Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 201 (2004) 333–343
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Damped gradient iteration and multigrid relaxation: tools for electronic structure calculations using orbital density-functionals

Stephan Kümmel *

Division Electronic Structure of Finite Systems, Max-Planck-Institut für Physik Komplexer Systeme, Nöthnitzer Strasse 38, D-01187 Dresden, Germany

Received 12 January 2004; received in revised form 24 May 2004; accepted 24 May 2004
Available online 6 July 2004

Abstract

A method for efficient solution of the Kohn–Sham equations, the central equations of density-functional electronic structure calculations, is presented. It combines the simplicity of the damped gradient iteration with the efficiency of multigrid techniques. The long-standing problem of calculating the optimized effective potential of Kohn–Sham theory for orbital density functionals has been solved with this algorithm.
© 2004 Elsevier Inc. All rights reserved.

PACS: 71.15.Mb; 31.15.Ew; 02.70.−c

Keywords: Electronic structure calculations; Density-functional theory; Orbital functionals; Multigrid methods; Finite difference methods; Gradient iteration

1. Kohn–Sham theory with orbital functionals

Density-functional theory (DFT) [1] has become one of the most widely used theories for calculating the electronic structure of molecules, clusters and solids. The centerpiece of Kohn–Sham DFT [2] are the Kohn–Sham equations

$$\left(\hat{h}_{s\sigma} - \varepsilon_{i\sigma}\right)\varphi_{i\sigma}(\mathbf{r}) = 0, \tag{1}$$

where $\hat{h}_{s\sigma} = -(\hbar^2/2m)\nabla^2 + v_{s\sigma}(\mathbf{r})$ is the Kohn–Sham Hamiltonian. The Kohn–Sham potential

* Tel.: +49-351-871-2123; fax: +49-351-871-1999.
E-mail address: kuemmel@mpipks-dresden.mpg.de.

0021-9991/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2004.05.018

$$v_{s\sigma}(\mathbf{r}) = v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc\sigma}(\mathbf{r}), \tag{2}$$

is the sum of the external potential $v(\mathbf{r})$, the Hartree potential $v_H(\mathbf{r}) = \int d^3r' e^2 n(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'|$, and the spin-dependent exchange-correlation potential

$$v_{xc\sigma}(\mathbf{r}) = \delta E_{xc}[n]/\delta n_\sigma(\mathbf{r}), \tag{3}$$

which is rigorously defined as the functional derivative of the exchange and correlation energy functional $E_{xc}[n]$ with respect to the spin density. From the self-consistent solutions of Eq. (1) the ground-state spin densities $n_\sigma(\mathbf{r}) = \sum_{i=1}^{N_\sigma} |\varphi_{i\sigma}(\mathbf{r})|^2$ and total density $n = n_\uparrow + n_\downarrow$ are calculated. The ground-state energy is then obtained from $E[n_\uparrow, n_\downarrow] = T_s[n_\uparrow, n_\downarrow] + E_H[n] + E_{xc}[n_\uparrow, n_\downarrow] + \int n(\mathbf{r}) v(\mathbf{r}) d^3r$.

The success of Kohn–Sham DFT rests on two columns. One column is the "Jacob's ladder" [3] of increasingly sophisticated approximations for $E_{xc}[n]$. While the local density approximation (LDA) [1,2] already lead to reasonable results for many solid-state systems, the advent of the generalized gradient approximations (GGAs) established DFT as a method for calculating the properties of many finite systems that so far had to be treated with wave-function based quantum-chemical methods. Meta-GGAs can achieve yet greater accuracy and the most recent form [4] leads to a uniformly improved description. However, all these semi-local density functionals are not fully self-interaction [1,5] free and the self-interaction error might be the single most important error in present day energy functionals. Curing this problem is therefore an important aspect of density-functional development.

The oldest successful self-interaction correction was the one of Perdew and Zunger [5] (PZ-SIC). More modern approaches achieve freedom from self-interaction by using exact exchange and iso-orbital indicators [6–8] and appear very promising [3,8,9]. However, both approaches have a drawback in practical applications: the resulting exchange-correlation functionals are orbital functionals, i.e., implicitly they are functionals of the density, but explicitly they are known only as functionals of the Kohn–Sham orbitals (see, e.g. [10,11] for detailed discussions). Thus, the exchange-correlation potential cannot be calculated by directly evaluating Eq. (3), but the functional derivative has to be obtained from the optimized effective potential (OEP) integral equation [10,12–14]

$$\sum_{i=1}^{N_\sigma} \int \varphi_{i\sigma}^*(\mathbf{r}') \left[v_{xc\sigma}(\mathbf{r}') - u_{xci\sigma}(\mathbf{r}') \right] \sum_{\substack{j=1 \\ j\neq i}}^{\infty} \frac{\varphi_{j\sigma}(\mathbf{r}')\varphi_{j\sigma}^*(\mathbf{r})}{\varepsilon_{i\sigma} - \varepsilon_{j\sigma}} d^3r' \varphi_{i\sigma}(\mathbf{r}) + \text{c.c.} = 0, \tag{4}$$

where

$$u_{xci\sigma}(\mathbf{r}) = \frac{1}{\varphi_{i\sigma}^*(\mathbf{r})} \frac{\delta E_{xc}[\{\varphi_{j\tau}\}]}{\delta \varphi_{i\sigma}(\mathbf{r})}. \tag{5}$$

Until recently, accurately solving this complicated integral equation for finite systems was only possible for effectively one-dimensional problems (spherical atoms [13,15] and jellium spheres [16]). Three-dimensional solutions based on Gaussian basis-set expansions [17–19] yield correct total energies but are numerically involved and do not produce the correct asymptotic behavior [20] of the potential (see [21] for a discussion of some of the numerical problems). Different approximations to the true OEP have been proposed [22–25]. Their common advantage is computational simplicity, but at the cost of introducing additional approximations beyond the ones already inherent in the energy functional itself.

An iterative approach to solving the OEP equation allows to calculate the *exact* exchange-correlation potential efficiently and accurately for any orbital functional [20]. It is based on the fact that Eq. (4) can be rewritten in the form [10,22]

$$v_{xc\sigma}(\mathbf{r}) = \frac{1}{2n_\sigma(\mathbf{r})} \sum_{i=1}^{N_\sigma} \left\{ |\varphi_{i\sigma}(\mathbf{r})|^2 \left[u_{xci\sigma}(\mathbf{r}) + (\bar{v}_{xci\sigma} - \bar{u}_{xci\sigma}) \right] - \frac{\hbar^2}{m} \nabla \cdot \left(\psi_{i\sigma}^*(\mathbf{r}) \nabla \varphi_{i\sigma}(\mathbf{r}) \right) \right\} + \text{c.c.} \tag{6}$$

Here, the $\psi^*_{i\sigma}(\mathbf{r})$ are the first-order perturbation theory orbital shifts defined by $\psi^*_{i\sigma}(\mathbf{r}) = -\sum_{j=1,\neq i}^{\infty} \langle\varphi_{i\sigma}|u_{\mathrm{xc}i\sigma} - v_{\mathrm{xc}\sigma}|\varphi_{i\sigma}\rangle\varphi^*_{j\sigma}(\mathbf{r})/(\varepsilon_{i\sigma} - \varepsilon_{j\sigma})$ and $\bar{v}_{\mathrm{xc}i\sigma} = \langle\varphi_{i\sigma}|v_{\mathrm{xc}\sigma}|\varphi_{i\sigma}\rangle$, $\bar{u}_{\mathrm{xc}i\sigma} = \langle\varphi_{i\sigma}|u_{\mathrm{cxc}i\sigma}|\varphi_{i\sigma}\rangle$ are orbital averages. The practical usefulness of Eq. (6) stems from the fact that the $\psi^*_i(\mathbf{r})$ can be calculated accurately and efficiently from the partial differential equations that result from first-order perturbation theory

$$(\hat{h}_{s\sigma} - \varepsilon_{i\sigma})\psi^*_{i\sigma}(\mathbf{r}) = -[v_{\mathrm{xc}\sigma}(\mathbf{r}) - u_{\mathrm{xc}i\sigma}(\mathbf{r}) - (\bar{v}_{\mathrm{xc}i\sigma} - \bar{u}_{\mathrm{xc}i\sigma})]\varphi^*_{i\sigma}(\mathbf{r}). \tag{7}$$

(For the sake of notational simplicity, non-degenerate orbitals are assumed. The extension to degenerate orbitals is straightforward [11].)

OEP can be calculated iteratively in the following way: for a given orbital functional $E_{\mathrm{xc}}[\{\varphi_i[n]\}]$, start with a suitable approximate expression for $v_{\mathrm{xc}}(\mathbf{r})$, e.g., the KLI-expression [22] which corresponds to neglecting all terms in Eq. (6) that involve the $\psi^*_{i\sigma}(\mathbf{r})$, and solve Eq. (1) self-consistently within this approximation. With the resulting $\varphi_i(\mathbf{r})$ the right-hand side of Eq. (7) can be evaluated and the equation solved. The thus obtained $\psi^*_{i\sigma}(\mathbf{r})$ can be inserted into Eq. (6) to obtain a new and better approximation to $v_{\mathrm{xc}}(\mathbf{r})$. (This step can also be modified such that dividing by the density, required for evaluating Eq. (6), can be avoided [20].) This procedure is repeated until self-consistency for the exact $v_{\mathrm{xc}}(\mathbf{r})$ (OEP) is reached. Convergence is quickly achieved, see below.

Detailed tests of this method have been done for spherical atoms using the true nuclear Coulomb potential [11,20]. For spherical atoms the OEP equation can also directly be solved [13] and total energies and Kohn–Sham eigenvalues from the iterative approach were compared to the ones obtained from the direct solution algorithm. The results agreed to all significant digits [11,20]. As a further check the exchange virial relation [11,26] has been investigated. It is violated by about 1% by the KLI approximation. The iterative construction reduces the error in the virial relation to 10^{-7}% and the tests thus show that the iterative algorithm indeed leads to the true OEP.

In this way, the problem of solving a very complicated integral equation has been reduced to solving systems of coupled partial differential equations. The combination of new, orbital based self-interaction free energy functionals [3,9] and an efficient algorithm for calculating the corresponding Kohn–Sham potential appears as a promising route to considerably extend the applicability of DFT in electronic structure calculations. For practical applications, one thus ''only'' needs efficient numerical algorithms for solving the underlying partial differential equations.

2. The damped gradient iteration

2.1. Overview

From the computational point of view, the second column on which the success of Kohn–Sham DFT rests are the various numerical techniques which have been developed to solve the Kohn–Sham equations. These can roughly be divided into two groups: methods that expand the relevant quantities in some basis set, typically plane waves or Gaussians, and methods that directly solve Eqs. (1)–(3) in real space. Here, the focus will be on the real space approaches. They are finding increasing interest [27–30] because they avoid the questions of basis-set completeness that can sometimes trouble Gaussian-based calculations for finite systems, they provide accurate and straightforward access to quantities like polarizabilities that would require special basis sets, they can efficiently be parallelized [28] and they are the method of choice for non-linear time-dependent calculations [30,31] that are required in the context of strong-field (e.g., ultra-short laser pulse) physics.

One particular way of solving equations like Eq. (1) is the damped gradient iteration of [32,33]. There, the method has been discussed in the context of nuclear Hartree–Fock calculations whereas here, the systems we have in mind are molecules and clusters described by some given energy density-functional. In

the following it will be discussed how the damped gradient iteration can be combined with multigrid techniques to achieve a straightforward and efficient algorithm for solving the Kohn–Sham and OEP equations. In addition, adaptive stepsize control will be discussed as an extension of the algorithm that increases the convergence rate.

2.2. Basic equations

Obviously, one can find the highest eigenvalue $\varepsilon_{\max}$ of the operator $\hat{h}_s$ (for notational simplicity the spin-index is dropped here) and its eigenfunction $\varphi_{\max}$ by the iteration

$$\varphi_{\max,v+1} = \hat{h}_s \varphi_{\max,v} / (\varepsilon_{\max,v} \sqrt{\langle \varphi_{\max,v} | \varphi_{\max,v} \rangle}),$$

where

$$\varepsilon_{\max,v} = \langle \varphi_{\max,v} | \hat{h}_s | \varphi_{\max,v} \rangle,$$

if the starting guess $\varphi_{\max,0}$ is not orthogonal to the true $\varphi_{\max}$: $\varphi_{\max,0}$ can be expanded in the eigenfunctions of $\hat{h}_s$, each application of $\hat{h}_s$ projects out the eigenvalue of each eigenfunction, and after a couple of iterations, $\varphi_{\max}$ dominates the expansion. Using a similar argument one can show that repeated application of the operator $\hat{o} = 1 - d_S(\hat{h}_s - \varepsilon_1)$ with subsequent normalization leads to the smallest eigenvalue of $\hat{h}_s$ if the real parameter d_S obeys $0 < d_S < 2/(\varepsilon_{\max} - \varepsilon_1)$. (Note that on a finite real space grid there is always a discrete largest eigenvalue since the maximum kinetic energy is limited by the grid spacing.) The latter algorithm is the "simple gradient iteration". It is robust and easy to program but for large scale applications it has the serious drawback of very slow convergence due to the appearance of $\varepsilon_{\max}$ in the denominator of the upper bound for the stepsize d_S. Although $\varepsilon_{\max}$ usually is of no interest in the actual calculation, its mere existence seriously limits the possible stepsize.

The damped gradient iteration remedies this shortcoming by formally replacing the parameter d_S by the operator $\hat{d}_{dg} = d/(\hat{t} + e)$, where d and e are positive real parameters and $\hat{t}$ is the kinetic energy operator. The reasoning for this replacement is that from the two contributions to $\varepsilon_{\max}$, i.e., potential and kinetic energy, the kinetic contribution by far dominates. If it can be factored out, a much larger steplength and thus faster convergence is possible. The stepping operator $\hat{d}$ achieves exactly this. The inverse $\hat{t}$ leads to a steplength that is limited ("damped") by the kinetic energy contributions actually present in the iterated orbital and no longer globally limited by the existence of a large $\varepsilon_{\max}$. The parameter e is added to stabilize the components with small kinetic energy, i.e., it represents the contribution of the potential to $\varepsilon_{\max}$. Consequently, e must be set to a value that roughly corresponds to the depth of the Kohn–Sham potential. Decreasing e leads to faster convergence but also decreases the stability of the iteration. The parameter d is the stepsize of the damped gradient iteration and "summarizes" the effects that ε_1 and d_S had in the simple iteration based on $\hat{o}$. With an additional orthogonalization step, eigenvalues and orbitals above the lowest one can be calculated. The full algorithm for the damped gradient iteration (step number v to $v+1$) thus reads:

$$\varphi_{i,0} = \varphi_{i,\text{initial guess}}, \tag{8}$$

$$\varphi'_{i,v} = \varphi_{i,v} - \frac{d}{\hat{t} + e} \left[\hat{h}_{s,v} - \langle \varphi_{i,v} | \hat{h}_{s,v} | \varphi_{i,v} \rangle \right] \varphi_{i,v}, \tag{9}$$

$$\varphi''_{i,v} = \varphi'_{i,v} - \sum_{j=1}^{i-1} \varphi_{j,v} \langle \varphi_{j,v} | \varphi'_{i,v} \rangle, \tag{10}$$

$$\varphi_{i,v+1} = \frac{\varphi''_{i,v}}{\sqrt{\langle \varphi''_{i,v} | \varphi''_{i,v} \rangle}}, \quad \text{for } i = 1, \ldots, N. \tag{11}$$

In this form one can now easily see why the damped gradient iteration is attractive for self-consistent calculations where only the few lowest occupied orbitals are needed: it does not try to reach self-consistency by repeatedly solving the eigenvalue problem of Eq. (1) (with the potential updated via Eq. (2)) by some diagonalization routine. Instead, the eigenvalue problem is solved iteratively only once and the self-consistency updates are performed along with the iteration that solves the eigenvalue problem. This is indicated in Eq. (9) by the index v on h_s. Thus, the damped gradient iteration minimizes the numerical burden from solving the eigenvalue problem.

2.3. Implementation using multigrids

The central step of the damped gradient iteration is Eq. (9). By introducing an intermediate function $\tilde{\varphi}_{i,v} = \frac{1}{\hat{t}+e}[\hat{h}_{s,v} - \langle \varphi_{i,v} | h_{s,v} | \varphi_{i,v} \rangle] \varphi_{i,v}$ it can be reduced to solving the inhomogeneous linear system

$$(\hat{t} + e)\tilde{\varphi}_{i,v} = (\hat{h}_{s,v} - \varepsilon_{i,v})\varphi_{i,v}, \tag{12}$$

for $\tilde{\varphi}_{i,v}$. (Note that all quantities on the right-hand side are known from the previous iteration step.) Solving this equation fast and efficiently is the objective of a practical implementation. The many advantages of multigrid techniques for solving partial differential equations are well known [27,34] and need not be repeated. The focus thus is on how Eq. (12) can be solved with this technique.

The first thing to be noted is that Eq. (12) looks similar to Poisson's equation of electrostatics: in Rydberg atomic units ($\hbar^2/(2m) = 1$, $e^2 = 2$), $-\hat{t}$ is just the Laplace operator Δ and Eq. (7) can be written in the shorthand notation

$$(\Delta - e)u = F, \tag{13}$$

where $u = -\tilde{\varphi}_{i,v}$ and F is used as an abbreviation for the right-hand side. The set of multigrid routines for solving Eq. (13) can therefore be developed taking into account experience from multigrid solvers [34] for the Poisson equation.

All equations including Eq. (13) are discretized on a three-dimensional cubic Cartesian grid since this allows for a suitable representation of the kinetic energy and all potentials that appear in the pseudopotential-based calculations of the electronic properties of the clusters that we are aiming at. The grid extension depends on the size of the system to be studied and for the clusters of interest here roughly ranges between 30 and 50 Bohr radii (a_0). The grid spacing dx ranges from about $0.25a_0$ for hydrogen with a hard pseudopotential [35] to about $0.7a_0$ for sodium with a soft pseudopotential [36]. Below this finest grid, the usual series of coarser grids is built, each having the same physical extension but twice the grid spacing than the next finer grid. Bilinear interpolation is used for the prolongation operations from the coarse to the fine grids. The restriction from fine to coarse grids is done by filling each point on the coarse grid by averaging over the corresponding point on the fine grid and its next, second next and third next neighbors, i.e., the smallest possible cube around the corresponding point on the fine grid. The points on the cube are weighted according to their spatial distance from the central point [37]. In this way one can go through the full multigrid algorithm, starting on the coarsest grid and working upwards. However, it yet remains to be

specified how the relaxation steps on each grid are to be performed, and this is where an important difference between Poisson's problem and Eq. (13) is encountered.

Due to its excellent smoothing properties and simplicity, Gauss–Seidel relaxation [34] is the method of choice for the relaxation steps if it can be made stable. The second derivatives appearing in Δ are represented by three and five point finite difference formulas on the smallest grids and close to grid boundaries and seven point formulas in all other cases in order to be able to work with a larger $\mathrm{d}x$. For the sake of notational simplicity, however, three point formulas will be used throughout in the following equations. Introducing indices j, l, m with $x = j\,\mathrm{d}x$, $y = l\,\mathrm{d}x$, $z = m\,\mathrm{d}x$ so that a point $u(x, y, z)$ is represented by u_{jlm}, the iteration step from n to $n+1$ for Eq. (13) reads [38]

$$u^{n+1}_{jlm} = u^n_{jlm} + \frac{\mathrm{d}t}{\mathrm{d}x^2}\left(u^n_{j+1lm} + u^n_{j-1lm} + u^n_{jl+1m} + u^n_{jl-1m} + u^n_{jlm+1} + u^n_{jlm-1} - 6u^n_{jlm}\right) - \mathrm{d}t\left(eu^n_{jlm} + F_{jlm}\right). \quad (14)$$

The real parameter $\mathrm{d}t$ appearing here is the relaxation step length. If it is chosen too large, the iteration Eq. (14) diverges. On the other hand it is highly desirable to use the largest possible value for $\mathrm{d}t$ in practical applications since otherwise the convergence is too slow. For the Poisson equation the maximum value is well known, $\mathrm{d}t_{\mathrm{P,max}} = \mathrm{d}x^2/6$. However, the additional term e makes Eq. (13) non-trivially different from Poisson's equation and there is thus no reason for why the Poisson value for $\mathrm{d}t_{\max}$ should still be appropriate. In fact, if it is used in the multigrid scheme for solving Eq. (13), the iteration quickly diverges and one finds that the relaxation steps on the coarsest grids (which are usually the most "well behaved" ones) diverge the fastest.

This behavior can be understood and the maximum value of $\mathrm{d}t$ found by subjecting Eq. (13) to von Neumann stability analysis [34]. The spirit of this analysis is that Eq. (14) is interpreted as the diffusion equation $\partial u/\partial t = (\Delta - e)u - F$ with forward Euler differencing for the time derivative. Iterating Eq. (14) thus corresponds to "relaxing" a starting guess to equilibrium. Substituting a plane wave $u^n_{jlm} = A^n \mathrm{expi}(k_j j\,\mathrm{d}x + k_l l\,\mathrm{d}x + k_m m\,\mathrm{d}x)$ into Eq. (14) will tell us how a general Fourier component will develop "in time" with discrete timesteps $n\,\mathrm{d}t$, i.e., under iteration of the differencing scheme of Eq. (14). The iteration is stable only if $|A| \leqslant 1$ for all k. Evaluating Eq. (14) for the plane wave after some straightforward algebra leads to

$$A = 1 + \mathrm{d}t\left[\frac{2}{\mathrm{d}x^2}\left(\sum_{i=1}^{3}\cos(k_i\,\mathrm{d}x) - 3\right) - e\right]. \quad (15)$$

Substituting this expression into the stability criterion then gives the desired maximum allowed time step

$$\mathrm{d}t \leqslant \frac{2}{\frac{12}{\mathrm{d}x^2} + e}. \quad (16)$$

(The analogous analysis for higher-order differencing schemes leads to $\mathrm{d}t = 2/[272/(45\,\mathrm{d}x^2) + e]$ for seven point finite differences and $\mathrm{d}t = 2/[16/(3\,\mathrm{d}x^2) + e]$ for five points.)

Eq. (16) reduces to the Poisson step for $e = 0$. For non-vanishing e, the maximum allowed step for Eq. (14) is always smaller than the Poisson step. This at first sight looks as if the computational effort for solving Eq. (12) might be considerably larger than the one for obtaining the electrostatic potential. However, in our applications e is usually of order 0.5 Ry. Thus, the e in the denominator matters most when $\mathrm{d}x$ is large, i.e., on the coarse grids with few points. On these grids, the smaller maximum steplength in the relaxation can be compensated by increasing the number of iterations, and the resulting increase in computational cost will be small since there are only few points in the grids. On the other hand, for the fine grid with many points the first term in the denominator dominates and the resulting maximum step is very close to the Poisson step.

In summary, the central equation of the damped gradient iteration can straightforwardly be solved with multigrid relaxation in analogy to the well studied case of Poisson's equation. The only necessary modifications are that the differencing scheme must be extended to include e and the maximum steplength must be chosen according to Eq. (16). The reduction in steplength is noticeable mainly on the coarsest grids and is compensated there by adding a few additional relaxation steps.

2.4. Adapative stepsize control

Thus, everything needed for implementing the algorithm of Eqs. (8)–(11) is at hand. However, a simple modification of the "traditional" algorithm [32,33] is helpful. Instead of using a fixed stepsize d as in [33], the convergence rate can be increased by allowing for a different stepsize for each orbital and adjusting it dynamically during the iteration. We tested different schemes and found that a simple pragmatic approach works well. The iteration is started with the same d for all orbitals, as in [33] (e.g., for the Na clusters with dx about $0.7a_0$, d is about 0.8). After each iteration of Eqs. (8)–(11) it is tested for each orbital whether the variance $[\langle h_{s,v}\varphi_{i,v}|h_{s,v}\varphi_{i,v}\rangle - \langle\varphi_{i,v}|h_{s,v}|\varphi_{i,v}\rangle^2]^{1/2}$ has been reduced by at least a certain threshold value, e.g., 10%. If not, the damped gradient stepsize d_i for this particular orbital is increased by a few percent. If one finds that the variance increases for an orbital, the corresponding d_i is decreased until the variance also starts to decrease and d_i is then kept at the largest stable value.

In this way one can compensate for the fact that the higher lying orbitals converge more slowly due to their greater number of nodes, i.e., higher kinetic energy which appears in the denominator of the stepping operator. As an additional bonus, lengthy experimentation to determine the optimum stepsize is avoided. In the cases studied, this simple measure reduces the number of damped gradient iteration steps necessary to reduce the variance to a certain value by nearly 50% in most cases. The little additional effort for the bookkeeping is thus well justified.

2.5. Employing orbital functionals

Everything said so far is enough to do Kohn–Sham calculations with explicit density functionals like LDA or GGAs. If orbital functionals are to be employed one needs to solve one more partial differential equation, Eq. (7), as discussed in Section 1. Stability analysis analogously to the one above shows that Eq. (7) is unconditionally unstable under Gauss–Seidel relaxation, and this does not come as a surprise since Eq. (7) is singular: if it is written in matrix form, the operator on the left-hand side cannot be inverted. It is obvious that there cannot be a unique solution because to any given solution ψ_i^* one can add an arbitrary multiple of φ_i to obtain another solution. However, the definition of ψ_i^* below Eq. (6) shows that the particular solution of Eq. (7) that we are seeking is the one orthogonal to φ_i. This particular solution can easily be calculated by the well known conjugate gradient method [34]. The conjugate gradient algorithm constructs the solution by adding a sequence of corrections to a starting guess. Basically, the first correction is constructed by letting the left-hand side operator act on the starting guess and subtracting the result from the right-hand side. Following corrections are calculated analogously [34]. It is easily verified that the right-hand side is orthogonal to φ_i, and thus conjugate gradient iteration automatically leads to the desired orthogonal solution. Small spurious non-orthogonal contributions due to numerical inaccuracies can be removed by explicitly orthogonalizing ψ_i^* to φ_i. For computationally demanding applications the conjugate gradient routine can be embedded in another multigrid environment for increased execution speed.

Finally, it is to be noted that the examples given in the following involve a three-dimensional Cartesian grid. On such a grid, the true nuclear potential with its Coulomb singularity cannot be represented accurately. For the present calculations this does not pose a problem since pseudopotentials are used. But it should be stressed that the method works as well when the full nuclear Coulomb potential is used. This has been demonstrated explicitly for atoms in [11,20] and combining our algorithm with the grid-based

all-electron molecular technique described in [10] should also be straightforward. Since the partial differential equations on which the algorithm is based are completely general, there is also no obvious reasons why solving them by some kind of basis-set expansion should be problematic.

3. Application: cluster calculations

3.1. Exact Kohn–Sham exchange for the sodium cluster Na_6

As a first application of the previously discussed algorithm the exact Kohn–Sham exchange potential for sodium clusters was calculated [20]. Since exact exchange is expected to be an important part of the next generation of density functionals because it cancels the Hartree self-interaction, this is an interesting example. The exact exchange energy of Kohn–Sham theory is defined as the Fock-Integral

$$E_{\mathrm{x}}[\{\varphi\}] = -\frac{e^2}{2}\sum_{\substack{i,j=1\\ \sigma=\uparrow,\downarrow}}^{N_\sigma}\int\int \frac{\varphi^*_{i\sigma}(\mathbf{r})\varphi^*_{j\sigma}(\mathbf{r}')\varphi_{j\sigma}(\mathbf{r})\varphi_{i\sigma}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3r'\mathrm{d}^3r, \tag{17}$$

evaluated with Kohn–Sham orbitals, i.e., with the orbitals coming from the local Kohn–Sham potential, *not* the Hartree–Fock orbitals [39]. The external potential is given by the ionic geometry that minimizes the total energy and each ion is described by a soft pseudopotential [36]. The starting guess for the Kohn–Sham orbitals is constructed from the atomic orbitals corresponding to the pseudopotential. For the cluster Na_6 that was chosen as a test case here, the total energy and Kohn–Sham eigenvalues are converged within 10^{-4} Ry (starting from the KLI-approximation) after four iterations of going back and forth between self-consistent solution of the Kohn–Sham equations with fixed v_{x} and updating v_{x} via the orbital-shift Eq. (7). It should be stressed that the Kohn–Sham eigenvalues are obtained stably without any need for a posteriori shifts, e.g., for Na_6 the eigenvalues of the valence electrons are (in Hartree atomic units) $\varepsilon_1^{\mathrm{KLI}} = -0.1871$ and $\varepsilon_2^{\mathrm{KLI}} = \varepsilon_3^{\mathrm{KLI}} = -0.1494$ in the KLI approximation and $\varepsilon_1 = -0.1861$ and $\varepsilon_2 = \varepsilon_3 = -0.1496$ for the OEP.

Fig. 1 shows the resulting $v_{\mathrm{x}}^{\mathrm{exact}}(\mathbf{r})$ for Na_6, plotted once along the x-axis and once along the z-axis. The striking result is that the potential goes to different asymptotic limits in different spatial directions. This

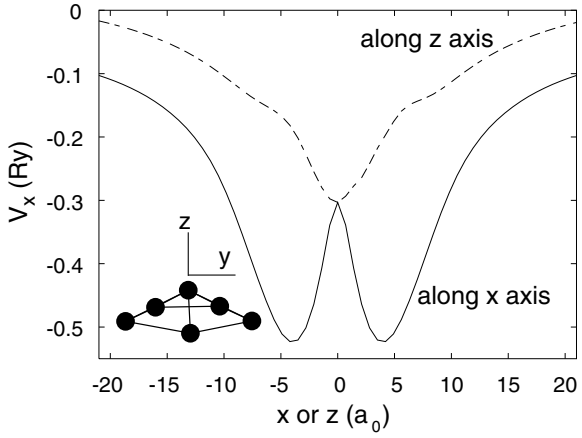

Fig. 1. The exact Kohn–Sham exchange potential for Na_6, calculated with the algorithm of Section 2. The ground-state cluster geometry and labeling of axes is indicated in the inset. Note that the exact exchange potential goes to different asymptotic limits in different spatial directions as a consequence of Eq. (6) [20,23].

behavior has been predicted theoretically and is a consequence of the orbital-average terms that appear in the first line of the right-hand side of Eq. (6) [23,20]: non-vanishing asymptotic constants are encountered whenever the highest occupied orbital has a nodal plane that reaches out to infinity. In Na_6 there are two degenerate highest occupied orbitals, one p_x and one p_y-like. Their nodal planes intersect on the z-axis and thus the exact exchange potential falls of to 0 like $-e^2/r$ everywhere except for on the z-axis. That this only recently discovered [20,23], counter-intuitive property of the exact Kohn–Sham exchange potential is correctly obtained is a reassuring test of our algorithm.

3.2. Chain of hydrogen atoms

Chain molecules (polymers) are a critical testing ground for conventional, semi-local functionals. It has been shown [40] that these functionals fail badly in particular for predicting the non-linear polarizabilities of such systems (they strongly overestimate them) whereas functionals incorporating exact exchange lead to a realistic description. As a paradigm test system, chains of hydrogen atoms have been investigated [40] comparing LDA, Hartree–Fock and Kohn–Sham exchange in the KLI approximation. It was found that compared to LDA, Kohn–Sham exchange in the KLI approximation greatly reduced the errors but did not lead to as realistic results for the second hyperpolarizabilities as Hartree–Fock theory: the KLI hyperpolarizabilities are still somewhat larger than the ones from Hartree–Fock theory. This is a puzzle because one expects exact Kohn–Sham exchange and Hartree–Fock exchange to give very similar answers. One hypothesis to explain the discrepancy is that for molecular chains there is a noticeable difference between the KLI-potential and the exact OEP. However, this to a certain extent is a daring hypothesis because it has been shown for many atoms that the KLI-potential and OEP-potential are very close [22]. In [40] the exact OEP could not be constructed and the hypothesis thus not tested.

Since our present implementation as discussed in Section 2 is based on cubic grids it is well suited to study compact systems like metal or semiconductor clusters and less suited for linear systems. Work is under way [41] to develop a code for systems with one dominant direction of extension and this should allow to calculate hyperpolarizabilities of chain molecules accurately. Nevertheless our present implementation already allows to qualitatively test the above mentioned hypothesis. Fig. 2 shows the exchange potential for a chain of six hydrogen atoms. The interatomic distances were chosen as in [40], i.e., alternating between 2 and $3a_0$. The broken line shows the KLI-potential and the full line the OEP. The general

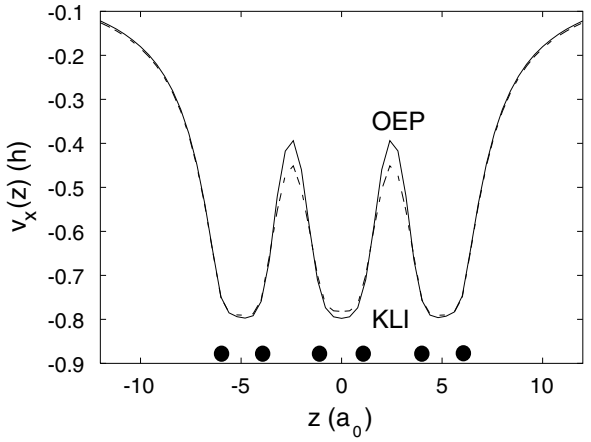

Fig. 2. The exact Kohn–Sham exchange potential (OEP, full line) and the KLI approximation to $v_x(\mathbf{r})$ (dashed line) for a chain of six hydrogen atoms. The circles indicate the positions of the atoms. See text for details.

trends in the potentials are similar but an important difference can be noted: the OEP has more pronounced minima and maxima than the KLI-potential. Thus one can expect that the OEP "traps" the electrons more strongly than the KLI-potential. Consequently, a stronger electrical field will be required to transfer electrons along the chain and it appears very plausible that this will reduce the hyperpolarizabilities.

4. Summary

A method for numerically solving the central equations of Kohn–Sham electronic structure theory was discussed with particular emphasis on orbital functionals. The method combines the transparency of the damped gradient iteration with the efficiency of multigrid techniques. Instead of treating the Kohn–Sham eigenvalue problem and the self-consistency iteration as separate steps and thus having to solve the eigenvalue problem several times (after each self-consistency update), the eigenvalue problem is solved only once. One iteration with interlaced self-consistency updates reaches self-consistency and the solution of the eigenvalue problem at the same time. The method of iteratively solving the OEP equation for orbital functionals via calculation of first-order orbital shifts can easily be combined with this method. In this way, the problem of *directly* solving the OEP integral equation – which proved to be practically impossible in the past – can elegantly be sidestepped and the exact OEP can nevertheless be calculated. As an example the three-dimensional exact exchange potential of the neutral six-atom sodium cluster was calculated and the peculiar property of non-vanishing asymptotic constants in the exact Kohn–Sham exchange potential demonstrated. As a second example, the exchange potential for a chain of hydrogen atoms was calculated and the observed differences between the exact potential and the KLI approximation set in relation to the question of accurately predicting the hyperpolarizabilities of chain molecules.

Acknowledgements

Financial support by the Deutsche Forschungsgemeinschaft under an Emmy Noether grant is gratefully acknowledged.

References

[1] R.M. Dreizler, E.K.U. Gross, Density Functional Theory, Springer, Berlin, 1990.
[2] W. Kohn, L.J. Sham, Phys. Rev. 140 (1965) A1133.
[3] J.P. Perdew, K. Schmidt, in: V. VanDoren, K. VanAlsenoy, P. Geerlings (Eds.), Density Functional Theory and its Applications to Materials, American Institute of Physics, 2001.
[4] J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, Phys. Rev. Lett. 91 (2003) 146401.
[5] J.P. Perdew, A. Zunger, Phys. Rev. B 23 (1981) 5048.
[6] A.D. Becke, Int. J. Quantum Chem. 23 (1985) 585.
[7] J.F. Dobson, J. Phys.: Condens. Matter 4 (1992) 7877.
[8] S. Kümmel, J.P. Perdew, Mol. Phys. 101 (2003) 1363.
[9] J. Tao, J.P. Perdew, S. Kümmel, V.N. Staroverov and G.E. Scuseria, work in progress.
[10] T. Grabo, T. Kreibich, S. Kurth, E.K.U. Gross, in: V. Anisimov (Ed.), Strong Coulomb Correlation in Electronic Structure: Beyond the Local Density Approximation, Gordon and Breach, Tokyo, 2000.
[11] S. Kümmel, J.P. Perdew, Phys. Rev. B 68 (2003) 035103.
[12] R.T. Sharp, G.K. Horton, Phys. Rev. 90 (1953) 317.
[13] J.D. Talman, W.F. Shadwick, Phys. Rev. A 14 (1976) 36.
[14] V. Sahni, J. Gruenebaum, J.P. Perdew, Phys. Rev. B 26 (1982) 4371.
[15] E. Engel, S.H. Vosko, Phys. Rev. A 47 (1993) 2800.
[16] E. Engel, S.H. Vosko, Phys. Rev. B 50 (1994) 10498.

[17] A. Görling, Phys. Rev. Lett. 83 (1999) 5459.
[18] S. Ivanov, S. Hirata, R.J. Bartlett, Phys. Rev. Lett. 83 (1999) 5455.
[19] W. Yang, Q. Wu, Phys. Rev. Lett. 89 (2002) 143002.
[20] S. Kümmel, J.P. Perdew, Phys. Rev. Lett. 90 (2003) 043004.
[21] S. Hirata, S. Ivanov, I. Grabowski, R. Bartlett, K. Burke, J. Talman, J. Chem. Phys. 115 (2001) 1635.
[22] J.B. Krieger, Y. Li, G.J. Iafrate, Phys. Rev. A 46 (1992) 5453.
[23] F. Della Sala, A. Görling, Phys. Rev. Lett. 89 (2002) 33003.
[24] O.V. Gritsenko, E.J. Baerends, Phys. Rev. A 64 (2001) 042506.
[25] C. Legrand, E. Suraud, P.-G. Reinhard, J. Phys. B 35 (2002) 1.
[26] M. Levy, J.P. Perdew, Phys. Rev. A 32 (1985) 2010.
[27] T.L. Beck, Rev. Mod. Phys. 72 (2001) 1041.
[28] A. Stathopoulos, S. Öğüt, Y. Saad, J.R. Chelikowsky, H. Kim, Comput. Sci. Eng. 2 (2000) 19.
[29] L. Kronik, I. Vasiliev, M. Jain, J.R. Chelikowsky, J. Chem. Phys. 115 (2001) 4322.
[30] C. Calvayrac, P.-G. Reinhard, E. Suraud, C.A. Ullrich, Phys. Rep. 337 (2000) 493.
[31] M.A.L. Marques, A. Castro, G.F. Bertsch, A. Rubio, Comput. Phys. Commun. 151 (2003) 60.
[32] P.-G. Reinhard, R.Y. Cusson, Nucl. Phys. A 378 (1982) 418.
[33] V. Blum, G. Lauritsch, J.A. Maruhn, P.-G. Reinhard, J. Comp. Phys. 100 (1992) 364.
[34] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in FORTRAN, Cambridge University Press, Cambridge, 1992.
[35] F. Gygi, Phys. Rev. B 48 (1993) 11692.
[36] S. Kümmel, M. Brack, P.-G. Reinhard, Phys. Rev. B 62 (2000) 7602;
Phys. Rev. B 63 (E) (2001) 129902.
[37] Together with the condition that the weights must sum up to 1, this leads to the weights 0.094768389, 0.047384195, 0.033505685, 0.027357277 for central, next, second next and third next neighbor. This simple restriction scheme worked satisfactorily in the cases we studied.
[38] Strictly speaking, the following equation denotes the Jacobi step. For the Gauss–Seidel step it is to be understood that the updated values of u can be used on the right hand side as soon as they become available.
[39] It is important to realize that the exact Kohn–Sham exchange energy and the Hartree–Fock exchange energy do not correspond to the same energy functional. For a recent discussion of this topic with many references to important earlier literature, see S. Ivanov, M. Levy, J. Chem. Phys. 119 (2003) 7087.
[40] S.J.A. van Gisbergen, P.R.T. Schipper, O.V. Gritsenko, E.J. Baerends, J.G. Snijders, B. Champagne, B. Kirtman, Phys. Rev. Lett. 83 (1999) 694;
M. Grüning, O.V. Gritsenko, E.J. Baerends, J. Chem. Phys. 116 (2002) 6435.
[41] S. Kümmel, L. Kronik, J.P. Perdew (work in progress).