PHYSICAL REVIEW B VOLUME 40, NUMBER 3 15 JULY 1989-II

Simple formula for the atomic forces in the augmented-plane-wave method

José M. Soler
International School for Advanced Studies, I-34014 Trieste, Italy
and Departamento de Física de la Materia Condensada (C-III), Universidad Autónoma de Madrid,*
E-28049 Madrid, Spain

Arthur R. Williams
IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598
(Received 30 January 1989)

We present a simple and explicit method for the computation of atomic forces in *ab initio* total-energy calculations using a basis of augmented plane waves (APW) and the local-density approximation for exchange and correlation. The force on an atom is given by integrals over its muffin-tin sphere only, which can be obtained easily in existing implementations of the linear APW method, for example. The extra computational cost of calculating the forces on all the atoms is negligible compared with that of performing one single self-consistency iteration step.

The ability to calculate the forces on the atoms, in *ab initio* electronic-structure calculations, is desirable for several obvious reasons, the first being that it allows the minimization of the total energy as a function of the ionic positions. In the last few years this task has been frequently performed by repeating a total-energy calculation for different atomic positions and making numerical derivatives. But this technique is very expensive because it requires at least one well-converged total-energy calculation for each atomic degree of freedom, and frequently many more. An even more interesting perspective is the technique of *ab initio* molecular dynamics, recently developed by Car and Parinello,[1] for which the efficient computation of all the forces at each step is a necessary prerequisite. On the other hand, once the electronic structure has been determined, the Hellmann-Feynman theorem[2] provides a simple, and in principle exact, means to compute the force on every nucleus. However, this means has proven in practice to be very inaccurate due to the practical necessity of using incomplete basis sets and charge densities that are not perfectly converged. The basic reason for this difficulty is that the core electrons produce enormous fields in the region of the nucleus, which can easily hide the chemically relevant fields produced by the valence electrons.[3] In practive, this difficulty has prevented the direct calculation of forces in most cases. An important exception is the plane-wave pseudopotential method.[4] The success of this method is based, on the one hand, on the "removal" of the core electrons and their fields by the use of a pseudopotential and, on the other hand, on the use of a basis set of plane waves which does not depend on the position of the atoms. However, despite its many merits, the plane-wave pseudopotential method has important limitations on the kind of systems it can describe. Some elements, and the transition metals in particular, have such strong pseudopotentials that a prohibitively large number of plane waves are required to represent the wave function in the pseudopotential region. In these cases a basis set of mixed,[5] localized[6] or augmented-type[7-9] orbitals is necessary. These basis sets depend on the atomic positions, changing when the atoms move, and this leads in practice to force corrections[10] difficult to calculate. The so-called "force theorem"[11] provides, in principle, an exact way to compute the forces in these cases, without dealing with the high-field cancellations involved in Hellmann-Feynman forces, but the application of this method has been confined mostly to the calculation of the pressure and, to our knowledge, no sufficiently practical formulas have been obtained for the forces on the atoms, similar to those existing for the plane-wave pseudopotential method. In this work we present such a practical formula for a basis of augmented plane waves (APW's).

For a better comprehension of what follows, we will refer to the standard linear APW (LAPW) method[7,8] even though the linear character of the LAPW is related to the solution of Schrödinger's equation and has nothing to do with the computation of the forces described here. In fact, we have presented[12] an APW scheme in which the exact wave functions inside the muffin-tin spheres are used, rather than their linearized counterparts. In the LAPW method, space is divided in two regions: interstitial space (and sometimes vacuum regions) and nonoverlapping muffin-tin spheres centered on every atom. Plane waves are used to represent the wave functions in the interstitial region. Each plane wave is "augmented" by matching it on the surface of the muffin-tin sphere of each atom α to a linear combination of spherical functions of the form $\phi_{\alpha l \sigma}(r) Y_{lm}(\hat{\mathbf{r}})$. Here $\{lm\}$ specifies the angular momentum and σ is an index with two possible values (ϕ_σ with $\sigma = 1,2$ correspond, respectively, to what is generally referred to as ϕ and $\dot{\phi}$). By choosing adequately the functions $\phi_{\alpha l \sigma}$ one ensures that all the wave functions are close to an exact solution of Schödinger's equation in the core region, orthogonal to the core states and with an energy in the range of interest. Although the usual point of view is to think of the augmentation of each individual plane wave, it is more convenient here to consider directly the augmentation of the electron wave functions. We will assume that we already know these wave functions $\tilde{\psi}_i(\mathbf{r})$ in the interstitial region, as given by their Fourier coefficients $\psi_{i\mathbf{G}}$:

$$\tilde{\psi}_i(\mathbf{r})=\sum_{\mathbf{G}}\psi_{i\mathbf{G}}e^{i(\mathbf{k}_i+\mathbf{G})\cdot\mathbf{r}} , \tag{1}$$

where $\mathbf{G}$ are the reciprocal-lattice vectors within some selected cutoff $k_{\max}$: $|\mathbf{k}_i+\mathbf{G}|\le k_{\max}$. $\mathbf{k}_i$ is a point within the first Brillouin zone and index i labels bands as well as $\mathbf{k}$ points. Then the augmented wave functions are

$$\psi_i(\mathbf{r})=\tilde{\psi}_i(\mathbf{r})+\sum_{\alpha}\Theta(s_\alpha-r_\alpha)\sum_{l=0}^{l_{\max}}\sum_{m=-l}^{l}[\psi_{i\alpha lm}(r_\alpha)-\tilde{\psi}_{i\alpha lm}(r_\alpha)]Y_{lm}(\hat{\mathbf{r}}_\alpha) , \tag{2}$$

where $\Theta(x)$ is the Heaviside unit-step function, s_α is the muffin-tin radius of atom α, and $\mathbf{r}_\alpha\equiv\mathbf{r}-\mathbf{R}_\alpha$, with $\mathbf{R}_\alpha$ the atom position. $\tilde{\psi}_{i\alpha lm}(r_\alpha)Y_{lm}(\hat{\mathbf{r}}_\alpha)$ are the terms of the expansion of $\tilde{\psi}_i(\mathbf{r})$ in spherical harmonics and

$$\psi_{i\alpha lm}(r_\alpha)=\sum_{\sigma=1}^{2}\psi_{i\alpha lm\sigma}\phi_{\alpha l\sigma}(r_\alpha) , \tag{3}$$

where $\psi_{i\alpha lm\sigma}$ are expansion coefficients determined by the matching conditions $\psi_{i\alpha lm}(s_\alpha)=\tilde{\psi}_{i\alpha lm}(s_\alpha)$ and $\psi'_{i\alpha lm}(s_\alpha)=\tilde{\psi}'_{i\alpha lm}(s_\alpha)$. Notice that both $\tilde{\psi}_i(\mathbf{r})$ and $\psi_i(\mathbf{r})$ are perfectly well defined in all space and identical in the interstitial region. Notice also that the components of $\tilde{\psi}(\mathbf{r})$ with angular momentum $l>l_{\max}$, which are not augmented, are nevertheless allowed to "penetrate" into the sphere.[13] The reason for doing this is that, because of their $\sim\mathbf{r}^l$ behavior, the high-l components of $\tilde{\psi}_i(\mathbf{r})$ and $\psi_i(\mathbf{r})$ are usually largest near the sphere surface, where they coincide in value and slope. Therefore, making $\psi_{i\alpha lm}(r_\alpha)=\tilde{\psi}_{i\alpha lm}(r_\alpha)$ is a much better approximation for these components than making them zero. Also, this ensures that the augmented wave functions $\psi_i(\mathbf{r})$ are perfectly continuous and smooth across the sphere boundary, independent of the value of $l_{\max}$.

The expansion coefficients $\psi_{i\alpha lm\sigma}$ are totally determined by the matching conditions, once we know $\tilde{\psi}_i(\mathbf{r})$ and $\mathbf{R}_\alpha$. Thus we can write the total energy E as a function of the only independent variables $\{\psi_{i\mathbf{G}},\mathbf{R}_\alpha\}$, subject to the orthonormality constraints. The force on atom α is given by the total derivative $dE/d\mathbf{R}_\alpha$ but, since E is stationary with respect to $\psi_{i\mathbf{G}}$, this equals the partial derivative $\partial E/\partial\mathbf{R}_\alpha$, except for the forces arising from the orthonormality constraints that we will address later. Thus, our first objective is to calculate this partial derivative, that is, the change in total energy when we move sphere α (and ion α) by an infinitesimal distance $\delta\mathbf{R}_\alpha$ *while keeping fixed the coefficients* $\psi_{i\mathbf{G}}$. In other words, we will move our muffin-tin sphere across the "sea" of fixed "floating" wave functions $\tilde{\psi}_i(\mathbf{r})$. It is convenient to think of the displacement of sphere α rather as an opposite displacement of the rest of the system while the ion α and its muffin-tin sphere remain stationary. Then the wave functions $\tilde{\psi}_i(\mathbf{r})$ are displaced and consequently change by $\delta\tilde{\psi}_i(\mathbf{r})=\delta\mathbf{R}_\alpha\cdot\nabla\tilde{\psi}_i(\mathbf{r})$. The augmented wave functions $\psi_i(\mathbf{r})$ also change inside our stationary sphere α because of the change of the matching conditions imposed by $\delta\tilde{\psi}_i(\mathbf{r})$ on the sphere surface. Within density-functional theory the total energy is given by the sum of the kinetic, electrostatic, and exchange-correlation (xc) energies $E=T+U+E_{\mathrm{xc}}$. The kinetic energy is

$$\begin{aligned}T&=\sum_i w_i\int_{\mathbb{R}^3}\psi_i^*(\mathbf{r})(-\nabla^2)\psi_i(\mathbf{r})d\mathbf{r}\\&=\sum_i w_i\left[\int_{\mathbb{R}^3}-\tilde{\psi}_i^*(\mathbf{r})\nabla^2\tilde{\psi}_i(\mathbf{r})d\mathbf{r}\right.\\&\qquad+\sum_\beta\int_{S_\beta}[-\psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})\\&\qquad\left.+\tilde{\psi}_i^*(\mathbf{r})\nabla^2\tilde{\psi}_i(\mathbf{r})]d\mathbf{r}\right] ,\end{aligned} \tag{4}$$

where w_i are occupation and $\mathbf{k}$-sampling weights and S_β represents the muffin-tin sphere of atom β. The sum on β extends to all atoms in the system. Atomic Rydberg units are used throughout the paper. When we move sphere α the change in kinetic energy is

$$\begin{aligned}\delta T=\sum_i w_i\int_{S_\beta}&[-\delta\psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})-\psi_i^*(\mathbf{r})\nabla^2\delta\psi_i(\mathbf{r})\\&+\delta\tilde{\psi}_i^*(\mathbf{r})\nabla^2\tilde{\psi}_i(\mathbf{r})+\tilde{\psi}_i^*(\mathbf{r})\nabla^2\delta\tilde{\psi}_i(\mathbf{r})]d\mathbf{r} .\end{aligned} \tag{5}$$

Making use of the Gauss theorem and of the fact that $\psi_i(\mathbf{r})=\tilde{\psi}_i(\mathbf{r})$ and $\delta\psi_i(\mathbf{r})=\delta\tilde{\psi}_i(\mathbf{r})$ at the sphere surface, one can transform Eq. (5) into

$$\begin{aligned}\delta T=\sum_i 2w_i\mathrm{Re}\int_{S_\alpha}&[-\delta\psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})\\&+\delta\tilde{\psi}_i^*(\mathbf{r})\nabla^2\tilde{\psi}_i(\mathbf{r})]d\mathbf{r} ,\end{aligned} \tag{6}$$

where Re is the real-part function.

In order to calculate the change in the electrostatic energy we will follow the ideas introduced by Weinert[14] for the solution of Poisson's equation in a crystal. In his method, the true charge density $n(\mathbf{r})$ (which includes ions and electrons) inside the muffin-tin spheres is replaced by a smoothed charge density $\tilde{n}(\mathbf{r})$. In each sphere $\tilde{n}(\mathbf{r})$ has the same multipole moments as $n(\mathbf{r})$ (including the monopole or total charge within the sphere). In the interstitial region $\tilde{n}(\mathbf{r})$ and $n(\mathbf{r})$ are identical. Because the potential created outside a sphere by the charge inside it depends only on its multipole moments, the potential $\tilde{V}(\mathbf{r})$ created by $\tilde{n}(\mathbf{r})$ is exact in the interstitial region and in particular on the surface of each sphere. Thus, after finding $\tilde V(\mathbf r)$ by conventional Fourier-transform solution of Poisson's equation, one can also find the correct potential $V(\mathbf r)$ inside the spheres using the Green's-function method[15] with the true charge density $n(\mathbf r)$ and the boundary condition $\tilde V(\mathbf r)=V(\mathbf r)$ on the surface. We will assume that $\tilde V(\mathbf r)$ and $V(\mathbf r)$ have been already calculated, as they actually are in the full-potential LAPW method.[16] Since the interstitial-interstitial and interstitial-spheres electrostatic interactions are correctly represented by the smoothed charge density $\tilde n(\mathbf r)$ (as well as the interaction between different spheres) we can write the total electrostatic energy as the self-energy of this smoothed charge plus a correction term for the self-energy of every sphere:

$$U=\tilde U+\sum_\beta (U_\beta-\tilde U_\beta)\ , \tag{7}$$

where

$$\tilde U=\tfrac{1}{2}\int_{\mathbb R^3}\tilde V(\mathbf r)\tilde n(\mathbf r)d\mathbf r\ , \tag{8}$$

$$U_\beta=\int_{S_\beta}V^\beta_{\rm ion}(\mathbf r)n_{\rm el}(\mathbf r)d\mathbf r+\tfrac{1}{2}\int_{S_\beta}V^\beta_{\rm el}(\mathbf r)n_{\rm el}(\mathbf r)d\mathbf r\ , \tag{9}$$

$$\tilde U_\beta=\tfrac{1}{2}\int_{S_\beta}\tilde V^\beta(\mathbf r)\tilde n(\mathbf r)d\mathbf r\ . \tag{10}$$

We have introduced the notation $V^\beta(\mathbf r)$ to mean "potential created by charges inside sphere β." The subscripts ion and el stand for ions and electrons, and no subscript means both. Thus, for instance,

$$V^\alpha_{\rm el}(\mathbf r)\equiv 2\int_{S_\alpha}\frac{n_{\rm el}(\mathbf r')}{|\mathbf r-\mathbf r'|}d\mathbf r'\ . \tag{11}$$

The smoothed charge density can be written as

$$\tilde n(\mathbf r)=\tilde n_{\rm el}(\mathbf r)+\sum_\beta \Delta\tilde n^\beta(\mathbf r)\ , \tag{12}$$

where

$$\tilde n_{\rm el}(\mathbf r)\equiv\sum_i w_i|\tilde\psi_i(\mathbf r)|^2 \tag{13}$$

and $\Delta\tilde n^\beta(\mathbf r)$ is an extra charge density added in each sphere to make the multipole moments of $\tilde n(\mathbf r)$ equal to those of $n(\mathbf r)$. When we move sphere α, only the term with $\beta=\alpha$ in Eq. (7) will change. Again, instead of displacing sphere α by $\delta\mathbf R_\alpha$ it is more convenient to think in terms of the displacement of the rest of the system by $-\delta\mathbf R_\alpha$, leaving the ion and sphere α stationary. Then the change in the density $\tilde n(\mathbf r)$ will be $\delta\mathbf R_\alpha\cdot\nabla[\tilde n(\mathbf r)-\Delta\tilde n^\alpha(\mathbf r)]+\delta\Delta\tilde n^\alpha(\mathbf r)$. The last term must be introduced because the displacement will produce a change of $\tilde n_{\rm el}(\mathbf r)$ and $n_{\rm el}(\mathbf r)$ inside the sphere and therefore $\Delta\tilde n^\alpha(\mathbf r)$ will have to be changed to match the change in the multipole moments. The change in the different terms of the electrostatic energy is, to first order in $\delta n(\mathbf r)$ and $\delta\tilde n(\mathbf r)$,

$$\begin{aligned}\delta\tilde U&=\int_{\mathbb R^3}\tilde V(\mathbf r)\delta\tilde n(\mathbf r)d\mathbf r\\&=\int_{\mathbb R^3}\tilde V(\mathbf r)\delta\mathbf R_\alpha\cdot\nabla\tilde n(\mathbf r)d\mathbf r-\int_{\mathbb R^3}\tilde V(\mathbf r)\delta\mathbf R_\alpha\cdot\nabla\Delta\tilde n^\alpha(\mathbf r)d\mathbf r\\&\quad+\int_{S_\alpha}\tilde V(\mathbf r)\delta\Delta\tilde n^\alpha(\mathbf r)d\mathbf r\ .\end{aligned} \tag{14}$$

The last integral has been restricted to sphere α because $\delta\Delta\tilde n^\alpha$ is nonzero only within this sphere. Taking into account that $\tilde V(\mathbf r)$ is the potential produced by $\tilde n(\mathbf r)$, the first integral can be integrated by parts and reduced to a zero surface integral at infinity. This result is obvious since this term is nothing but the first-order change in electrostatic energy produced by a displacement of the whole charge $\tilde n(\mathbf r)$. The second integral can also be integrated by parts and we get

$$\delta\tilde U=\int_{\mathbb R^3}\delta\mathbf R_\alpha\cdot\nabla\tilde V(\mathbf r)\Delta\tilde n^\alpha(\mathbf r)d\mathbf r+\int_{S_\alpha}\tilde V(\mathbf r)\delta\Delta\tilde n^\alpha(\mathbf r)d\mathbf r\ . \tag{15}$$

Also,

$$\begin{aligned}\delta U_\alpha&=\int_{S_\alpha}[V^\alpha_{\rm ion}(\mathbf r)+V^\alpha_{\rm el}(\mathbf r)]\delta n_{\rm el}(\mathbf r)d\mathbf r\\&=\int_{S_\alpha}V(\mathbf r)\delta n_{\rm el}(\mathbf r)d\mathbf r-\int_{S_\alpha}V^{\rm out}(\mathbf r)\delta n_{\rm el}(\mathbf r)d\mathbf r\ ,\end{aligned} \tag{16}$$

where $V^{\rm out}(\mathbf r)$ means "potential created by all charges outside sphere α" and we have taken into account that $V(\mathbf r)=V^\alpha(\mathbf r)+V^{\rm out}(\mathbf r)$, and

$$\begin{aligned}\delta\tilde U_\alpha&=\int_{S_\alpha}\tilde V^\alpha(\mathbf r)\delta\tilde n(\mathbf r)d\mathbf r\\&=\int_{S_\alpha}\tilde V(\mathbf r)\delta\tilde n(\mathbf r)d\mathbf r-\int_{S_\alpha}\tilde V^{\rm out}(\mathbf r)\delta\tilde n(\mathbf r)d\mathbf r\\&=\int_{S_\alpha}\tilde V(\mathbf r)\delta\mathbf R_\alpha\cdot\nabla\tilde n_{\rm el}(\mathbf r)d\mathbf r+\int_{S_\alpha}\tilde V(\mathbf r)\delta\Delta\tilde n^\alpha(\mathbf r)d\mathbf r\\&\quad-\int_{S_\alpha}\tilde V^{\rm out}(\mathbf r)\delta\tilde n(\mathbf r)d\mathbf r\ ,\end{aligned} \tag{17}$$

where we have used $\delta\tilde n(\mathbf r)=\delta\mathbf R_\alpha\cdot\nabla\tilde n_{\rm el}(\mathbf r)+\delta\Delta\tilde n^\alpha(\mathbf r)$ because $\Delta\tilde n^\alpha(\mathbf r)$ remains static with sphere α when we displace the rest of the system. Now we sum $\delta U=\delta\tilde U+\delta U_\alpha-\delta\tilde U_\alpha$. In doing so we take into account that $\delta\tilde n(\mathbf r)$ and $\delta n(\mathbf r)$ have the same multipole moments and also that $\tilde V^{\rm out}(\mathbf r)=V^{\rm out}(\mathbf r)$ inside sphere α. Thus the last terms in Eqs. (16) and (17) are equal. If we now call $\delta\tilde n_{\rm el}(\mathbf r)=\delta\mathbf R_\alpha\cdot\nabla\tilde n_{\rm el}(\mathbf r)$ and $\delta\tilde V_{\rm el}(\mathbf r)=\delta\mathbf R_\alpha\cdot\nabla\tilde V_{\rm el}(\mathbf r)$ we finally obtain

$$\begin{aligned}\delta U&=\int_{\mathbb R^3}\delta\tilde V(\mathbf r)\Delta\tilde n^\alpha(\mathbf r)d\mathbf r\\&\quad+\int_{S_\alpha}[V(\mathbf r)\delta n_{\rm el}(\mathbf r)-\tilde V(\mathbf r)\delta\tilde n_{\rm el}(\mathbf r)]d\mathbf r\ .\end{aligned} \tag{18}$$

The exchange-correlation energy can be written in the local-density approximation as

$$\begin{aligned}E_{\rm xc}&=\int_{\mathbb R^3}\varepsilon_{\rm xc}[\tilde n_{\rm el}(\mathbf r)]\tilde n_{\rm el}(\mathbf r)d\mathbf r\\&\quad+\sum_\beta\int_{S_\alpha}\{\varepsilon_{\rm xc}[n_{\rm el}(\mathbf r)]n_{\rm el}(\mathbf r)-\varepsilon_{\rm xc}[\tilde n_{\rm el}(\mathbf r)]\tilde n_{\rm el}(\mathbf r)\}d\mathbf r\ .\end{aligned} \tag{19}$$

Then, to first order in $\delta n_{\rm el}$ and $\delta\tilde n_{\rm el}$

$$\delta E_{\rm xc}=\int_{S_\alpha}[\mu_{\rm xc}(\mathbf r)\delta n_{\rm el}(\mathbf r)-\tilde\mu_{\rm xc}(\mathbf r)\delta\tilde n_{\rm el}(\mathbf r)]d\mathbf r\ , \tag{20}$$

where $\mu_{\rm xc}\equiv d(n\varepsilon_{\rm xc})/dn$. Now we use

$$\delta n_{\rm el}(\mathbf r)=\sum_i 2w_i{\rm Re}[\delta\psi_i^*(\mathbf r)\psi_i(\mathbf r)] \tag{21}$$

to write the change in total energy produced by the displacement $\delta\mathbf R_\alpha$ as

$$\delta E=\int_{\mathbb{R}^3}\delta\tilde{V}(\mathbf{r})\Delta\tilde{n}^{\alpha}(\mathbf{r})d\mathbf{r}+\sum_i 2w_i\mathrm{Re}\int_{S_\alpha}\{\delta\psi_i^*(\mathbf{r})[-\nabla^2+V(\mathbf{r})+\mu_{\mathrm{xc}}(\mathbf{r})]\psi_i(\mathbf{r})-\delta\tilde{\psi}^{i\,*}(\mathbf{r})[-\nabla^2+\tilde{V}(\mathbf{r})+\tilde{\mu}_{\mathrm{xc}}(\mathbf{r})]\tilde{\psi}_i(\mathbf{r})\}d\mathbf{r}\ . \tag{22}$$

This is not yet the final result for the force because the change of the wave functions inside sphere α produced by its displacement destroys the orthonormality, even if we do not change the interstitial wave functions. Therefore, we have to change also the coefficients $\psi_{i\mathrm{G}}$ in order to reorthonormalize the wave functions after the displacement $\delta\mathbf{R}_\alpha$. Thus we write

$$\delta' E=\delta E+\sum_{i,\mathbf{G}}\left[\delta\psi_{i\mathbf{G}}^*\frac{\partial E}{\partial\psi_{i\mathbf{G}}^*}+\frac{\partial E}{\partial\psi_{i\mathbf{G}}}\delta\psi_{i\mathbf{G}}\right], \tag{23}$$

where the variations $\delta\psi_{i\mathbf{G}}$ must be chosen to solve the reorthonormalization conditions

$$\delta\langle\psi_i|\mathbf{S}|\psi_j\rangle+\sum_{\mathbf{G}}(\delta\psi_{i\mathbf{G}}^*\langle\mathbf{G}|\mathbf{S}|\psi_j\rangle+\langle\psi_i|\mathbf{S}|\mathbf{G}\rangle\delta\psi_{j\mathbf{G}})=0\ . \tag{24}$$

where $\mathbf{S}$ is the overlap operator, $|\mathbf{G}\rangle$ is an *augmented* plane wave, and $\delta\langle\psi_i|\mathbf{S}|\psi_j\rangle$ is the change in $\langle\psi_i|\mathbf{S}|\psi_j\rangle$ produced by the displacement of the sphere. Equations (24) can be solved by making

$$\delta\psi_{i\mathbf{G}}=\sum_j\delta\psi_{ij}\psi_{j\mathbf{G}}\delta_{\mathbf{k}_i\mathbf{k}_j}\ . \tag{25}$$

Here $\delta_{\mathbf{k}_i\mathbf{k}_j}$ is a delta function to indicate that only wave functions with the same $\mathbf{k}$ are orthonormalized and $\delta\psi_{ij}$ are new expansion coefficients to be found. Replacing Eq. (25) into Eq. (24) and imposing the symmetric orthogonalization condition $\delta\psi_{ji}=\delta\psi_{ij}^*$ one obtains $\delta\psi_{ij}^*=-\frac{1}{2}\delta\langle\psi_i|\mathbf{S}|\psi_j\rangle$. Substituting this into Eq. (25) and the resulting $\delta\psi_{i\mathbf{G}}$ in Eq. (23) and using $\partial E/\partial\psi_{i\mathbf{G}}^*=w_i\langle\mathbf{G}|\mathbf{H}|\psi_i\rangle$ we get

$$\delta' E=\delta E-\sum_{i,j}\delta_{\mathbf{k}_i\mathbf{k}_j}\frac{w_i+w_j}{2}\langle\psi_j|\mathbf{H}|\psi_i\rangle\delta\langle\psi_i|\mathbf{S}|\psi_j\rangle\ . \tag{26}$$

Now we can write

$$\langle\psi_i|\mathbf{S}|\psi_j\rangle=\int_{\mathbb{R}^3}\tilde{\psi}_i^*(\mathbf{r})\tilde{\psi}_j(\mathbf{r})d\mathbf{r}+\sum_\beta\int_{S_\beta}[\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})-\tilde{\psi}_i^*(\mathbf{r})\tilde{\psi}_j(\mathbf{r})]d\mathbf{r} \tag{27}$$

and it is then simple to transform Eq. (26) into

$$\delta' E=\delta E-\sum_{ij}\delta_{\mathbf{k}_i\mathbf{k}_j}\frac{w_i+w_j}{2}2\,\mathrm{Re}\left[\langle\psi_j|\mathbf{H}|\psi_i\rangle\int_{S_\alpha}[\delta\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})-\delta\tilde{\psi}_{\mathrm{i}}^*(\mathbf{r})\tilde{\psi}_{\mathrm{j}}(\mathbf{r})]\mathrm{d}\mathbf{r}\right]\ . \tag{28}$$

Notice that the $\delta\psi_{i\mathrm{G}}$ introduced in Eq. (23) for reorthonormalization have become "virtual" in Eq. (28) in the sense that, although they are implicitly included, they need not be computed explicitly in any way. Thus, all the $\delta\psi_i^*(\mathbf{r})$ in Eq. (28) refer exclusively to changes produced by the displacement of the sphere, while keeping all the $\psi_{i\mathrm{G}}$ fixed. If the wave functions $\psi_i(\mathbf{r})$ are converged eigenfunctions of the Hamiltonian (in the variational sense) then $\langle\psi_i|\mathbf{H}|\psi_j\rangle=\varepsilon_i\delta_{ij}$ and only the diagonal terms in Eq. (28) are nonzero[17]. Dividing by $\delta\mathbf{R}_\alpha$ we obtain the final result for the force,

$$\begin{aligned}\mathbf{F}_\alpha=&-\int_{\mathbb{R}^3}\nabla\tilde{V}(\mathbf{r})\Delta\tilde{n}^{\alpha}(\mathbf{r})d\mathbf{r}\\&-\sum_i 2w_i\mathrm{Re}\int_{S_\alpha}\left[\frac{\partial\psi_i^*(\mathbf{r})}{\partial\mathbf{R}_\alpha}[-\nabla^2+V(\mathbf{r})+\mu_{\mathrm{xc}}(\mathbf{r})-\varepsilon_i]\psi_i(\mathbf{r})-\frac{\partial\tilde{\psi}_{\mathrm{i}}^*(\mathbf{r})}{\partial\mathbf{R}_\alpha}[-\nabla^2+\tilde{\mathrm{V}}(\mathbf{r})+\tilde{\mu}_{\mathrm{xc}}(\mathbf{r})-\varepsilon_{\mathrm{i}}]\tilde{\psi}_{\mathrm{i}}(\mathbf{r})\right]\mathrm{d}\mathbf{r}\end{aligned} \tag{29}$$

or, more schematically,

$$\begin{aligned}\mathbf{F}_\alpha=&-\int_{\mathbb{R}^3}\nabla\tilde{V}(\mathbf{r})\Delta\tilde{n}^{\alpha}(\mathbf{r})d\mathbf{r}\\&-\sum_i 2w_i\mathrm{Re}[\langle\nabla_\alpha\psi_i|(\mathbf{H}-\varepsilon_i\mathbf{S})|\psi_i\rangle_\alpha\\&-\langle\nabla_\alpha\tilde{\psi}_i|(\mathbf{H}-\varepsilon_i\mathbf{S})|\tilde{\psi}_i\rangle_\alpha]\ ,\end{aligned} \tag{30}$$

where $\langle\ \rangle_\alpha$ means "integral in sphere α." For simplicity we have not included in these expressions the nondiagonal terms of Eqs. (26) and (28). These terms are likely to be rather small in practice, even for nonconverged wave functions. And even with these terms included, the forces calculated with nonconverged wave functions and potentials will be only approximate (as is the total energy). Thus we suspect that an efficient energy minimization can be done simultaneously for ion positions and electron wave functions even ignoring these nondiagonal terms. If this hypothesis would prove to be incorrect, it is clear that the inclusion of these terms would not represent any serious complication. Also, the nondiagonal terms are probably necessary in a molecular-dynamics simulation[1] for good stability in the integration of the equations of motion.[18]

Equation (30) is remarkably simple and practical. The dependence of $\tilde{\psi}_i^*(\mathbf{r})$ on $\mathbf{R}_\alpha$ comes only from structure factors of the form $\langle\mathbf{G}|\mathbf{R}_\alpha\rangle\propto\exp[-i(\mathbf{k}+\mathbf{G})\cdot\mathbf{R}_\alpha]$. Notice also that, in the case of a well-converged basis [i.e., for a sufficiently high $k_{\max}$ in Eq. (1)], Schrödinger's equation holds everywhere in space and the first term in the square brackets in Eq. (30) cancels. Thus, in this case, only nonaugmented functions enter in the computation of the force. However, in the practical case of a basis not perfectly converged we believe that it is better to keep this term, making a rigid shift of the potential inside the sphere.[11] Then the core density and the functions $\phi_{al\sigma}(r_\alpha)$ in Eq. (3) are fixed and only the expansion coefficients $\psi_{i\alpha lm\sigma}$ depend on $\mathbf{R}_\alpha$ through the same structure factors $\langle \mathbf{G}|\mathbf{R}_\alpha\rangle$ because of the change in the matching conditions imposed by $\tilde{\psi}_i(\mathbf{r})$ on the sphere boundary. Thus, the gradient $\partial/\partial\mathbf{R}_\alpha$ simply implies a product of each structure factor by $-i(\mathbf{k}+\mathbf{G})$. The part which multiplies $\mathbf{k}$ is pure imaginary and one obtains

$$\mathrm{Re}\langle \nabla_\alpha\psi_i|\mathbf{H}|\psi_i\rangle_\alpha=\mathrm{Im}\sum_{\mathbf{G}}\mathbf{G}\psi_{i\mathbf{G}}^{*}\langle\mathbf{G}|\mathbf{H}|\psi_i\rangle_\alpha\ , \tag{31}$$

and performing a similar operation with the other terms in Eq. (30) one finally obtains

$$\begin{aligned}\mathbf{F}_\alpha=&-\mathrm{Im}\sum_{\mathbf{G}}^{G_{\max}}\mathbf{G}\tilde{V}_{\mathbf{G}}^{*}\Delta\tilde{n}_{\mathbf{G}}^{\alpha}\\ &-\sum_i 2w_i\,\mathrm{Im}\sum_{\mathbf{G}}^{k_{\max}}\mathbf{G}\psi_{i\mathbf{G}}^{*}[\langle\mathbf{G}|(\mathbf{H}-\tilde{\mathbf{H}})|\psi_i\rangle_\alpha\\ &\qquad+\varepsilon_i\langle\mathbf{G}|(\mathbf{S}-\tilde{\mathbf{S}})|\psi_i\rangle_\alpha]\ ,\end{aligned} \tag{32}$$

where $\tilde{V}_{\mathbf{G}}$ and $\Delta\tilde{n}_{\mathbf{G}}^{\alpha}$ are the Fourier coefficients of $\tilde{V}(\mathbf{r})$ and $\Delta\tilde{n}^{\alpha}(\mathbf{r})$, respectively, and $G_{\max}$ is the cutoff for the expansion of the smooth density and potential. We have introduced the notation $\tilde{\mathbf{S}}$ meaning "overlap without augmenting" and $\tilde{\mathbf{H}}$ is also defined to apply between unaugmented wave functions. The matrix elements in Eq. (32) may be written as

$$\langle\mathbf{G}|(\mathbf{H}-\tilde{\mathbf{H}})|\psi_i\rangle_\alpha=\sum_{\mathbf{G}'}\langle\mathbf{G}|(\mathbf{H}-\tilde{\mathbf{H}})|\mathbf{G}'\rangle_\alpha\psi_{i\mathbf{G}'} \tag{33}$$

and the same for $\langle\mathbf{G}|\mathbf{S}-\tilde{\mathbf{S}}|\psi_i\rangle_\alpha$. The matrix elements $\langle\mathbf{G}|(\mathbf{H}-\tilde{\mathbf{H}})|\mathbf{G}'\rangle_\alpha$ and $\langle\mathbf{G}|(\mathbf{S}-\tilde{\mathbf{S}})|\mathbf{G}'\rangle_\alpha$ are currently computed in the LAPW method as parts of the secular matrix elements $\langle\mathbf{G}|\mathbf{H}|\mathbf{G}'\rangle$ and $\langle\mathbf{G}|\mathbf{S}|\mathbf{G}'\rangle$. Therefore it should be trivial to compute all the forces at every self-consistency step with a negligible increase in computation time.

We have assumed that the exact potentials $\tilde{V}(\mathbf{r})$ and $V(\mathbf{r})$ are calculated and used to compute the matrix elements $\langle\mathbf{G}|(\mathbf{H}-\tilde{\mathbf{H}})|\mathbf{G}'\rangle_\alpha$. But notice that approximations to the shape of the potential inside the spheres will benefit from the same cancellations usually argued for the total energy. Thus, the nonspherical (NS) part of the potential only enters as $V_{\mathrm{NS}}(\mathbf{r})n_{\mathrm{el}}(\mathbf{r})-\tilde{V}_{\mathrm{NS}}(\mathbf{r})\tilde{n}_{\mathrm{el}}(\mathbf{r})$. Since $V_{\mathrm{NS}}(\mathbf{r})$ and $\tilde{V}_{\mathrm{NS}}(\mathbf{r})$, as well as $n_{\mathrm{el}}(\mathbf{r})$ and $\tilde{n}_{\mathrm{el}}(\mathbf{r})$, are equal in value and slope at the sphere surface (where these quantities are usually largest) an important cancellation occurs between the terms neglected in a spherical-muffin-tin approximation.

We have recently obtained preliminary results[19] for frozen-phonon frequencies in semiconductors and transition metals, and for the bond length and bond stiffness in diatomic molecules, using the formula presented in this work. These preliminary results suggest that the forces calculated with the APW method are at least as accurate as those obtained with the pseudopotential-plane-wave method and converge considerably faster. More thorough calculations are currently under way.

We would like to acknowledge useful discussions with R. Car and M. Parrinello. One of us (J.M.S.) wants to thank the IBM Zürich and Thomas J. Watson Research Laboratories for invitations during which this work was initiated and to the Comisión Asesora de Investigatión Científica y Técnica (CAICyT), Madrid, Spain for financial support through Project No. PB86-0109.

---

[*] Permanent and mailing address.

[1] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[2] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[3] N. D. Lang and A. R. Williams, Phys. Rev. Lett. **34**, 531 (1975).

[4] J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979); **13**, 3095(E) (1980).

[5] K.-M. Ho, C.-L. Fu, and B. N. Harmon, Phys. Rev. B **28**, 6687 (1983).

[6] J. R. Chelikowsky and S. G. Louie, Phys. Rev. B **29**, 3470 (1984).

[7] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[8] D. D. Koelling and G. O. Arbman, J. Phys. F **5**, 2041 (1975).

[9] A. R. Williams, J. Kübler, and C. D. Gelatt, Jr., Phys. Rev. B **19**, 6094 (1979).

[10] P. Pulay, Mol. Phys. **17**, 197 (1969).

[11] O. K. Andersen, H. L. Skriver, and H. Nohl, Pure Appl. Chem. **52**, 93 (1979).

[12] A. R. Williams and J. M. Soler, Bull. Am. Phys. Soc. **32**, 562 (1987).

[13] L. Kleinman and R. Shurtleff, Phys. Rev. **188**, 1111 (1969).

[14] M. Weinert, J. Math. Phys. **22**, 2433 (1981).

[15] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).

[16] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).

[17] For Eqs. (26) and (28) to be variationally correct, the Hamiltonian must be consistent with the wave functions. Thus, the matrix elements $\langle\psi_i|\mathbf{H}|\psi_j\rangle$ are those prior to diagonalization.

[18] R. Car and M. Parrinello (private communication).

[19] J. M. Soler and A. R. Williams (unpublished).