Pergamon

Solid State Communications, Vol. 99, No. 2, pp. 57–61, 1996
Copyright © 1996 Published by Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0038-1098/96 $12.00 + .00

PII S0038-1098(96)00291-8

ADAPTIVE COORDINATE, REAL-SPACE ELECTRONIC STRUCTURE CALCULATIONS ON PARALLEL COMPUTERS

Gil Zumbach, N. A. Modine and Efthimios Kaxiras
Department of Physics, Harvard University, Cambridge MA 02138, U.S.A.

(*Received* 28 *February* 1996; *accepted* 8 *May* 1996 *by S. Louie*)

We present a method for electronic structure calculations that retains all of the advantages of real space, but addresses the major weakness of a regular grid, i.e., its inability to treat some regions of space with more resolution than others. The computations are carried out on a *regular* mesh in *curvilinear space*, which allows natural and efficient decomposition on parallel computers, and effective use of iterative numerical methods. A novel feature is the use of error analysis to optimize the curvilinear grid for highly inhomogeneous electronic distributions. We report accurate all-electron calculations for H_2, O, and O_2. Copyright ©1996 Published by Elsevier Science Ltd.

Keywords: D. electronic states.

1. INTRODUCTION

AB INITIO electronic structure calculations are computationally very challenging because the singular Coulomb potential of the nuclei results in highly localized core wavefunctions with cusps at the nuclear positions. Even when pseudopotentials are used to eliminate the core electrons, it is often desirable to treat valence electrons with highly localized wavefunctions (e.g., 1s, 2p, 3d, or 4f valence electrons), on the same footing as delocalized ones. Typical implementations use a basis in a (one-particle) Hilbert space, the choice of which requires a tradeoff between simplicity and fast convergence of physical quantities with basis size. The simplest basis consists of plane waves. Its main drawback is uniform precision, leading to slow convergence for inhomogeneous systems such as atoms, molecules, clusters, or solid surfaces. On the other hand, basis sets such as linearized augmented plane waves (LAPW) or linearized muffin tin orbitals (LMTO) can be tailored to specific physical problems and therefore have excellent convergence properties. However, they tend to lead to complex equations. A promising alternative is a real space approach. All terms are local except for the Laplacian, which is short range. The resulting sparse Hamiltonian allows effective use of iterative algorithms, which vastly reduce both memory and time requirements. A sparse Hamiltonian is also a prerequisite for any $O(N)$ treatment of electronic structure.

Harnessing the computational power of massively parallel architectures imposes additional constraints on the choice of basis. To achieve good load balance, computational complexity and memory requirements must be evenly divided among processors, a task made very difficult by complex bases like LAPW and LMTO. Another important consideration is the minimization and localization of communication between processors. Since Fourier transforms (the underlying operations in a plane wave basis) require communication between all processors, even plane waves are not an efficient basis in this respect. In contrast, a *regular* grid in real space is a very natural choice for a massively parallel computer architecture: assigning an equal section of the grid to each processor provides good load balance, minimizes interprocessor communication, and produces communication patterns that are both local and conflict-free.

A number of groups [1–5] have recently reported real space electronic structure computations using a regular grid. When applied to an inhomogeneous problem, a regular grid in real space suffers from the same drawback as a plane wave basis, i.e., it has the same resolution in every region of space. This problem has been circumvented by locally adding more basis ele-

---

[1] This work was supported by the Office of Naval Research grant N00014-93-1-0190.

ments using wavelets [6, 7], finite-elements [8, 9], and multi-grid methods [10]. Irrespective of the formalism, such local enhancement interferes with the natural mapping onto a parallel machine and makes efficient parallelization more challenging.

Progress toward overcoming the limitations of plane wave bases through the introduction of a curvilinear coordinate system has also been reported recently [11–13]. Pioneered by Gygi [11], this adaptive plane wave approach enhances precision in important regions of space by adapting the change of coordinates to each physical problem. The adaptive plane wave approach eliminates the major drawback of the standard plane wave basis, but lacks the simplicity, sparseness, and natural parallelization properties of real space algorithms.

In the present work, we combine the advantages of real space calculations and those of adaptive coordinates in a scheme for performing electronic structure calculations within the context of density functional theory using either the local density approximation (LDA) [14, 15], or the generalized gradient approximation (GGA) [16]. Our scheme is efficient and accurate for systems with *very inhomogeneous charge distributions* and takes full advantage of massively parallel computer architectures. The central idea is to work on a regular grid, but in a curvilinear space $\boldsymbol{\xi}$. The change of coordinates $\boldsymbol{x}(\boldsymbol{\xi})$, generates a *single* grid in real space $\boldsymbol{x}$, which is finer where high precision is needed. This general idea is very well established in the study of fluid flow and heat transfer [17–19]. More recently, Gygi and Galli [20] have reported an implementation of real-space adaptive-coordinate electronic structure calculations using pseudopotentials. We refer to our method as the Adaptive Coordinate, Real-space, Electronic Structure (ACRES) algorithm. It has the following advantages:

1. It can achieve an essentially optimal distribution of grid points. This is accomplished by a versatile choice of the curvilinear coordinates, through which a grid of fixed size is adapted to provide resolution commensurate with the physics. The only cost of the adaptation is the introduction of a nontrivial metric $g^{\alpha\beta}(\boldsymbol{\xi})$.
2. The coordinate transform is chosen in order to minimize the discretization error. The idea is to determine a good set of curvilinear coordinates through the use of error analysis. This differs from Gygi's approach [11] in the adaptive plane wave scheme, where the variational principle makes it possible to find the optimal change of coordinates through an energy minimization.
3. Since the communications pattern remains exactly the same as that of a regular grid in real space, highly efficient parallelization is trivially accomplished.
4. The sparsity of the equations allows us to take advantage of iterative algorithms. This makes it possible to employ rather large grids, and consequently to investigate complex systems. The computational time scales as $N \times n_e$ where N is the total number of points in the 3-dimensional grid and n_e the number of electrons.

2. THEORY AND IMPLEMENTATION

The real space coordinates $x^i(\xi^\alpha; P^m)$ depend on the curvilinear coordinates ξ^α and on some set of parameters P^m that allow us to tune the change of coordinates to a particular problem [21]. The Jacobian is given by

$$J^i_\alpha(\xi; P) = \partial x^i / \partial \xi^\alpha \qquad (1)$$

with $|J| = \det J$ its determinant. The trivial metric $g^{ij} = \delta_{ij}$ in real space corresponds to the metric $g^{\alpha\beta} = J^{-1}{}^{\alpha}_{i} J^{-1}{}^{\beta}_{i}$ in curvilinear coordinates (summation over repeated indices is implied). The Laplacian operator in curvilinear space is

$$\Delta = \frac{1}{|J|}\partial_\alpha \left(|J| g^{\alpha\beta} \partial_\beta\right), \qquad (2)$$

and the integrals are transformed according to $\int d^3x = \int d^3\xi \ |J|$. The Coulomb potential is found by solving the Poisson equation [discretized in curvilinear coordinates by means of the Laplacian, Eq. (2)] with the sum of electronic and nuclear charges as the source.

The equations are discretized in a box of linear size Λ_i, using a finite difference scheme on a regular grid in curvilinear space $\boldsymbol{\xi}$ with N_i points in each direction. Any boundary conditions, including the phase shifts required in performing multiple k-point calculations for solids, can easily be implemented in this approach. In the following, we use periodic boundary conditions.

Implementation of the method presents certain challenges associated with the choice of discretization schemes. The most important ones, and the manner in which we resolved them, are discussed here briefly:

Equations that are equivalent in the continuum limit are not necessarily equivalent after discretization. For example, there exist several expressions for the Laplacian which are equivalent in the continuum limit. From physical and computational considerations, it is desirable to have a self-adjoint discretization of the Laplacian. The expression given in Eq. (2) is self-adjoint after discretization if, for a fixed pair of indices (α, β), the finite difference operators used to represent ∂_α and ∂_β are identical.

The order of the finite difference approximation for the derivatives is very important. The lowest order, two-point symmetric derivative is insufficient and does not give good results. Our experience indicates that we need to use a symmetric discretization for the derivatives with at least four points (second order).

The discrete representation of the nuclear charge is equally important. For an atom with atomic number Z at position $\boldsymbol{R}$, the nuclear charge is $\rho(\boldsymbol{\xi}) = Z\,\delta(\boldsymbol{\xi};\boldsymbol{R})$ where $\delta(\boldsymbol{\xi};\boldsymbol{R})$ is a representation of a Dirac δ function at $\boldsymbol{R}$ on the regular grid in $\boldsymbol{\xi}$ space. In addition to the normalization condition, an important constraint on the representation of the δ function is

$$\int d\boldsymbol{\xi}\,|J|\,\delta(\boldsymbol{\xi};\boldsymbol{R})\,\boldsymbol{x}(\boldsymbol{\xi}) = \boldsymbol{R}. \tag{3}$$

Constraints on higher moments could also be imposed [22], but we did not find this necessary. We found the most useful representation to be a Gaussian.

$$\delta(\boldsymbol{\xi};\boldsymbol{R}) \propto \exp\left(-|\boldsymbol{\xi}-\boldsymbol{\xi}_0|^2 / 2\sigma^2\Delta\xi^2\right) \tag{4}$$

with $\Delta\xi$ the regular grid spacing, σ an adjustable parameter, and $\boldsymbol{\xi}_0$ chosen to satisfy Eq. (3). This choice reduces the translational invariance problems discussed below.

The presence of the grid breaks translational invariance. We call the distance between an atomic center and the nearest grid point the *offset*. The energy depends on the offset, and this effect can be quite large due to the Coulomb singularity. The dependence is reduced by strong adaptation, which makes the cell of the real space grid very small near the atomic sites. The Gaussian representation of the δ function further reduces the dependence of the energy on the offset because it results in a smoother transfer of charge as the position of an atom changes. The combined use of strong adaptation and a Gaussian δ function eliminates the translation invariance problem.

A final challenge is the actual choice of curvilinear coordinates. A necessary condition for the mapping between $\boldsymbol{x}$ and $\boldsymbol{\xi}$ is that it be single valued, i.e., the grid in x space must not be folded. As the Laplacian involves the derivative of the metric, and the metric is computed from the Jacobian, the mapping must be twice continuously differentiable on a 3D domain with periodic boundary conditions in order to ensure smoothness. It is also desirable that the mapping be spherically symmetric around an atom. We use a two level coordinate transformation $\boldsymbol{x}(\boldsymbol{\xi};P)$ with a *global backdrop* useful for simulating isolated structures, and further *local adaptation* around each atom position. The global backdrop is a simple independent transformation along each axis $x^i = x^i(\xi^i)$ and creates a flat central region with a high density of grid points and a surrounding region with a decreasing density of grid points. The local adaptation creates a spherical deformation of the grid around each atomic center $\boldsymbol{R}_\nu$, with the amount of adaptation A_ν and the size of the adapted region ρ_ν as independent variables. For a given $\det J(\boldsymbol{R}_\nu)$ and ρ_ν, the final results are not very sensitive to the details of the formula for $\boldsymbol{x}(\boldsymbol{\xi})$ [11]. The computations presented below were carried out with the simple form for the local adaptation

$$\boldsymbol{x}(\boldsymbol{\xi};P) = \boldsymbol{\xi} - \sum_\nu A_\nu\,(\boldsymbol{\xi}-\boldsymbol{R}_\nu)\, \exp\left(-\frac{|\boldsymbol{\xi}-\boldsymbol{R}_\nu|^2}{2\,\sigma^2(A_\nu,\rho_\nu)}\right) \tag{5}$$

with $\sigma(A,\rho)$ chosen so that ρ gives the real space width of the adapted region.

In order to generate a nearly optimal mesh for a given physical problem, we define a *merit functional* $m(f;P)$ and choose the parameters P that minimize this quantity. To motivate our choice of a merit functional, consider an estimate for the error in a periodic, one-dimensional integral $I(f) = \int dx\,f(x) = \int d\xi\,\tilde{f}(\xi;P)$ with $\tilde{f}(\xi;P) = |J|(\xi;P)\,f(x(\xi;P))$. The integral is computed numerically on a regular mesh in ξ coordinates

$$I_N(f;P) = \sum_i \Delta\xi\,\tilde{f}(\xi_i;P) \tag{6}$$

with mesh spacing $\Delta\xi$. We estimate the error [23] by using I_N as an approximant for I and comparing its value to $I_{N/2}$. With $N/2$ points, the rectangular element of integration is $\delta I_{N/2} = 2\Delta\xi\,\tilde{f}(\xi_i;P)$. With N points, the same region contributes $\delta I_N = \Delta\xi\,[\tilde{f}(\xi_{i-1};P)/2 + \tilde{f}(\xi_i;P) + \tilde{f}(\xi_{i+1};P)/2]$ to the integral. An estimate of the error contributed by this region is given by

$$\delta e(f) = \delta I_N - \delta I_{N/2} = \Delta\xi^3\,\tilde{f}_i''/2. \tag{7}$$

To avoid cancellation of errors from different regions of space, we take our merit functional to be

$$m(f;P) = \left(\sum_i (\delta e)^2\right)^{1/2} = \frac{1}{2}\,\Delta\xi^{5/2}\left(\sum_i \Delta\xi\,(\tilde{f}_i'')^2\right)^{1/2}. \tag{8}$$

This provides a measure of how well the grid represents the function f and is suitable for minimization with respect to P. The above idea is easily generalizable to three-dimensional integrals. The last step is to pick an integrand f so as to obtain an *a priori* estimate of the optimal grid parameters by minimizing $m(f;P)$. By experimenting on several atoms we have found that the exchange-correlation energy density provides an adequate indicator. Due to large cancellations forced by the Kohn-Sham eigenvalue equation, this term gives the leading factor for the error in the total energy.

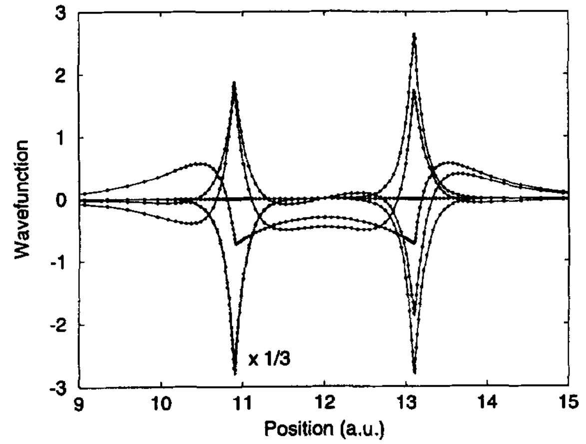

Fig. 1. Occupied wave functions of the O_2 molecule, along a line through the centers of the atoms. The π bonding and anti-bonding wave functions collapse onto the horizontal axis (they have nodes through the atomic centers). The 1*s* bonding and anti-bonding states were scaled by a factor of 1/3 so they could be displayed on the same scale.

Table 1. Calculated bond length a_0 and vibrational frequency ω for H_2 and O_2, minimum energy E_0 for H_2, and atomic energy E_{at} of O. The zero-point vibrational energy was subtracted from the experimental total energy of H_2.

	ACRES	Other DFT Results	Experiment
H_2 [LDA]			
a_0 (a.u.)	1.448	1.446[a]	1.401[b]
ω (cm^{-1})	4192	4207[a]	4401[b]
E_0 (Ry)	-2.276	−2.27[c]	−2.349[b]
H_2 [GGA]			
a_0 (a.u.)	1.416	1.413[d]	1.401[b]
ω (cm^{-1})	4381	4373[d]	4401[b]
E_0 (Ry)	-2.340	−2.34[e]	−2.349[b]
O [LDA]			
E_{at} (Ry)	-148.870	−148.938[f]	−150.027[g]
O [GGA]			
E_{at} (Ry)	-149.912	−149.994[h]	−150.027[g]
O_2 [LDA]			
a_0 (a.u.)	2.32	2.30[a]	2.28[b]
ω (cm^{-1})	1661	1642[a]	1580[b]
O_2 [GGA]			
a_0 (a.u.)	2.34	2.34[d]	2.28[b]
ω (cm^{-1})	1557	1518[d]	1580[b]

[a] Reference [26], Spin polarized, S-VWN functional
[b] Reference [27]
[c] Reference [28], Spin polarized, Perdew-Wang functional
[d] Reference [26], Spin polarized, B-LYP functional
[e] Reference [28], Spin polarized, PW GGA-II functional
[f] Reference [29], Spin unpolarized, Perdew-Zunger functional
[g] Reference [30], Spin unpolarized ground state, $2p^4\ {}^1D$
[h] Reference [29], Spin unpolarized, PW91 functional

3. RESULTS AND DISCUSSION

Using the approach described above, we have performed DFT/LDA [15] and DFT/GGA [16] electronic structure calculations on the CM-5 massively parallel supercomputer. Within this approach, all-electron computations involving atoms in the first row of the periodic table are feasible. We have also implemented the norm-conserving nonlocal pseudopotentials of Bachelet et al. [24], and the Kleinman-Bylander procedure that renders the nonlocal components separable [25].

For the all-electron calculations, the adaptation of the grid is determined by the requirement that the density of grid points near the atomic cores is sufficient to accurately represent the $1/r$ divergence of the Coulomb potential. For example, Fig. 1 shows the occupied wave functions of the O_2 molecule along a line through the centers of the two atoms.

The enhancement of the grid resolution throughout the regions where the electronic wave functions are large and the very strong enhancement close to the nuclei allow accurate representation of the smooth tails of the wave functions as well as the cusps and nodes near the nuclei.

For a more quantitative comparison to other theoretical results and to experiment, Table 1 shows our calculated results for H_2, O, and O_2. It is clear from this comparison that our results are in good agreement with other theoretical work using similar methods. We find a relatively slow deviation from converged values as the number of grid points is reduced. For example, the results for H_2 in Table 1 were obtained using a 12 × 12 × 24 a.u. box and a 128 × 128 × 256 grid. The bond length, vibrational frequency, and binding energy change by approximately 0.1 % with a grid of 64 × 64 × 128 points, and by approximately 1% with a grid of 32 × 32 × 64 points.

In order to compare ACRES to a more traditional method, we computed the total energy of a H atom using several sizes of unadapted, regular grids. The resulting fractional errors in the total energy and the computational times are compared to corresponding ACRES results in Table 2. The relevant comparison is between results of comparable accuracy: the ACRES computation with a 32 × 32 × 32 point grid achieved slightly better accuracy (at a cost of 18 seconds) than the regular grid computation with 128 × 128 × 128 points, which cost 300 seconds. Therefore, for an equivalent level of accuracy, ACRES greatly reduces computational time and memory requirements relative to a regular grid method. In fact, it would be extremely difficult to do accurate all electron computations for the systems discussed in Table 1 without

Table 2. Convergence with respect to grid size of the total energy of a H atom in a 12 × 12 × 12 a.u. box. The errors are computed relative to the largest (most accurate) ACRES calculation

	Regular Grid		ACRES	
grid size	error	time (s)	error	time (s)
32^3	9%	11	0.9%	18
64^3	3%	32	0.1%	59
128^3	1%	300	-	442

adaptation of the grid.

Acknowledgements—We are grateful to F. Gygi for sharing his thoughts on adaptive methods.

REFERENCES

1. J. R. Chelikowsky, N. Troullier, and Y. Saad, Phys. Rev. Lett. **72**, 1240 (1994); J. R. Chelikowsky, N. Troullier, K. Wu, and Y. Saad, Phys. Rev. B **50**, 11355 (1994).
2. J. Bernholc, J.-Y. Yi, and D. J. Sullivan, Faraday Discuss. **92**, 217 (1991); E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **52**, R5471 (1995).
3. S. Baroni and P. Giannozzi, Europhys. Lett. **17**, 547 (1992).
4. K. A. Iyer, M. P. Merrick, and T. L. Beck, J. Chem. Phys. **103**, 227 (1995).
5. T. Hoshi, M. Arai, and T. Fujiwara, Phys. Rev. B **52**, R5459 (1995).
6. K. Cho, T. A. Arias, J. D. Joannopoulos, and P. K. Lam, Phys. Rev. Lett. **71**, 1808 (1993).
7. S. Q. Wei and M. Y. Chou, preprint.
8. S. R. White, J. W. Wilkins, and M. P. Teter, Phys. Rev. B **39**, 5819 (1989).
9. E. Tsuchida and M. Tsukada, Sol. St. Comm. **94**, 5 (1995); Phys. Rev. B **52**, 5573 (1995).
10. E. J. Bylaska *et al.*, in *Proc. 6th SIAM Conf. Parallel Processing for Sci. Comput.* (San Francisco, 1995).
11. F. Gygi, Europhys. Lett. **19**, 617 (1992); Phys. Rev. B **48**, 11692 (1993); **51**, 11190 (1995); Private communication.
12. D. R. Hamann, Phys. Rev. B **51**, 7337 (1995); **51**, 9508 (1995).
13. A. Devenyi, K. Cho, T. A. Arias, and J. D. Joannopoulos, Phys. Rev. B **49**, 13373 (1994).
14. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. Sham, *ibid.* **140**, A1133 (1965).
15. We use the exchange-correlation local functional proposed by J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
16. J. P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991).
17. W. H. Frey, Int. J. Numer. Methods Eng. **11**, 1653 (1977).
18. C. D. Mobley and R. J. Stewart, J. Comput. Phys. **34**, 124 (1980).
19. W. C. Thacker, Int. J. Numer. Methods Eng. **15**, 1335 (1980), and references therein.
20. F. Gygi and G. Galli, Phys. Rev. B **52**, R2229 (1995).
21. Roman (Greek) letters indicate real (curvilinear) coordinate indices, with the standard convention of upper (lower) indices corresponding to contravariant (covariant) components.
22. R. W. Hockney and J. W. Eastwood, *Computer Simulations Using Particles* (Adam Hilger, Philadelphia, PA, 1988).
23. The second Euler-Maclaurin summation is only an asymptotic formula and cannot be used to estimate the error for a discretized integral of a C^∞ periodic function, where it identically gives zero independent of the grid.
24. G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
25. L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
26. B. G. Johnson, P. M. W. Gill, and J. A. Pople, J. Chem. Phys. **98**, 5612 (1993).
27. K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure* (Van Nostrand Reinhold Company, New York, 1979), Vol. IV.
28. J. P. Perdew *et al.*, Phys. Rev. B **46**, 6671 (1992).
29. Y.-M. Juan and E. Kaxiras, Phys. Rev. B **48**, 14944 (1993); Private communication.
30. C. E. Moore, *Atomic Energy Levels* (U. S. Government Printing Office, Washington, 1971), Vol. I.