All-Electron Mixed-Basis Calculation with Conjugated Gradient Method to Optimize Structure of Copper Clusters

Materials Transactions, JIM, Vol. 40, No. 11 (1999), pp. 1205 to 1208
Special Issue on Advances in Computational Materials Science and Engineering

All-Electron Mixed-Basis Calculation with Conjugated Gradient Method to Optimize Structure of Copper Clusters

Y. C. Bae†, Kaoru Ohno and Yoshiyuki Kawazoe

Institute for Materials Research (IMR), Tohoku University, Sendai 980-8577, Japan

In this paper, we implement the conjugate gradient method in the all-electron mixed-basis approach. The convergence of the electronic states is greatly improved by using the conjugate gradient method compared to the steepest descent method. We show a result for the case of Cu_4.

(Received May 24, 1999; In Final Form September 17, 1999)

Keywords: *ab initio, all-electron mixed-basis calculation, conjugate gradient method, electronic-structure, electronic states*

I. Introduction

Ab initio molecular dynamics[1][2] based on the local density approximation (LDA)[3] and the adiabatic approximation[4] has attracted considerable attention as a conceptually new method, which is capable to describe dynamically the stability and reactivity of any objects including clusters, surfaces and bulk materials at finite temperatures, in principle, without using any parameters. In electronic-structure calculations of transition-metal systems, where 3d valence orbitals are relatively localized, the standard pseudopotential approach combined with plane wave (PW) expansion[5] becomes heavier than alternative approaches such as muffin-tin (including APW, KKR, LMTO, and similar methods) and LCAO (linear combination of atomic orbitals) methods[6].

The application of muffin-tin method to the molecular dynamics of clusters and surfaces is, however, not easy because the way of partitioning the system into Wigner-Seitz cells changes time to time[7], although several formulations and calculations can be found rather recently[8]-[11]. Ohno *et al.* have developed the all-electron mixed-basis approach which is applicable to the molecular dynamics of objects in any atomic environments[12].

The mixed-basis means the combination of both plane waves (PWs) and atomic orbitals (AOs) is used as basis functions. The introduction of AOs reduces considerably the computational load of the PW-expansion method[13]. In particular, in the all-electron mixed-basis method, not only valence AOs but also core AOs are incorporated to describe all-electrons without using pseudopotentials. Historically, all-electron mixed basis method was first formulated by Bendt and Zunger[14], who incorporated just core AOs in the PW-expansion method. However, the method[6][12] used in the present study is independent and different from Bendt and Zunger's, because ours incorporates also valence AOs in the basis set. Although the number of basis functions n_{basis} needed for describing the electronic states with the same accuracy reduces significantly by using the mixed-basis, the computational load still increases as n_{basis}^3, and there is a limitation in the application of the all-electron mixed-basis approach both in the system size and in the simulation time period, as is similar to other *ab initio* methods. When n_{basis} becomes order of 10^4, the usual matrix diagonalization by means of, for example, Householder's method, becomes quite time consuming and the use of alternative algorithms is inevitable.

One of the efficient algorithms to solve the self-consistent Kohn-Sham equation is the conjugate gradient (CG) method[2][15]. The present paper deals, for the first time, with the incorporation of the CG method in the all-electron mixed-basis approach. As an example we will present a result on the structural optimization of Cu_4, which is the simplest nontrivial transition-metal cluster and can be treated with a spin-unpolarized calculation.

The paper is organized as follows: In Section 2, the CG method that is used in the present study is explained in detail. The comparison with the steepest descent (SD) method and the numerical result for Cu_4 is presented in Section 3. Lastly, Section 4 is devoted to the summary.

II. Level-to-Level CG Method

1. SD method

Before going into the explanation of the CG method, let us consider first the easiest SD method. The SD method is to move the electronic states toward the SD direction in the energy surface. The time evolution equ-

† Graduate Course Student, Tohoku University.

ation for the SD method reads[16]

$$\mu \frac{d\psi_i}{dt}=\zeta_i; \tag{1}$$

with

$$\zeta_i=-(H-\lambda_i)\psi_i, \tag{2}$$

and

$$\lambda_i=\langle\psi_i|H|\psi_i\rangle. \tag{3}$$

Here ψ is the wavefunction of the ith level, H the Hamiltonian, ζ_i the SD direction, and λ_i the energy expectation value. The time derivative of the above equation is replaced by a finite difference,

$$\frac{d\psi_i}{dt}=\frac{\psi_i(t+\Delta t)-\psi_i(t)}{\Delta t}. \tag{4}$$

2. CG method

The CG method is a related but more elaborate method in which electronic states are converged more quickly towards the eigenstates. In this method, the eigenstates are searched in the direction that is conjugate to all previous directions. If the energy surface is purely parabolic, it is guaranteed that the minimum is reached after only two iterations; see **Fig. 1**. We adopt here the level-to-level optimization first proposed by Payne *et al.*[2] Single level is updated twice by this method combined with the line minimization described below subsection 3. (For the line minimization, Payne *et al.* have proposed a way to calculate the second derivative of the total energy with respect to the change in the conjugate direction of the wavefunctions. However, this requires recomputation of the total energy at each change of one electronic level and becomes heavy if the number of electronic levels increases. In place of using their technique, here we will adopt an alternative method proposed by King-Smith and Vanderbilt[17].) When one special level is optimized, wavefunctions of all the other levels are held fixed. After this optimization for one special level finishes, similar optimizations are performed one-by-one for all levels. Since these optimizations are independent for every level, the resulting wavefunctions are not mutually consistent. In order to improve this consistency, the same procedure is repeated five times with a fixed Hamiltonian. Finally, in order to achieve the full self-consistency of the Kohn-Sham equation, the whole process is repeated with updating the charge density, potential and Hamiltonian. Hereafter, we will call this whole process updating once the charge density, etc. "one iteration". In our later application to the structural optimization of a Cu_4 cluster, we will perform enough iterations, until the difference between the total energies at subsequent steps becomes smaller than 1.0×10^{-6} Hartree. Now we explain the single-level optimization[2]. Let us assume that the ith level is optimized and all the other levels are fixed. First, in order to make the SD direction ζ_i orthogonal to all the states ψ_i, we calculate

$$\zeta_i'^m=\zeta_i^m-\sum_{j\neq i}\langle\psi_j|\zeta_i^m\rangle\psi_j-\langle\psi_i^m|\zeta_i^m\rangle\psi_i^m, \tag{5}$$

where the superscript m denotes the iteration index; $m=1, 2$.

Then we introduce the function φ_i^m ($m=1, 2$) defined by

$$\varphi_i^m=\zeta_i'^m+\gamma_i^m\varphi_i^{m-1} \tag{6}$$

with $\gamma_i^1=0$ and

$$\gamma_i^2=\frac{\langle\zeta_i'^2|\zeta_i'^2\rangle}{\langle\zeta_i'^1|\zeta_i'^1\rangle}. \tag{7}$$

Finally, the CG direction $\varphi_i'^m$ is obtained by orthonormalizing $\varphi_i'^m$ as follows:

$$\varphi_i''^m=\varphi_i^m-\langle\psi_i^m|\varphi_i^m\rangle\psi_i^m \tag{8}$$

$$\varphi_i'^m=\frac{\varphi_i''^m}{\langle\varphi_i''^m|\varphi_i''^m\rangle}. \tag{9}$$

3. Line minimization

Once the CG direction is determined in the way explained in the previous section, it is necessary to determine the new wavefunction as

$$\psi_i^{m+1}=\alpha\psi_i^m+\beta\varphi_i'^m, \tag{10}$$

where α and β should be determined so as to minimize the total energy. This procedure is called the line minimization. This is not straightforward when the Hamiltonian depends explicitly on the wavefunctions and the minimum of the total energy has to be determined self-consistently. However, if the Hamiltonian is fixed and only the ith eigenvalue λ_i has to be minimized, the line minimization can be rewritten in a 2×2 matrix eigenvalue problem[17]:

$$\begin{pmatrix}\langle\varphi_i|H|\varphi_i\rangle & \langle\varphi_i|H|\psi_i\rangle\\ \langle\psi_i|H|\varphi_i\rangle & \langle\psi_i|H|\psi_i\rangle\end{pmatrix}\begin{pmatrix}a_\pm\\ b_\pm\end{pmatrix}=\lambda_\pm\begin{pmatrix}a_\pm\\ b_\pm\end{pmatrix} \tag{11}$$

and α and β are obtained as the elements a_- and b_- of the eigenvector corresponding to the smaller eigenvalue λ_-.

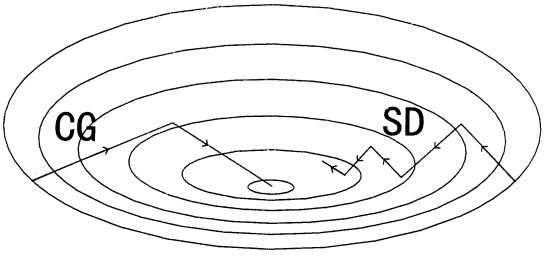

Fig. 1 Schematic illustration of steepest descent (SD) and conjugate gradient (CG) methods of convergence to the center of anisotropic harmonic potential. SD method requires many steps to converge. CG method allows complete convergence only in two steps.

III. Results

First, we compare the results of CG and SD methods.

At once, we calculated eigenvalue and eigenvector self-consistency. Using the same atomic positions (held fixed) and the same Hamiltonian (also held fixed), we applied separately CG and SD methods to obtain the ground state for the electronic states of Cu_4. The cutoff energy for PW's is chosen as 7.6 Ry. Moreover, 1s, 2s, 2p, 3s, 3p AO's are used directly from the atomic code, and 3d AO is also used but modified from that of an isolated atom, *i.e.*, it is truncated by subtracting a smooth quadratic function which has the same value and the same slope at the cutoff radius r_c, which is the radius of the nonoverlapping atomic sphere defined at the outset. The result of the calculation is shown in **Fig. 2**. Here, the abscissa is the iteration number n and the ordinate is the difference of the total energy, ΔE. The absolute minimum of the total energy, which is used as a reference energy, is determined by extrapolating the total energy as a function of $1/n$ in the limit $1/n \to 0$. From Fig. 2, the result of the CG method (solid curve) converges much faster than the result of the SD method (broken curve). In this Figure, the result of the CG method seems to converge already after 10 iterations ($n=10$). However, if we draw the behavior of ΔE in a log-log scale as in **Fig. 3**, the electronic states still move steadily toward the energy minimum. From this result, we can conclude that the CG method works well also in the all-electron mixed-basis program.

Next, we carried out the structural optimization of Cu_4 by using the CG method for the electronic state convergence. Also the same all-electron mixed-basis code was used. For the atomic motion, a simple SD algorithm was adopted for simplicity, although the convergence of the atomic position was rather slow. The result is shown in **Fig. 4**. Although we do not insist any distortion, the initial structure starts distorting spontaneously. We used the same cutoff energy for PW's and the same set of AO's as above. We start from the ideal square of 0.24 nm × 0.24 nm. Concerning the total energy, it decreases monotonically toward the 450th step, and has a local minimum there. Although there is slight humps in the total energy around the 300th and 650th steps, the total energy decreases almost monotonically everywhere. Around the 1500th step, the total energy seems to converge well. On the other hand, for the atomic positions, the ideal square becomes rectangular at first. The shorter bond in this rectangle decreases monotonically toward the 600th step and then increases again. They converge to 0.236–0.237 nm. On the other hand, one of the diagonal bonds increases monotonically and the other one decreases monotonically. Both of them converge around the 1600th step and the resulting shape is a rhombus.

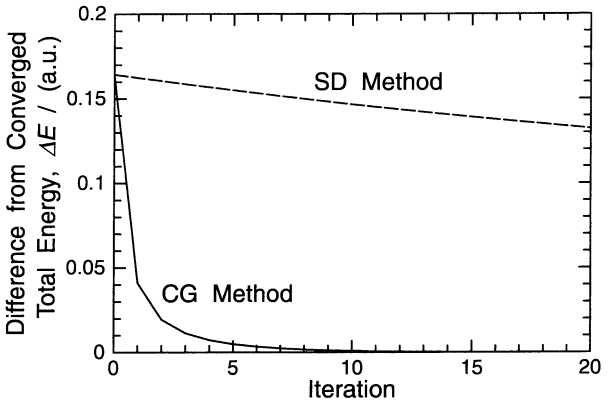

Fig. 2 Error in the total energy of Cu_4 vs iteration number by CG (solid curve) and SD methods (broken curve).

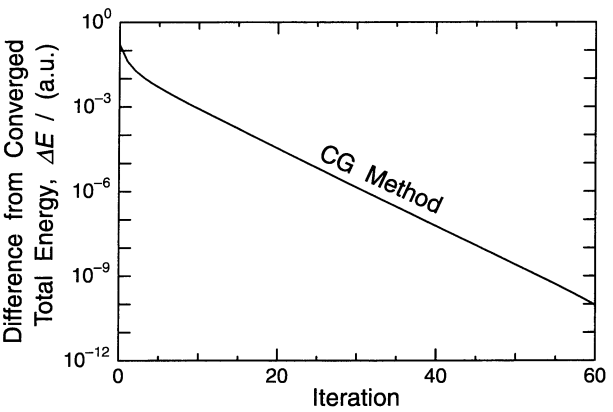

Fig. 3 Log-scale error in the total energy of Cu_4 vs iteration number by CG method (solid curve).

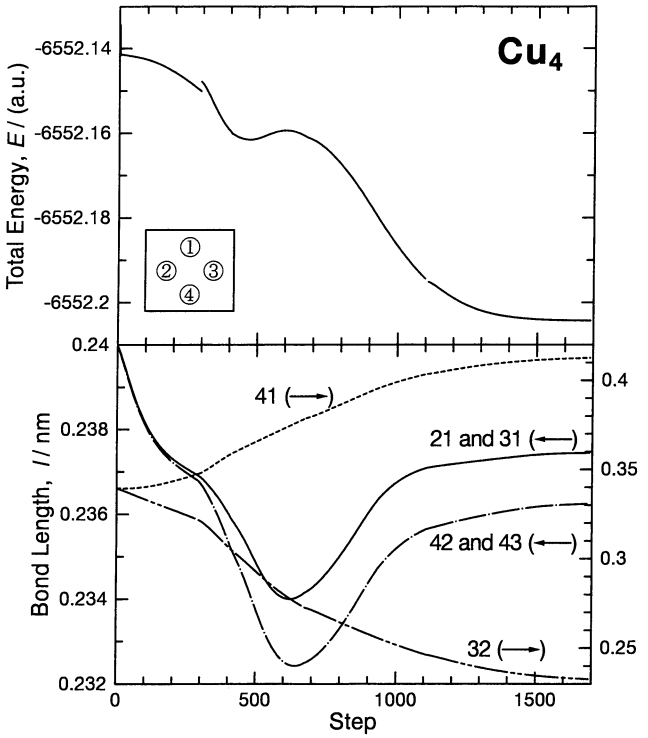

Fig. 4 Total energy of Cu_4 (top); the bond lengths between two Cu atoms, which are indicated by a pair of integers followed by an arrow (bottom) vs atomic iteration number. Starting position of Cu_4 is illustrated on top picture. The arrows show the axis direction. The "Step" in the abscissa denotes the step-number of updating atomic positions. Before each update of atomic positions, enough iterations are performed for the convergence of the electronic structure by means of the present CG method.

IV. Summary

In this paper, we have implemented the CG method in our all-electron mixed-basis program and found that the convergence of the electronic states is much improved by using the CG method compared to the SD method. We have shown a result of the CG for the case of Cu_4.

In order to speed up the atomic relaxation, it is desirable to introduce a more efficient algorithm also for the updations of atomic positions. We are now trying to implement the Broyden method[18] for that purpose. Also the Broyden method can be used to mix the electronic charge density of previous steps appropriately.

Acknowlegements

The authors would like to thank the technical staffs (Nakanomyo, Akiyama, Ichinoseki, Itoh, Wada, Miura, and Sato) in the supercomputing center at IMR for their continuous support of the supercomputing facility.

REFERENCES

(1) R. Car and M. Parrinello: Phys. Rev. Lett., **55** (1985), 2471–2474.
(2) M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos: Rev. Mod. Phys., **64** (1992), 1045–1097.
(3) W. Kohn and L. J. Sham: Phys. Rev., **140** (1965), 1133–1138.
(4) M. Born: Z. Physik., **40** (1927), 167–192; M. Born and J. R. Oppenheimer: Ann. Physik., **84** (1927), 457–484.
(5) L. Kleinman and D. M. Bylander: Phys. Rev. Lett., **48** (1982), 1425–1428.
(6) K. Ohno, K. Esfarjani and Y. Kawazoe: *Computational Materials Science: From Ab Initio to Monte Carlo Methods*, Springer Series on *Solid-State Sciences*, Springer-Verlag, Berlin, Heidelberg, (1999), p. 1–342.
(7) D. J. Singh: *Planewaves, Pseudopotentials, and the LAPW Method*, Kluwer Acadeic Press, Amsterdam, (1994).
(8) J. M. Soler and A. R. Williams: Phys. Rev. B, **40** (1989), 1560–1564; **42** (1990), 9728–9731; **47** (1993), 6784–6786.
(9) R. Yu, D. Singh and H. Krakauer: Phys. Rev. B, **43** (1991), 6411–6422.
(10) S. Goedecker and K. Mashke: Phys. Rev. B, **45** (1992), 1597–1604.
(11) P. E. Blöchl: Phys. Rev. B, **50** (1994), 17953–17979.
(12) K. Ohno, Y. Maruyama, H. Kamiyama, E. Bei, K. Shiga, Z.-Q. Li, K. Esfarjani and Y. Kawazoe: in *Mesoscopic Dynamics of Fracture: Computational Materials Design*, Vol. 1, edited by H. Kitagawa, T. Aihara, Jr. and Y. Kawazoe, Advances in Materials Research Series, Springer, Berlin, Heidelberg, (1998) pp. 210–219.
(13) S. G. Louie, K.-M. Ho and M. L. Cohen: Phys. Rev. B, **19** (1979), 1774–1782.
(14) P. Bendt and A. Zunger: Phys. Rev. B, **26** (1982), 3114–3137.
(15) W. H. Press, S. A. Teukolsky, W. T. Vettering, B. R. Fiannery: *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, New York, (1992), pp. 413–418.
(16) M. C. Payne, M. Needels and J. D. Joannopoulos: Phys. Rev. B, **37** (1988), 8138–8144.
(17) R. D. King-Smith and D. Vanderbilt: Phys. Rev. B, **49** (1994), 5828–5844.
(18) C. G. Broyden: Math. Comput., **19** (1965), 577–593.