Elimination of the linearization error in *GW* calculations based on the linearized augmented-plane-wave method

Christoph Friedrich,* Arno Schindlmayr, and Stefan Blügel
Institut für Festkörperforschung, Forschungszentrum Jülich, 52425 Jülich, Germany

Takao Kotani
Department of Chemical and Materials Engineering, Arizona State University, Tempe, Arizona 85287-6006, USA
(Received 7 February 2006; revised manuscript received 24 May 2006; published 7 July 2006)

This paper investigates the influence of the basis set on the *GW* self-energy correction in the full-potential linearized augmented-plane-wave (LAPW) approach and similar linearized all-electron methods. A systematic improvement is achieved by including local orbitals that are defined as second and higher energy derivatives of solutions to the radial scalar-relativistic Dirac equation and thus constitute a natural extension of the LAPW basis set. Within this approach linearization errors can be eliminated, and the basis set becomes complete. While the exchange contribution to the self-energy is little affected by the increased basis-set flexibility, the correlation contribution benefits from the better description of the unoccupied states, as do the quasiparticle energies. The resulting band gaps remain relatively unaffected, however; for Si we find an increase of 0.03 eV.

DOI: 10.1103/PhysRevB.74.045104 PACS number(s): 71.15.Qe, 71.45.Gm, 71.20.Mq

I. INTRODUCTION

Electronic excitation energies may be obtained from the solution of the quasiparticle equation of many-body perturbation theory. This equation contains a nonlocal and frequency-dependent operator, the self-energy $\Sigma^{xc}(\mathbf{r},\mathbf{r}';\epsilon)$, which, in principle, incorporates all electronic exchange and correlation effects. As it cannot be treated exactly for real systems, practical implementations typically use the *GW* approximation,[1] which has become increasingly popular for electronic-structure calculations of excited states in recent years and yields band structures in good quantitative agreement with experimental spectroscopy for a wide range of materials.

Due to its technical simplicity, the first implementations were based on the pseudopotential plane-wave approach. In spite of several approximations in the numerical treatment, which were necessary because of the lack of computer power in the 1980s, initial results were very promising. Hybertsen and Louie[2] as well as Godby *et al.*[3] showed that the calculated band gap of Si fell within a margin of about 0.1 eV from the experimental value. Shortly afterwards the same authors reported band gaps for several other semiconducting materials that turned out to be equally accurate.[4,5] After these pioneering studies the *GW* approximation was applied to a variety of semiconductors, insulators, and metals with great success.[6]

So far, most codes still rely on the pseudopotential approximation, which restricts the range of materials that can be examined. Transition-metal compounds and oxides, in particular, cannot be treated efficiently in this approach. Two early all-electron calculations using the *GW* approximation were done by Hamada *et al.*[7] for Si and by Aryasetiawan[8] for Ni, both within the linearized augmented-plane-wave (LAPW) method. However, only very recently were further full-potential implementations reported, based on the LAPW (Refs. 9 and 10) the linearized muffin-tin orbital (LMTO) (Refs. 11 and 12), the projector-augmented-wave (PAW) (Refs. 13 and 14) and the Korringa-Kohn-Rostoker[15] method together with applications to a larger variety of systems.

Compared to earlier pseudopotential results, it was found, however, that LAPW, LMTO, and PAW calculations appeared to yield systematically smaller band gaps for semiconductors and, in many cases, a worse agreement with experiment.[9,11–14] Ku and Eguiluz[9] hence argued that, in contrast to all-electron methods, the pseudopotential calculations benefited from a fortuitious error cancellation between the pseudopotential approximation and the neglect of vertex corrections in the *GW* approximation. Although their calculations were subsequently criticized for not being converged with respect to the number of bands,[16,17] other all-electron calculations showed a similar underestimation of the band gap.[11–14] In a different attempt to make one step towards an all-electron treatment, Tiago *et al.*[16] relaxed the pseudopotential approximation by constructing a pseudopotential only for the 1*s* state of Si while treating the 2*s* and 2*p* states as valence. Surprisingly, the resulting band gaps did not deviate substantially from the previous pseudopotential results, which made the conjecture of Ku and Eguiluz doubtful. In order to resolve this conflict, it is imperative to carefully analyze and compare the numerical approximations made in the two approaches.

Of course, deviations are expected for several reasons. First, the pseudized wave functions differ from their true counterparts and modify the matrix elements of the self-energy. Second, the core electrons are not included in the construction of the nonlocal self-energy if pseudopotentials are used, which may lead to errors in the core-valence exchange contribution. On the other hand, the single-particle wave functions (and the corresponding energies) of both pseudopotential and linearized all-electron approaches, such as LAPW or LMTO, become more and more inaccurate at higher energies. The LAPW basis set, on which we concentrate in the following, is defined by an expansion around fixed energy parameters that yield accurate wave functions only in their neighborhood, i.e., the valence band. Of course,

this does not affect calculations within density-functional theory[18] (DFT), which only makes use of the occupied states. The *GW* self-energy, however, depends on the unoccupied states up to high energies through the Green function G as well as the screened Coulomb interaction W. The inappropriate description of these states might, therefore, cause errors in the self-energy correction.

The construction of the pseudopotentials guarantees an accurate wave function and energy only for the ground state of a given angular momentum but leads to deviations for higher-lying states. While the deficiency in the pseudopotential approach is inherent in the pseudopotential construction, in the LAPW method it must be attributed to the inadequacy of the basis set for high-lying unoccupied states and can be overcome by increasing the basis-set flexibility. It is the purpose of this paper to elucidate the influence of the basis set on the *GW* results by systematically extending it towards basis-set completeness. We achieve this by adding local orbitals defined as second and higher energy derivatives of solutions of the radial scalar-relativistic Dirac equation. There are several alternative approaches: Bross and Fehrenbach[19] use spline functions to augment the basis set, Krasovskii *et al.*[20] employ conventional local orbitals located in the conduction bands together with their energy derivatives. The advantage of the present approach is that no special consideration about the energy parameters is needed. Furthermore, with the order of the derivatives in each angular-momentum channel, it contains well defined convergence parameters that allow a systematic attainment of basis-set completeness.

This paper is organized as follows. In Sec. II we provide a brief introduction to the *GW* and LAPW methods. In Sec. III we describe our extension of the LAPW basis set in detail. As an illustration, in Sec. IV the effect of the basis-set extension on the *GW* results for Si is discussed. We show that the extension of the basis set and the convergence with respect to the number of bands both reduce the discrepancy with the pseudopotential and the experimental band gap. Unless stated otherwise, we use Hartree atomic units.

II. METHODS

A. *GW* approximation

Within many-body perturbation theory, the quasiparticle wave functions $\psi_{n\mathbf{k}\sigma}(\mathbf{r})$ and energies $\epsilon_{n\mathbf{k}\sigma}$ are obtained from the solution of the quasiparticle equation

$$\left(-\frac{1}{2}\nabla^2 + V^{\mathrm{ext}}(\mathbf{r}) + V^{\mathrm{H}}(\mathbf{r})\right)\psi_{n\mathbf{k}\sigma}(\mathbf{r}) + \int \Sigma_\sigma^{\mathrm{xc}}(\mathbf{r},\mathbf{r}';\epsilon_{n\mathbf{k}\sigma})\psi_{n\mathbf{k}\sigma}(\mathbf{r}')d^3r' = \epsilon_{n\mathbf{k}\sigma}\psi_{n\mathbf{k}\sigma}(\mathbf{r}), \quad (1)$$

where $V^{\mathrm{ext}}(\mathbf{r})$, $V^{\mathrm{H}}(\mathbf{r})$, and $\Sigma_\sigma^{\mathrm{xc}}(\mathbf{r},\mathbf{r}';\epsilon_{n\mathbf{k}\sigma})$ are the external potential created by the crystal field and other applied static fields, the Hartree potential, and the exchange-correlation self-energy, respectively. The quantum numbers n, $\mathbf{k}$, and σ signify the band, wave vector, and spin. Like the majority of existing implementations, our code exploits the formal similarity to the Kohn-Sham (KS) equation to obtain approximate energies within first-order perturbation theory

$$\epsilon_{n\mathbf{k}\sigma} \approx \epsilon_{n\mathbf{k}\sigma}^{\mathrm{KS}} + Z_{n\mathbf{k}\sigma}\langle\varphi_{n\mathbf{k}\sigma}|\Sigma_\sigma^{\mathrm{xc}}(\epsilon_{n\mathbf{k}\sigma}^{\mathrm{KS}}) - V_\sigma^{\mathrm{xc}}|\varphi_{n\mathbf{k}\sigma}\rangle, \quad (2)$$

where $V_\sigma^{\mathrm{xc}}(\mathbf{r})$ is the local exchange-correlation potential, $\varphi_{n\mathbf{k}\sigma}(\mathbf{r})$ the Kohn-Sham wave function, and the quasiparticle renormalization factor is given by

$$Z_{n\mathbf{k}\sigma} = \left(1 - \langle\varphi_{n\mathbf{k}\sigma}|\frac{\partial\Sigma_\sigma^{\mathrm{xc}}}{\partial\epsilon}(\epsilon_{n\mathbf{k}\sigma}^{\mathrm{KS}})|\varphi_{n\mathbf{k}\sigma}\rangle\right)^{-1} \leqslant 1. \quad (3)$$

In the following we will always use the shorthand notation $\langle\Sigma^{\mathrm{xc}}\rangle = \langle\varphi_{n\mathbf{k}\sigma}|\Sigma_\sigma^{\mathrm{xc}}(\epsilon_{n\mathbf{k}\sigma}^{\mathrm{KS}})|\varphi_{n\mathbf{k}\sigma}\rangle$.

We employ the *GW* approximation for the self-energy, symbolically written as $\Sigma^{\mathrm{xc}} = iGW$, where G is the Kohn-Sham Green function

$$G_\sigma(\mathbf{r},\mathbf{r}';\epsilon) = \sum_{n,\mathbf{k}}^{\mathrm{occ}} \frac{\varphi_{n\mathbf{k}\sigma}(\mathbf{r})\varphi_{n\mathbf{k}\sigma}^*(\mathbf{r}')}{\epsilon - \epsilon_{n\mathbf{k}\sigma} - i\delta} + \sum_{n,\mathbf{k}}^{\mathrm{unocc}} \frac{\varphi_{n\mathbf{k}\sigma}(\mathbf{r})\varphi_{n\mathbf{k}\sigma}^*(\mathbf{r}')}{\epsilon - \epsilon_{n\mathbf{k}\sigma} + i\delta} \quad (4)$$

(δ is an infinitesimal positive number) and W the dynamically screened Coulomb interaction. The latter is calculated from $W = v + vPW$, where v is the bare Coulomb potential and P the polarization function

$$P_\sigma(\mathbf{r},\mathbf{r}';\epsilon) = \sum_{n,\mathbf{k}}^{\mathrm{occ}}\sum_{n',\mathbf{k}'}^{\mathrm{unocc}} \frac{2(\epsilon_{n'\mathbf{k}'\sigma} - \epsilon_{n\mathbf{k}\sigma} - i\delta)}{\epsilon^2 - (\epsilon_{n'\mathbf{k}'\sigma} - \epsilon_{n\mathbf{k}\sigma} - i\delta)^2} \times \varphi_{n\mathbf{k}\sigma}^*(\mathbf{r})\varphi_{n'\mathbf{k}'\sigma}(\mathbf{r})\varphi_{n'\mathbf{k}'\sigma}^*(\mathbf{r}')\varphi_{n\mathbf{k}\sigma}(\mathbf{r}') \quad (5)$$

in the random-phase approximation. In practice, the self-energy is decomposed into exchange and correlation contributions

$$\Sigma^{\mathrm{xc}} = \Sigma^{\mathrm{x}} + \Sigma^{\mathrm{c}} = iGv + iG(W - v). \quad (6)$$

The exchange contribution only depends on occupied states, whereas unoccupied states up to high energies, typically 100–200 eV above the Fermi energy, are needed for an accurate evaluation of the correlation contribution.

B. LAPW basis set

The all-electron APW method[21] as well as the related LAPW method[22] rely on a decomposition of space into muffin-tin (MT) spheres, centered at the atomic nuclei, and the interstitial region. The core-electron wave functions, which are (mostly) confined to the muffin-tin spheres, are directly obtained from a solution of the fully relativistic Dirac equation. Here only the spherical part of the effective potential is retained. For the valence electrons a basis set $\{\phi_{\mathbf{k}+\mathbf{G}}^\sigma(\mathbf{r})\}$ is constructed. Its basis functions, the so-called augmented plane waves, are defined everywhere in space. The smoothness of the potential in the interstitial region motivates the use of plane waves with $|\mathbf{k}+\mathbf{G}| \leqslant K_{\max}$, where $K_{\max}$ is a convergence parameter. In the MT spheres, on the other hand, the potential is peaked at the nuclei and predominantly spherical. In the APW method, which is usually implemented in combination with a shape approximation of

spherically symmetric potentials inside the MT spheres, one uses numerical solutions $u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r)$ of the radial Schrödinger equation

$$\hat{h}_{l\sigma} r u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r) = \epsilon_{l\sigma} r u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r) \tag{7}$$

with the Kohn-Sham Hamiltonian

$$\hat{h}_{l\sigma} = -\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{l(l+1)}{2r^2} + V_{\sigma}^{\mathrm{eff}}(r) \tag{8}$$

and the relevant effective potential $V_{\sigma}^{\mathrm{eff}}(r)$. For simplicity we give the nonrelativistic equations here; the scalar-relativistic versions are shown in the Appendix. The atom index is suppressed throughout. Augmenting the interstitial plane waves with linear combinations of the $u_{lm\sigma}^{(0)}(\epsilon_{l\sigma},\mathbf{r}) = u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r)Y_{lm}(\hat{\mathbf{r}})$, taking into account the continuity at the MT sphere boundaries, yields the basis functions $\phi_{\mathbf{k+G}}^{\sigma}(\mathbf{r})$. In the APW method the energy parameters $\epsilon_{l\sigma}$ are identical to the band energies $\epsilon_{n\mathbf{k}\sigma}$. For a spherically symmetric MT potential this condition guarantees that the Kohn-Sham wave functions $\varphi_{n\mathbf{k}\sigma}(\mathbf{r})$ can be obtained exactly (subject to convergence with the l cutoff and $K_{\max}$), but it leads to an expensive nonlinear eigenvalue problem. In the LAPW method this difficulty is circumvented by fixing the $\epsilon_{l\sigma}$ at suitable energies in the valence-band region, thereby avoiding the expensive self-consistency condition with respect to the band energies and making a linear solution of the Kohn-Sham eigenvalue equation possible. Apart from solutions $u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r)$ of the radial Schrödinger equation (7), the energy derivatives $u_{l\sigma}^{(1)}(\epsilon_{l\sigma},r) = \partial u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r)/\partial\epsilon$ are also employed to increase the basis-set flexibility in the MT spheres. They are obtained from

$$\hat{h}_{l\sigma} r u_{l\sigma}^{(1)}(\epsilon_{l\sigma},r) = \epsilon_{l\sigma} r u_{l\sigma}^{(1)}(\epsilon_{l\sigma},r) + r u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r), \tag{9a}$$

$$\frac{d}{d\epsilon}\langle u_{l\sigma}^{(0)}(\epsilon_{l\sigma}) | u_{l\sigma}^{(0)}(\epsilon_{l\sigma})\rangle = 2\langle u_{l\sigma}^{(0)}(\epsilon_{l\sigma}) | u_{l\sigma}^{(1)}(\epsilon_{l\sigma})\rangle = 0. \tag{9b}$$

The two sets of radial functions are combined with the interstitial plane waves in such a way that not only the basis functions but also their radial derivatives are continuous at the MT sphere boundaries. This procedure yields the LAPW basis set for the valence electrons

$$\phi_{\mathbf{k+G}}^{\sigma}(\mathbf{r}) = \begin{cases} \dfrac{1}{\sqrt{\Omega}} e^{i(\mathbf{k+G})\cdot\mathbf{r}}, & \text{if } \mathbf{r} \notin \mathrm{MT}, \\ \displaystyle\sum_{l,m}\sum_{\nu=0}^{1} a_{lm\sigma\nu}^{\mathbf{k+G}} u_{lm\sigma}^{(\nu)}(\epsilon_{l\sigma},\mathbf{r}), & \text{if } \mathbf{r} \in \mathrm{MT}, \end{cases} \tag{10}$$

where the coefficients $a_{lm\sigma\nu}^{\mathbf{k+G}}$ are uniquely determined by the matching conditions and Ω is the unit-cell volume. The inclusion of the energy derivatives implies a linear approximation for the radial functions

$$u_{l\sigma}(\epsilon_{n\mathbf{k}\sigma},r) \approx u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r) + (\epsilon_{n\mathbf{k}\sigma} - \epsilon_{l\sigma}) u_{l\sigma}^{(1)}(\epsilon_{l\sigma},r), \tag{11}$$

thus introducing a linearization error, which grows with the deviation of the band energy $\epsilon_{n\mathbf{k}\sigma}$ from the parameters $\epsilon_{l\sigma}$ and becomes especially relevant for high-lying states. In full-potential LAPW implementations the radial functions are still derived from the spherical part of the MT potential. Therefore, the actual wave functions cannot be constructed from the $u_{lm\sigma}^{(0)}(\epsilon_{l\sigma},\mathbf{r})$ alone, but as the nonspherical modulation is typically small, it is commonly accepted that the energy derivatives add enough basis-set flexibility to describe the valence electrons in the full potential accurately.[23]

In order to quantify the error incurred by the approximation (11), we compare the Kohn-Sham eigenvalue spectra with and without the linearization of the radial functions. In the second case we determine the band energies $\epsilon_{n\mathbf{k}\sigma}$ iteratively by setting the energy parameters $\epsilon_{l\sigma}$ equal to the $\epsilon_{n\mathbf{k}\sigma}$ until self-consistency is reached. For a spherically symmetric MT potential this procedure is equivalent to the APW method, because the contribution of the energy derivatives in our scheme vanishes at the self-consistency point $\epsilon_{l\sigma}=\epsilon_{n\mathbf{k}\sigma}$, and the description of the wave functions is thus identical: linear combinations of $u_{lm\sigma}^{(0)}(\epsilon_{l\sigma},\mathbf{r})$ inside the MT spheres and plane waves in the interstitial region. It should be noted that although the APW basis functions exhibit a derivative discontinuity at the MT sphere boundaries, the wave functions themselves are smooth. Furthermore, the increased basis-set flexibility achieved by the additional radial functions makes it possible to apply our method to the full crystal potential. In this case the energy derivatives yield a nonvanishing but small contribution that reflects the nonspherical modulation.

In Fig. 1(a) we compare the resulting Kohn-Sham band structure with that from a standard full-potential LAPW calculation for the prototype semiconductor Si. Both band structures are evaluated within the local-density approximation for the same effective potential, which is obtained from the LAPW self-consistency loop. The calculation is carried out with the experimental lattice constant of 10.26 Bohr, a muffin-tin radius of 2.16 Bohr, $K_{\max}=4.0$ Ha, an angular-momentum cutoff of 8, and 512 $\mathbf{k}$ points in the full Brillouin zone. We use the FLEUR code.[24] In Fig. 1(b) we make a similar comparison with the pseudopotential plane-wave method, which is the most frequently used approach in the context of the *GW* approximation. For this calculation we use a standard norm-conserving Hamann pseudopotential[25] and an energy cutoff of 9 Ha for the plane waves, the other parameters are the same as in LAPW. In both cases deviations start to occur at around 20 eV above the Fermi energy, i.e., around the twentieth band. This is far below the number of bands normally included in a *GW* calculation to ensure convergence of the self-energy.

In the next section we discuss our extension of the LAPW basis set, which allows a systematic reduction and ultimate elimination of the linearization error. In the basis-set limit the basis becomes complete and all states, including high-lying conduction bands, are accurately described.

III. THE LAPW BASIS-SET EXTENSION

In order to enhance the basis-set flexibility, it is not sufficient to simply increase the plane-wave cutoff $K_{\max}$, as is the case in pseudopotential calculations, since this only improves the basis set in the interstitial region but not within the muffin-tin spheres. In fact, Krasovskii[26] has shown that a

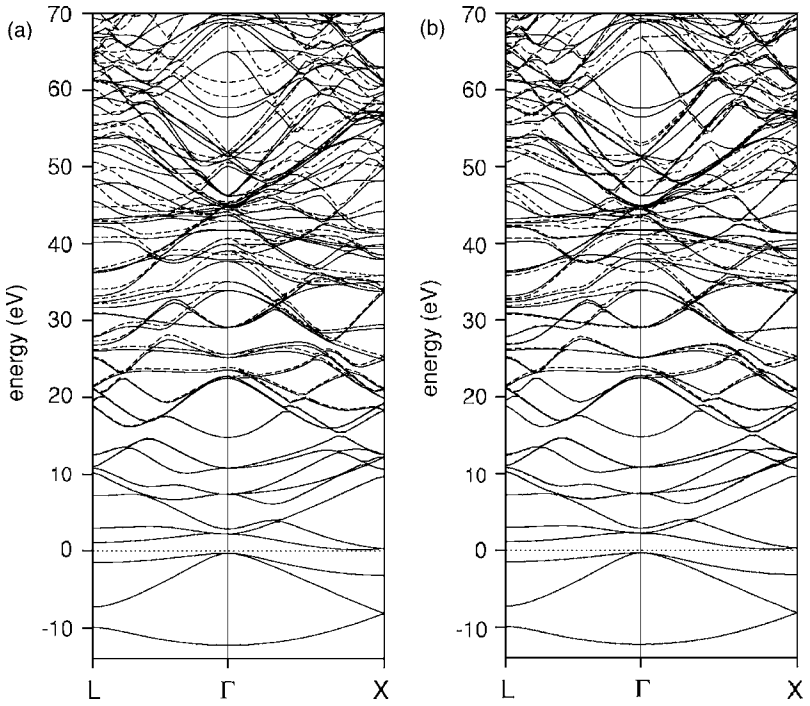

FIG. 1. Comparison of (a) full-potential LAPW and (b) pseudopotential band structures for silicon in the local-density approximation (dashed lines) with results from our nonlinear scheme in which the energy parameters are determined self-consistently (solid lines). In both cases deviations start to appear at around 20 eV above the Fermi energy (0 eV).

fairly small $K_{\max}$ is sufficient to provide enough flexibility in the interstitial region, while the MT part of the basis set quickly deteriorates the more the wave-function energies deviate from the parameters $\epsilon_{l\sigma}$. This inadequacy cannot be overcome with a higher $K_{\max}$.

In a straightforward extension of the LAPW approach one can include second energy derivatives in the Taylor expansion (11). In this case one would need a third matching condition, though, in order to uniquely define the augmented plane waves (10), e.g., the continuity of the second radial derivative at the MT sphere boundaries. These more stringent conditions are known to lead to a less efficient basis set, i.e., one with a slower convergence with respect to $K_{\max}$, however.[27] To circumvent this problem we introduce the higher energy derivatives as additional basis functions in the form of local orbitals,[27] which vanish outside and on the MT sphere boundaries and, therefore, do not need to be matched to plane waves. In general, the radial part of a local orbital is constructed as a linear combination of three radial functions with the conditions of vanishing value and derivative at the MT sphere boundary as well as normalization. In our approach these three functions are $u_{l\sigma}^{(0)}(\epsilon_{l\sigma},r)$, $u_{l\sigma}^{(1)}(\epsilon_{l\sigma},r)$, and one of $u_l^{(\nu)}(\epsilon_{l\sigma},r)$ with $\nu \geqslant 2$. The latter is the νth energy derivative, which can be obtained in analogy to Eq. (9) from

$$\hat{h}_{l\sigma} r u_{l\sigma}^{(\nu)}(\epsilon_{l\sigma},r) = \epsilon_{l\sigma} r u_{l\sigma}^{(\nu)}(\epsilon_{l\sigma},r) + \nu r u_{l\sigma}^{(\nu-1)}(\epsilon_{l\sigma},r), \quad (12a)$$

$$\frac{d^\nu}{d\epsilon^\nu}\langle u_{l\sigma}^{(0)}(\epsilon_{l\sigma})|u_{l\sigma}^{(0)}(\epsilon_{l\sigma})\rangle = 0 \quad (12b)$$

(see the Appendix for the scalar-relativistic treatment).

We now prove that the functions $u^{(\nu)}(\epsilon,r)$ are (i) linearly independent of each other and (ii) orthogonal to solutions $u^{(0)}(\epsilon_0,r)$ of Eq. (7) with $\epsilon_0 \neq \epsilon$ that vanish outside and on the MT sphere boundary. For simplicity the indices l and σ are omitted here.

The first statement guarantees that each derivative increases the variational freedom and does not lead to an overcomplete basis set. For $u^{(0)}(\epsilon,r)$ and $u^{(1)}(\epsilon,r)$ this follows trivially from condition (9b). Now assume that $u^{(\nu)}(\epsilon,r)$ for some $\nu \geqslant 2$ can be written as a linear combination

$$u^{(\nu)}(\epsilon,r) = \sum_{\mu=0}^{\nu-1} c_\mu u^{(\mu)}(\epsilon,r) \quad (13)$$

of the linearly independent functions $u^{(\mu)}(\epsilon,r)$ with $0 \leqslant \mu \leqslant \nu-1$. Then the application of Eqs. (7), (9), and (12) leads to

$$\nu r u^{(\nu-1)}(\epsilon,r) = (\hat{h}-\epsilon) r u^{(\nu)}(\epsilon,r) = \sum_{\mu=0}^{\nu-2} \mu c_{\mu+1} r u^{(\mu)}(\epsilon,r), \quad (14)$$

which contradicts the assumption of linear independence of the $u^{(\mu)}(\epsilon,r)$ with $0 \leqslant \mu \leqslant \nu-1$. Therefore, the derivatives of orders $0,\ldots,\nu$ must all be linearly independent.

The second statement guarantees that orbitals $u^{(\nu)}(\epsilon,r)$ constructed from the valence basis are orthogonal to the core states $u^{(0)}(\epsilon_0,r)$. It should be noted, however, that shallow semicore states do not vanish sufficiently outside the MT spheres and should, therefore, be treated with (conventional) local orbitals at suitable energies. For the special case $\nu=0$ the second statement follows from $(\hat{h}-\epsilon_0) r u^{(0)}(\epsilon_0,r) = (\hat{h}-\epsilon) r u^{(0)}(\epsilon,r) = 0$ with $\epsilon \neq \epsilon_0$ and integration by parts, giving

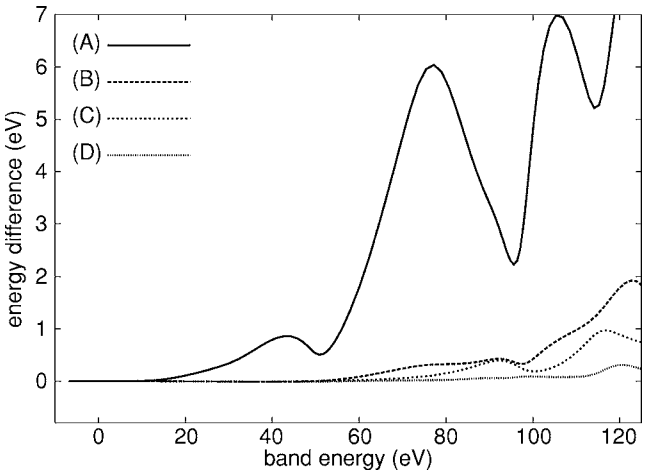

FIG. 2. Deviation of the Kohn-Sham eigenvalues from the results of our nonlinear reference calculation at the Γ point of Si as a function of energy for the basis sets (A), (B), (C), and (D) (for details see the text). The energy parameters are optimized for the valence bands and identical in all calculations. The eigenvalues are given with respect to the Fermi level. The curves are smoothed using a Bzier algorithm for clarity.

$$
\begin{aligned}
&(\epsilon-\epsilon_0)\langle u^{(0)}(\epsilon_0)|u^{(0)}(\epsilon)\rangle \\
&\quad=\int_0^R [ru^{(0)}(\epsilon_0,r)(\hat{h}ru^{(0)}(\epsilon,r)) \\
&\qquad-(\hat{h}ru^{(0)}(\epsilon_0,r))ru^{(0)}(\epsilon,r)]dr \\
&\quad=\frac{R^2}{2}[u^{(0)}(\epsilon,R)u^{(0)\prime}(\epsilon_0,R)-u^{(0)}(\epsilon_0,R)u^{(0)\prime}(\epsilon,R)] \\
&\quad=0, \qquad (15)
\end{aligned}
$$

where R is the MT radius. For $\nu \geqslant 1$ this expression instead reads

$$
\begin{aligned}
&(\epsilon-\epsilon_0)\langle u^{(0)}(\epsilon_0)|u^{(\nu)}(\epsilon)\rangle \\
&\quad=\frac{R^2}{2}[u^{(\nu)}(\epsilon,R)u^{(0)\prime}(\epsilon_0,R)-u^{(0)}(\epsilon_0,R)u^{(\nu)\prime}(\epsilon,R)] \\
&\qquad-\nu\langle u^{(0)}(\epsilon_0)|u^{(\nu-1)}(\epsilon)\rangle \\
&\quad=0, \qquad (16)
\end{aligned}
$$

where the last equality follows from induction.

When applied to Si, the example shown in Fig. 1, already the inclusion of second-derivative local s, p, d, and f orbitals yields agreement with the reference band structure from our nonlinear scheme, on the scale of the figure, up to 60 eV above the Fermi energy. The addition of third derivatives and g functions pushes this limit up to around 80 eV. The improvement over conventional LAPW in the description of the unoccupied states is clearly seen in Fig. 2, which shows the deviation of the single-particle energies from the corresponding reference values for (A) the conventional LAPW basis set, (B) with second-derivative local orbitals for $l \leqslant 3$, (C) with second-derivative local orbitals for $l \leqslant 4$, and (D) with second- and third-derivative local orbitals for $l \leqslant 4$ at the Γ point of Si as a function of the single-particle energies. The inclusion of just the second-derivative local orbitals for all $l \leqslant 3$ in (B) gives energies within 0.2 eV of the reference values up to 60 eV above the Fermi energy. Adding local orbitals with $l=4$ in (C) and third derivatives in (D) further improves the agreement. The advantage over similar approaches to improve the flexibility of the basis set (e.g., Ref. 20) is that systematic convergence can be achieved without special assumptions about the energy parameters of the additional local orbitals. As we find the results to be relatively independent of their position, one can simply use the LAPW energy parameters, which are located at the center of gravity of the valence band. The convergence towards basis-set completeness is then controlled by two simple parameters, the maximum quantum number l of the local orbitals and the order of derivatives ν (together with the l-cutoff and $K_{\max}$ of the augmented plane waves).

IV. APPLICATION TO *GW*

Once the wave functions are determined in the basis (10), the polarization function (5) and related quantities are represented in terms of a mixed basis designed for the expansion of products of eigenfunctions.[11] It consists of plane waves with a cutoff of 2.7 Bohr^{-1} in the interstitial region and products of two radial basis functions inside the MT spheres, where the first is related to an occupied and the second to an unoccupied state. For occupied states $u_l^{(1)}$ contributions and higher energy derivatives can be neglected, since the parameter ϵ_l is chosen to be the center of gravity of the occupied l-like density of states ($\epsilon_0=-7.9$ eV, $\epsilon_1=-3.2$ eV, $\epsilon_2=-3.4$ eV, $\epsilon_{l\geqslant 3}=-4.8$ eV relative to the top of the valence band). For unoccupied states the inclusion of $u_l^{(\nu)}$ with $\nu \geqslant 1$ as well as higher angular momenta can be important, because the corresponding wave-function coefficients become large for higher-lying states. In our calculations the mixed basis is constructed from $u_l^{(0)}$ with $l \leqslant 2$ for occupied as well as $u_l^{(0)}$ with $0 \leqslant l \leqslant 6$ and $u_l^{(\mu)}$ with $1 \leqslant \mu \leqslant \nu$ and $l \leqslant L$ [(A) $\nu=1$, $L=3$; (B) $\nu=2$, $L=3$; (C) $\nu=2$, $L=4$; (D) $\nu=3$, $L=4$] for unoccupied states. The omission of the other radial functions results in a negligible error not exceeding 0.1 meV in the correlation contribution $\langle \Sigma^c \rangle$ for all states considered here. The self-energy is evaluated with an angular-momentum cutoff of 6 and 216 **k** points in the full Brillouin zone. All other parameters are the same as in the underlying DFT calculation.

Changes in the expectation values of the exchange term $\langle \Sigma^x \rangle$ and the exchange-correlation potential $\langle V^{xc} \rangle$, which are both independent of the unoccupied states, are small and compensate each other; the main effect on the quasiparticle energies is due to the basis-set dependence of the correlation contribution to the self-energy. In Fig. 3 we show $\langle \Sigma^c \rangle$ for the valence-band maximum of Si at Γ as a function of the number of bands included in the Green function (4) and the polarization function (5) for the different basis sets (see Sec. III). Up to the fiftieth band the curves are nearly identical but then begin to deviate due to the linearization error. As the denominator in (5) reduces the weight of higher-lying states while their linearization error grows at the same time, the

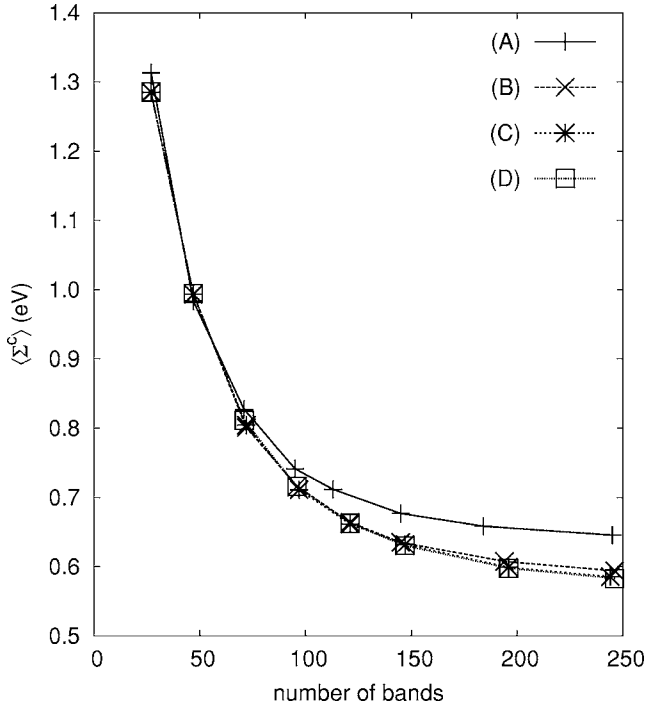

FIG. 3. Expectation value of the correlation contribution to the self-energy at the valence-band maximum of Si at Γ as a function of the number of bands. The inclusion of second-derivative local orbitals for $l \leqslant 3$ (B) changes the asymptotic value by more than 0.05 eV with respect to the conventional LAPW basis set (A). Further extensions of the basis give only minor corrections.

largest increase of the difference between the curves is seen for intermediate numbers of bands between 60 and 150. The curves converge rapidly with respect to the basis-set size: already with the inclusion of second-derivative local orbitals (B) convergence is reached to within 0.01 eV.

The corresponding curves for other states look qualitatively similar. Tables I and II give the values of $\langle\Sigma^c\rangle$ and the resulting quasiparticle energies ϵ calculated with 245 bands for the different basis sets. As a reference, the valence-band maximum in the underlying Kohn-Sham calculation is set to zero. All quasiparticle energies tend towards smaller values in the basis-set limit. Consequently, the effect on relative transition energies is smaller but of the same order. In Fig. 4

TABLE I. Expectation values of the correlation contribution $\langle\Sigma^c\rangle$ to the self-energy correction for the basis sets (A), (B), (C), and (D) (see text) together with the differences $\Delta\langle\Sigma^c\rangle$ between basis sets (A) and (D). All values are in eV.

	$\langle\Sigma^c\rangle$				
	(A)	(B)	(C)	(D)	$\Delta\langle\Sigma^c\rangle$
$\Gamma_{25'v}$	0.646	0.594	0.585	0.583	–0.063
Γ_{15c}	–4.183	–4.224	–4.230	–4.233	–0.050
X_{4v}	1.824	1.784	1.778	1.776	–0.049
X_{1c}	–3.749	–3.776	–3.780	–3.782	–0.033
$L_{3'v}$	1.139	1.091	1.083	1.081	–0.058
L_{1c}	–3.911	–3.960	–3.967	–3.971	–0.060

TABLE II. Quasiparticle energies ϵ for the basis sets (A), (B), (C), and (D) (see text) together with the differences $\Delta\epsilon$ between basis sets (A) and (D). All values are in eV.

	ϵ				
	(A)	(B)	(C)	(D)	$\Delta\epsilon$
$\Gamma_{25'v}$	–0.646	–0.685	–0.692	–0.694	–0.048
Γ_{15c}	2.541	2.514	2.509	2.509	–0.032
X_{4v}	–3.579	–3.608	–3.613	–3.615	–0.036
X_{1c}	0.521	0.502	0.499	0.498	–0.023
$L_{3'v}$	–1.883	–1.919	–1.925	–1.927	–0.044
L_{1c}	1.467	1.432	1.427	1.424	–0.043

we show the behavior of the indirect band gap as an example. The results are lowered by 0.02 eV if the **k**-point mesh is fully converged, and by another 0.02 eV if screening due to the 2*p* electrons is included in the correlation self-energy. The convergence with respect to the number of bands and the basis-set extension both increase the calculated band gap and thus narrow the distance to the experimental value of 1.17 eV. The final deviation is comparable to that of typical pseudopotential calculations, which tend to a slight overestimation.[2,3] Although a certain discrepancy with respect to the pseudopotential results still remains, we conclude that it is, in fact, smaller than suggested by previous calculations.[9,11,13,14] The stronger underestimation in these studies must be attributed at least in part to an incomplete convergence with respect to the number of bands (e.g., 24 bands in Ref. 9) in combination with the linearization error. In the case of Si we find that the latter accounts for less than 0.03 eV, but we cannot rule out that it is larger in other

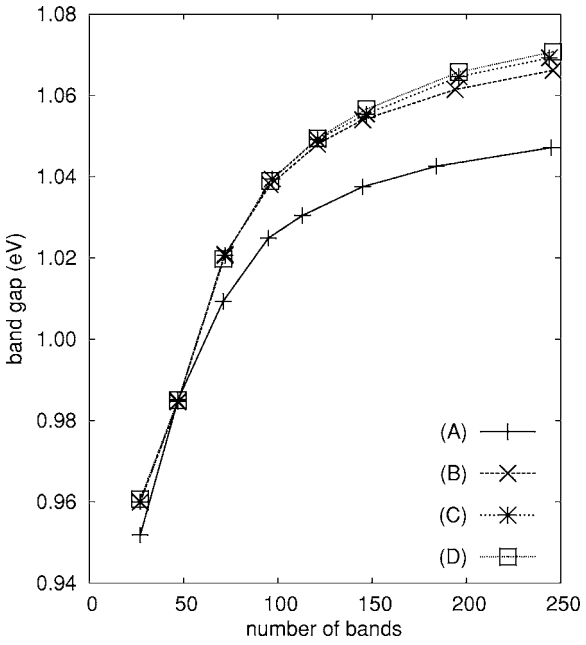

FIG. 4. Indirect band gap of Si as a function of the number of bands. The inclusion of second-derivative local orbitals for $l \leqslant 3$ (B) provides the largest step towards convergence with respect to the basis set.

systems and must be taken into account in order to obtain reliable *GW* results.

Very recently van Schilfgaarde *et al.*[28] also reexamined the convergence of the self-energy and the quasiparticle energies in Si, employing the same *GW* algorithm[11] as in this work. However, the eigenfunctions were generated by the full-potential LMTO method with basis sets ranging from 50 to 185 orbitals, including additional MT orbitals located in the conduction-band region. They obtained a $\Gamma_{25'v}$–X_{1c} gap of 1.15 eV, similar to our results in Table II; the slight discrepancy is due to the different way of generating the basis functions used to construct the self-energy. With the corrections for **k**-point convergence and screening due to the $2p$ electrons, their best estimate for the $\Gamma_{25'v}$–X_{1c} gap, if the *GW* approximation is evaluated with LDA eigenstates, is 1.10 eV. Furthermore, Shishkin and Kresse reported quasiparticle band gaps obtained with a new PAW implementation in quantitative agreement with our results.[29]

V. CONCLUSIONS

In this work we investigated the influence of the basis-set accuracy in linearized all-electron methods on the *GW* self-energy correction. We showed that the addition of local orbitals defined as second and higher energy derivatives of solutions of the radial Schrödinger (or scalar-relativistic Dirac) equation constitutes an extension of the LAPW basis set that allows a systematic improvement towards basis-set completeness. In the case of silicon basis-set convergence of the *GW* results was essentially reached with the inclusion of second-derivative local orbitals. It is the basis-set dependence of $\langle \Sigma^c \rangle$ that is responsible for changes in the quasiparticle energies. As all of them decrease towards basis-set completeness, the effect on the relative transition energies is smaller than on the quasiparticle energies themselves. The fundamental band gap increases by less than 0.03 eV. This makes the LAPW method a suitable quantitative reference scheme for all-electron *GW* calculations. A stronger basis-set dependence in other systems or in the correction of higher bands is possible, however. If convergence with respect to the number of bands is also taken into account, the *GW* band gap turns out to be larger and closer to pseudopotential values than previously reported all-electron results, although a certain discrepancy still remains.

ACKNOWLEDGMENTS

The authors acknowledge valuable discussions with Gustav Bihlmayer and Mark van Schilfgaarde as well as financial support from the Deutsche Forschungsgemeinschaft through the Priority Program 1145.

APPENDIX: SCALAR-RELATIVISTIC EQUATIONS

The radial scalar-relativistic Dirac equations for the large and small components p and q of a free electron in a spherical potential at energy ϵ are given by

$$p'(\epsilon,r) = 2M(\epsilon,r)q(\epsilon,r) + \frac{1}{r}p(\epsilon,r), \tag{A1}$$

$$q'(\epsilon,r) = -\frac{1}{r}q(\epsilon,r) + w(\epsilon,r)p(\epsilon,r) \tag{A2}$$

with $M(\epsilon,r)=1+[\epsilon-V^{\mathrm{eff}}(r)]/(2c^2)$ and $w(\epsilon,r)=[l(l+1)]/[2M(\epsilon,r)r^2]+V^{\mathrm{eff}}(r)-\epsilon$, where $V^{\mathrm{eff}}(r)$ is the spherical part of the Kohn-Sham effective potential.[30] Their νth energy derivatives are given by

$$p^{(\nu)'}(\epsilon,r) = 2M(\epsilon,r)q^{(\nu)}(\epsilon,r) + \frac{1}{r}p^{(\nu)}(\epsilon,r) + \frac{\nu}{c^2}q^{(\nu-1)}(\epsilon,r), \tag{A3}$$

$$q^{(\nu)'}(\epsilon,r) = -\frac{1}{r}q^{(\nu)}(\epsilon,r) + w(\epsilon,r)p^{(\nu)}(\epsilon,r) + \nu w^{(1)}(\epsilon,r) \times p^{(\nu-1)}(\epsilon,r) + \sum_{\mu=2}^{\nu}\binom{\nu}{\mu}w^{(\mu)}(\epsilon,r)p^{(\nu-\mu)}(\epsilon,r), \tag{A4}$$

where $M^{(1)}(\epsilon,r)=1/(2c^2)$ has been used. The νth energy derivative $w^{(\nu)}(\epsilon,r)$ has the expression

$$w^{(\nu)}(\epsilon,r) = (-1)^{\nu}\frac{\nu!\,l(l+1)}{2^{\nu+1}M(\epsilon,r)^{\nu+1}r^2c^{2\nu}} - \delta_{1\nu}. \tag{A5}$$

In the nonrelativistic limit these formulas correspond to Eq. (12).

*Electronic address: c.friedrich@fz-juelich.de

[1] L. Hedin, Phys. Rev. **139**, A796 (1965).
[2] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).
[3] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).
[4] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[5] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **35**, 4170 (1987).
[6] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, in *Solid State Physics*, edited by H. Ehrenreich and F. Spaepen (Academic, New York, 2000), Vol. 54, p. 1, and references therein.
[7] N. Hamada, M. Hwang, and A. J. Freeman, Phys. Rev. B **41**, 3620 (1990).
[8] F. Aryasetiawan, Phys. Rev. B **46**, 13051 (1992).
[9] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).
[10] M. Usuda, N. Hamada, T. Kotani, and M. van Schilfgaarde, Phys. Rev. B **66**, 125101 (2002).
[11] T. Kotani and M. van Schilfgaarde, Solid State Commun. **121**, 461 (2002).
[12] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).

[13] B. Arnaud and M. Alouani, Phys. Rev. B **62**, 4464 (2000).
[14] S. Lebègue, B. Arnaud, M. Alouani, and P. E. Bloechl, Phys. Rev. B **67**, 155208 (2003).
[15] A. Ernst, M. Lüders, P. Bruno, W. M. Temmerman, and Z. Szotek (unpublished).
[16] M. L. Tiago, S. Ismail-Beigi, and S. G. Louie, Phys. Rev. B **69**, 125212 (2004).
[17] K. Delaney, P. García-González, A. Rubio, P. Rinke, and R. W. Godby, Phys. Rev. Lett. **93**, 249701 (2004); W. Ku and A. G. Eguiluz, *ibid.* **93**, 249702 (2004).
[18] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965).
[19] H. Bross and G. M. Fehrenbach, Z. Phys. B: Condens. Matter **81**, 233 (1990).
[20] E. E. Krasovskii, A. N. Yaresko, and V. N. Antonov, J. Electron Spectrosc. Relat. Phenom. **68**, 157 (1994).
[21] J. C. Slater, Phys. Rev. **51**, 846 (1937); Adv. Quantum Chem. **1**, 35 (1964).
[22] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975); D. D. Koelling and G. O. Arbman, J. Phys. F: Met. Phys. **5**, 2041 (1975).
[23] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981); M. Weinert, E. Wimmer, and A. J. Freeman, *ibid.* **26**, 4571 (1982).
[24] http://www.flapw.de
[25] D. R. Hamann, Phys. Rev. B **40**, 2980 (1989).
[26] E. E. Krasovskii, Phys. Rev. B **56**, 12866 (1997).
[27] D. Singh, Phys. Rev. B **43**, 6388 (1991).
[28] M. van Schilfgaarde, T. Kotani, and S. V. Faleev, cond-mat/0508295 (unpublished).
[29] M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).
[30] D. J. Singh, *Planewaves, Pseudopotentials and the LAPW Method* (Kluwer, Dordrecht, 1994).