PHYSICAL REVIEW B VOLUME 40, NUMBER 18 15 DECEMBER 1989-II

Solution of Schrödinger's equation for large systems

Michael P. Teter
Applied Process Research, Corning Glass Works, Corning, New York 14831
and Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501

Michael C. Payne
Cavendish Laboratory, Madingley Road, Cambridge, CB3 0HE, United Kingdom

Douglas C. Allan
Applied Process Research, Corning Glass Works, Corning, New York 14831

(Received 24 April 1989; revised manuscript received 7 August 1989)

Iterative diagonalization of the Hamiltonian matrix is required to solve very large electronic-structure problems. Present algorithms are limited in their convergence rates at low wave numbers by stability problems associated with large changes in the Hartree potential, and at high wave numbers with large changes in the kinetic energy. A new method is described which includes the effect of density changes on the potentials and properly scales the changes in kinetic energy. The use of this method has increased the rate of convergence by over an order of magnitude for large problems.

I. INTRODUCTION

The motivation for this work has its foundations in attempts to model periodic cells containing many atoms while using a plane-wave basis. The number of plane waves required is moderate when modeling a material such as silicon, but is much larger when some of the atoms are first-row or transition-metal atoms. The nonlocal pseudopotentials associated with these atoms are quite deep and require high–wave-number plane waves to describe the resulting wave functions. The dual requirements of simultaneously describing a large cell volume and a high kinetic energy lead to an explosive growth in the number of basis functions. The need for a new algorithm became apparent when simulations had to be abandoned due to computational difficulties related to the large number of plane waves, as well as the charge instabilities associated with large systems.

We outline an adaptation to systems which have orthonormality constraints of an iterative conjugate-gradient method known to be efficient in very large minimization problems. The practical use of this method has been limited by the difficulty of determining the exact linear combinations of vectors which minimize the total energy. We describe a simple new technique for minimizing the total energy in a self-consistent manner.

II. BACKGROUND

The method of self-consistent fields commonly used in quantum-mechanical simulation is gracefully nonlinear. In the usual formulation, the wave function of the system to be studied is expressed as a product of single-particle eigenstates. The energy of the system is a function of a set of coefficients of basis functions. While there exist many stationary points in the energy, for unpolarized electrons there are no false minima. As long as the ground state is not orthogonal to the starting point, the exact ground state is reachable from any starting set of coefficients by following a path of decreasing energy. Since the energy is at least a quartic function of the coefficients, no direct method of solution exists and approximate solutions must be iterated until they no longer change. To begin the standard procedure, an initial electron-electron potential is determined by some sort of guess and a Hamiltonian matrix is generated. The Hamiltonian matrix is diagonalized and the lowest eigenvectors are occupied. An electron density and corresponding electron-electron potential are generated from the solutions, and the process begins again. If one simply uses the output electron density from one iteration as the input density to the next, it is found that the process oscillates and each new electron density has overcorrected that from the previous iteration. This overcorrection is typically handled by averaging the input and output densities in some manner. As one studies physically larger systems, it is observed that the oscillation becomes unstable and the new solutions overcorrect the previous ones in ways that simple density averaging cannot easily overcome. Higher-order schemes exist to surmount these difficulties, and their discussion forms a subject by itself.[1] Most of these methods are reasonably well conceived and are useful for the simulation of systems of medium size. Some of the methods have been designed to partially overcome the lack of self-consistency in the diagonalization of the Hamiltonian matrix by simple wave-vector-dependent damping schemes, but are approximate. Others model the dielectric response of the material under study,[2] using information gained from previous diagonalizations of the converging Hamiltonian matrix. The effort required to model the dielectric matrix is a small part of the effort required to diagonalize the Hamiltonian matrix. While it is unquestionably better to damp the density changes differently for different wave vectors than to simply mix densities, the knowledge of the dielectric response of the system under study must become much more detailed for rapid convergence of extremely large systems which have very many unstable density wave vectors. These methods have had a great deal of success in overcoming previously intractable instabilities.

When the system becomes very large, the Hamiltonian matrix can no longer be stored in the computer memory, and must be diagonalized by iterative methods. Pioneering efforts are those of Davidson,[3] Haydock,[4] Bendt, Wood and Zunger,[5,6] and Car and Parrinello.[7] For most early studies, however, direct diagonalization of the Hamiltonian matrix was possible. The success of these early efforts led to attempts to solve larger and larger problems which included full relaxation of the ionic coordinates. In order to allow ionic relaxation, Car and Parrinello[7] developed a molecular-dynamics approach to diagonalization by developing fictitious equations of motion for the plane-wave coefficients and combining them with equations of motion for the ions. Both ions and electrons were relaxed together in this scheme. The discovery that this technique scaled better than direct diagonalization with increasing numbers of plane waves made the approach interesting. Payne *et al.*[8] developed a semianalytic method of integrating the equations of motion which for some systems allowed an order of magnitude larger time step to be taken. This method was used successfully by Allan and Teter to predict a fully relaxed crystal structure of SiO_2.[9] Williams and Soler[10] suggested that since there were no false minima in the electronic problem, it would be more efficient to adapt the Payne method to integrating Schrödinger's equation in imaginary time rather than solving the fictitious second-order equations of Car and Parinello. The Williams and Soler adaptation of the method of Payne *et al.* was considered to be the method of choice for iterative solution.

All of the iterative methods use vectors which are products of the Hamiltonian matrix and trial vectors of plane-wave coefficients. The methods differ in the manner in which they use the product vectors to minimize the total energy. The single most important feature of the iterative approach to diagonalization is that these vectors can be formed extremely efficiently within the Hohenberg-Kohn-Sham local-density approximation (LDA). When a plane-wave basis is used, the Hamiltonian matrix can be split into parts which are diagonal either in momentum space or in real space. Consequently, the entire Hamiltonian matrix need never be stored. The kinetic-energy operator acting on the vector is diagonal in momentum space. Usually the potential is multiplied by the trial vector in real space; the result is Fourier transformed and then combined in momentum space with the contribution from the kinetic-energy operator. Thus the formation of the vector, which is the product of the Hamiltonian matrix and a trial vector, scales approximately linearly in the number of plane waves.

Typically the electron-electron interaction potential is updated after changing all of the single-particle wave functions. The iterative techniques tend to be better behaved than direct diagonalization for the density overcorrection instability because any single iteration contains only a partial reaction to the potential. For very large systems, however, the problem of the electron density overreacting to changes in the potential again becomes the limiting factor in solving the problem. One may choose to model the dielectric response of the material in order to properly damp the density oscillations as in the previously mentioned work of Ho, Ihm, and Joannopoulos[2] which requires several diagonalizations of the Hamiltonian matrix with fixed potentials. Since diagonalization is the main computational cost associated with very large simulations, one would like to avoid having to perform this task several times. In order to avoid multiple diagonalizations, a simple scheme which accounts for the self-consistent response of the system under study while iteratively minimizing the energy of the system is required.

One may also try to defeat the instability by making smaller changes during the iterative procedure, but the maximum stable change to a trial eigenvector goes inversely as the square of the largest dimension of the problem. This scaling will be explained later. The onset of this instability can also be delayed by making smaller density changes by iterating each trial vector separately and updating the Hartree and exchange-correlation potentials after each band is changed. Actual trials of this procedure show a significant increase in the maximum stable time step for a problem with a very large unit cell, but the extra computational work of updating the density and potentials after each band update yields no net gain in efficiency. If band-by-band updating of the potentials is to be successful, more intelligent use must be made of the additional information.

The standard eigenvector equations are derived from the condition that the first derivative of the total energy with respect to each of the plane-wave coefficients is zero. In order to have the proper linear expression for the gradient, the total energy must be correctly expanded to second order in the coefficients. This includes the changes in the total-energy operator due to density changes as well. The usual matrix equations have assumed a linear expansion in the total-energy operator with changes in the coefficients, yielding the single-particle Hamiltonian operator. This is the correct single-particle operator only if the density does not change. The proper second-order expansion possesses an extra set of terms that describes the second-order variation of the total-energy operator. The neglect of these terms causes the overcorrection. Although these terms are well understood, they are universally neglected because of the practical difficulties of including them.

Any iterative technique confronts two questions. In which direction should changes be made to a vector of band coefficients? How much of the change vector should be added to the original vector? There are ideal answers to each question. The ideal direction should be the difference between the trial vector and the exact eigenvector, and the amount should be just sufficient to cancel the error in the original vector.

Unfortunately, one cannot determine the error content of an arbitrary trial vector. The exact solutions of Schrödinger's equation minimize the total energy of the system under the constraints of orthonormalization. One can therefore calculate the gradient of the total energy with respect to the plane-wave coefficients within the trial vector under the condition that the vector remain normalized, and orthogonal to the other trial vectors. When the gradient vector is not zero, it is the only practical measure of error in the original trial vector. It also represents the vector of changes to subtract from the trial vector to which will result in the most rapid initial reduction in the total energy.

The proper amount of an arbitrary vector of changes to subtract from the original trial vector is that amount which minimizes the total energy of the system. Since the exact magnitude of change which will minimize the energy is difficult to determine, most common iterative solution methods use a simpler scheme.

Investigation of the exact behavior of the total-energy variation with changes in the coefficients of a single band shows that a simple assumption of the form of this variation can be made. This paper derives a procedure for parametrizing this form which allows, for the first time, proper inclusion of all of the second-order effects of coefficient changes on the total energy. The approximation gives not only the correct second-order expansion for small changes in the band coefficients but also allows accurate extrapolation of the results for large changes.

Given an arbitrary vector of changes to a vector of band coefficients, determining the proper amount of this vector to subtract in order to minimize the total energy then becomes easy. When one has a good method of determining a nearly optimal change vector, one obtains an efficient self-consistent band-by-band updating procedure.

A preconditioned conjugate gradient procedure is then described which generates effective vectors of changes, and incorporates this form of minimization. This process updates the approximations for each band, and minimizes the total-energy contribution from that band subject to the constraints that the band remains orthogonal to all other bands and normalized. The preconditioning is necessary to ensure that the computational effort to solve a given problem scales favorably with increasing numbers of plane waves.

This method of solution converges more rapidly than solving for all bands simultaneously by direct diagonalization. This efficiency exists for two reasons. The iteration scales better than direct diagonalization with increasing number of plane waves. The density instabilities associated with direct diagonalization do not occur. This approach requires one gradient evaluation and one estimate of the change in the Hartree and exchange-correlation energies for each one-dimensional energy minimization. It properly includes the changing Hamiltonian matrix in the minimization step. There are then several distinct advantages: it is the fastest iterative diagonalization method currently known for large quantum-mechanical problems; every iteration lowers the total energy; there are no electron-density instabilities; it requires no problem-dependent parameters such as times steps or fictitious masses and it is numerically stable.

III. SECOND-ORDER EXPANSION OF THE TOTAL ENERGY

Assume a plane-wave basis set $\{e^{2\pi i \mathbf{G}\cdot\mathbf{r}}\}$ in a box of volume Ω and sides $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$. The $\mathbf{G}$'s are reciprocal-lattice vectors such that $\mathbf{G}\cdot\mathbf{l}_i = n_i$, where n_i is an integer. Initially, only $\mathbf{k}=0$ will be treated. This method is usually used on multiple $\mathbf{k}$ points, and the assumption in the derivation is not a limiting factor.

Given a set of doubly occupied bands $\{\psi_n(r)\}$ which are expressed as a sum of plane waves,

$$\psi_n(r) = \sum_G C_{Gn} e^{2\pi i \mathbf{G}\cdot\mathbf{r}} .$$

One may write the energy per unit cell of the system under the usual LDA assumptions as

$$E = 2\sum_n \sum_G \sum_{G'} C^*_{Gn} C_{G'n} \langle \mathbf{G}|E_{\text{op}}|\mathbf{G}'\rangle - 2\sum_m \sum_n \lambda_{mn} \left[\sum_G C^*_{Gm} C_{Gn} - \delta_{mn} \right] ,$$

where the λ_{mn}'s constrain the bands to be orthonormal.

The total-energy operator E_{op} is

$$E_{\text{op}} = -\tfrac{1}{2}\nabla^2 + V_{\text{ion}}(r) + \tfrac{1}{2}\int \frac{\rho(r')}{|r-r'|} d^3r' + \epsilon_{xc}(\rho(r)) .$$

Forming a second-order Taylor expansion of the total energy around a set of band coefficients,

$$E = E_0 + \sum_n \sum_G \frac{\partial E}{\partial C^*_{Gn}} \delta C^*_{Gn} + \tfrac{1}{2} \sum_n \sum_m \sum_G \sum_{G'} \delta C^*_{Gn} \frac{\partial^2 E}{\partial C^*_{Gn}\, \partial C_{G'm}} \delta C_{G'm} .$$

The Gnth component of the gradient of E is

$$\nabla E_{Gn} = \frac{\partial E}{\partial C^*_{Gn}} + \sum_m \sum_{G'} \frac{\partial^2 E}{\partial C^*_{Gn}\, \partial C_{G'm}} \delta C_{G'm} .$$

The derivatives appearing in the gradient expression include not only the explicit dependence of the energy on the C's, but also the dependence of the total energy on the density which depends implicitly on the C's. It is the effect of the derivatives on the density dependence of the Hartree and exchange-correlation energies which determines the difference between the single-particle Hamiltonian operator H_{op} and the total-energy operator E_{op}.

$$H_{\text{op}} = -\tfrac{1}{2}\nabla^2 + V_{\text{ion}}(r) + \mu_{xc}(r) + \int \frac{\rho(r')}{|r-r'|} d^3r' ,$$

where

$$\mu_{xc}(r) = \epsilon_{xc}(r) + \rho(r) \frac{\partial \epsilon_{xc}(r)}{\rho(r)} .$$

The Hamiltonian matrix H may be formed by taking plane-wave matrix elements using the Hamiltonian operator.

In the Taylor expansion, there is a corresponding change in the Hamiltonian due to the second derivatives of the Hartree and exchange-correlation energies with respect to density.

The additional terms reflect changes in the Hartree and exchange-correlation potentials with variation in the C's. If one tries to take this variation fully into account the problem becomes much larger than the original eigenvalue-eigenvector problem since all the bands become coupled through the changes in the electron-electron potential. On the other hand, ignoring this variation in the potential leads to slow convergence due to the charge oscillation which was described earlier.

The additional terms are dominated by changes in the Hartree potential at small wave numbers. The Fourier transform of the Hartree potential at a given wave number is proportional to the Fourier transform of the electron density at that wave number divided by the square of the magnitude of the wave number. Thus changes in the electron density can yield extremely large changes in the Hartree potential as the size of the system increases, and the size of the minimum wave number correspondingly decreases. Ignoring the changes in the charge density produced by changes in the wave function produces larger and larger errors in the Hartree potential as the minimum wave number becomes smaller. As a trial wave function interacts with its corresponding potential, it loses magnitude in regions of high potential and increases in regions of low potential. Ignoring the changes in potential as the wave function varies causes the solution process to overpopulate the low-potential regions and underpopulate the high ones. Upon recalculation of the potential, the high- and low-potential regions have interchanged, and repetition of this process yields the corresponding low–wave-number oscillation of charge in the unit cell, commonly referred to as charge "sloshing."

IV. BAND-BY-BAND MINIMIZATION OF THE TOTAL ENERGY

Assume that one wishes to minimize the total energy of the system by changing the mth band subject to the constraint that it remains orthogonal to all other bands and also remains normalized. After n iterations, one has a current estimate for the mth eigenvector, $\mathbf{C}_m^n$, and a normalized vector of changes $\mathbf{D}^n$ which is orthogonal to all of the bands including $\mathbf{C}_m^n$. The problem separates into two parts: how to determine the optimum $\mathbf{D}^n$ and how much of $\mathbf{D}^n$ to add to $\mathbf{C}_m^n$ to minimize the energy. Because $\mathbf{D}^n$ is orthogonal to all of the bands, the orthogonality constraint is satisfied with any linear combination of the two vectors. To maintain normality, a simple convention is to add the vectors according to the scheme

$$\mathbf{C}_m^{n+1}=\mathbf{C}_m^n\cos\theta+\mathbf{D}^n\sin\theta \ .$$

For small θ, the total energy may be expanded in a simple quadratic function of θ, and, when $\mathbf{C}_m^n$ is close to a minimum, this form is sufficient. When $\mathbf{C}_m^n$ is far from the converged solution, however, large values of θ may be required. A better form for the total energy as a function of θ is necessary. Since the total energy depends upon the square of the wave function, the total energy is periodic with period π. Thus an exact parameterization for the energy as a function of θ is

$$E(\theta)=E_{\text{avg}}+\sum_{j=1}^{\infty} E_{cj}\cos(2j\theta)+E_{sj}\sin(2j\theta) \ .$$

Since no false minima exist in the total energy within the unpolarized version of the LDA, the high-frequency contributions must be small. Numerical experimentation shows that the total contribution to the sum from j greater than 1 is typically more than an order of magnitude smaller than the $j=1$ contribution. Figure 1 shows a plot of the total energy of an 8-atom silicon cell as a function of θ when one trial vector is added. The dots are the exact total energy and the line is the the $j=1$ contribution. Consequently a very good approximation of the functional form of E over the entire range of θ is

$$E(\theta)=E_{\text{avg}}+E_{c1}\cos(2\theta)+E_{s1}\sin(2\theta) \ .$$

Three pieces of information are then needed to determine an approximate minimum of $E(\theta)$. It is important to understand that the only part of the total energy which must be evaluated is that part which depends on θ. This includes the Hartree energy, the exchange-correlation energy, the kinetic energy of band m, and the electron-ion energy of band m. After grouping the last two into a partial Hamiltonian matrix H_1 which is independent of density, the part of the total energy which depends upon θ is

$$E(\theta)=f_m[(\mathbf{C}_m^{n*}\cos\theta+\mathbf{D}^{n*}\sin\theta)\cdot H_1\cdot(\mathbf{C}_m^n\cos\theta+\mathbf{D}^n\sin\theta)] + \tfrac{1}{2}\int\int\frac{\rho(r)\rho(r')}{|r-r'|}d^3r\,d^3r' + \int\rho(r)\epsilon_{xc}(r)d^3r \ ,$$

where f_m is the occupancy of band m, usually 2, and would include a **k**-point weight if there were more than one **k** point.

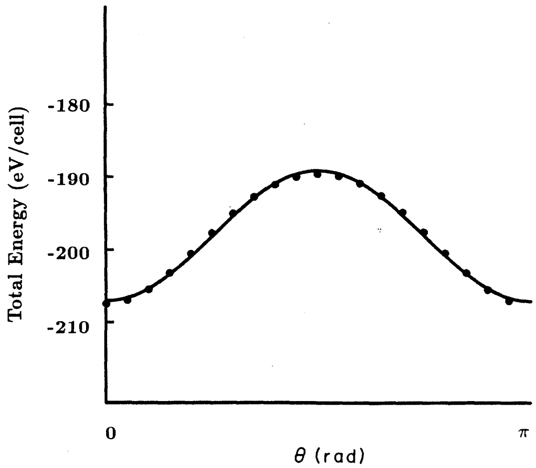

FIG. 1. The total energy of an 8-atom silicon cube in eV/(unit cell) is plotted as a function of θ. The dots represent the exact calculation and the line is the lowest harmonic contribution.

This can be evaluated easily at $\theta=0$ to give one of the three pieces of information needed. If $\rho(r)$ is expressed in terms of θ and the above result is differentiated and evaluated at $\theta=0$, one obtains

$$\left.\frac{\partial E(\theta)}{\partial\theta}\right|_{\theta=0^\circ}=f_m(\mathbf{C}^{n*}\cdot H\cdot\mathbf{D}^n+\mathbf{D}^{n*}\cdot H\cdot\mathbf{C}^n)$$

$$=2f_m\mathrm{Re}(\mathbf{D}^{n*}\cdot H\cdot\mathbf{C}^n)\ ,$$

where H is the single-particle Hamiltonian matrix whose operator was defined previously. This is the second required piece of information. Since $H\cdot\mathbf{C}^n$ must be generated to obtain $\mathbf{D}^n$, only a simple dot product is required to generate the first derivative of the total energy with respect to θ. If the second derivative of the energy with respect to θ were evaluated at $\theta=0$, one would have the required third piece of information. A count of the operations required to do this calculation implies that a better procedure is to evaluate $E(\theta)$ self-consistently at some small θ. A good compromise is $\pi/300$. This value of θ is not so small as to yield significant rounding error, but close enough to zero that the curvature estimate at $\theta=0$ is accurate. This estimate of accurate curvature around $\theta=0$ is especially important when the trial vector is nearly converged.

E at $\pi/300$ can be computed by updating the trial vector for band m, calculating the contribution from the density-independent Hamiltonian and the changes in the Hartree and exchange-correlation energies. The density-independent Hamiltonian is a linear operator. If one begins the nth iteration for band m with $H_1\cdot\mathbf{C}_m^n$, the result of applying H_1 to $\mathbf{D}^n$ will allow the linear combination of the two vectors to be calculated both at $\pi/300$ and the final value of θ which minimizes the energy. $H_1\cdot\mathbf{C}_m^{n+1}$ will then be available at the end of the nth iteration. Therefore the density-independent Hamiltonian matrix must be applied once at the beginning of the band iteration and once more for each update. This is a significant savings since the electron-ion interaction can be computationally intensive if nonlocal pseudopotentials are used. The only parts of the energy which must be calculated twice are the Hartree and exchange-correlation energies, and to calculate these, only one band density must be updated. The only additional work over an algorithm which updates the potential after each band update is to calculate the change in the Hartree and exchange-correlation energies at $\theta=\pi/300$, which is an easy computation. The gain over other band-by-band algorithms is the minimization of the total energy within the step, not merely total energy reduction. The increase in computational costs to calculate the changes in this fashion comes largely from the increase in the number of Fourier transforms. The number of Fourier transforms per band goes from approximately three in a method which only updates the potential after all the bands within a **k** point are updated to six per band when the potential is updated after every band change, and the changes are calculated self-consistently. Since Fourier transforms only make up a fraction of the total computational cost, the actual increase per band update is approximately 40%.

This procedure gives reasonable performance when the vectors chosen are the gradients of the total energy, resulting in a simple steepest-descents algorithm. When the vectors are chosen to be conjugate to one another, the rate of convergence increases substantially. The performance degrades, however, as the number of plane waves is increased. Basically, the highest-wave-number plane waves contribute a disproportionate share to the gradient because of their high kinetic-energy content. Payne *et al.*[8] solved this problem by allowing a semianalytic integration of the gradient which caused the high-wave-number components to decay most rapidly and limited their effect. By preconditioning the gradient one achieves the same effect by limiting the contribution of the high-wave-number parts of the gradient.

In the conjugate-gradient method, one minimizes a function by first computing the gradient of the function. The gradient is then used to establish a vector of parameter changes which is added to the vector of parameters in such an amount as to minimize the function. This process is repeated, but each new vector is forced to be conjugate to previous change vectors.[11] In practice, for nearly quadratic functions, the relative efficiency of the method depends upon two factors. The first is the bandwidth of the problem. If the eigenvalue spectrum is too wide, the method will converge very slowly. This may be easily demonstrated. Assume a nearly converged trial vector for band m, $\mathbf{C}_m$, which is made up of a sum of eigenvectors of H:

$$\mathbf{C}_m=\mathbf{E}_m+\sum_{i\ (\neq m)}\epsilon_i\mathbf{E}_i\ .$$

The sum represents the error in $\mathbf{C}_m$. Assuming that the squares of the ϵ's are negligible, the measure of error which one actually can compute is the residual vector

$$H\cdot\mathbf{C}_m-\lambda_m\mathbf{C}_m=\sum_{i\ (\neq m)}\epsilon_i(\lambda_i-\lambda_m)\mathbf{E}_i\ ,$$

where

$$\lambda_m=\mathbf{C}_m\cdot H\cdot\mathbf{C}_m\ .$$

If the λ_i's span a broad range, there is no amount of this residual vector which may be added to the trial eigenvector to eliminate the error in $\mathbf{C}_m$. The exact error vector may be recovered by multiplying the residual vector by the inverse of the Hamiltonian matrix shifted by λ_m. This is not feasible, but if we can multiply the residual by an approximate inverse, the problem becomes much better conditioned. This is the concept of preconditioning. The approximate inverse need not be very good, but it is important that it be a positive definite operator so the projection of the conditioned gradient on the actual gradient is never zero. With increasing wave number, the Hamiltonian matrix becomes diagonally dominant, with the kinetic energy overwhelming any potential terms. Each high-wave-number plane wave is nearly an eigenfunction of the Hamiltonian matrix. If we denote the kinetic energy of a plane wave by

$$E_{\mathrm{kin}}(\mathbf{G})=2\pi^2\mathbf{G}^2\ .$$

and the expectation value of this for the mth band,

$$E^{m}_{\mathrm{kin}}=\sum_{G}\mathbf{C}^{*}_{Gm}\cdot\mathbf{C}_{Gm}E_{\mathrm{kin}}(\mathbf{G})\ .$$

At **G**'s high enough so that the kinetic energy dominates the contributions from the potential, and assuming that the contributions of the potential to the eigenvalues roughly cancel, the *G*th component of the residual vector described earlier becomes approximately

$$\sum_{i\ (\neq m)}\epsilon_i[E_{\mathrm{kin}}(\mathbf{G})-E^{m}_{\mathrm{kin}}]\mathbf{E}_{iG}\ .$$

A typical example of the requirement of high-wave-number planewaves occurs in the simulation of $\mathrm{SiO_2}$ using nonlocal pseudopotentials.[8] Kinetic-energy cutoffs as high as 25 hartrees are used to completely converge the system even though the occupied eigenvalues are only a few hartrees in magnitude.

One can then precondition the residual vector by multiplying each component by a function of $|\mathbf{G}|$ which becomes proportional to $1/[E_{\mathrm{kin}}(\mathbf{G})-E^{m}_{\mathrm{kin}}]$ as the magnitude of **G** becomes large. The reason for the elimination of the potential from this operator is to ensure a positive definite operator. This multiplication roughly removes the effect of the original Hamiltonian operation on the residual leaving nearly the original error term for high wave numbers. Thus most of the errors in the high-wave-number coefficients are eliminated together, and the method scales extremely well with increasing number of plane waves. At low wave numbers, the Hamiltonian matrix is not diagonally dominant, and there is no simple way to multiply the residual vector to yield the error vector. Consequently, one simply lets the function remain constant at energies below E^{m}_{kin}.

The second factor which determines the relative efficiency of the conjugate-gradient method is the computational effort required to perform minimizations with arbitrary vectors of changes to the band coefficients. As has been shown, the total energy is a somewhat more complex function of θ than a simple quadratic.

Previous attempts by the authors to use the conjugate-gradient method to minimize the energy succeeded in generating eigenvectors for simple problems in a few iterations, but proved impractical because the method scaled badly with increasing numbers of plane waves and also the exact one-dimensional minimizations were too costly. For large steps, the nonquadratic nature of the energy response required several gradient evaluations to achieve a minimum energy. These gradient evaluations could be used more efficiently in other algorithms which scaled better with increasing numbers of plane waves.

Štich *et al.*[12] describe a conjugate-gradient method in which the minimizations are less expansive because they do not update the Hamiltonian matrix when they minimize. This method suffers from the previously mentioned problem of overcorrection because the curvature is underestimated. The method also has no preconditioning, so when the maximum allowable kinetic energy is increased, its convergence rate will suffer.

A short description of the preconditioned conjugate-gradient algorithm applied to constrained systems is now provided. There are two vectors, one related to the current gradient after preconditioning which shall be called $\mathbf{G}^{n}$ and another vector which is the previous correction $\mathbf{F}^{n-1}$. Initially $\mathbf{F}^{0}$ is zero. One forms $\mathbf{G}^{n}$ by the following process. Take the product of the single-particle Hamiltonian matrix and the current estimate of the band coefficients. One then projects the current band out of this vector. If this vector were orthogonal to the remaining bands, it would be the gradient of the total energy with respect to the present band coefficients, fulfilling the constraints of orthonormality. These orthogonalizations are accomplished implicitly through a subspace diagonalization step which will be described later. If the resulting vector is zero, the band coefficients formed an eigenvector and one is finished with that band. The residual vector is then multiplied by a diagonal conditioning matrix which reduces the high-wave-number components while leaving the low wave numbers untouched. This constitutes the preconditioning. The conditioning matrix *K* is

$$K_{GG'}=\delta_{GG'}\frac{27+18x+12x^{2}+8x^{3}}{27+18x+12x^{2}+8x^{3}+16x^{4}}\ ,$$

where

$$x=\frac{E_{\mathrm{kin}}(\mathbf{G})}{E^{m}_{\mathrm{kin}}}\ .$$

K approaches $x=0$ with a value of 1 and has zero first, second, and third derivatives. This guarantees that the low-wave-number components of the gradient are left unchanged. Above $x=1$, *K* approaches $1/[2(x-1)]$ with an asymptotic expansion correct to fourth order in $1/x$. The factor of $\frac{1}{2}$ in *K* is used to expedite the joining of the low-*x* and high-*x* expansions. This freedom is allowed since the high-wave-number form of *K* is only defined to within a multiplicative constant. At small wave numbers, the algorithm is a simple conjugate gradient algorithm. At high $|\mathbf{G}|$, however, the residual vector has been transformed to resemble the original error vector. The effect of the preconditioning, therefore, is to cause all of the high-wave-number components to converge at the nearly the same rate. If the preconditioning is not done, then most of the time is spent making small changes to the high-wave-number components of the trial eigenvector.

The trial eigenvectors of all of the bands, including the current band, are then projected out of this filtered vector, the result is negated and becomes $\mathbf{G}^{n}$. The *n*th conditioned residual vector is

$$\mathbf{R}^{n}=-(H\cdot\mathbf{C}^{n}_{m}-\lambda^{n}_{m}C^{n}_{m})\cdot K\ ,$$

where

$$\lambda^{n}_{m}=\mathbf{C}^{*n}_{m}\cdot H\cdot C^{n}_{m}\ .$$

The vector passed to the one-dimensional minimization process must be orthogonal to all of the bands so

$$\mathbf{G}^{n}=\mathbf{R}^{n}-\sum_{k\ (\neq m)}\alpha^{n}_{mk}\mathbf{C}_{k}-\alpha_{mm}\mathbf{C}^{n}_{m}\ ,$$

where

$$\alpha^{n}_{mk}\mathbf{C}^{*}_{k}\cdot\mathbf{R}^{n}\ .$$

Typically double precision arithmetic is used to guarantee that after all of this manipulation, the resulting vector contains more information than rounding error. Then the vector corresponding to the new direction along which to minimize the total energy is

$$\mathbf{F}^n=\mathbf{G}^n+\gamma^n\mathbf{F}^{n-1}\ ,$$

where

$$\gamma^n=\frac{\mathbf{G}^{n*}\cdot\mathbf{G}^n}{\mathbf{G}^{n-1*}\cdot\mathbf{G}^{n-1}}$$

and

$$\gamma^1=0\ .$$

Textbooks on the preconditioned conjugate-gradient method contain a factor of the inverse of K in the definition of γ to ensure conjugacy. We have not found this to be necessary. A plausibility argument is that K and its inverse are 1 for low wave numbers, and the high-wave-number errors are eliminated on the first few passes. The minimization is supposed to be performed using the vector $\mathbf{F}^n$ which is no longer orthogonal to $\mathbf{C}_m^n$ because of the addition of some of the previous vector. The minimization procedure described earlier requires a normalized vector which is orthogonal to our current band estimate. A component parallel to the current band vector cannot be added to the band vector with any lasting effect since such an addition changes only the magnitude of the vector and the vector must remain normalized. The minimization is therefore performed using another vector $\mathbf{D}^n$ which is formed by projecting the band vector $\mathbf{C}_m^n$ from $\mathbf{F}^n$ and normalizing. One achieves the same result as if the minimization had used $\mathbf{F}^n$. The vector $\mathbf{D}^n$ is used for the minimization. $\mathbf{F}^n$ is saved and is used along with $\mathbf{G}^{n+1}$ to generate $\mathbf{F}^{n+1}$.

After a pass through the bands has been completed, in which each band has been iterated four or five times before proceeding to the next band, one then forms the subspace Hamiltonian matrix using the orthogonal filled bands as basis functions. The elements of this matrix are simply the Lagrange multipliers for orthonormality. The eigenvectors of this small matrix will give the linear combinations of the trial vectors which are eigenvectors in their subspace. This similarity transformation changes neither the energy nor the density of the system. This process not only solves any problems associated with band ordering, but also guarantees that any subsequent change to one band which lowers its energy will be orthogonal to all of the other bands. This is the implicit orthogonalization of the gradient described earlier. If the subspace diagonalization step is not performed, the gradients must be orthogonalized to all of the bands before preconditioning as well as after. After convergence without subspace diagonalization, the resulting vectors will in general constitute a similarity transformation of the eigenvectors.

To use the algorithm with multiple $\mathbf{k}$ points, simply treat each as an independent minimization problem with the proviso that each is processed by one pass through the above algorithm before any are repeated. In the preconditioning step replace $\mathbf{G}^2$ with $(\mathbf{k}+\mathbf{G})^2$. There is no degradation in the efficiency in that the number of passes to convergence is reasonably independent of the number of $\mathbf{k}$ points. The total computational effort, however, goes directly with the number of $\mathbf{k}$ points. The advantage of using many $\mathbf{k}$ points is that an infinite material may be modeled with a small periodic cell. The advantage of this technique is that a large cell may be modeled requiring many fewer $\mathbf{k}$ points, but allowing significant variations from the symmetry of the perfect material such as frozen phonons or point defects. The SiO_2 example in the following section on results was done with four special points and a 12-atom supercell, but the method is designed to be useable for periodic cells containing over 100 atoms.

V. RESULTS

Two simulations of silicon were performed which are known to have problems at either low or high wave numbers. Four methods were used in the simulations. The original method was a time intergration of Schrödinger's equation in imaginary time using the semianalytic approach of Payne *et al.* applied to the first-order scheme of Williams and Soler. In this method, the Hamiltonian is updated only after all of the bands are updated. The other three methods, steepest descents, conjugate gradient, and preconditioned conjugate gradient, all used the one-dimensional energy minimization technique described earlier with three different ways of chosing the vector of changes to be added to the trial eigenvector. They also update the Hamiltonian after each band update. The steepest-descents method used the gradient vector orthogonalized to all the bands. The conjugate-gradient method used the orthogonalized gradient vector made conjugate to earlier change vectors. The preconditioned-conjugate-gradient method added the preconditioning described earlier to the gradient vector before applying the conjugate-gradient method. Initial trial vectors were filled with random numbers and then orthonormalized. This extreme starting point is a deliberate test of the robustness of the algorithms since it ensures that every eigenvector of the Hamiltonian matrix will contribute to the error. Finally, a simulation was performed of SiO_2 using only the original and preconditioned-conjugate-gradient methods.

The first simulation was an 8-atom cube of silicon using plane wave up to 16 hartrees in kinetic energy. The normal maximum kinetic energy used to model silicon lies in the 4–6-hartree region. Many other materials require much higher kinetic-energy plane waves (SiO_2 requires plane waves of approximately 15–25 hartrees) and so it is of interest to observe the robustness of a particular solution method with respect to increasing number of plane waves. Figure 2 shows the result of applying the original method of Williams and Soler and three others. As may be seen from Fig. 2, each subsequent method was a substantial improvement on its predecessor. The scales are *different* for the Williams-Soler method and those using the self-consistent energy minimization. The number of iterations plotted is 20 for each of the methods using

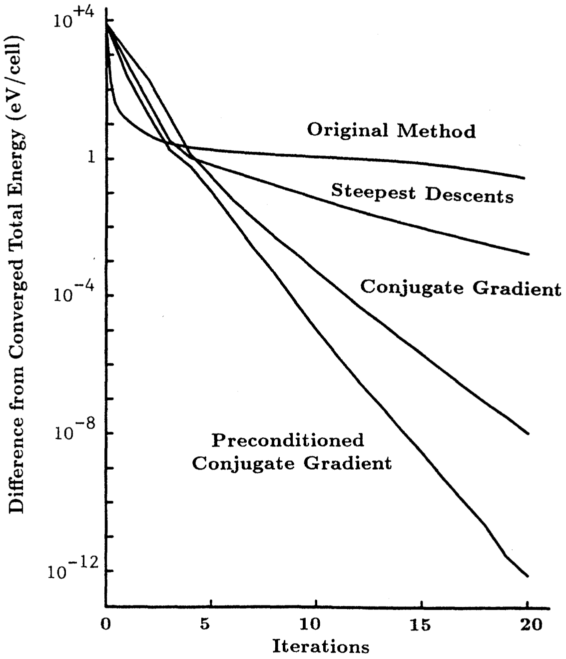

FIG. 2. The total energy error in eV/(unit cell) of an 8-atom silicon cube with a 16-hartree kinetic-energy cutoff vs iteration number for various methods. The scales are different for the different methods. The number of iterations of the original Williams-Soler method has been divided by 5 to allow comparison at the same level of computational effort.

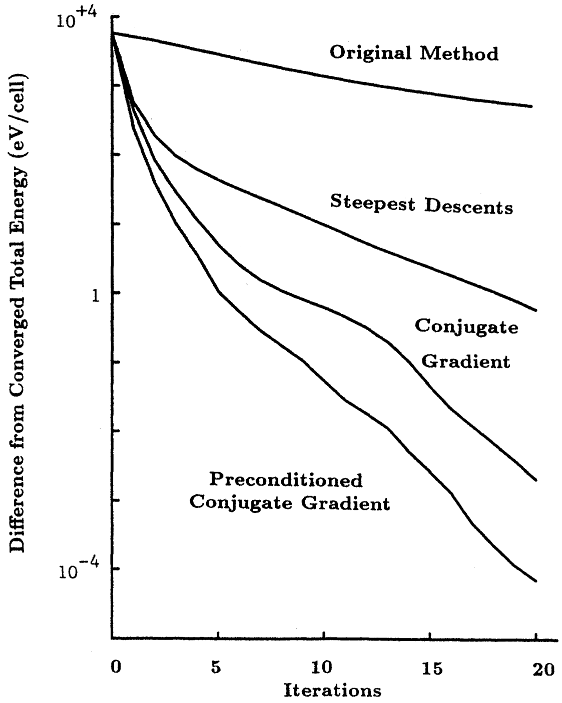

FIG. 3. The total-energy error in eV of a row of 12 silicon unit cells with a 4-hartree kinetic-energy cutoff vs iteration number for various methods. The methods are scaled as in Fig. 2.

the one-dimensional minimization technique, but 100 iterations for the original method. The original method applied the gradient operator once per time step. The minimization methods would typically update each band four times per iteration. The extra calculation to ensure self-consistency takes some additional time, so the choice of scales is to make the computation time roughly equal.

The next simulation involves a row of 12 silicon unit cells with 24 atoms and an energy cutoff of 4 hartrees. This was originally modeled to investigate the self-consistent response of silicon to low-wave-number potential perturbations. The convergence was extremely slow using the original method. To avoid the charge oscillations the time step had to be cut by a factor of 150 over that used in the 2-atom unit cell. Thousands of iterations were necessary for convergence. As may be seen in Fig. 3, all of the self-consistent energy minimization methods constitute an enormous improvement over the original technique and again the preconditioned-conjugate-gradient method is significantly more efficient than the others.

The last simulation was of SiO_2 and consisted of a 12-atom unit cell of α-cristobalite. The limit on the kinetic energy of the plane waves was 16 hartrees. Figure 4 shows the result of the calculation. Again 100 iterations of the original method were to be plotted versus 20 of the preconditioned conjugate gradient to roughly normalize the computational effort. After only 11 iterations, the preconditioned-conjugate-gradient method had reached the limit of rounding error of the computation of total

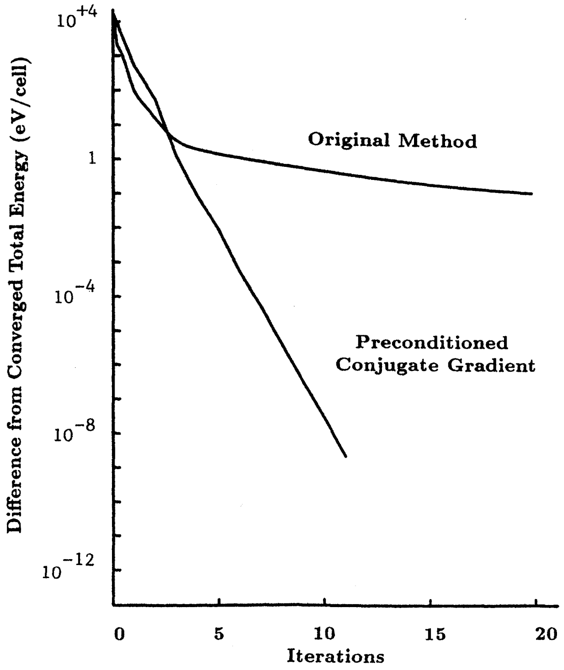

FIG. 4. The total-energy error in eV of a 12-atom cell of the α-cristobalite form of SiO_2 with a 16-hartree kinetic-energy cutoff vs iteration number for various methods. The methods are scaled as in Fig. 2.

energy. Since analytic expressions for the gradients were being used, the errors in the wave function continued to be lowered even though the total energy could not be computed more accurately.

VI. SUMMARY

A method has been outlined which gives an efficient implementation of the conjugate gradient method applied to the iterative computation of eigenvectors of the self-consistent Hamiltonian matrix. A simple and efficient procedure has been outlined which is correct to second order in total energy for changes in the band coefficients and hence completely self-consistent. This technique is accurate for large changes in the band coefficients and also is free from the charge "sloshing" instabilities normally associated with the solution of very large problems. The gradient is preconditioned to make the convergence rate reasonably independent of the number of plane waves in the problem. The relative efficiency of this method has been compared to other methods and increases in rates of convergence greater than an order of magnitude have been observed in large problems.

ACKNOWLEDGMENTS

The authors jointly thank the Cornell Theory Center, the Cornell Electronic Structure Group of the Laboratory of Atomic and Solid State Physics, Corning Glass Works, and the National Science Foundation for resources supplied during the development of this work. One of us (M.C.P.) thanks the Royal Society (United Kingdom) for financial support.

---

[1] D. H. Vanderbilt and S. G. Louie, Phys. Rev. B **30**, 6118 (1984).

[2] Kai-Ming Ho, J. Ihm, and J. D. Joannopoulos, Phys. Rev. B **25**, 4260 (1982).

[3] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).

[4] R. Haydock, in *Solid State Physics,* edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1980), Vol. 35, p. 215.

[5] P. Bendt and A. Zunger, Solar Energy Research Institute Technical Report TP-212-1698, 1982.

[6] D. M. Wood and A. Zunger, J. Phys. A **18**, 1343 (1985).

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

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

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

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

[11] M. R. Hestenes and E. Stiefel, J. Res. Nat. Bur. Stand. **49**, 409 (1952).

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