Towards a linear-scaling algorithm for electronic structure calculations with the tight-binding Korringa–Kohn–Rostoker Green function method

2008 J. Phys.: Condens. Matter 20 294215

(http://iopscience.iop.org/0953-8984/20/29/294215)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **20** (2008) 294215 (8pp) doi:10.1088/0953-8984/20/29/294215

Towards a linear-scaling algorithm for electronic structure calculations with the tight-binding Korringa–Kohn–Rostoker Green function method

Rudolf Zeller

Institut für Festkörperforschung, Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany

E-mail: Ru.Zeller@fz-juelich.de

Received 13 February 2008
Published 24 June 2008
Online at stacks.iop.org/JPhysCM/20/294215

Abstract

The proposed algorithm is based on the exponential decay of the screened structure constants in the tight-binding (TB) Korringa–Kohn–Rostoker (KKR) Green function method and on a spatial truncation of the Green function in the spirit of Kohn's principle of nearsightedness of electronic matter. The KKR matrix equations are solved iteratively and non-zero electronic temperatures are used to accelerate the iterations. The dependence of the total energy accuracy on the size of the truncation region was investigated for large Cu and Pd supercells and it was found that total energy errors smaller than 2 meV could be achieved if the truncation region contained a few thousand atoms.

1. Introduction

Electronic structure calculations within density-functional theory are widely used to investigate structural, electronic and magnetic properties of condensed matter. Systems with a few hundred atoms can be treated routinely today, but larger systems with thousands of atoms require enormous computer resources, if standard techniques are used to solve the density-functional equations, since computing memory and time increase with second and third power of the number N of atoms. In recent years much effort has been spent on overcoming these $O(N^2)$ and $O(N^3)$ bottlenecks. As a result a number of computer codes have appeared, for example SIESTA [1], CONQUEST [2], ONETEP [3] and OPEN-MX [4], suitable for linear-scaling $O(N)$ calculations for semiconducting and insulating systems. In these systems the spatial decay of the density matrix is exponential [5], a property which is used in the programs, with a decay constant that increases with size of the band gap. In metallic systems the density matrix decays algebraically at zero electronic temperature and an exponential decay is only achieved if non-zero temperatures are used in the calculations. However, for reasonable temperatures, which maintain the physical properties, the decay constant is smaller than in band-gap systems and it is not clear whether the exponential decay is fast enough for accurate linear-scaling calculations for metallic systems.

The present study aims to gain insight into the questions of how the Korringa–Kohn–Rostoker (KKR) Green function method in its tight-binding (TB) form [6] together with nearsightedness of electronic matter [7, 8] can be used for linear-scaling calculations, how accurate results can be obtained and how much non-zero electronic temperatures can reduce the computational effort. The use of electronic nearsightedness is not new in the KKR and the related linear muffin-tin orbital (LMTO) methods. For many years these methods have been using this principle successfully, for example for description of impurity embedding in host crystals [9–11] and treatment of large systems in the locally self-consistent (LS) multiple-scattering and Green function methods [12–14]. Whereas impurity calculations exploit screening of the electronic potential by neglecting its long-range change, the locally self-consistent multiple-scattering (LSMS) and locally self-consistent Green function (LSGF) methods are based on the idea that the electron density near an atomic site is almost unaffected by the potential far away from this site. The potential outside of a local interaction zone around the atom, for which the density is calculated, is then either completely neglected (as in the LSMS method) or replaced by an effective medium potential (as in the LSGF

method). Since each atom and its surrounding interaction zone are treated independently, the effort in the LS methods scales linearly with the number of atoms. A disadvantage, however, is that the accuracy is limited by the size of the local interaction zone [15] and that the computational effort increases with the third power of the number of atoms in the zone.

The present work is based on the observation that the LS methods only exploit electronic nearsightedness, whereas methods like SIESTA, CONQUEST, ONETEP and OPEN-MX utilize, in addition, the fact that localized basis functions and iterative solution lead to considerable saving of computer resources. Within the LSMS method some authors have attempted to accelerate the calculations by iteration [16] and by sparsity of the TB-KKR matrix [15, 17], but for interaction zones, which are usually restricted to less than a hundred atoms, the TB-KKR matrix is not really sparse enough and often too many iterations are needed so that direct solutions are cheaper than iterative ones. A different approach to exploit sparsity and iteration is suggested by Smirnov and Johnson [18]. They propose to solve the KKR equations for the entire system by iteration and find computing times per atomic site that scale as $O(N^{1+\epsilon})$ with $\epsilon < 0.2$. They argue that for increasing system size the required number of **k** points in reciprocal space scales as N^{-1} and conclude that total energy calculations are possible with $O(N^{1+\epsilon})$ scaling if the number of **k** points is reduced appropriately. However, if the system is so large that one **k** point is enough, or for real space calculations, this argument breaks down and the computational effort is proportional to $N^{2+\epsilon}$.

In this paper it will be shown that the approach of Smirnov and Johnson can be supplemented with a scheme that uses the decay of the density matrix and can be applied to calculate accurate total energies for large metallic systems with actual linear-scaling effort. To explain the idea and to illustrate its usefulness the article is organized as follows. Section 2 contains the basic equations of the KKR Green function method. Section 3 describes the methodology with explanation of how sparsity is achieved, iterative solutions are accomplished and nearsightedness is utilized. In section 4 total energy results are presented which illustrate how accuracy improves with the number of iterations and how it is affected by the approximation used to obtain linear scaling. Sections 5 and 6 contain a discussion and summary.

2. Basic equations

In wavefunction-based formulations of density-functional theory the density is usually obtained by

$$n(\mathbf{r}) = \sum_{i,occ} |\Psi_i(\mathbf{r})|^2 \tag{1}$$

as a sum over occupied eigenfunctions $\Psi_i(\mathbf{r})$ that obey the Kohn–Sham equation

$$[-\nabla_{\mathbf{r}}^2 + V(\mathbf{r}) - E_i]\Psi_i(\mathbf{r}) = 0 \tag{2}$$

for the effective Kohn–Sham potential $V(\mathbf{r})$. Compared to that, in the KKR Green function method the density is obtained by

$$n(\mathbf{r}) = -\frac{2}{\pi}\mathrm{Im}\int_{-\infty}^{E_F} \mathrm{d}E\, G(\mathbf{r}, \mathbf{r}; E) \tag{3}$$

as an energy integral over the Green function that obeys the differential equation

$$\left[-\nabla_{\mathbf{r}}^2 + V(\mathbf{r}) - E\right] G(\mathbf{r}, \mathbf{r}'; E) = -\delta(\mathbf{r} - \mathbf{r}') \tag{4}$$

with the appropriate boundary condition $G(\mathbf{r}, \mathbf{r}'; E) \to 0$ for $|\mathbf{r} - \mathbf{r}'| \to \infty$. Atomic units ($\hbar^2/2m = e^2/2 = 1$) are used throughout this article, the factor 2 arises from spin degeneracy, which is assumed to simplify the following equations, and E_F denotes the Fermi level. Instead of solving (4) directly, which is obviously more complicated than (2), it is easier to solve the Dyson equation

$$G(\mathbf{r}, \mathbf{r}'; E) = G^r(\mathbf{r}, \mathbf{r}'; E) + \int \mathrm{d}\mathbf{r}'' G^r(\mathbf{r}, \mathbf{r}''; E) \times \left[V(\mathbf{r}'') - V^r(\mathbf{r}'')\right] G(\mathbf{r}'', \mathbf{r}'; E), \tag{5}$$

where V^r is the potential of a reference system, for which the Green function G^r is assumed to be known. The equivalence of (4) and (5) is easily verified, if the operator $-\nabla_{\mathbf{r}}^2 + V^r(\mathbf{r}) - E$ is applied on both sides of (5). Compared with (4) only the potential difference appears in (5), but not the kinetic energy operator $-\nabla_{\mathbf{r}}^2$, which is already taken into account in the reference system. Thus in a stepwise manner the Green function of a complicated system, for example an atom adsorbed on a surface, can be obtained starting from the analytically known Green function of free space by intermediate use of the ideal bulk and ideal surface as reference systems. In each step, periodicity and locality of the potential in one, two or three dimensions can be exploited [19]. Artificial reference systems can also be used, for instance one with a repulsive potential, which is of particular importance for the present work, since the Green function of a suitably chosen repulsive reference system [6] decays exponentially for energies E used in density-functional calculations.

For the numerical treatment it is convenient to divide the integral in (5) into integrals over non-overlapping space-filling cells around the atomic positions $\mathbf{R}^n$ and to apply multiple-scattering theory to separate the calculation into single-scattering and multiple-scattering events. In cell-centred coordinates $\mathbf{r}$ and $\mathbf{r}'$ the multiple-scattering representation of the Green function can be written [19] as

$$G(\mathbf{r} + \mathbf{R}^n, \mathbf{r}' + \mathbf{R}^{n'}; E) = \delta^{nn'} G_s^n(\mathbf{r}, \mathbf{r}'; E) + \sum_{LL'} R_L^n(\mathbf{r}; E) G_{LL'}^{nn'}(E) R_{L'}^{n'}(\mathbf{r}'; E). \tag{6}$$

Here L stands for the pair l and m of angular-momentum indices and $G_s^n(\mathbf{r}, \mathbf{r}'; E)$ and $R_L^n(\mathbf{r}; E)$ denote single-scattering Green function and wavefunctions. They are defined by the integral equations

$$G_s^n(\mathbf{r}, \mathbf{r}'; E) = G^0(\mathbf{r}, \mathbf{r}'; E) + \int_n \mathrm{d}\mathbf{r}''\, G^0(\mathbf{r}, \mathbf{r}''; E) V^n(\mathbf{r}'') G_s^n(\mathbf{r}'', \mathbf{r}'; E) \tag{7}$$

$$R_L^n(\mathbf{r}; E) = Y_L(\hat{\mathbf{r}})\, j_l(r\sqrt{E}) + \int_n \mathrm{d}\mathbf{r}'' G^0(\mathbf{r}, \mathbf{r}''; E) V^n(\mathbf{r}'') R_L^n(\mathbf{r}''; E), \tag{8}$$

where integration is over cell n and $V^n(\mathbf{r})$ denotes the potential restricted to cell n.

$$G^0(\mathbf{r}, \mathbf{r}'; E) = -\frac{\exp(\mathrm{i}\sqrt{E}|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|} \tag{9}$$

is the Green function for free space with vanishing potential $V(\mathbf{r}) \equiv 0$ and $Y_L = Y_{lm}$ denote spherical harmonics and j_l spherical Bessel functions. For $\mathbf{r}$ inside cell n the density (3) is given by

$$n(\mathbf{r}+\mathbf{R}^n) = -\frac{2}{\pi}\,\mathrm{Im}\int_{-\infty}^{E_\mathrm{F}} \mathrm{d}E\, G(\mathbf{r}+\mathbf{R}^n, \mathbf{r}+\mathbf{R}^n; E), \quad (10)$$

which according to (6) depends on the single-site solutions (7) and (8) and on the on-site ($n = n'$) Green function matrix elements $G^{nn}_{LL'}(E)$. Here it is important to realize that solution of (7) and (8) requires only $\mathbf{r}$ and $\mathbf{r}'$ values in cell n because the potential $V^n(\mathbf{r})$ vanishes outside cell n. This means that (7) and (8) can be solved independently for each site n with obviously linear-scaling effort.

The combination of the single-scattering events is accomplished by the Green function matrix elements, which obey the matrix equation

$$G^{nn'}_{LL'}(E) = G^{r,nn'}_{LL'}(E) + \sum_{n''L''L'''} G^{r,nn''}_{LL''}(E)\Delta t^{n''}_{L''L'''}(E) G^{n''n'}_{L'''L'}(E), \quad (11)$$

and provide all information about the multiple-scattering problem. This system of linear equations represents an algebraic Dyson equation with complex symmetric (but not Hermitian) matrices of dimension $(l_\mathrm{max}+1)^2 N$ if angular-momentum contributions with $l \leqslant l_\mathrm{max}$ are used. The Green function matrix elements $G^{r,nn'}_{LL'}(E)$ are the ones of the reference system and $\Delta t^n_{LL'}(E)$ is a t-matrix difference, which can be calculated by

$$\Delta t^n_{LL'}(E) = \int_n \mathrm{d}\mathbf{r}\, j_l(r\sqrt{E}) Y_L(\hat{\mathbf{r}}) V^n(\mathbf{r}) R^n_{L'}(\mathbf{r}; E) - \int_n \mathrm{d}\mathbf{r}\, j_l(r\sqrt{E}) Y_L(\hat{\mathbf{r}}) V^{r,n}(\mathbf{r}) R^{r,n}_{L'}(\mathbf{r}; E) \quad (12)$$

independently for each site n. Here $R^{r,n}_L(\mathbf{r}; E)$ denotes single-scattering solutions of the reference system defined in analogy to (8) with the potential $V^n(\mathbf{r})$ replaced by the reference potential $V^{r,n}(\mathbf{r})$. For periodic crystals the algebraic Dyson equation (11) is solved in reciprocal space by Fourier transformation. This leads to

$$G^{\nu\nu'}_{LL'}(\mathbf{k}; E) = G^{r,\nu\nu'}_{LL'}(\mathbf{k}; E) + \sum_{\nu''L''L'''} G^{r,\nu\nu''}_{LL''}(\mathbf{k}; E) \times \Delta t^{\nu''}_{L''L'''}(E) G^{\nu''\nu'}_{L'''L'}(\mathbf{k}; E), \quad (13)$$

where $\mathbf{k}$ is a wavevector in the Brillouin zone and ν, ν' denote basis sites in the unit cell. The back transformation to obtain the Green function matrix elements in real space is provided by integration over the Brillouin zone.

Note that the linear equations (11) and (13) are more easily solved than (5). Instead of many grid points for the spatial variables $\mathbf{r}$ and $\mathbf{r}'$ only a few angular-momentum indices L and L' are needed. Experience has shown that usually an angular-momentum cut-off $l_\mathrm{max} = 3$ is enough for total energy calculations. For N sites the matrix dimension in (11) and (13) is then $(l_\mathrm{max}+1)^2 N = 16N$. The angular-momentum representation is advantageous because it is well adapted to treat nearly spherical, atomic-like potentials along with the $1/r$ singularity in the vicinity of the nuclear centres and the rather flat potential between the atoms.

For evaluation of (10) it is important that the integration can be done on a contour in the complex energy plane [20], avoiding the singularities of the Green function that appear on the real axis as poles and branch cuts arising from discrete and continuous states. A useful contour, motivated by finite-temperature density-functional theory [21, 22], is obtained if the density is calculated by

$$n(\mathbf{r}+\mathbf{R}^n) = -\frac{2}{\pi}\,\mathrm{Im}\int_{-\infty}^{\infty} \mathrm{d}E\, f(E, E_\mathrm{F}, T) \times G(\mathbf{r}+\mathbf{R}^n, \mathbf{r}+\mathbf{R}^n; E), \quad (14)$$

where $f(E, E_\mathrm{F}, T) = (1+\exp(\beta(E-E_\mathrm{F})))^{-1}$ is the Fermi–Dirac function for inverse temperature $\beta = (kT)^{-1}$. For the present study the contour starts on the negative real energy axis at energy E_0 below the valence and above the core states. From E_0 the contour goes parallel to the imaginary axis up to $E_0 + 2Ji\pi kT$, where J is a chosen small integer. From there the contour goes parallel to the real axis to infinity. On this line the Fermi–Dirac function changes rapidly only near $E_\mathrm{F} + 2Ji\pi kT$ and it is a real function because of $f(E+2Ji\pi kT, E_\mathrm{F}, T) = f(E, E_\mathrm{F}, T)$. The integrals on the straight lines, which were done by Gauss integration with a total number of 25 mesh points, must be supplemented with the residues at the first J Matsubara energies $E_j = E_\mathrm{F} + (2j-1)i\pi kT$ with $j = 1, 2, \ldots, J$. For the temperatures $T = 400$, 800 and 1600 K used below, J was chosen as 15, 7 and 4. Note that the contour does not include the contribution of the core states, which was calculated by treating the core states in an atomic-like fashion and adding it to the contribution of the valence states.

3. Methodology

Because of the site-diagonality of the Δt matrix it is useful to rewrite $G = G^r + G^r \Delta t G$ as $(1 - G^r\Delta t)G = G^r = -(1 - G^r\Delta t)(\Delta t)^{-1} + (\Delta t)^{-1}$, which by multiplication with $(1 - G^r\Delta t)^{-1}$ leads to $G = -(\Delta t)^{-1} + (\Delta t)^{-1}[(\Delta t)^{-1} - G^r]^{-1}(\Delta t)^{-1}$. Thus the site-diagonal Green function matrix elements required in (14) can be calculated in O(N) operations from the inverse of the screened KKR matrix $M = (\Delta t)^{-1} - G^r$. The bottleneck in the calculation of M^{-1} by direct solution is the factorization step, which decomposes the matrix M into a product of lower and upper triangular matrices. This step requires O(N^3) floating-point operations for dense matrices. For sparse matrices the operation count is smaller, but an efficient implementation is difficult, in particular for parallel computing with many processors. In iterative methods the factorization step is avoided and replaced by matrix-vector products which easily take advantage of sparsity. It will now be described how sparsity was obtained, how iterations were performed and how nearsightedness was utilized by a simple spatial truncation of the Green function of the system.

3.1. Sparsity

The standard reference system in the KKR method is free space. Here the Green function matrix elements $G^{0,nn'}_{LL'}(E)$,

traditionally called structure constants, are known analytically, but decay unfavourably slowly with distance between site n and n'. Because of the slow decay all elements in the KKR matrix $t^{-1} - G^0$ contribute and iterative solution is too expensive. Since O(N) on-site matrix elements $G^{nn}_{LL'}(E)$ are needed for the density and for each n dense matrix-vector products with O(N^2) operations are used, the overall scaling is O(N^3) with a prefactor given by the number $N_{\rm it}$ of iterations which is considerably larger than the prefactor for direct solution. Thus an iterative solution is advantageous only if a reference system can be found with rapidly decaying structure constants G^r which can be neglected beyond a chosen distance to obtain a sparse matrix $(\Delta t)^{-1} - G^r$ instead of the dense matrix $t^{-1} - G^0$. Such structure constants can be obtained by the concept of screening, originally used for development of the TB-LMTO method [23] and later applied [24] to transform the KKR method [24] into a TB form with exponentially decaying structure constants.

The calculation of these screened structure constants, first determined by fitting [24, 25], is considerably simplified if the screening transformation is understood in a physically transparent way either in terms of a hard sphere solid [26] or a repulsive reference system [6]. A useful reference system with rapidly decaying structure constants in the energy range relevant for density-functional calculations consists of an infinite array of repulsive potentials (here chosen with a constant height of 8 Ryd) which are confined to non-overlapping muffin-tin spheres around the sites $\mathbf{R}^n$. The TB Green function matrix elements (screened structure constants) of the reference system obey the Dyson equation

$$G^{r,nn'}_{LL'}(E) = G^{0,nn'}_{LL'}(E) + \sum_{n''L''L'''} G^{0,nn''}_{LL''}(E) t^{r,n''}_{L''L'''}(E) G^{r,n''n'}_{L'''L'}(E) \qquad (15)$$

which requires linear-scaling effort because it is solved separately for each site n' by use of a finite cluster of $N_{\rm cl}$ repulsive potentials. The restriction to the potentials in the vicinity of $\mathbf{R}^{n'}$ is possible since more distant potentials do not contribute if the rapidly decaying TB-KKR structure constants $G^{r,nn'}_{LL'}(E)$ are neglected for sites n not contained in the cluster. This neglect leads to matrix dimension $(l_{\max}+1)^2 N_{\rm cl}$ in (15), to a sparse TB-KKR matrix with sparsity degree $N_{\rm cl}/N$ and to O($N_{\rm it} N_{\rm cl} N^2$) scaling for the iterative solution of (11).

3.2. Iteration

The inverse X of the TB-KKR matrix $M = (\Delta t)^{-1} - G^r$ satisfies $\Delta t M X = \Delta t$ which, by inserting M, can be rewritten as $X - \Delta t G^r X = \Delta t$. This equation is used here to derive the iteration scheme

$$X^{(i+1)} = \Delta t + \Delta t G^r X^{(i)}. \qquad (16)$$

Here it is important to realize that (16) does not couple different columns of X so that each column can be iterated independently which is suited ideally for massively parallel computing. Unfortunately (16), which corresponds to the Born iteration of scattering theory, often diverges and must be replaced by a more complicated, convergence producing scheme. One such scheme, which was used successfully in the current investigation, is Anderson mixing [27, 28], which prepares the input to iteration $i+1$ by an optimal linear combination of input and output vectors of all previous iterations. A disadvantage of Anderson mixing, which is sometimes also used to accelerate the density-functional self-consistency cycle, is that the required memory increases with iteration number since information of all previous iterations is kept and used. A second scheme which produced convergent iterations for all materials studied so far, that is for Cu, Pd, Ni, Si and GaN, is the quasi-minimal residual (QMR) method [29, 30] in its transpose free version. This method, which has been applied in previous multiple-scattering calculations [16, 18], requires one to store only a few iteration vectors. Therefore it was preferred for the very large systems considered below, although Anderson mixing often needed fewer matrix-vector products. It should be pointed out that because of the singularities of the Green function on the real energy axis the use of complex energies is unavoidable and that the required number $N_{\rm it}$ of iterations increases with decreasing imaginary part of E [18].

3.3. Truncation

The spatial decay of the Green function $G(\mathbf{r}, \mathbf{r}'; E)$ and the spatial decay of the finite-temperature single-particle density matrix

$$\rho(\mathbf{r}, \mathbf{r}') = -\frac{1}{\pi} \operatorname{Im} \int_{-\infty}^{\infty} \mathrm{d}E\, f(E, E_{\rm F}, T) G(\mathbf{r}, \mathbf{r}'; E) \qquad (17)$$

are connected. Evaluation of (17) by complex energy integration as described above shows that the decay of the density matrix is dominated by the decay of the Green function at the first Matsubara energy $E = E_{\rm F} + {\rm i}\pi kT$. Thus, if the temperature is not too small, a neglect of the Green function for large distances $|\mathbf{r} - \mathbf{r}'|$ corresponds to a neglect of the density matrix for similar distances. This similarity suggests that a spatial truncation of the Green function may lead to N scaling. The truncation of the Green function was implemented by dividing the system around each considered atom into a central region C and the rest R. In block notation (16) could then be partitioned into

$$\begin{bmatrix} X^{(i+1)}_{\rm CC} \\ X^{(i+1)}_{\rm RC} \end{bmatrix} = \begin{bmatrix} \Delta t_C \\ 0 \end{bmatrix} + \begin{bmatrix} \Delta t_{\rm C} G^r_{\rm CC} & \Delta t_{\rm C} G^r_{\rm CR} \\ \Delta t_{\rm R} G^r_{\rm RC} & \Delta t_{\rm R} G^r_{\rm RR} \end{bmatrix} \begin{bmatrix} X^{(i)}_{\rm CC} \\ X^{(i)}_{\rm RC} \end{bmatrix} \qquad (18)$$

and into a similar equation for $X_{\rm CR}$ and $X_{\rm RR}$, not needed, however, to calculate the density of the considered atom, for which only the $(l_{\max}+1)^2$ first columns and rows of $X_{\rm CC}$ are required. Note that for each atom the region C was constructed such that the considered atom was situated in the centre of C.

The proposed spatial truncation of the inverse $X = M^{-1}$ of the TB-KKR matrix consists in the neglect of $X_{\rm RC}$ on both sides of (18). Then the equation for $X_{\rm CC}$ simplifies to

$$X^{(i+1)}_{\rm CC} = \Delta t_{\rm C} + \Delta t_{\rm C} G^r_{\rm CC} X^{(i)}_{\rm CC} \qquad (19)$$

with fixed dimensions $(l_{\max}+1)^2 N_{\rm tr}$ determined by the number $N_{\rm tr}$ of atoms in the truncation region C. Compared with (16) the effort to solve (19) is reduced by a factor $N_{\rm tr}/N$. The total effort for all N atoms is then proportional to $N_{\rm it} N_{\rm cl} N_{\rm tr} N$. This increases linearly with N provided that $N_{\rm it}$ approaches a constant value for large N. This condition is satisfied, as will be shown in the next section.

4. Results

To investigate the behaviour of the proposed algorithm, in particular how the total energy $E_{\rm tot}$ converges with the number $N_{\rm it}$ of iterations and how it is affected by the number $N_{\rm tr}$ of atoms in the truncation region, large Cu and Pd supercells were considered as model systems. The accuracy of $E_{\rm tot}$ depends of course also on other numerical parameters, for instance the angular-momentum cut-off $l_{\max}$, the number $N_{\rm cl}$ of repulsive potentials used to calculate the TB-KKR structure constants and the number of sampling points in the Brillouin zone. The results presented below were obtained with the minimal choice $l_{\max} = 2$, $N_{\rm cl} = 13$ and a single point $\mathbf{k} = (1/4, 1/4, 1/4) \times 2\pi/a$ in the irreducible part of the Brillouin zone, where a is the lattice constant of the supercell. This choice, which permitted very large truncation regions, was justified since tests done with $l_{\max} = 3$, $N_{\rm cl} = 55$ and several $\mathbf{k}$ point meshes showed rather similar dependences of $E_{\rm tot}$ on $N_{\rm it}$ and $N_{\rm tr}$.

The supercells were constructed by repeating a simple cubic unit cell with four atoms, arranged in the appropriate face-centred-cubic geometry, 16, 32 or 48 times in all three space directions. The calculation of the self-consistent potential for the supercells was accomplished by the concept of equivalent $\mathbf{k}$ point meshes. As discussed by Chetty *et al* [31], the simple cubic cell with four atoms leads to identical density and total energy results if equivalent $\mathbf{k}$ points are used in the Brillouin zones of this cell and the supercell with many atoms. Thus the self-consistent potential for the supercells could be calculated by use of 816, 5984 and 19600 points in the irreducible part of the Brillouin zone of the simple cubic unit cell. This feature together with the fact that all atoms are equivalent in the model systems represented an enormous simplification for the current investigation. Contrary to realistic systems, where (16) or (19) and the density integral (14) must be calculated for all atoms, here only one atom had to treated so that all the results presented below could be obtained with a desktop computer even for the very large supercells considered.

We first studied how the total energy depends on the tolerance criterion used to stop the QMR iterations and how many QMR iterations are needed. For this study a $16 \times 16 \times 16$ supercell with $N = 16\,384$ atoms was used, the temperature was chosen as $T = 800$ K, which corresponds to an imaginary part $\pi kT = 0.217$ eV of the first Matsubara energy, and the criterion for the relative residual norm $||r||$ was varied between 10^{-3} and 10^{-8}. Table 1 contains the results for the total energy error $\Delta E_{\rm tot}$ and the number $N_{\rm it}$ of iterations needed at the first Matsubara energy where $N_{\rm it}$ has its largest value. The table shows that $\Delta E_{\rm tot}$ decreases rapidly with the specified tolerance threshold and that approximately a constant number of iterations is needed to reduce $||r||$ and $\Delta E_{\rm tot}$ by an order of magnitude. The faster convergence for Pd is probably caused by the hybridization between s and d states near $E_{\rm F}$, which leads to a smaller contribution of longer ranged s states in Pd than in Cu.

Table 1. Total energy error $\Delta E_{\rm tot}$ achieved (in meV per atom) and number of iterations $N_{\rm it}$ needed at the first Matsubara energy $E_{\rm F} + {\rm i}\pi kT$ if the QMR iterations are stopped at a relative residual norm $||r||$. Here $N_{\rm it}$ is obtained by averaging over the nine independently iterated L components.

	Cu		Pd	
$\|\|r\|\|$	$\Delta E_{\rm tot}$	$N_{\rm it}$	$\Delta E_{\rm tot}$	$N_{\rm it}$
10^{-3}	5.3740	403	2.3790	234
10^{-4}	0.3456	528	0.4179	315
10^{-5}	0.0055	670	0.0167	397
10^{-6}	0.0003	814	0.0015	463
10^{-7}	0.0000	946	0.0001	540
10^{-8}	0.0000	1047	0.0000	610

We then studied how $N_{\rm it}$ depends on system size. This is an important question because the computing time is proportional to $N_{\rm it}$ and thus too large values of $N_{\rm it}$ disfavour iterative solutions compared with direct solutions. Smirnov and Johnson [18], who use supercells containing up to 2048 atoms, found that the computing time per atomic site scales as $N^{1+\epsilon}$, which indicates that $N_{\rm it}$ increases as N^{ϵ}, where ϵ depends on the position of the energy mesh point in the complex plane. For the much larger systems investigated here a different behaviour of $N_{\rm it}$ was found. $N_{\rm it}$ does not increase according to a power law, but approaches a constant value for large system size. This is illustrated in figure 1, where values for $N_{\rm it}$ necessary at the first Matsubara energy are shown as a function of the number $N_{\rm tr}$ of atoms in truncation regions that were constructed by using more and more neighbour shells of atoms around the central site with the idea that always one more shell in the close-packed (110) direction should be included. The curves shown in figure 1 were obtained by fitting the calculated values of $N_{\rm it}$ to the function

$$N_{\rm it} = N_{\rm it}^{\infty} \alpha \exp(-\gamma N_{\rm tr}^{1/3}) \tag{20}$$

with three temperature dependent parameters $N_{\rm it}$, α and γ. The results for the fit parameters are given in table 2. Figure 1 shows that the dependence of $N_{\rm it}$ on $N_{\rm tr}$ is rather well described by the exponential behaviour (20) with a constant limit $N_{\rm it}^{\infty}$ for $N_{\rm tr} \to \infty$. This means that for large systems $N_{\rm it}$ practically does not increase with system size and that the computational effort, which is proportional to $N_{\rm it} N_{\rm cl} N_{\rm tr} N$, shows true linear scaling because only the factor N increases with system size.

The most important question in the current investigation was how the spatial truncation of the Green function, which generates the linear-scaling behaviour, affects the accuracy of the total energy. Contrary to figure 1, where only the first Matsubara energy point was considered, total energy calculations require us to use all mesh points in the complex energy plane. To save computing time the total energies were not calculated for the $48 \times 48 \times 48$ supercell used in figure 1, but for the smaller $32 \times 32 \times 32$ supercell with 131 072 atoms.

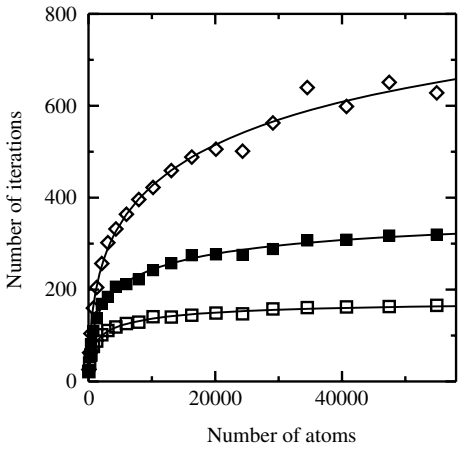

Figure 1. Number $N_{\rm it}$ of QMR iterations as a function of the number $N_{\rm tr}$ of atoms contained in the truncation region in a $48 \times 48 \times 48$ Pd supercell. Solid and open squares are for $T = 800$ and 1600 K, the diamonds for $T = 400$ K. The lines are fitted as described in the text.

Table 2. Parameters determined by fitting the number of iterations required to obtain a relative residual norm $||r|| < 10^{-6}$ at the Matsubara energy $E_{\rm F} + {\rm i}\pi kT$ to a function of the form (20).

	Cu			Pd		
T (K)	$N_{\rm it}^{\infty}$	α	γ	$N_{\rm it}^{\infty}$	α	γ
400	2318	2469	−0.0173	999	1101	−0.0303
800	503	646	−0.0688	368	430	−0.0575
1600	220	314	−0.1095	173	207	−0.0808

The tolerance criterion for the QMR relative residual norm was specified as $||r|| = 10^{-6}$ which, according to table 1, gives well converged total energies. The effect of truncation on the total energy accuracy is shown in figure 2, where the error $\Delta E_{\rm tot}$ is plotted as a function of the number $N_{\rm tr}$ of atoms in the truncation region for $N_{\rm tr}$ values between 767 and 34 521 for Cu and between 1289 and 34 521 for Pd. Results of $\Delta E_{\rm tot}$ for smaller values of $N_{\rm tr}$ did not exhibit a similar clear trend to the one displayed in figure 2. For $N_{\rm tr} = 55$ the calculated total energy error was approximately 70 meV for Cu and Pd. For $N_{\rm tr} = 177$ it was approximately 5 meV for Cu and 20 meV for Pd and for $N_{\rm tr} = 381$ approximately 5 meV for both metals. Figure 2 illustrates that the proposed linear-scaling algorithm can be used for total energy calculations in large metallic systems and that the total energy error can be reduced below 2 meV if truncation regions with a few thousand atoms are used. Figure 2 also shows that higher electronic temperature does not reduce the truncation error substantially except for rather large truncation regions. This is probably a consequence of the fact that the algebraic decay of the Green function (and density matrix) dominates the additional exponential decay caused by temperature up to truncation regions with approximately 20 000 atoms which corresponds to a truncation distance of approximately ten times the face-centred-cubic lattice constant.

It should be mentioned here that for the investigation of the truncation effect it was not necessary to calculate the total energies self-consistently because the modified energy functional [32]

$$\tilde{E}[n(\mathbf{r})] = E[n(\mathbf{r})] - E_{\rm F}\left(\int n(\mathbf{r}) - N_{\rm el}\right) \qquad (21)$$

was used instead of the standard energy functional $E[n(\mathbf{r})]$. The functional (21) is extremal for density variations, even if they do not conserve the total number of electrons $N_{\rm el}$, and has been demonstrated to yield accurate results in impurity calculations [32], where charge neutrality is reached only within 0.1 to 0.01 electrons in embedding regions consisting of one to four shells of neighbouring atoms.

5. Discussion

The results presented above indicate that the proposed algorithm based on the sparsity of the TB-KKR matrix is suitable for large metallic systems. For applications it is

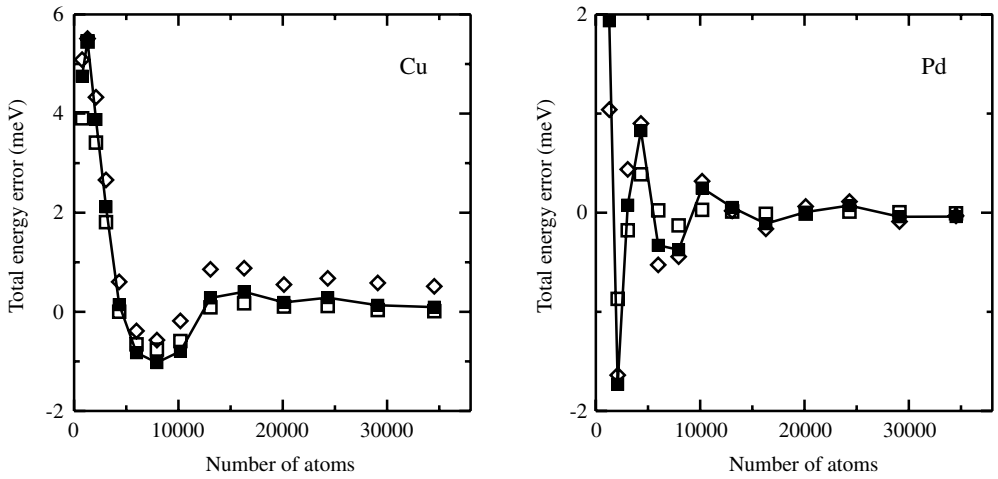

Figure 2. Total energy error per atom as a function of the number of atoms contained in the truncation region. Solid and open squares are for $T = 800$ and 1600 K, diamonds for $T = 400$ K. The lines, which connect the results for $T = 800$ K, serve as a guide for the eye. Note that different energy scales are used for Cu and Pd.

important to know whether it performs better than the direct solution of (11) or (13). This depends on the system size which can be estimated by

$$N_{\mathrm{E}} N^3 = 3 N_{\mathrm{it}} N_{\mathrm{cl}} N_{\mathrm{tr}} N. \tag{22}$$

Here N_{E} is the number of complex energy points used for the density (14) and the factor 3 takes into account that (16) or (19) must be repeated approximately three times for each QMR iteration. For iterative solutions the factor N_{E} can be left off because the main effort arises at the first Matsubara energy. For typical values $N_{\mathrm{E}} = 40$, $N_{\mathrm{cl}} = 50$ and $N_{\mathrm{it}} = 300$ and first considering $N_{\mathrm{tr}} = N$, evaluation of (22) shows that iteration is favoured for more than approximately 1000 atoms. This estimate is probably too high because in the current investigation zero vectors were used to start the QMR iterations instead of physically motivated ones and because preconditioning applied to (16) and (19) can reduce N_{it} although the preconditioners tested so far [18] were not really successful. It can also be expected that an iterative solution would have a better parallel performance because of the less demanding communication requirements. The truncation of the Green function always reduces the computational effort provided that the choice $N_{\mathrm{tr}} < N$ leads to total energy errors which can be tolerated.

The behaviour of the algorithm was illustrated for large Cu and Pd supercells chosen as a model system where only one atom had to be considered in the iterative solution. Thus a desktop computer with 2 Gbyte of memory was sufficient to treat even $48 \times 48 \times 48$ supercells with truncation regions containing up to 54 953 atoms. For realistic systems, however, the computing time increases with the number of inequivalent atoms and massively parallel computing is unavoidable. A high parallel efficiency can be expected because communication among the processors only arises for the exchange of TB-KKR structure constants and t matrices and for the determination of the Hartree potential, Fermi level and possible parameters used in advanced self-consistency mixing schemes. In order to study these issues in a realistic system, work is in progress to reorganize the currently used KKR computer code.

Very likely, other physical properties than the total energy can also be calculated. Since calculation of forces in the KKR method is straightforward [33] and calculated atomic displacements are found in good agreement with extended x-ray absorption fine structure (EXAFS) measurements [33] and *ab initio* pseudopotential molecular dynamics calculations [34], it would be interesting to see how spatial truncation of the Green function affects forces and displacements. The algorithm is also not restricted to occupied states and metallic systems. Since the exponential decay of the TB-KKR structure constants occurs up to approximately 10 to 20 eV above E_{F}, unoccupied states can be treated in this energy range. In semiconductors and insulators, however, in particular with the intention of utilizing the naturally faster spatial decay of the density matrix, an old problem of the KKR Green function method must be solved. The problem is that E_{F} is not found in the gap between valence and conduction states if the usual angular-momentum cut-off is applied to calculate the Green function (6). To circumvent the problem, recently a procedure based on Lloyd's formula has been suggested [35], but it is not yet clear how Lloyd's formula can be used together with the spatial truncation. Finally, it should be pointed out that the proposed algorithm can be used not only for the Schrödinger equation, but very likely also for the Dirac equation because the extension of the TB-KKR method to fully relativistic calculations is quite straightforward [36].

6. Summary

An algorithm based on the TB-KKR Green function method was presented which is useful to solve the Schrödinger equation in the energy range relevant for density-functional calculations for large metallic systems with linear-scaling computational effort. It was shown how linear scaling is obtained by combined use of sparsity of the TB-KKR matrix, iterative solution of the KKR matrix equations and a spatial truncation of the Green function which corresponds to the use of a restricted density matrix range in other linear-scaling methods. For the iterations the quasi-minimal residual (QMR) method in its transpose free form was applied and no convergence problems were found. The algorithm was tested for large Cu and Pd supercells chosen as model systems and it was shown that total energy errors smaller than 2 meV could be obtained if the truncation region contained a few thousand atoms which corresponds to a truncation distance of approximately five times the elemental lattice constant of Cu and Pd. It was demonstrated that finite electronic temperature considerably reduces the required number of iterations whereas it improves the accuracy of total energies appreciably only for rather large truncation regions. It was explained that the iterations treat each atomic site independently of all other sites, which makes the algorithm particularly suitable for massively parallel computing.

References

[1] Soler J M, Artacho E, Gale J D, Garcia A, Junquera J, Ordejon P and Sanchez-Portal D 2002 *J. Phys.: Condens. Matter* **14** 2745

[2] Bowler D R, Miyazaki T and Gillan M J 2002 *J. Phys.: Condens. Matter* **14** 2781

[3] Skylaris C-K, Haynes P D, Mostofi A A and Payne M C 2005 *J. Chem. Phys.* **122** 084119

[4] Ozaki T and Kino H 2005 *Phys. Rev.* B **72** 045121

[5] Goedecker S 1999 *Rev. Mod. Phys.* **71** 1085

[6] Zeller R, Dederichs P H, Újfalussy B, Szunyogh L and Weinberger P 1995 *Phys. Rev.* B **52** 8807

[7] Prodan E and Kohn W 2005 *Proc. Natl. Acad. Sci. USA* **102** 11635

[8] Kohn W 1996 *Phys. Rev. Lett.* **76** 3168

[9] Zeller R and Braspenning P J 1982 *Solid State Commun.* **42** 701

[10] Braspenning P J, Zeller R, Lodder A and Dederichs P H 1984 *Phys. Rev.* B **29** 703

[11] Gunnarsson O, Jepsen O and Andersen O K 1983 *Phys. Rev.* B **27** 7144

[12] Wang Y, Stocks G M, Shelton W A, Nicholson D M, Szotek Z and Temmerman W M 1995 *Phys. Rev. Lett.* **75** 2867

[13] Abrikosov I A, Niklasson A M, Simak S I, Johansson B, Ruban A V and Skriver H L 1996 *Phys. Rev. Lett.* **76** 4203

[14] Abrikosov I A, Simak S I, Johansson B, Ruban A V and Skriver H L 1997 *Phys. Rev.* B **56** 9319
[15] Smirnov A V and Johnson D D 2001 *Phys. Rev.* B **64** 235129
[16] Nachtigal N M, Shelton W A and Stocks G M 1994 *ORNL Technical Report* ORNL/TM-12873
[17] Petit L, Beiden S V, Temmerman W M, Szotek Z, Stocks G M and Gehring G A 2000 *J. Phys.: Condens. Matter* **12** 8439
[18] Smirnov A V and Johnson D D 2002 *Comput. Phys. Commun.* **148** 74
[19] Papanikolaou N, Zeller R and Dederichs P H 2002 *J. Phys.: Condens. Matter* **14** 2799
[20] Zeller R, Deutz J and Dederichs P H 1982 *Solid State Commun.* **44** 993
[21] Mermin N D 1965 *Phys. Rev.* **137** A1441
[22] Wildberger K, Lang P, Zeller R and Dederichs P H 1995 *Phys. Rev.* B **52** 11502
[23] Andersen O K and Jepsen O 1984 *Phys. Rev. Lett.* **53** 2571
[24] Andersen O K, Postnikov A V and Savrasov S Yu 1992 *Application of Multiple Scattering Theory to Materials Science (MRS Symp. Proc. No. 253)* ed W H Butler, P H Dederichs, A Gonis and R L Weaver (Pittsburgh, PA: Materials Research Society) p 37
[25] Szunyogh L, Újfalussy B, Weinberger P and Kollar J 1994 *Phys. Rev.* B **49** 2721
[26] Andersen O K, Jepsen O and Krier G 1994 *Lectures on Methods of Electronic Structure Calculations* ed V Kumar, O K Andersen and A Mookerjee (Singapore: World Scientific)
[27] Anderson D G 1965 *J. Assoc. Comput. Mach.* **12** 547
[28] Eyert V 1996 *J. Comput. Phys.* **124** 271
[29] Freund R W and Nachtigal N M 1991 *Numer. Math.* **60** 315
[30] Freund R W 1993 *SIAM J. Sci. Comput.* **14** 470
[31] Chetty N, Weinert M, Rahman T S and Davenport J W 1995 *Phys. Rev.* B **52** 6313
[32] Drittler B, Weinert M, Zeller R and Dederichs P H 1989 *Phys. Rev.* B **39** 930
[33] Papanikolaou N, Zeller R, Dederichs P H and Stefanou N 1997 *Phys. Rev.* B **55** 4157
[34] Settels A, Schroeder K, Korhonen T, Papanikolaou N, Aretz M, Zeller R and Dederichs P H 2000 *Solid State Commun.* **113** 239
[35] Zeller R 2008 *J. Phys.: Condens. Matter* **20** 035220
[36] Szunyogh L, Újfalussy B, Weinberger P and Kollar J 1994 *J. Phys.: Condens. Matter* **6** 3801