution of a one-particle Schrödinger equation has been given in Ref. 2.

We now derive the expressions of the ionic forces. The ith component of the ionic force on atom n is defined by

$$F^i_n = -\frac{\partial E}{\partial R^i_n}. \qquad (20)$$

Clearly, the electronic kinetic energy does not contribute to the ionic forces, since there is no explicit dependence of the basis set on the atomic positions $\mathbf{R}_n$. The ionic forces, therefore, reduce to

$$F^i_n = -\frac{\partial\langle V\rangle}{\partial R^i_n} = -\int \frac{\partial V_n(\mathbf{x}-\mathbf{R}_n)}{\partial R^i_n}\,\rho(\mathbf{x})\, d^3x \qquad (21)$$

which can also be written

$$F^i_n = \int \frac{\partial V_n(\mathbf{x}-\mathbf{R}_n)}{\partial x^i}\,\rho(\mathbf{x})\, d^3x\,. \qquad (22)$$

This expression is closely related to the forces on coordinates obtained from the potential energy in Eq. (19). Using the expressions in Appendix A, it is easily shown that forces on atoms can be derived from the forces on coordinates in the following way:

$$F^i_n = \left(\frac{\partial\langle V_n\rangle}{\partial x^i_{\mathbf{Q}}}\right)_{\mathbf{Q}=0}, \qquad (23)$$

where $\langle V_n\rangle$ is the potential energy associated with atom n. Thus, the computation of the ionic forces requires no additional effort once the forces on coordinates have been calculated. We also note that the calculation of the generalized forces can be made efficient through extensive use of fast Fourier transform algorithms.

D. Elastic energy of the coordinates

In a typical calculation involving plane waves in adaptive coordinates, it is generally found that the coordinates adjust during the calculation so as to increase the effective energy cutoff in the regions where solutions vary rapidly. In the initial phase of the adaptation of the coordinates, the effective energy cutoff rapidly reaches a large value in these regions, which insures a correct description of the electronic wave functions. This rapid increase is usually followed by a much slower adaptation phase, where the effective energy cutoff increases very little, and where the total energy is not improved significantly. At the end of this second phase, the coordinates reach their equilibrium configuration. However, all physical properties of interest are already converged at the end of the first phase. In order to avoid the slow convergence of the second phase, we introduce a means of controlling the degree of adaptation of the coordinates. Although this is not *necessary*, it allows us to reach convergence more rapidly. A practical way of achieving this is to add to the Lagrangian (14) an elastic energy term E_{elastic} which depends only on the coordinates $x^p(\xi)$, and measures to what extent they are deformed with respect to Euclidean coordinates. The elastic energy is defined such that any deviation from Euclidean coordinates makes it increase. This effectively amounts to introducing a restoring force that "pulls" the coordinates back to the Euclidean limit. Note that, since the expression of the elastic energy is independent of atomic positions and of the wave-function coefficients, the unbiased character of the adaptation of the basis set is preserved. This approach has clear connections with nonlinear elasticity theory.[7] Following considerations of elementary rheology,[8] and noting the analogy of the Riemannian metric tensor g_{ij} with the right Cauchy-Green tensor, we require that the elastic energy associated with the coordinates depends only on the principal invariants of g_{ij} or of its inverse g^{ij}. These invariants are $\mathrm{tr}g_{ij}$, $\det g_{ij}$, and the quantity $\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1$, where the λ_i's are the eigenvalues of g_{ij}. Similar expressions derived from the inverse g^{ij} can also be used. We consider only $\det g^{ij}$ and $\mathrm{tr}g^{ij}$ since, as will be shown below, they lead to a simple physical interpretation of the elastic energy. We define

$$\begin{aligned} E_{\text{elastic}}[\{x^p(\xi)\}] &= \mu_{\text{comp}} \int \det g^{ij}\, d^3\xi \\ &\quad + \mu_{\text{shear}} \int \mathrm{tr} g^{ij}\, d^3\xi, \end{aligned} \qquad (24)$$

where the first term describes an isotropic compression energy and the second term is a shear deformation energy. The constants μ_{comp} and μ_{shear} can be chosen so as to control the maximum value that the effective cutoff can reach. In practice, it turns out that μ_{comp} can be set to zero, and that the maximum effective cutoff can be controlled with μ_{shear} of the order of 10^{-4}.

III. DFT AND NONLOCAL POTENTIALS

A. Hartree and exchange-correlation potentials

In this section, we extend the formalism of plane waves in adaptive coordinates to density-functional theory and to the use of norm-conserving, nonlocal pseudopotentials. The electronic density is now defined in terms of the solutions $\psi_n(\mathbf{x})$ of the Kohn-Sham equations by

$$\rho(\mathbf{x}) = \sum_n |\psi_n(\mathbf{x})|^2, \tag{25}$$

where the index n runs over all occupied states. We write the Hartree energy as

$$\begin{aligned} E_H &= \frac{1}{2}\int v_C(\mathbf{x}-\mathbf{x}')\ \rho(\mathbf{x})d^3x\ \rho(\mathbf{x}')d^3x'. \\ &= \frac{1}{2}\int v_H(\mathbf{x})\ \rho(\mathbf{x})d^3x, \end{aligned} \tag{26}$$

where $v_C(\mathbf{x}) = 1/|\mathbf{x}|$ is the Coulomb potential, and $v_H(\mathbf{x})$ is the solution of the Poisson equation

$$\Delta v_H(\mathbf{x}) = 4\pi\rho(\mathbf{x}). \tag{27}$$

The Poisson equation cannot be solved as easily as in the plane-wave basis. Indeed, it can be seen from Eq. (7) that the coordinate-free Laplacian has off-diagonal elements in the $\chi_\mathbf{k}$ basis. However, the calculation of the Laplacian of a given function expressed in the $\chi_\mathbf{k}$ basis can be carried out entirely through products of functions in both real space and $\chi_\mathbf{k}$ space, and by using fast Fourier transform algorithms to go from one representation to the other. This facilitates considerably the calculation of $\Delta v_H(\mathbf{x})$. We therefore use an iterative procedure, the conjugate gradient method, to solve the Poisson equation.[9] It should be noted that the iterative solution of the Poisson equation, even though it differs from the direct solution obtained using conventional plane waves, does not introduce any computational effort which scales worse than other parts of the calculation. Also, it has been pointed out that an exact solution of the Poisson equation is not necessary at every step of a minimization (or MD) procedure.[10,11] Instead, the Hartree potential can be included in the set of dynamical variables, and a fictitious mass can be associated with it. This approach leads to substantial savings of computing time in the determination of the equilibrium geometry of a molecule using plane waves in adaptive coordinates.

Functional derivatives of the Hartree energy with respect to coordinates can be obtained by considering the variation of Hartree energy δE_H induced by an infinitesimal change of coordinates δx^p. Using again the invariance of the measure $d\rho = \rho(\mathbf{x})d^3x$, we find

$$\delta E_H = \frac{1}{2}\int \rho(\mathbf{x})d^3x\ \rho(\mathbf{x}')d^3x'\ \delta v_C(\mathbf{x}-\mathbf{x}')\,. \tag{28}$$

The variation of the Coulomb potential is

$$\begin{aligned} \delta v_C(\mathbf{x}-\mathbf{x}') &= \frac{\partial v_C(\mathbf{x}-\mathbf{x}')}{\partial(x^p - x^{p'})}(\delta x^p - \delta x^{p'}) \\ &= -\frac{(x^p - x^{p'})}{|\mathbf{x}-\mathbf{x}'|^3}(\delta x^p - \delta x^{p'})\,. \end{aligned} \tag{29}$$

Using the symmetry under the exchange $\mathbf{x} \leftrightarrow \mathbf{x}'$ and the fact that

$$\frac{\partial v_C(\mathbf{x}-\mathbf{x}')}{\partial x^p} = -\frac{\partial v_C(\mathbf{x}-\mathbf{x}')}{\partial x^{p'}} \tag{30}$$

we finally get

$$\begin{aligned} \delta E_H &= \int \frac{\partial \delta v_C(\mathbf{x}-\mathbf{x}')}{\partial x^p}\ \rho(\mathbf{x})d^3x\ \rho(\mathbf{x}')d^3x' \\ &= \int \frac{\partial v_H(\mathbf{x})}{\partial x^p}\delta x^p\ \rho(\mathbf{x})d^3x, \end{aligned} \tag{31}$$

where the absence of factor 1/2 should be noted.

The exchange-correlation energy in the local-density approximation is expressed as

$$E_{\mathrm{xc}} = \int \epsilon_{\mathrm{xc}}[\rho(\mathbf{x})]\rho(\mathbf{x})d^3x, \tag{32}$$

where $\epsilon_{\mathrm{xc}}[\rho(\mathbf{x})]$ is the exchange-correlation energy density. The variation of E_{xc} caused by an infinitesimal coordinate transformation δx^p is given by

$$\begin{aligned} \delta E_{\mathrm{xc}} &= \int \delta\{\epsilon_{\mathrm{xc}}[\rho(\mathbf{x})]\rho(\mathbf{x})\}d^3x \\ &\quad + \int \epsilon_{\mathrm{xc}}[\rho(\mathbf{x})]\rho(\mathbf{x})\ \delta\{d^3x\} \\ &= \int v_{\mathrm{xc}}(\mathbf{x})\delta\rho(\mathbf{x})d^3x \\ &\quad + \int \epsilon_{\mathrm{xc}}[\rho(\mathbf{x})]\rho(\mathbf{x})\ \delta\{d^3x\}, \end{aligned} \tag{33}$$

where $v_{\mathrm{xc}}(\mathbf{x})$ is the exchange-correlation potential. Using

$$\delta(\rho(\mathbf{x})d^3x) = 0 = \delta\rho(\mathbf{x})d^3x + \rho(\mathbf{x})\delta d^3x \tag{34}$$

we get

$$\delta E_{\mathrm{xc}} = \int (\{\epsilon_{\mathrm{xc}}[\rho(\mathbf{x})] - v_{\mathrm{xc}}(\mathbf{x})\}\rho(\mathbf{x}))\delta d^3x. \tag{35}$$

This can be further simplified by using the identity

$$\delta d^3x = \frac{\partial \delta x^p}{\partial x^p}d^3x. \tag{36}$$

Integrating by parts and using

$$\frac{\partial}{\partial x^p}\{\epsilon_{\mathrm{xc}}[\rho(\mathbf{x})]\rho(\mathbf{x})\} = v_{\mathrm{xc}}(\mathbf{x})\frac{\partial\rho(\mathbf{x})}{\partial x^p} \tag{37}$$

we finally get

$$\delta E_{\mathrm{xc}} = \int \frac{\partial v_{\mathrm{xc}}(\mathbf{x})}{\partial x^p}\delta x^p\ \rho(\mathbf{x})\,d^3x. \tag{38}$$

B. Nonlocal pseudopotentials

We consider the energy associated with separable nonlocal pseudopotentials as formulated by Kleinman and Bylander,[12] which can be written

$$E_{\mathrm{nl}} = \sum_l \frac{|\langle\psi|\phi_l v_l\rangle|^2}{\langle\phi_l|v_l|\phi_l\rangle}, \tag{39}$$

where ϕ_l is the valence atomic orbital of angular momentum l, and v_l is the nonlocal part of the pseudopotential associated with angular momentum l. Functional derivatives of E_{nl} involve the expression

$$\delta\langle\psi|v_l|\phi_l\rangle = \delta\int\psi(\mathbf{x})v_l(\mathbf{x})\phi_l(\mathbf{x})d^3x. \tag{40}$$

Since the functions $v_l(\mathbf{x})$ and $\phi_l(\mathbf{x})$ do not depend on the derivatives of $\mathbf{x}(\xi)$, we have

$$\begin{aligned}\delta\langle\psi|v_l|\phi_l\rangle &= \int\delta\{\psi(\mathbf{x})d^3x\}v_l(\mathbf{x})\phi_l(\mathbf{x})\\ &\quad+\int\psi(\mathbf{x})d^3x\frac{\partial}{\partial x^p}[v_l(\mathbf{x})\phi_l(\mathbf{x})]\delta x^p.\end{aligned} \tag{41}$$

While the evaluation of the second term in Eq. (41) poses no particular problem, the first part involves the functional derivatives

$$\begin{aligned}\frac{\delta}{\delta x^p}\left|\frac{\partial x}{\partial\xi}\right|^{\frac{1}{2}} &= \frac{\partial}{\partial x^p}\left|\frac{\partial x}{\partial\xi}\right|^{\frac{1}{2}} - \frac{\partial}{\partial\xi^q}\frac{\partial}{\partial\frac{\partial x^p}{\partial\xi^q}}\left|\frac{\partial x}{\partial\xi}\right|^{\frac{1}{2}}\\ &= \frac{\partial}{\partial x^p}\left|\frac{\partial x}{\partial\xi}\right|^{\frac{1}{2}} - \frac{1}{2}\frac{\partial}{\partial\xi^q}\frac{\partial\xi^q}{\partial x^p}.\end{aligned} \tag{42}$$

IV. APPLICATIONS TO CO AND H$_2$O

In this section, we present calculations of the equilibrium structure of CO and H$_2$O using plane waves in adaptive coordinates. We used the norm-conserving nonlocal pseudopotentials of Bachelet, Hamann, and Schlüter[13] for the description of oxygen. Potentials derived by Giannozzi,[14] the parameters of which are given in Ref. 15 and in Appendix B, were used for carbon and hydrogen. The exchange-correlation energy functional was that of Ceperley and Alder,[16] as parametrized by Perdew and Zunger.[17]

The molecules were placed at the center of a cubic unit cell of lattice parameter $a = 10$ (a.u.). The calculations were started with Euclidean coordinates, and a homogeneous plane-wave energy cutoff of 20 Ry, corresponding to 1503 basis functions. The total energy was then minimized using both the steepest-descent algorithm and a simulated-annealing approach. The atomic positions and the coordinates $\mathbf{x}(\xi)$ were held fixed in the initial stage of the calculation. After the electronic energy had come close to a minimum, the coordinates and the atomic positions were successively relaxed, and the minimization was pursued until all degrees of freedom converged. Coordinates were represented with 57 plane waves in Eq. (13), which corresponds to an energy cutoff of 2 Ry.

A. Results for CO

The calculation of the equilibrium bond length of CO was carried out with $\mu_{\mathrm{comp}} = 0$ and $\mu_{\mathrm{shear}} = 0.001$. After full relaxation of all degrees of freedom, the effective plane-wave energy cutoff $E_{\mathrm{cut}}(\mathbf{x})$ reached a value $E_{\mathrm{cut}}^{\mathrm{max}} = 85.2$ Ry at the position of the oxygen atom. The equilibrium bond length was $d_{\mathrm{CO}} = 2.138$ (a.u.), and the vibrational frequency estimated from a quadratic fit to the force in the vicinity of the equilibrium position was $\omega_e = 2125\ \mathrm{cm}^{-1}$. These values are in excellent agreement with results obtained with other methods.[18] In order to confirm the high accuracy of the results, we carried out calculations using the conventional plane-wave method and plane-wave energy cutoffs of 80 Ry, 100 Ry, and 120 Ry. The results are summarized in Table I and show that the calculation done with plane waves in adaptive coordinates is equivalent to the conventional plane-wave calculation at 100 Ry, even though the maximum effective energy cutoff reached a value of only 85.2 Ry. This confirms that $E_{\mathrm{cut}}^{\mathrm{max}}$ underestimates the energy cutoff of an equivalent calculation using conventional plane waves (see Sec. II B). Other results obtained at lower energy cutoffs also show this trend, i.e., that a calculation performed with plane waves in adaptive coordinates in which the maximum effective energy cutoff reaches the value $E_{\mathrm{cut}}^{\mathrm{max}}$ is equivalent to a conventional calculation with a cutoff *larger* than $E_{\mathrm{cut}}^{\mathrm{max}}$. We note that the number of basis functions used in adaptive coordinates (1503) is more than ten times smaller than the number of plane waves needed in an equivalent calculation using plane waves (16879 with a cutoff of 100 Ry). This reduction in the number of degrees of freedom also implies a large reduction in the number of iterations needed to attain convergence. In order to illustrate the increase in effective energy cutoff in the unit cell, we show in Fig. 1 the image of a regular grid under the map $\xi \mapsto \mathbf{x}(\xi)$. The concentration of the coordinates in the vicinity of the oxygen atom is clearly apparent, and leads to an increase in effective energy cutoff by a factor of 4.2 at that point. The increase in effective energy cutoff at the position of the carbon atom is less important and only reaches 1.75.

B. Results for H$_2$O

Calculations of the equilibrium structure of H$_2$O were performed under the same conditions as those used for CO. A first calculation was carried out with $\mu_{\mathrm{comp}} = 0$ and $\mu_{\mathrm{shear}} = 5 \times 10^{-4}$. The resulting maximum local energy cutoff after full relaxation of all degrees of free-

TABLE I. Equilibrium bond length of CO obtained using plane waves in adaptive coordinates, compared with the results of conventional plane-wave calculations. $E_{\mathrm{cut}}^{\mathrm{max}}$ indicates the maximum effective energy cutoff in the unit cell, E_{cut}^{0} is the initial energy cutoff in Euclidean coordinates, and N_{PW} is the number of basis functions used.

$E_{\mathrm{cut}}^{\mathrm{max}}$ (Ry)	E_{cut}^{0} (Ry)	N_{PW}	d_{CO} (a.u.)	Basis set
85.2	20	1503	2.138	adaptive
80	80	12053	2.148	conventional
100	100	16879	2.137	conventional
120	120	22119	2.132	conventional

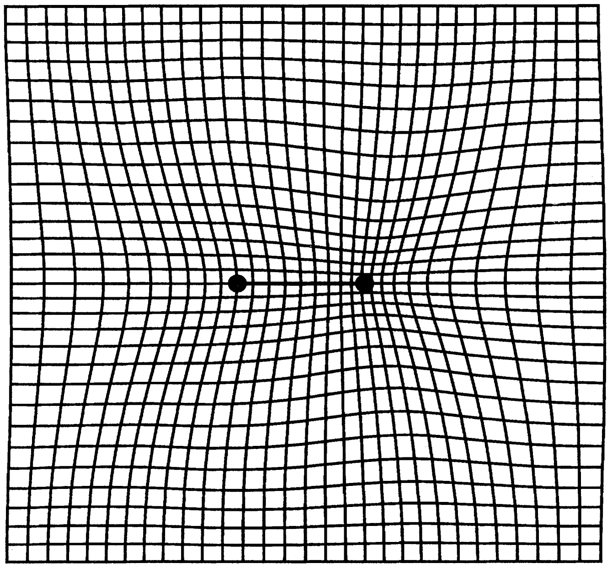

FIG. 1. Image of a regularly spaced rectangular grid in ξ space under the map $x(\xi)$, calculated in a plane containing the CO bond. The left and right dots represent the carbon and oxygen atoms, respectively.

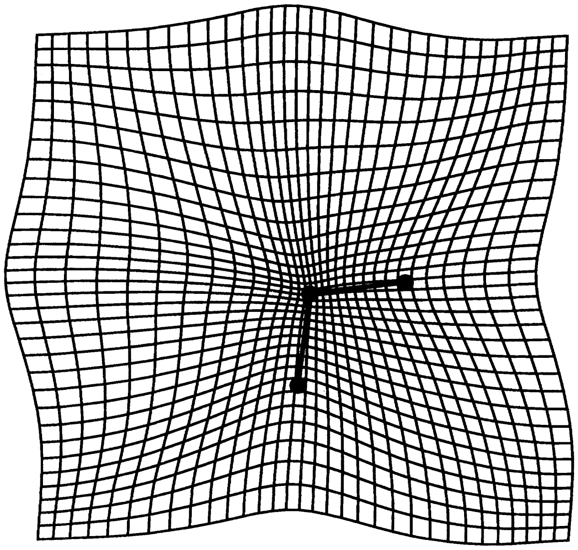

FIG. 2. Image of a regularly spaced rectangular grid in ξ space under the map $x(\xi)$, calculated in the plane of the H_2O molecule.

dom was 98 Ry at the position of the oxygen atom. The resulting OH distance was $d_{\rm OH} = 1.839$ (a.u.) and the angle between the two bonds was $\theta_{\rm HOH} = 104.3$ (deg.). These values are close to the results of similar calculations performed with conventional plane waves, and energy cutoffs of 80 Ry (Ref. 19) and 88 Ry.[20] We find our OH bond length to be slightly smaller than the ones obtained in Ref. 19 [$d_{\rm OH} = 1.865$ (a.u.)] and in Ref. 20 [$d_{\rm OH} = 1.855$ (a.u.)], which is consistent with the trend of $d_{\rm OH}$ as the energy cutoff increases. Our HOH bond angle is slightly larger than those in Ref. 19 [$\theta_{\rm HOH} = 102.1$ (deg.)] and in Ref. 20 [$\theta_{\rm HOH} = 102.7$ (deg.)] which is again consistent with the increase of the bond angle observed in conventional calculations as the energy cutoff is increased.

In a second calculation, the elastic constraint on coordinates was partially released by lowering the value of $\mu_{\rm shear}$ to 2×10^{-4}, which let the local effective energy cutoff reach 105 Ry at the position of the oxygen atom. This caused the OH distance to decrease to $d_{\rm OH} = 1.825$ (a.u.) and the HOH angle to increase up to $\theta_{\rm HOH} = 104.9$ deg.

Finally, a calculation was performed in which the value of $\mu_{\rm shear}$ was reduced to 2×10^{-5}. This allowed the local effective energy cutoff to reach 172 Ry at the position of the oxygen atom, and yielded $d_{\rm OH} = 1.825$ (a.u.) and $\theta_{\rm HOH} = 105.0$ deg. Figure 2 shows the corresponding relaxation of coordinates around the H_2O molecule. No conventional calculation using a similar energy cutoff was available for comparison. We note, however, that such a calculation would require more than 38000 plane waves, i.e., more than 25 times the number of basis functions needed in the adaptive coordinate calculation. These calculations on H_2O illustrate how the elastic energy is used to control the effective energy cutoff during the relaxation of coordinates.

V. CONCLUSIONS

We have presented a method for the calculation of electronic structure within density-functional theory and including nonlocal pseudopotentials. The efficiency of the plane-wave basis is increased dramatically by letting coordinates adapt to the physics of the problem under study, and by letting the effective plane-wave energy cutoff vary in space in an unbiased but controlled way. We have also demonstrated that plane waves in adaptive coordinates can yield highly accurate equilibrium geometries of molecules, using only a small number of basis functions. Further developments of this method include its application to *ab initio* molecular-dynamics simulations.

ACKNOWLEDGMENTS

It is a pleasure to thank G. Galli and K. Laasonen for many fruitful discussions, and D. R. Hamann for a critical reading of the manuscript.

APPENDIX A: FORCES ON THE COORDINATES

The forces on the Fourier coefficients of the coordinates $x^p_{\bf Q}$ are obtained by considering the variation of energy δE induced by a change of coordinates δx^p,

$$\delta E = \frac{\hbar^2}{2m}\int \delta\left\{\psi^*({\bf x})(-\Delta)\psi({\bf x})d^3x\right\} + \int \delta\left\{V({\bf x})\rho({\bf x})d^3x\right\} . \qquad (A1)$$

The first integrand can be expanded into

$$\delta\left\{\psi^*(\mathbf{x})(-\Delta)\psi(\mathbf{x})d^3x\right\} = \frac{1}{\Omega}\sum_{\mathbf{k}\mathbf{k}'} c^*_{\mathbf{k}}c_{\mathbf{k}'}\exp[i(\mathbf{k}'-\mathbf{k})\xi] \times\{(k_i+iA_i)\delta g^{ij}(k'_j-iA_j)+[i(k'_i-k_i)+2A_i]g^{ij}\delta A_j\}d^3\xi\,, \qquad \text{(A2)}$$

where

$$\delta g^{ij} = \frac{\partial g^{ij}}{\partial\frac{\partial x^p}{\partial\xi^q}}\frac{\partial}{\partial\xi^q}\delta x^p, \qquad \text{(A3)}$$

and

$$\delta A_i = \frac{\partial A_i}{\partial\frac{\partial x^p}{\partial\xi^q}}\frac{\partial}{\partial\xi^q}\delta x^p + \frac{\partial A_i}{\partial\frac{\partial^2 x^p}{\partial\xi^q\partial\xi^r}}\frac{\partial^2}{\partial\xi^q\partial\xi^r}\delta x^p, \qquad \text{(A4)}$$

which are evaluated using the following identities derived from the definition of g_{ij} and A_i:

$$\frac{\partial g^{ij}}{\partial\frac{\partial x^p}{\partial\xi^q}} = -\left(\frac{\partial\xi^i}{\partial x^p}g^{qj}+\frac{\partial\xi^j}{\partial x^p}g^{qi}\right)\,, \qquad \text{(A5)}$$

$$\frac{\partial A_i}{\partial\frac{\partial x^p}{\partial\xi^q}} = \frac{1}{2}\frac{\partial}{\partial\xi^i}\frac{\partial\xi^q}{\partial x^p}\,, \qquad \text{(A6)}$$

$$\frac{\partial A_i}{\partial\frac{\partial^2 x^p}{\partial\xi^q\xi^r}} = \frac{1}{4}\left(\delta^i_q\frac{\partial\xi^r}{\partial x^p}+\delta^i_r\frac{\partial\xi^q}{\partial x^p}\right)\,. \qquad \text{(A7)}$$

Note that there exist many equivalent expressions for these derivatives, which can be obtained from each other by multiplication by the metric tensor or its inverse. The forms given here are slightly shorter, though exactly equivalent to, the expressions given in Ref. 2.

The variation of the potential energy is

$$\delta\langle V\rangle = \int\delta\{V(\mathbf{x})\rho(\mathbf{x})\,d^3x\} \qquad \text{(A8)}$$

where $\rho(\mathbf{x}) = |\psi(\mathbf{x})|^2$ is the electronic density. Using the invariance of $d\rho = \rho(\mathbf{x})d^3x$, under coordinate transformations, we have

$$\delta\left\{V(\mathbf{x})\rho(\mathbf{x})\,d^3x\right\} = \frac{\partial V(\mathbf{x})}{\partial x^p}\delta x^p\,\rho(\mathbf{x})\,d^3x\,. \qquad \text{(A9)}$$

The force on the Fourier coefficient $x^p_{\mathbf{Q}}$ is obtained by setting $\delta x^p = \epsilon\exp[-i\mathbf{Q}\xi]$. We then have

$$\frac{\partial E}{\partial x^p_{\mathbf{Q}}} = \frac{\partial}{\partial\epsilon}\delta E\Big|_{\epsilon=0}\,. \qquad \text{(A10)}$$

Setting $\mathbf{Q} = 0$ in this expression and replacing the potential $V(\mathbf{x})$ by the potential $V_n(\mathbf{x})$ associated with atom n, we recover the expression for the force on atom n. We note that, since the sum of the forces on all atoms is zero, the total force on the $\mathbf{Q} = 0$ component of the coordinates vanishes, which prevents any global translation of the coordinates.

APPENDIX B: PARAMETERS OF THE HYDROGEN PSEUDOPOTENTIAL

The local pseudopotential derived by Giannozzi[14] for hydrogen has the form

$$V(r) = -\frac{1}{r}\mathrm{erf}\left(\frac{r}{r_{c1}}\right) + (a+br^2)\exp\left[-\left(\frac{r}{r_{c2}}\right)^2\right]\,, \qquad \text{(B1)}$$

where the parameters expressed in atomic units are $r_{c1} = 0.25$, $a = -1.9287$, $b = 0.3374$, and $r_{c2} = 0.284$.

---

[*] Present address: Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA), PHB Ecublens, CH-1015 Lausanne, Switzerland.

[1] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985). For a review, see, e.g., G. Galli and M. Parrinello, in *Computer Simulation in Materials Science*, edited by M. Meyer and V. Pontikis (Kluwer, Dordrecht, 1991), pp. 283–304.

[2] F. Gygi, Europhys. Lett. **19**, 617 (1992).

[3] B. Doubrovine, S. Novikov, and A. Fomenko, *Géométrie Contemporaine, Méthodes et Applications, Vols. 1 & 2* (Mir, Moscow, 1982).

[4] R. D. Richtmyer, *Principles of Advanced Mathematical Physics, Vol. 2* (Springer, New York, 1981).

[5] F. Peter and H. Weyl, Math. Ann. **97**, 737 (1927).

[6] N. J. Vilenkin, *Special Functions and the Theory of Group Representations*, Translations of Mathematical Monographs Vol. 22 (American Mathematical Society, Providence, RI, 1968).

[7] C. Truesdell and W. Noll, *The Non-Linear Field Theories of Mechanics*, Vol. III/3 of *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1965).

[8] H. Markovitz, *Rheology*, in *Physics Vade Mecum*, edited by H. L. Anderson (American Institute of Physics, New York, 1981).

[9] Other methods, such as the multigrid method, may be more efficient than the conjugate gradient method in that particular application. See, e.g., W. Hackbusch, *Multi-Grid Methods and Applications*, Springer Series in Computational Mathematics Vol. 4 (Springer, Berlin, 1985).

[10] R. Car and M. Parrinello, Solid State Commun. **62**, 403 (1987).

[11] R. Virkkunen and R. M. Nieminen, in *Many-Atom Interactions in Solids*, edited by R. M. Nieminen, M. J. Puska, and M. J. Manninen (Springer, Berlin, 1990).

[12] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[13] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[14] P. Giannozzi (unpublished).

[15] W. Andreoni, D. Scharf, and P. Giannozzi, Chem. Phys. Lett. **173**, 449 (1990).

[16] D. Ceperley and B. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[17] J. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[18] P. A. Serena, A. Baratoff, and J. M. Soler (unpublished).

[19] K. Laasonen, F. Csajka, and M. Parrinello, Chem. Phys. Lett. **194**, 172 (1992).

[20] F. Csajka (unpublished).