Theory of large-scale matrix computation and applications in electronic structure calculation

2008 J. Phys.: Condens. Matter 20 294202

(http://iopscience.iop.org/0953-8984/20/29/294202)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **20** (2008) 294202 (7pp) doi:10.1088/0953-8984/20/29/294202

Theory of large-scale matrix computation and applications in electronic structure calculation

T Fujiwara[1,2], **T Hoshi**[2,3] **and S Yamamoto**[2]

[1] Center for Research and Development of Higher Education, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan
[2] Core Research for Evolutional Science and Technology, Japan Science and Technology Agency (CREST-JST), Japan
[3] Department of Applied Mathematics and Physics, Tottori University, Tottori 680-8550, Japan

Received 21 January 2008
Published 24 June 2008
Online at stacks.iop.org/JPhysCM/20/294202

Abstract

We review the methods that we recently developed for making large-scale electronic structure calculations, both using one-electron theory and using many-electron theory. The methods are based on the density matrix representation, together with the Wannier state representation and the Krylov subspace method, using one-electron theory of systems on a scale of a few tens of nanometers. The hybrid method of quantum mechanical molecular dynamical simulation is explained. The Krylov subspace method, the CG (conjugate gradient) method and the shifted COCG (conjugate orthogonal conjugate gradient) method can be applied in the investigation of the ground state and the excitation spectra using many-electron theory. The mathematical foundation of the Krylov subspace method for large-scale matrix computation is focused on, and the key techniques of the shifted COCG method, i.e. using the collinear residual and 'seed-switching', are explained. A wide variety of applications of the extended novel algorithm are also explained. These include studies of fracture formation and propagation, liquid carbon, and formation processes of gold nanowires, together with the application to the extended Hubbard model.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Large-scale matrix computation is crucial in electronic structure theory, both for one-electron theory for large-scale systems and for many-body theory for strongly interacting electron systems. The interplay of the electronic structure and nanoscale atomic structure plays an essential role in physical properties of nanostructure materials and the order-N algorithm has been extensively investigated. The size of the Hilbert space grows exponentially with linear increase of the system size in many-electron problems.

Ten very important algorithms were invented in the 20th century [1, 2]. These algorithms are the Krylov subspace method, the QR algorithm, the Householder algorithm, the fast Fourier transformation (FFT) etc; most of them involve matrix algebra and an order-N algorithm. The FFT algorithm is one of the bases of the local density approximation (LDA) for density functional theory (DFT), and the Lanczos method, one of the Krylov subspace methods, is one of the bases for many-body electron theory. The efficiency of the modern Krylov subspace method seems not to be widely known in the field of electron theory, both for LDA and for many-electron theory.

For one-electron theory or DFT, the most important states are the states near the Fermi energy or the band gap. Then the standard mathematical tool is the diagonalization of the Hamiltonian matrix. This may be a serious difficulty for large-scale systems. For many-electron theory, the difficulty is the huge size of the Hamiltonian matrix and the resultant memory size and computational time. These are exactly the targets of the field of large-scale matrix computation mentioned above.

In this paper, we report our recent activity in (1) developing a quantum mechanical molecular dynamical (MD) simulation method with exact diagonalization, the Wannier states representation and the Krylov subspace method for nanoscale systems up to a few tens of nm in size and (2) the investigation

of the many-electron problem, i.e. the degenerate orbital extended Hubbard Hamiltonian of size 6.4×10^7, with the Krylov subspace method. We explain the key aspects in one-electron theory for a large-scale system and the many-electron theory in section 2. Section 3 is devoted to the explanation of the Krylov subspace methods. Several applications are reviewed in section 4 and a summary is given in section 5.

2. One-electron theory versus many-electron theory

2.1. One-electron spectra of large-scale systems

2.1.1. Density matrix formulation. The LDA calculation is based on variational principles and usually on eigenfunction representation of the ground state. However, the eigenfunctions are not always necessary in actual calculations nor useful in numerical investigation of large-scale systems. Instead, one can construct the formulation with the one-body density matrix [3]. Any physical property can be represented using the density matrix ρ as

$$\langle \hat{X} \rangle = \mathrm{Tr}\,[\rho \hat{X}] = \sum_{ij} \rho_{ij} \hat{X}_{ij}, \tag{1}$$

where $\hat{X}$ is an operator for the physical property X and i and j denote atomic sites and orbitals. The energy and forces acting on an individual atom can be calculated by replacing $\hat{X}$ by the Hamiltonian or its derivative. Therefore, one needs only (i, j) elements of the density matrix ρ corresponding to non-zero X_{ij} but not all elements. The density matrix ρ_{ij} is given as

$$\rho = \sum_{\alpha}^{(\mathrm{occ})} |\phi_\alpha\rangle\langle\phi_\alpha|, \tag{2}$$

where $|\phi_\alpha\rangle$ are the eigenstates or the Wannier states and the summation is restricted to within the occupied states. It can also be written as

$$\rho_{ij} = -\frac{1}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}\varepsilon\, \mathrm{Im}\, G_{ij}(\varepsilon) f\left(\frac{\varepsilon - \mu}{k_\mathrm{B} T}\right), \tag{3}$$

where G_{ij} is the Green's function defined as

$$G_{ij}(\varepsilon) = [(\varepsilon + \mathrm{i}\delta - \hat{H})^{-1}]_{ij}. \tag{4}$$

Here, μ, k_B, T and f are the chemical potential, the Boltzmann constant, the temperature and the Fermi–Dirac distribution function, respectively.

We have developed a set of computational methods for electronic structure calculations, i.e. the generalized Wannier state method [4–6], the Krylov subspace method (the subspace diagonalization method [7] and the shifted COCG method [8]) and the generalized Wannier state solver with parallelism [9]. These methods are ones for calculating the one-body density matrix and/or the Green's function for a given Hamiltonian. Calculation was carried out using the tight-binding formalism of the Hamiltonian. These methods can be used in a hybrid way as is explained in section 2.1.5 [10].

2.1.2. Wannier state representation. The order-N algorithm can be constructed for semiconductors and insulators on the basis of the Wannier state representation. The generalized Wannier states are localized wavefunctions in condensed matter obtained by the unitary transformation of occupied eigenstates [11, 12, 4], and also obtained in an iterative way, starting from a trial localized wavefunctions, with a mapped eigenvalue equation [4]

$$H_{\mathrm{WS}}^{(i)} |\phi_i^{(\mathrm{WS})}\rangle = \varepsilon_{\mathrm{WS}}^{(i)} |\phi_i^{(\mathrm{WS})}\rangle, \tag{5}$$

where

$$H_{\mathrm{WS}}^{(\mathrm{i})} \equiv H + 2\eta_\mathrm{s}\bar{\rho}_i - H\bar{\rho}_i - \bar{\rho}_i H \tag{6}$$

$$\bar{\rho}_i \equiv \rho - |\phi_i^{(\mathrm{WS})}\rangle\langle\phi_i^{(\mathrm{WS})}| = \sum_{j(\neq i)}^{\mathrm{occ}} |\phi_j^{(\mathrm{WS})}\rangle\langle\phi_j^{(\mathrm{WS})}|, \tag{7}$$

and the energy parameter η_s should be much larger than the highest occupied level. Once one obtains the Wannier states, the density matrix can be easily constructed from equation (2) and the force acting on each atom can be calculated. We observed that the bond forming and breaking processes in the localized Wannier states are well described as changes between a bonding and a non-bonding orbital. The Wannier states depend upon the local environment and the above iterative procedure is suitable for MD simulation.

2.1.3. The Krylov subspace method. For metallic systems, the Krylov subspace method is very useful for achieving computational efficiency (accuracy and speed) [7, 8]. The Green's function can be calculated in the Krylov subspace and one calculates the density matrix using equation (3). Details are explained in section 3. The Krylov subspace method is, of course, applicable to semiconductors and insulators, too.

2.1.4. Comparison among solver methods and order-N character. Figure 1 demonstrates our methods for 10^2–10^7 atoms with and without parallel computation [5, 6, 10], where the computational time is shown for the standard eigenstate solver (EIG) and the solver methods that we developed: the Wannier state solver with variational procedure (WS-VR), the Wannier state solver with perturbation procedure (WS-PT) and the Krylov subspace solver with subspace diagonalization (KR-SD). Parallel computations are achieved by the Open MP technique (http://www.openmp.org). The Hamiltonian forms used here are the Slater–Koster form ones for silicon [13] and carbon [14], and we use the linear muffin-tin orbital (LMTO) theory [15] in the form for the first order ($H^{(1)}$) for copper. For the data in figure 1, except those from the eigenstate solver, the computational cost is 'order N' or linearly proportional to the system size (N), up to ten million atoms, and shows a satisfactory performance in parallel computation. The computational performance of the Wannier state methods can be faster, at best by several hundred times, than that of the Krylov subspace method (see figure 1, for example), particularly if a dominant part of the wavefunctions are well localized. Now the program package (ELSES: extra-large-scale electronic structure calculation) is being prepared [16].

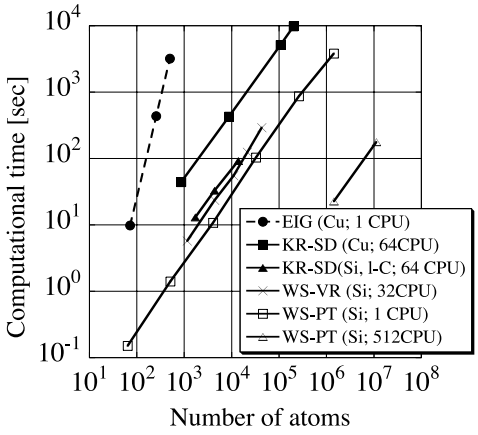

Figure 1. The computational time as a function of the number of atoms (N) [5, 6, 10]. The time was measured for metallic (fcc Cu and liquid C) and insulating (bulk Si) systems with up to 11 315 021 atoms, using conventional eigenstate calculation (EIG) and using our methods for large systems: the KR-SD, WS-VR and WS-PT methods. See the original papers [5, 6, 10] for the details of the parallel computation.

2.1.5. The multiple-solver method. Since our method is based on the density matrix formulation, we can construct another very important method 'the multiple-solver method'. The basic idea is a division of the Hilbert space:

$$\rho = \rho_{\mathrm{A}} + \rho_{\mathrm{b}}, \qquad \rho_{\mathrm{A}}\rho_{\mathrm{B}} = 0, \tag{8}$$

and the calculation can be done independently on the different parts A and B. The importance of this is that this hybrid method is completely within the quantum mechanical framework. Then we can use this hybrid scheme as the multiple-solver method for nanoscale systems and the calculated results do not cause any artificial discontinuity of physical quantities which can always occur in the hybrid scheme based on naive simple division of the physical space. The choice or hybrid of the solvers is important for an optimal calculation with a proper balance between accuracy and computational cost [5, 10].

2.2. The ground state of many-electron theory and excitation spectra

In many-electron theory, we usually treat large matrices and the calculation becomes more and more difficult, since the occupation freedom of one site grows exponentially and the matrix size is extraordinarily large [17]. In order to get the precise eigenenergy and eigenvector of the ground state, one should use the Lanczos method and the CG method (the inverse iteration method) simultaneously. The Lanczos method is useful for getting the approximate eigenenergy and eigenvectors of the ground state. However, the orthogonality of the basis vectors generated is broken at low iteration steps and the precision of the ground state energy and wavefunction could not be preserved. Then we use the CG method (the inverse iteration method) to improve them. After estimating the accuracy of the calculation (the norm of the residual vector), we should repeat this procedure until enough accuracy is obtained.

After we obtain the wavefunction of the ground state, we should analyze the properties of excitations for a wide range of energy. For this purpose, the shift property of the COCG method, i.e. the shifted COCG method, can be a powerful tool. The advantages of the shifted COCG method are an efficient algorithm for solving shifted linear equations, the ability to monitor errors during the iterative calculation and the robustness. Thus, it is very suitable for addressing problems in the many-electron problem area [18]. A detailed explanation is given in section 3.

3. The Krylov subspace method

3.1. The Krylov subspace

We consider the simultaneous linear equations

$$[(\varepsilon + \mathrm{i}\delta)\mathbf{1} - \hat{H}]\,|x_j\rangle = |j\rangle, \tag{9}$$

for a given vector $|j\rangle$, real numbers ε and δ. $\mathbf{1}$ is the unit matrix. When $\hat{H}$ is a huge $N \times N$ matrix, the inverse of $\hat{H}$ or $[(\varepsilon+\mathrm{i}\delta)\mathbf{1}-\hat{H}]$ is not easily obtained or is impossible to obtain, and the iterative method becomes a useful concept. One can obtain an approximate eigenvector $|x_j\rangle$ in a subspace spanned by vectors $\{\hat{H}^n|j\rangle\}$:

$$K_\nu(\hat{H}, |j\rangle) \equiv \mathrm{span}\{|j\rangle,\ \hat{H}|j\rangle,\ \hat{H}^2|j\rangle, \ldots,\ \hat{H}^{\nu-1}|j\rangle\}. \tag{10}$$

This subspace $K_\nu(\hat{H}, |j\rangle)$ is called the Krylov subspace. The basic theorem of the Krylov subspace is the invariance of the subspace under a scalar shift $\sigma\mathbf{1}$;

$$K_\nu(\hat{H}, |j\rangle) = K_\nu(\sigma\mathbf{1} + \hat{H}, |j\rangle). \tag{11}$$

Lanczos found a new powerful way to generate an orthogonal basis for such a subspace when the matrix is symmetric [19]. Hestenes and Stiefel proposed an elegant method, known as the conjugate gradient (CG) method, for systems that are both symmetric and positive definite [20].

3.2. Subspace diagonalization

The first method is to find eigenvectors $\{|w_\alpha^{(j)}\rangle\}$ approximated in $K_\nu(\hat{H}, |j\rangle)$ by diagonalizing the reduced Hamiltonian matrix

$$H^{K_\nu(\hat{H},|j\rangle)} = \{\langle K_n^{(j)}|\hat{H}|K_m^{(j)}\rangle\}, \tag{12}$$

where $\{|K_m^{(j)}\rangle | m = 1, \ldots, \nu\}$ is the orthogonalized basis set of the Krylov subspace $K_\nu(\hat{H}, |j\rangle)$, which satisfies the three-term recurrence relation, and is constructed by the Lanczos process or the Gram–Schmidt process. For this subspace we can calculate the density matrix very easily [7].

The subspace diagonalization method may be accurate enough for several purposes for one-electron spectra of large-scale systems and calculation of total and local densities of states. However, the orthogonality would be broken if we used a larger number of the subspace dimensions, for the basis vectors satisfying the three-term recurrence relation [7, 8]. Therefore, the numerical accuracy may be limited when one needs finer structure of spectra and we should extend the methodology to the shifted COCG method.

For many-electron theory, the Lanczos method is widely used for obtaining the eigenenergy and many-electron wavefunction of the ground state using the exact diagonalization method. The accuracy can be greatly improved when we use the CG method.

3.3. The shifted COCG method and 'seed-switching' technique

When the matrix $(\varepsilon_0\mathbf{1} - \hat{H})$ is real symmetric, one can use the CG method for an iterative solution of the simultaneous linear equation $(\varepsilon_0\mathbf{1} - \hat{H})\mathbf{x} = \mathbf{b}$. One should introduce the infinitesimally small (but finite) imaginary number $\mathrm{i}\delta$ for the Green's function and the matrix $(\varepsilon_0 + \mathrm{i}\delta)\mathbf{1} - \hat{H}$ is complex symmetric. Then we can use the conjugate orthogonal conjugate gradient (COCG) method for solving the equation $\{(\varepsilon_0 + \mathrm{i}\delta)\mathbf{1} - \hat{H}\}\mathbf{x} = \mathbf{b}$ [8].

Since the energy parameters ε are arbitrarily given or continuously changing over a wide energy range, one should solve also the shifted linear equations

$$[(\varepsilon_0 + \sigma + \mathrm{i}\delta)\mathbf{1} - \hat{H}]\,|x_j^{(\sigma)}\rangle = |j\rangle, \tag{13}$$

with a fixed energy (seed) ε_0. The energy shift parameter σ can even be complex. A shifted COCG method was constructed [8, 18] in which the theorem of the collinear residual [21] for the shifted linear systems is applied to the COCG method. The essential property is based on the basic invariance theorem of the Krylov subspace (equation (11)) under an energy shift $\varepsilon_0 + \sigma$ from ε_0. Therefore, the Krylov subspace for the equation $[(\varepsilon_0 + \sigma + \mathrm{i}\delta)\mathbf{1} - \hat{H}]\,|x_j^{(\sigma)}\rangle = |j\rangle$ can be generated from that of $[(\varepsilon_0 + \mathrm{i}\delta)\mathbf{1} - \hat{H}]\,|x_j\rangle = |j\rangle$ for a selected seed energy ε_0. The very important fact is that this shift procedure is a scalar linear calculation. The essential cost for solving equation (9) should be paid only for the seed energy ε_0; the rest is a scalar linear calculation which has negligible cost.

The choice of seed energy is not unique and sometimes the calculations cannot be finished under required criteria. Then one should continuously change the energy parameter and choose a new seed energy $\varepsilon + \mathrm{i}\eta$ again. The essentially important point is that we can continue the calculation with a new seed energy, keeping the calculated information for the former seed energy. This is another very important property called 'seed-switching' [22, 18].

3.4. Accuracy control with the residual vector and the robustness of the shifted COCG method

It is essentially important to know the accuracy of the solution during the iteration procedure and we can monitor the convergence behavior of the iterative solutions of the Krylov subspace method.

The residual vector can be defined both in the subspace diagonalization and in the shifted COCG method [8] as

$$|r_j^{(\nu)}\rangle = (\varepsilon + \mathrm{i}\delta - \hat{H})|x_j^{(\nu)}\rangle - |j\rangle, \tag{14}$$

where $|x_j^{(\nu)}\rangle$ is the νth iterative solution. This residual vector can be monitored during the iterative calculation and we can stop the iterative procedure, without fixing the dimension of the Krylov subspace, once one can obtain the required accuracy. The norm of the residual vector can give the upper limit of accuracy of the Green's function itself [18].

The shifted COCG method is numerically robust and one can reduce the norm of the residual vector to the machine accuracy. Therefore, the shifted COCG method may be used to calculate accurate or fine densities of electronic states in one-electron spectra for large-scale systems or fine excitation spectra for many-electron problems.

4. Applications

4.1. Application to nanoscale systems

4.1.1. Formation and propagation of fractures in silicon crystal. In this subsection, we present an application study of our simulation: we look at fracture formation and propagation phenomena for Si nanoscale crystal [5, 6], where the Hamiltonian is given as a tight-binding representation [13]. The calculation was carried out by Wannier state methods with up to 10^5 atoms.

In the dynamical fracture formation process on the (001) plane, two bonds are broken and an asymmetric dimer (2×1 periodicity on the resultant (001) surface) is formed after thermal motions in a time of about 0.4 ps. First, bonds are broken successively in an atom array of the dimer bonds on the plane along one of the [110] directions. Along the asymmetric dimer bonds formed, the inter-atomic distance is shortened due to the bonding character formed. The distortion energy is accumulated and, then, other bonds along a parallel atom array, but not the same one, are broken. This fracture propagation (perpendicular to the direction of the asymmetric dimer bonds formed) is governed by the accumulated distortion energy. Our calculation can represent this mechanism of surface breaking on the (001) planes of Si crystals [5].

We also studied for 14 nm scale Si crystals the easy-propagating plane of fracture [6]. It is widely known that the easy-propagating plane of fracture in Si is a (110) or (111) plane. In the case of fracture on the (111) plane, the (111)-(2×1) surface reconstruction appears (the Pandey structure [23]) and several steps are formed. The fracture propagation plane is explained not by the energy of established stable surfaces but by that of ideal or transient surface structure without reconstruction. In a MD process for larger systems with 14 nm length, even if a fracture propagation starts on a (001) plane, the plane of the fracture propagation changes to a (111) or (110) plane. Figure 2 shows examples of the simulation results.

4.1.2. Liquid carbon. Liquid carbon with 13 824 atoms was simulated with the Krylov subspace method [10]. The density and the temperature are set at $\rho = 2.0\ \mathrm{g\,cm^{-3}}$ and $T = 6000$ K. The time interval of a MD step is $\Delta t = 1$ fs and the subspace dimension and the number of interacting atoms are chosen to be $\nu = 30$ and $N_{\mathrm{PR}} = 200$, respectively. Figure 3(a) shows the resultant pair correlation (PC) function with comparison of the conventional eigenstate method for 216

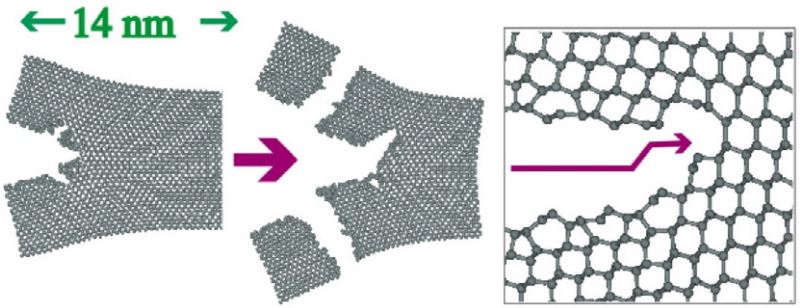

Figure 2. Silicon cleavage dynamics [6]. A 14 nm scale simulation result shows the bending of the cleavage path from an unstable (001)-like plane into experimentally observed (111)-like and (110)-like planes. The right panel shows a process of step formation on the Si(111)-2 × 1 surface.

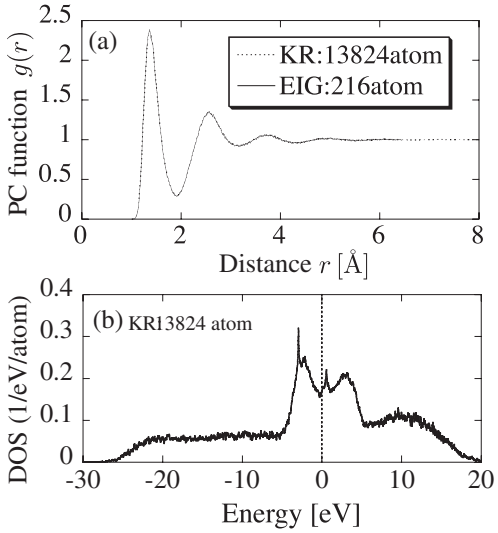

Figure 3. Molecular dynamical simulation of liquid carbon from the MD simulation, with the Krylov subspace method. (a) The pair correlation function and (b) the electron density of states.

atoms and we should note that the two graphs are identical. Figure 3(b) shows the electronic density of states (DOS) of a system with 13 824 atoms, from the Green's function, obtained by the Krylov subspace method. The DOS calculation was achieved with the controlling parameters at a heavier computational cost ($\nu = 300$ and $N_{\mathrm{PR}} = 1000$) and $\eta = 0.05$ eV. Since the present Hamiltonian includes only s and p orbitals, the resultant DOS misses a structure in higher energy regions. The resultant DOS shows the characteristic profile of liquid carbon, e.g. a narrow π band appears between -5 and $+5$ eV as for carbon nanotubes. The π bond in the liquid phase is imperfect and non-bonding (atomic) p states appear as a sharp peak near the chemical potential ($\varepsilon \simeq 0.6$ eV).

4.1.3. Helical multishell structure of gold nanowire. Another application is to the process of formation of helical multishell gold nanowires [24]. Gold nanowires obtained by the TEM thinning process have helical multishell structures along the original [110] axis with helicity, and the outermost shell is a (111)-like atomic sheet [25]. The difference in number of atoms between the outermost and the next outermost shells is 7, called the 'magic number', except for cases of five and seven atoms on the outermost shell.

We proposed a two-stage model of formation of Au nanowires where the driving force for the helicity is the atom row slip. At the first stage, the outermost shell is dissociated from the inner shell to rotate freely. At the second stage, an atom row on the outermost shell slips and the (001) faces on the rod surface transform into (111) surfaces.

We verified the above two-stage model by using MD simulations with a tight-binding Hamiltonian [26], starting from an 'ideal' nanowire of stacking (110) sections of the fcc lattice. The calculation was carried out using the eigenstate solver with about 80–1020 atoms. Here, we show results for 143 atoms in figure 4(a) [24] and those for larger systems of 1020 atoms in (b). The total energy decreases almost monotonically after 1000 MD steps (1 MD step = 1 fs). First, the surface atoms dissociate from the inner shell and, then, can move rather freely. From 2000 to 5000 MD steps, the (001) sheet reconstructs into a hexagonal (111)-like surface with an atom row slip deformation, and the helical structure on the surface appears. The inner shell rotates as the atom row slip occurs. Analysis of the electronic structure shows that the mechanism is governed, in both stages, by the d band electrons extending over the (111)-like surface, where the center of gravity of the d band locates in the lower energy side. The helical nanowires appear only among metals with a wider d band, e.g. in Au and Pt but not in Ag and Cu. Helicity is introduced by the surface reconstruction or the atom row slip on the (001) sheet, because the triangular (111)-like sheet is more preferable for d orbitals extending over the surface. The d bandwidths for platinum and gold are commonly wider than those for lighter elements, Ag and Cu, and the calculated result explains why platinum nanowire can also be formed with helicity.

4.2. Application to many-electron systems: the excitation spectrum of the multi-orbital extended Hubbard Hamiltonian on a two-dimensional square lattice

The transition metal oxides have been paid a great deal of attention due to their various physical properties which are drastically changed and controllable by external fields

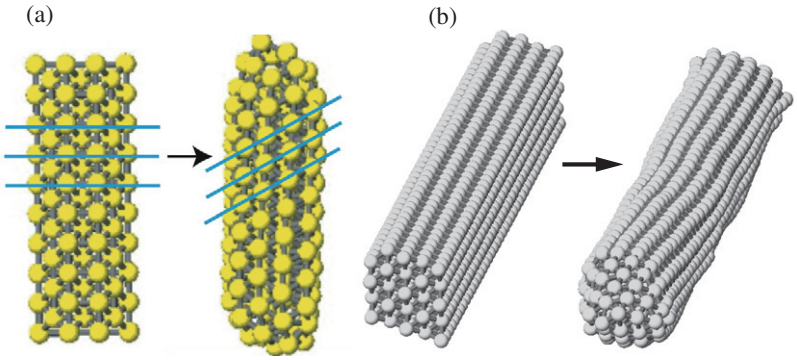

Figure 4. The process of formation of the helical multishell structure in gold nanowires. Non-helical structure (left) is transformed into a helical one (right). (a) Process of formation of an (11–4) helical multishell gold nanowire [24]. The symbol (11–4) means that the numbers of atoms are 11 and 4 in the outer and inner shells, respectively. Atom row slip along the wire axis introduces helicity, as shown by three lines. (b) Transformation of the structure of the longer gold nanowire of 1020 Au atoms; the resultant (15–8–1) nanowire contains several defects.

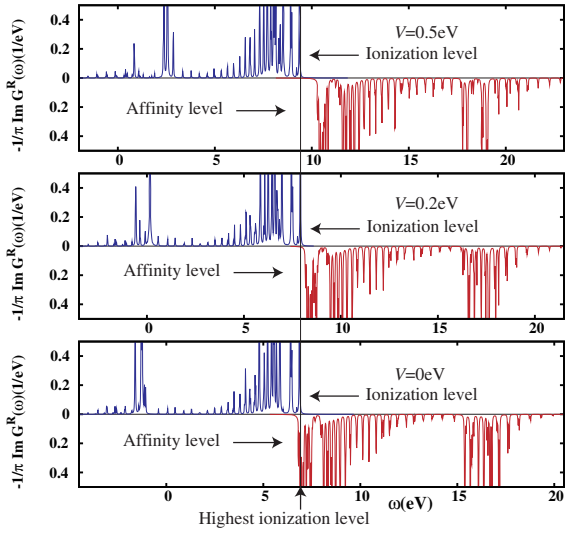

Figure 5. The excitation spectra, electron ionization and affinity of the doubly degenerate extended Hubbard model on an $\sqrt{8} \times \sqrt{8}$ two-dimensional square lattice [17]. The matrix size of the reduced subspace of the total spin $S_z = 0$ is 64 128 064. The excitation gap is formed with increasing inter-site Coulomb interaction V.

or doping. Here we show an application of the shifted COCG method to the extended Hubbard model with doubly degenerate orbital and the inter-site Coulomb interaction on a two-dimensional square lattice [17]. This is a model for $La_{\frac{3}{2}}Sr_{\frac{1}{2}}NiO_4$ and we used a finite unit of $N = 8$ sites and the total number of electrons $N_e = \frac{3}{2}N = 12$.

We focused our attention on the Hilbert space of total spin $S_z = 0$ and the matrix size is $({}_{16}C_6)^2 = 64\,128\,064$. The difficulties for many-electron systems are (1) the large dimension of the Hamiltonian matrix which grows exponentially with number of atoms increasing linearly and (2) the very small energy intervals between adjacent eigenenergies which cause difficulty in the separation of the respective eigenvectors. This difficulty leads to the requirement of a fast, reliable and stable calculation algorithm for large matrices.

Figure 5 shows the excitation spectra of electron ionization and affinity levels, and the energy gap between these two corresponds to the excitation gap. Normally, the Hubbard model of a non-degenerate orbital, in the case of integer occupation, gives the insulating gap due to the on-site Coulomb interaction and, in contrast, in the case of non-integer occupation, the system is a metal. Here, in the doubly degenerate case, the charge stripe order with an insulator gap is formed due to the inter-site Coulomb interaction V and, on top of that, a spin stripe is formed with anisotropy of electron hoppings [17].

The crucial point is that we should keep very high accuracy of the computation for judging the 'gap', compared with the 'level interval' in finite systems, and that the iteration convergence should be controlled during the iterative calculation. Therefore, the capabilities of convergence (accuracy) monitoring and robustness are seriously important, and the shifted COCG method can resolve this difficulty.

5. Conclusions

We have reviewed our methods recently developed for performing large-scale electronic structure calculation, applied for both one-electron theory and many-electron theory. For large-scale systems of about the scale of tens of nm, one can use several solver methods simultaneously as a multi-solver method. We also explained differences between the two theories from the viewpoint of large-scale matrix computation. Then we presented examples of the applications for nanoscale systems, the formation and propagation of fracture in large silicon crystals, MD simulation for liquid carbon, and the formation of helical multishell structure for gold nanowires, and an example of a many-electron problem, the orbitally degenerate extended Hubbard model. In these applications, we stressed the importance of the hybrid scheme of multiple-solver methods and the novel computational algorithm.

Acknowledgment

The numerical calculation was partly carried out using the supercomputer facilities of the Institute for Solid State Physics, University of Tokyo.

References

[1] Dongarra J and Sullivan F 2000 *Comput. Sci. Eng.* **2** 22
[2] Cipra B A 2000 *SIAM News* **33** (4)
[3] Kohn W 1996 *Phys. Rev. Lett.* **76** 3168
[4] Hoshi T and Fujiwara T 2000 *J. Phys. Soc. Japan* **69** 3773
[5] Hoshi T and Fujiwara T 2003 *J. Phys. Soc. Japan* **72** 2429
[6] Hoshi T, Iguchi Y and Fujiwara T 2005 *Phys. Rev.* B **72** 075323
[7] Takayama R, Hoshi T and Fujiwara T 2004 *J. Phys. Soc. Japan* **73** 1519
[8] Takayama R, Hoshi T, Sogabe T, Zhang S-L and Fujiwara T 2006 *Phys. Rev.* B **73** 165108
[9] Geshi M, Hoshi T and Fujiwara T 2004 *J. Phys. Soc. Japan* **72** 2880
[10] Hoshi T and Fujiwara T 2006 *J. Phys.: Condens. Matter* **18** 10787
[11] Mauri F, Galli G and Car R 1993 *Phys. Rev.* B **47** 9973
[12] Marzari N and Vanderbilt D 1997 *Phys. Rev.* B **56** 12847
[13] Kwon I, Biswas R, Wang C Z, Ho K M and Soukoulis C M 1994 *Phys. Rev.* B **49** 7242
[14] Xu C H, Wang C Z, Chan C T and Ho K M 1992 *J. Phys.: Condens. Matter* **4** 6047
[15] Andersen O K and Jepsen O 1984 *Phys. Rev. Lett.* **53** 2571
[16] http://elses.jp
[17] Yamamoto S, Fujiwara T and Hatsugai Y 2007 *Phys. Rev.* B **76** 165114
[18] Yamamoto S, Sogabe T, Hoshi T, Zhang S-L and Fujiwara T 2008 in preparation
[19] Lanczos C 1950 *J. Res. Natl Bur. Stand.* **45** 225
Lanczos C 1952 *J. Res. Natl Bur. Stand.* **49** 33
[20] Hestenes M R and Stiefel E 1952 *J. Res. Natl Bur. Stand.* **49** 409
[21] Frommer A 2003 *Computing* **70** 87
[22] Sogabe T, Hoshi T, Zhang S-L and Fujiwara T 2007 *Frontiers of Computational Science* ed Y Kaneda, H Kawamura and M Sasai (Berlin: Springer) pp 189–95
[23] Pandey K C 1981 *Phys. Rev. Lett.* **47** 1913
[24] Iguchi Y, Hoshi T and Fujiwara T 2007 *Phys. Rev. Lett.* **99** 125507
[25] Kondo Y and Takayanagi K 2000 *Science* **289** 606
[26] Kirchhoff F, Mehl M J, Papanicolaus N I, Papaconstantpoulos D A and Khan F S 2001 *Phys. Rev.* B **63** 195101