RAPID COMMUNICATIONS

PHYSICAL REVIEW B VOLUME 42, NUMBER 15 15 NOVEMBER 1990-II

Augmented-plane-wave forces

José M. Soler*
Departamento de Física de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain
and IBM Research Division, Zürich Research Laboratory, CH-8803 Rüschlikon, Switzerland

Arthur R. Williams
IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598

(Received 9 July 1990)

Results are presented that demonstrate the effectiveness of a calculational method of electronic-structure theory. The method combines the power (tractable basis-set size) and flexibility (transition and first-row elements) of the augmented-plane-wave method with the computational efficiency of the Car-Parrinello method of molecular dynamics and total-energy minimization. Equilibrium geometry and vibrational frequencies in agreement with experiment are presented for Si, to demonstrate agreement with existing methods and for Cu, N_2, and H_2O to demonstrate the broader applicability of the approach.

Over the last fifteen years, *ab initio* methods based on the density-functional formalism and the local-density approximation for exchange and correlation have accumulated an impressive record of calculations.[1,2] These include not only electronic and structural properties of molecules, clusters, surfaces, and bulk solids but also thermal and kinetic properties as well. For many properties of such systems, numerical accuracies of 10% or better are systematically obtained and, with the continuous improvements in computer codes and hardware, these techniques may soon become an important tool in material and drug design. But, in order to make the calculations practical, it is frequently necessary to obtain not only the total energy but also the forces on the atoms. This is particularly important for quantum-mechanical molecular dynamics[3] and for the optimization of complex structures.[4-6] Unfortunately, some of the most powerful methods for total-energy calculations have serious difficulties obtaining the atomic forces directly. This is partly because in these methods electron wave functions are expanded in basis sets which depend on the atomic positions: when the atoms move, the energy variations associated with the change of the basis functions mask the relevant forces. This problem does not occur in the plane-wave pseudopotential (PWPP) method,[7] which gives the atomic forces directly and with high precision. In this method, the basis set is formed only by plane waves, which are of course independent of the atomic positions. This basis set has the additional advantages of being completely general and easy to improve systematically and it also makes the codes simpler and easier to vectorize. The price paid for this generality is high, however: Since plane waves treat all regions of space equally, the same precision is used where the wave functions vary rapidly or slowly and where the electron density is high or negligible. As a consequence, enormous basis sets are required (typically 1 or 2 orders of magnitude larger than localized basis sets), and they are especially inefficient when there are large interstitial or vacuum regions. Furthermore, for transition metals and other elements with localized orbitals, the pseudopotentials are very deep and calculations become impractical. Thus, only a rather limited fraction of elements and systems is accessible to this method and, consequently, to direct force calculations.

On the other hand, the augmented-plane-wave (APW) method[8,9] requires computations that are quite similar in form to those arising in the PWPP method but offers the advantage of making all the elements of the Periodic Table equally treatable. Like the PWPP method, it describes the electron-ion interaction in terms of an effective coupling among low-frequency plane waves. But, instead of explicitly using an effective (pseudo)potential to describe such a coupling, it "augments" the wave functions, close to the atoms, with exact solutions of the spherically averaged potential. These augmented wave functions are then used to calculate the effective matrix elements. A key difference between the two approaches is the requirement of norm-conserving pseudopotentials that the linear combination of plane waves constituting the "pseudo"-wave-function possess the full norm (charge) of the all-electron wave function. This requires the pseudo-wave-function to vary more sharply than its APW counterpart and to converge more slowly with increasing plane-wave frequency. Very recently, a scheme to release the norm-conservation condition was proposed,[10,11] bringing the PWPP method even closer to our APW method.

APW formulations also vary in their detail. Among them ours is the most similar to the PWPP method in that only the low-angular-momentum components of the wave function and the spherically symmetric component of the potential are augmented. As in the PWPP method, high-angular-momentum components of the wave functions and nonsymmetric components of the potential are described by the pseudo-wave-functions and the pseudodensity (the part of the wave functions expanded in plane waves and their square).

Although the APW method requires many fewer plane waves than the PWPP method, its basis set is still much larger than typical basis sets of localized orbitals.[9,12,13] Since, in the standard Rayleigh-Ritz procedure, the computation time typically increases as the square or the cube of the dimension of the secular matrix, the difference in efficiency is still very large, Car and Parrinello[3] have solved the bottleneck of the large plane-wave basis sets (in the case of the PWPP method) by formulating the electronic problem in a way that the number of plane waves N enters only as $N \ln N$. In addition, their method has the crucial advantage of making quantum-mechanical molecular dynamics feasible.[3] Both things together make the increased basis set a more than reasonable price to pay. In our APW method, we have used the same basic ideas to make the computation time increase only as $N \ln N$ with the number of plane waves. Furthermore, we have presented recently,[14] a simple formula for the atomic forces in the APW method. We have now implemented this formula, together with a number of improvements over the existing APW methods, to make a very efficient and versatile total energy and force method.

Our APW method will be described in more detail elsewhere. In brief, the electron wave functions $|\psi_i\rangle$, initially determined by their plane-wave expansion coefficients $\langle \mathbf{G}|\psi_i\rangle$, are expanded also in spherical harmonics and augmented in each atomic sphere. The mean energies $\varepsilon_i \equiv \langle\psi_i|H|\psi_i\rangle / \langle\psi_i|S|\psi_i\rangle$ and the derivatives

$$\langle \mathbf{G}|H|\psi_i\rangle \equiv \partial\langle\psi_i|H|\psi_i\rangle / \partial\langle \mathbf{G}|\psi_i\rangle^*$$

and

$$\langle \mathbf{G}|S|\psi_i\rangle \equiv \partial\langle\psi_i|S|\psi_i\rangle / \partial\langle \mathbf{G}|\psi_i\rangle^*$$

are computed. Then the expansion coefficients are modified according to the formula[15]

$$\delta\langle \mathbf{G}|\psi_i\rangle = -\frac{1-e^{-D_{\mathbf{G}}\delta t}}{D_{\mathbf{G}}}\langle \mathbf{G}|(H-\varepsilon_i S)|\psi_i\rangle . \tag{1}$$

δt is an "imaginary-time" step parameter and the second derivatives $D_{\mathbf{G}} \equiv \partial^2\varepsilon_i / \partial|\langle \mathbf{G}|\psi_i\rangle|^2$ are approximated by $G^2 \times$(interstitial volume). Equation (1) is an extension for a first-order differential equation of that proposed by Payne *et al.*,[16] and acts effectively as a preconditioned steepest descent method. After reorthonormalizing the wave functions by the Gram-Schmidt procedure, the iteration is repeated until the "forces" $\langle \mathbf{G}|(H-\varepsilon_i S)|\psi_i\rangle$ become small enough. Although more efficient iteration schemes have been proposed,[17,18] the above formula works adequately in the cases we have studied. Notice that only electron wave functions, rather than individual plane waves, are augmented. Thus, no matrix elements of the form $\langle \mathbf{G}|H|\mathbf{G}'\rangle$ are computed nor stored, this being an essential feature of the Car-Parrinello[3] approach. Since we do not construct a secular matrix, we do not need a common basis set for all wave functions and therefore we need not make a linear approximation for the energy dependence of the augmented wave functions (linerized APW, or LAPW, method[9]). More specifically, instead of using the same linearization energies ε_l for all wave functions, each $|\psi_i\rangle$ is linearized at $\varepsilon_l = \varepsilon_i$, thus effectively making an automatic "infinite panneling."[19]

As mentioned earlier, another important feature of our method is that the $l > l_{\max}$ components of the "interstitial" wave functions (which are not augmented), are allowed to penetrate into the muffin-tin spheres.[14,20] This ensures the exact continuity of the wave functions in value and slope for every angular momentum l, and allows for a substantial reduction of $l_{\max}$. In fact $l_{\max} = 2$ gives an accuracy of better than 1 mRy in the total energy for all the cases that we have studied and $l_{\max} = 1$ was sufficient to get the same precision with first row molecules. Also, the nonspherical components of the potential inside the atomic spheres are treated using their plane wave expansion rather than augmenting them as in the full-potential LAPW method (FLAPW).[21-24] This approximation can be tested by reducing the muffin-tin radius (and increasing the plane wave cutoff accordingly) and it also gives very small errors in all the cases studied. The initial wave functions are constructed by diagonalizing an atomic muffin-tin Hamiltonian (with a superposition of atomic densities in the interstitial region) in the subspace of a minimal LCAO basis set (with each atomic orbital fitted by a single Gaussian, expanded in APW's).

Table I presents the frequencies of various phonon modes of Si, compared with experiment and with those obtained analogously with the PWPP method.[7] We have used the same supercells of Ref. 7 and between 18 and 44 special **k** points[25] in an irreducible wedge of the Brillouin zone (IBZ) of the supercells. All the frequencies are obtained from the first derivatives of the forces (rather than from the second derivatives of the total energy), using atomic displacements of 1% of the lattice constant. A muffin-tin radius of 2.0 a.u. and a plane-wave cutoff of 10 Ry were used, which give a convergence of the total energy better than 1 mRy. It can be seen that the results compare with experiment as well as those of the PWPP method. In Table I we also compare with experiment several phonon frequencies of Cu, for which the PWPP method is not practical. In this case we used $r_{\mathrm{mt}} = 2.3$ a.u., $E_{\mathrm{cut}} = 15$ Ry, and between 36 and 121 special points in the IBZ of the two-atom supercells.[25] Again, there is good agreement with experiment.

Efficiency depends critically on the speed of the convergence of the atomic forces with wave function accuracy. Figure 1 shows the interatomic force in the nitrogen molecule for different self-consistency iterations. This is an

TABLE I. Experimental and calculated phonon frequencies for silicon and copper. Experimental values for Si and Cu are taken from Refs. 7 and 26, respectively. PWPP refers to Yin and Cohen (Ref. 7). Frequencies are in THz, lattice constants in a.u. Deviations from experiment are given in parenthesis.

	Expt.	APW		PWPP	
Silicon lattice constant	10.26	10.21	(−0%)	10.30	(0%)
LTO(Γ)	15.53	15.30	(−2%)	15.14	(−3%)
LOA(X)	12.32	12.11	(−2%)	11.98	(−3%)
TO(X)	13.90	13.67	(−2%)	13.51	(−3%)
TA(X)	4.49	4.41	(−2%)	4.37	(−3%)
Copper lattice constant	6.82	6.70	(−2%)		
L(X)	7.25	7.44	(3%)		
T(X)	5.13	5.27	(3%)		
L(L)	7.30	7.57	(4%)		
T(L)	3.42	3.58	(5%)		

especially demanding system for the APW method. First, molecules in general represent an extreme case of "inefficiency" for a plane-wave basis set because most of the volume is empty in the unit cell. Furthermore, nitrogen has the shortest and strongest bond among first-row molecules (10% shorter and 90% stronger than oxygen). Finally, to be able to study the potential curve, we have used a muffin-tin radius of only 0.95 a.u., which is less than half the equilibrium distance. To make the interaction between repeated molecules small enough, we used a simple cubic unit cell with a side of 8 a.u. and one special **k** point. Under these conditions, the total energy is converged to 1 mRy with respect to the lattice constant but we need a plane-wave cutoff of 60 Ry and over 4000 plane waves to obtain such a precision. However, because of our Car-Parrinello-like approach, this calculation is not especially demanding in terms of computer resources (much less demanding than that of the Cu phonons, for example), showing the great potential of this method is molecular calculations. It can be seen in Fig. 1 that the calculated forces converge very fast and that they agree perfectly with the derivative of the total energy after a few iterations. Notice that reasonable forces are available already at the second iteration, when the total energy is still 8 eV away from its final value. In fact, even at the first iteration, when the total energy is 3.5 Ry above its converged value, the force cancels at a distance only 6% smaller than the final equilibrium distance. The values obtained for the bond length, binding energy, and vibrational frequency are presented in Table II and compare very well with experimental values. Only the binding energy is substantially overestimated, this being a well-known problem of the local-density approximation.

TABLE II. Experimental and calculated parameters for nitrogen and water molecules. Deviations from experiment are given in parenthesis.

	Expt.	APW	
N_2 binding energy (eV)	9.9	11.5	(16%)
Bond length (a.u.)	2.07	2.07	(0%)
Vibration frequency (cm^{-1})	2358	2400	(2%)
H_2O binding energy (eV)	9.5	11.6	(22%)
Bond length (a.u.)	1.81	1.83	(1%)
Bond angle (deg)	104.5	102.4	(−2%)
Vibration frequencies:	1595	1610	(1%)
(cm^{-1})	3657	3600	(−2%)
	3756	3670	(−2%)

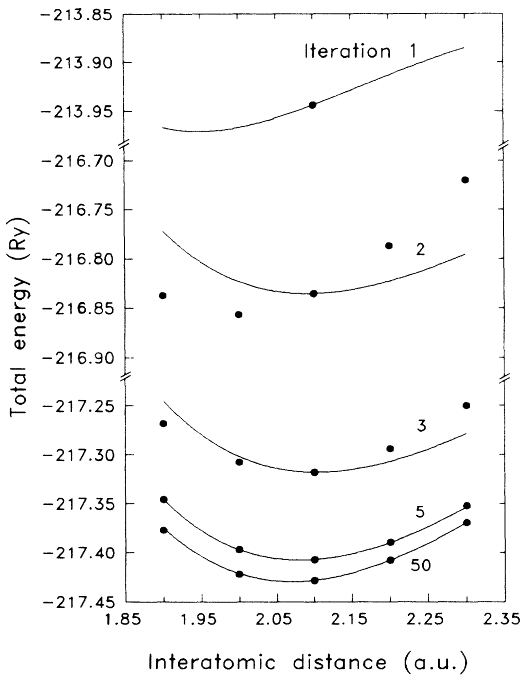

FIG. 1. Total energy of N_2 at different self-consistency iterations, as a function of interatomic distance d. The solid circles are the calculated total energies for five distances (for the first iteration, other solid circles are out of the range of the figure). Solid curves give the total energies obtained by integrating the calculated forces. These are previously fitted to a cubic polinomial from the five calculated points. The curves have been adjusted vertically to the calculated total energy at $d = 2.0$ a.u.

The rapid convergence of the forces suggests the feasibility of a simultaneous relaxation of nuclear and electronic degrees of freedom. We show such a simultaneous relaxation in Table III for a simple but nontrivial system: the waver molecule. Starting rather far from its equilibrium structure, we move the hydrogen atoms according to their forces (with the steepest descent method), keeping the oxygen atom at the origin. The unit cell was the same as for the nitrogen molecule. A plane-wave cutoff of 30 Ry was used during the minimization. This was then increased to 50 Ry in order to obtain perfectly converged forces, total energies, and the Hessian matrix[27] at the final positions. From these we find the bond length, bond angle, binding energy, and vibrational frequencies, which are given in Table II and also compare very well with the experimental values.

In conclusion, we have presented a very accurate, fast and flexible *ab initio* method to study molecules, clusters, and bulk systems with elements of the whole Periodic Table. Atomic forces are obtained directly, opening the door to optimization of complex structures and to *ab initio* molecular dynamics for transition metals.

TABLE III. Evolution of the bond length (in a.u.) and bond angle (in deg.) of the water molecule and of the total force of each H atom (in mRy/a.u.) during a simultaneous optimization of electron wave functions and atomic positions.

Iteration	Bond length	Bond angle	Force
5	2.000	90.0	89.3
10	1.943	91.4	37.9
20	1.885	94.4	17.7
40	1.861	98.6	8.2
60	1.855	100.9	4.1
80	1.854	102.0	2.1
100	1.854	102.4	1.1

We want to acknowledge useful conversations with D. C. Allan, O. K. Andersen, E. Artacho, R. Car, P. J. H. Denteneer, J. Kübler, M. Methfessel, M. Parrinello, M. C. Payne, K. Schwarz, and M. T. Teter. One of us (J.M.S.) wants to acknowledge the hospitality of the IBM Thomas J. Watson Reserch Laboratory and the International School for Advanced Studies (Trieste, Italy) where part of the method was developed. This work has been supported by Grant No. PB86-0109 from the Comision Interministerial de Ciencia y Tecnologia.

---

*Electronic address: JSOLER@EMDUAM11.

[1] *Atomic Scale Calculations in Materials Science,* edited by, J. Tersoff, D. Vanderbilt, and V. Vitek, MRS Symposia Proceedings No. 141 (Materials Research Society, Pittsburgh, 1989).

[2] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

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

[4] M. C. Payne, P. D. Bristowe, and J. D. Joannopoulos, Phys. Rev. Lett. **58**, 1348 (1987).

[5] D. C. Allan and M. P. Teter, Phys. Rev. Lett. **59**, 1136 (1987).

[6] Pietro Ballone, Wanda Andreoni, Roberto Car, and Michelle Parrinello, Phys. Rev. Lett. **60**, 271 (1988).

[7] M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 3259 (1982).

[8] J. C. Slater, Phys. Rev. **51**, 846 (1937).

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

[10] David Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[11] P. E. Blöchel (unpublished).

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

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

[14] José M. Soler and Arthur R. Williams, Phys. Rev. B **40**, 1560 (1989).

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

[16] M. C. Payne, J. D. Joannopoulos, D. C. Allan, M. P. Teter, and D. H. Vanderbilt, Phys. Rev. Lett. **56**, 2656 (1986).

[17] I. Stich, R. Car, M. Parrinello, and S. Baroni, Phys. Rev. B **39**, 4997 (1989).

[18] Michael P. Teter, Michael C. Payne, and Douglas C. Allan, Phys. Rev. B **40**, 12255 (1989).

[19] The derivation of the force formula in Ref. 14 does not depend on the choice of the linearization energies.

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

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

[22] L. F. Mattheiss and D. R. Hamann, Phys. Rev. B **33**, 823 (1986).

[23] P. Blaha, K. Schwarz, P. Sorantin, and S. B. Trickey, Comput. Phys. Commun. **59**, 399 (1990).

[24] We include only the spherically symmetric part of H-$\tilde{H}$ in Eq. (32) of Ref. 14. Consistently, the extra compensating charge considered in this reference is added only for the monopole term.

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

[26] R. M. Nicklow, G. Gilat, H. G. Smith, L. J. Raubenheimer, and M. K. Wilkinson, Phys. Rev. **164**, 922 (1967).

[27] That is, the matrix of second derivatives of the total energy as a function of atomic positions. This was computed by a series of total-energy and force calculations, each run with one atom displaced by ±0.1 a.u. in one of the cartesian directions. Notice that the availability of the forces makes the computation increase linearly rather than with the square of the number of degrees of freedom.