

Density-Functional Treatment of an Exactly Solvable Semiconductor Model

O. Gunnarsson

Max-Planck-Institut für Festkörperforschung, D-7000 Stuttgart 80, Federal Republic of Germany

and

K. Schönhammer

Institut für Theoretische Physik, Universität Göttingen, D-3400 Göttingen, Federal Republic of Germany

(Received 23 September 1985)

The band-gap problem is studied for an exactly solvable semiconductor model. The exact exchange-correlation potential is constructed and its discontinuity is studied. Depending on the parameter range, a large or a small discontinuity is obtained. The "scissor-operator" method is compared with the local-density approximation for an impurity calculation. In general, the local-density approximation is closer to the exact solution. Some general features of the density-functional formalism are illustrated.

PACS numbers: 71.45.Gm, 71.25.Rk, 71.55.-i

The local-density (LD) approximation in the density-functional (DF) formalism¹ gives a substantial underestimate of the band gap in semiconductors.² This is a major problem in impurity calculations, since the impurity states depend sensitively on the size of the band gap. A crucial question is therefore if and how one should correct for this band-gap problem in LD calculations. It was recently realized by Perdew and Levy³ and by Sham and Schlüter⁴ that the true band gap defined below need not be equal to the energy eigenvalue gap in an exact DF calculation, because of a discontinuity in the exchange-correlation (XC) potential as a function of the number of electrons.⁵ This raises the question if the underestimate of the band gap in the LD approximation is mainly due to

this discontinuity or to the LD approximation itself.

To address these issues, we consider an exactly solvable model with and without an impurity. For this model we introduce a DF formalism and a LD approximation. By comparing the true band gap with the eigenvalue gap in the exact and approximate DF calculations, we determine the importance of the discontinuity for the model. The exact results for the model with an impurity are compared with the LD results and with the results obtained after the band gap has been "corrected." This provides a test of methods for solving the band-gap problem.

As a model we use a finite linear chain of M atoms with two nondegenerate levels, which may be understood as s and p levels. Thus we consider the Hamiltonian

$$\begin{aligned}
 H = \sum_{i=1}^M \sum_{\alpha=1}^2 \epsilon_{i\alpha} \hat{n}_{i\alpha} + U \sum_{i=1}^M \hat{n}_{i1} \hat{n}_{i2} + \sum_{i=2}^M (t_1 \psi_{i,1}^\dagger \psi_{i-1,1} + t_2 \psi_{i,2}^\dagger \psi_{i-1,2} + t_3 \psi_{i,2}^\dagger \psi_{i-1,1} - t_3 \psi_{i,1}^\dagger \psi_{i-1,2} + \text{H.c.}) \\
 + (t_4 \psi_{M,2}^\dagger \psi_{1,1} + t_5 \psi_{1,2}^\dagger \psi_{M,1} + \text{H.c.}), \quad (1)
 \end{aligned}$$

where $\epsilon_{i\alpha}$ is the energy of the α level (i.e., s or p) on the i th atom. The second term describes the Coulomb interaction U between two electrons on the same atom. If $t_1 < 0$, $t_2 > 0$, and $t_3 > 0$, we can interpret the terms containing t_1 , t_2 , and t_3 as s - s , p - p , and s - p hopping, respectively. The terms containing t_4 and t_5 , which connect the end points, have been introduced to remove some degeneracies. Alternatively, we can set $t_1 = t_2$, $t_3 = t_4 = t_5 = 0$, and $\epsilon_{i\alpha} = 0$. Then the model reduces to a Hubbard chain, with α being a spin index. To describe an impurity on the site $i = i_0 = (M+1)/2$ in the model (1), we make the replacement $\epsilon_{i_0\alpha} \rightarrow \epsilon_{i_0\alpha} - \Delta V$.

To solve this model exactly, we consider all possible

configurations with N electrons and calculate the corresponding Hamiltonian matrix. For $N = M = 9$ the matrix size is $48\,620 \times 48\,620$. By using a site representation for the configurations, we obtain a very sparse matrix and the lowest eigenvalues and eigenvectors can be obtained rather easily. This gives the occupation numbers $n_{i\alpha}$ for the exact ground state. In this model the energy levels $\epsilon_{i\alpha}$ play the role of the external potential $v_{\text{ext}}(r)$ in the normal formulation of DF theory. In the "DF theory" for our model (1) the occupation numbers $n_{i\alpha}$ are therefore used as the basic variables instead of the density $n(r)$. To obtain the exact XC potential $v_{i\alpha}^{\text{xc}}$, we replace $\epsilon_{i\alpha}$ by an effective

potential $v_{i\alpha}^{\text{eff}}$, which is varied until the exact $n_{i\alpha}$ are obtained for the noninteracting system. This defines the exact

$$v_{i\alpha}^{\text{xc}} = v_{i\alpha}^{\text{eff}} - \epsilon_{i\alpha} - U(n_{i1} + n_{i2})$$

apart from a constant. This constant is fixed by the requirement that the highest occupied eigenvalue is equal to the exact first ionization potential.^{5,6} The systems which have been studied are v -representable in the sense that we could find numerically a $v_{i\alpha}^{\text{eff}}$ which gives the $n_{i\alpha}$ prescribed.

We define a LD approximation as

$$E_{\text{xc}}^{\text{LD}}[n_{i\alpha}] = -CU \sum_{i=1}^M (n_{i1} + n_{i2})^{4/3}, \quad (2)$$

which contains the density to the $\frac{4}{3}$ power, a dependence typical for local approximations.⁷ The prefactor $C = 2^{-4/3}$ is fixed so that (2) is exact for a filled band ($N = 2M$).

We now demonstrate within our model DF theory that a discontinuity in the XC potential can contribute to the true band gap

$$E_g = E_0(M+1) + E_0(M-1) - 2E_0(M), \quad (3)$$

where $E_0(N)$ is the ground-state energy of a system

with N electrons. We consider a Hubbard chain with periodic boundary conditions in the limit $N = M \rightarrow \infty$. Then the exact $n_{i\alpha} = \frac{1}{2}$ (no spin polarization) and v^{eff} is a constant. The exact DF treatment therefore does not give an energy eigenvalue gap. The half-filled Hubbard chain is, however, a Mott insulator.⁸ The gap is therefore entirely due to the XC potential discontinuity,⁹ in agreement with earlier arguments.^{5,6} Recently, an explicit formula [Eq. (14) in Ref. 4] for the exchange-correlation potential discontinuity was proposed. We have demonstrated that this formula vanishes identically.¹⁰ The example above shows, however, that the discontinuity is not, in general, zero, and that therefore the formula in Ref. 4 is not correct.¹⁰

With parameters more appropriate for a semiconductor ($t_1 < 0$, $t_2 > 0$, $t_3 \neq 0$, $\epsilon_{i2} > \epsilon_{i1}$), the exact DF eigenvalue gap is close to the exact gap E_g , while the LD eigenvalue gap is substantially smaller. Thus, within the same model we can obtain the situation where a poor LD gap is primarily due to the LD approximation itself, as well as the situation where it is due to the discontinuity. Below we study both these cases.

Table I shows results for the Hubbard chain as a function of U . The exact XC energy E_{xc} is compared with the difference

$$\tilde{E}_{\text{int}} = E_{\text{int}} - E_{\text{elstat}} = U \sum_i \langle \hat{n}_{i1} \hat{n}_{i2} \rangle - U \sum_i (\langle \hat{n}_{i1} \rangle + \langle \hat{n}_{i2} \rangle)^2 / 2$$

between the interaction energy and the "electrostatic" energy, where the angular brackets indicate the expectation value for the exact ground state. We also show the "kinetic" energy, which is defined as the expectation value of the hopping terms in Eq. (1). T_{int} is calculated for the exact solution of (1) and T_{nonint} is obtained for a noninteracting system ($U=0$) with the same occupation numbers. $|E_{\text{xc}}|$ is appreciably smaller than $|\tilde{E}_{\text{int}}|$, since $E_{\text{xc}} = \tilde{E}_{\text{int}} + T_{\text{int}} - T_{\text{nonint}}$ by definition¹ contains a kinetic energy contribution. This contribution is positive, since the correlation of the electrons in the interacting system leads to an increase in the kinetic energy. \tilde{E}_{xc} can be expressed in terms of $\tilde{E}_{\text{int}}(U)$ as¹¹

$$E_{\text{xc}} = \int_0^U dU' \tilde{E}_{\text{int}}(U') / U'.$$

In Table I the integration is performed with use of the trapezoidal formula with the step $dU' = 1$, which gives good agreement with the exact E_{xc} . The LD result $E_{\text{xc}}^{\text{LD}}$ for the XC energy is more negative than the exact result.

The Hubbard chain with a substitutional impurity is considered in Table II. In addition to the exact and LD results, we show results of a calculation ("Sciss") using the "scissor operator,"¹² which shifts the unoccupied levels of the system without an impurity by an amount E_s , so that the LD eigenvalue gap agrees with the exact gap. Finally, Table II shows results of a calculation ("Sciss corr"), where the scissor operator is used to generate the occupation numbers, but the energy is evaluated by use of a pure LD functional

TABLE I. Results for a Hubbard chain. The parameters are $M=9$, $t_1=t_2=t_5=-1$, $t_3=t_4=0$, $\epsilon_{i\alpha}=0$, and $\Delta V=0$. All energies are in electronvolts.

U	\tilde{E}_{int}	E_{xc}	$E_{\text{xc}}^{\text{LD}}$	T_{int}	T_{nonint}	$\int dU' \tilde{E}_{\text{int}}/U'$	E_g
2	-5.89	-5.20	-7.14	-10.42	-11.11	-5.20	0.61
3	-9.93	-8.33	-10.72	-9.50	-11.11	-8.33	0.96
4	-14.64	-11.80	-14.29	-8.24	-11.11	-11.80	1.52
5	-19.30	-15.54	-17.86	-7.33	-11.11	-15.53	2.24

TABLE II. Results for a Hubbard chain with an impurity with $\Delta V = -3$ and $U = 5$. The other parameters are the same as in Table I. All energies are in electronvolts.

	E_B	Δn^B	I_p	Δn^I
Exact	0.91	0.81	-2.38	-0.58
LD	1.24	0.54	-2.38	-0.13
Sciss	-0.74	0.49	-4.22	-0.18
Sciss corr	1.21	0.49	-2.40	-0.18

without the scissor operator.^{13,14} We consider a donor ($N = M + 1$ for the neutral impurity system), since the effect of the scissor operator is smaller for an acceptor. E_B is the energy difference between the systems without and with an impurity, and Δn^B is the increase of the charge on the site where the impurity is introduced. $I_p = E_0(M) - E_0(M + 1)$ is the ionization potential of the impurity system, and Δn^I is the change in the ground-state charge on the impurity site when the impurity is ionized. The energies (E_B, I_p) are better described in the LD approximation than with the scissor operator. The energy of the $(M + 1)$ -electron system is higher when the scissor operator is used, because at least one electron feels the repulsive nonlocal "scissor operator potential." The exact XC potential is very attractive on the impurity site for the $(M + 1)$ -electron system. Neither the LD nor the scissor-operator approximations can describe this, and both give too small values of Δn^B . Both these approximations fail to give a proper description of Δn^I . The corrected scissor-operator method and the LD approximation give similar results.

We now consider the model (1) with parameters more appropriate for a semiconductor. Figure 1 shows the eigenvalues in the exact DF formalism (ϵ_i DF) and the LD approximation (ϵ_i LD) for a system without an impurity. The LD eigenvalue spectrum has been shifted so that the top of the valence band agrees with the exact DF calculation. There is a good agreement between the exact energy gap E_g and the exact DF eigenvalue gap (deviation 0.5%), which shows that the discontinuity is very small in this case. The poor LD eigenvalue gap is therefore mainly due to the LD approximation itself. Small values for the discontinuity were also obtained for the other sets of "semiconductor" parameters studied as well as for some modifications in the form of the Hamiltonian (1).

In the exact DF formalism the highest occupied eigenvalue gives the ionization potential,^{5,6} while the other eigenvalues in general do not give excitation energies.¹ It is nevertheless interesting to compare the eigenvalues with the energies of some excited states, $|E_n(M - 1)\rangle$ and $|E_n(M + 1)\rangle$, which evolve from a

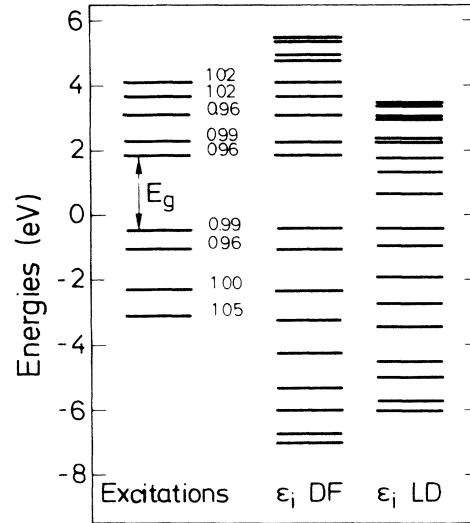


FIG. 1. The eigenvalue spectra (ϵ_i) and some exact excitation energies for a semiconductor model without an impurity. The parameters are $M = N = 9$, $t_1 = -1.8$, $t_2 = 1.0$, $t_3 = -1.2$, $t_4 = -t_5 = t_3$, $U = 4$, $\epsilon_{i1} = -4$, and $\epsilon_{i2} = 0$. All energies are in electronvolts.

noninteracting state if an electron is removed or added without the creation of electron-hole pairs. To select such states we introduce

$$\hat{\tau} = \sum_{i=1}^M (\tau_{i1}\psi_{i1} + \tau_{i2}\psi_{i2})$$

and consider some states¹⁵ for which

$$M_n^- = \langle E_n(M - 1) | \tau^- | E_0(M) \rangle$$

and

$$M_n^+ = \langle E_n(M + 1) | \tau^+ | E_0(M) \rangle$$

are large. The corresponding energy differences, $E_0(M) - E_n(M - 1)$ and $E_n(M + 1) - E_0(M)$, are shown under "Excitations." These energies, for instance, provide the values of I_p and E_g . The figure shows that there is a remarkable agreement between the exact energy differences and the exact DF eigenvalues. The LD hole spectrum [$E_n(M - 1)$] is slightly compressed relative to the exact results, as tends to be the case in realistic band-structure calculations. Although the eigenvectors obtained in the DF formalism in general cannot be used to calculate matrix elements,¹ we have formed Slater determinants and evaluated M_n^- and M_n^+ . The numbers at the levels give the ratios to the exact results. Again we find a surprising agreement. For the Hubbard chain, on the other hand, there is a large disagreement between the DF eigenvalues and the exact excitation energies as well as between the corresponding matrix elements. This is consistent with the observation that for systems with strong correlation effects, e.g., many rare-earth

TABLE III. Results for the semiconductor model with an impurity with $\Delta V = -4$. The remaining parameters are the same as in Fig. 1. All energies are in electronvolts.

	E_B	Δn^B	$\epsilon_b - \epsilon_v$	Δn^I
Exact	4.13	0.78	1.07	0.47
LD	3.84	0.53	1.45	0.09
Sciss	2.55	0.50	2.57	0.14
Sciss corr	3.81	0.50	1.44	0.14

compounds, the LD eigenvalues give an unsatisfactory description of excitation energies, while for many other systems the description is better than originally expected.

Table III shows results for an acceptor substitutional impurity in the semiconductor chain. We present results for E_B , Δn^B , and Δn^I and for the separation

$$\epsilon_b - \epsilon_v = E_0(M+1) + E_0(M-1) - 2E_0(M)$$

of the impurity level from the top of the valence band. As for the Hubbard chain, we find that the LD approximation gives better results for the binding energy E_B than the scissor-operator method. In the LD approximation, the gap state ϵ_b is somewhat too far above the top of the valence band. Since the band gap is too small, the gap state is too close to the bottom of the conduction band, and the state is too extended. This is illustrated by the small value of Δn^I . In the scissor-operator method the conduction band is raised by 1.2 eV. Since, however, the gap state is mainly derived from the conduction band, it tends to follow the conduction band and it remains too extended. The value of Δn^I is therefore not substantially better than in the LD approximation. For the chain without an impurity, the exact v^{xc} is in this case about 3 eV higher for the essentially unoccupied state (p level) than for the s level. When the impurity is introduced also, the p state on the impurity site obtains a large occupancy in the exact solution. The exact v^{xc} is then similar for the s and p states on the impurity site. As far as the mainly occupied states are concerned, the LD approximation therefore gives a better simulation of the exact v^{xc} than the scissor operator.

In conclusion, we have studied a finite, one-dimensional model with a short-range interaction. The Hubbard limit of this model demonstrates that the discontinuity in the XC potential can give a large contribution to the band gap. The XC discontinuity is, however, small in this model, if parameters more appropriate for a semiconductor are used. We have compared the LD approximation with the scissor-operator

method for impurity calculations and found that for the model considered the LD approximation generally gives better agreement with the exact solution.

We want to thank O. K. Andersen, W. Hanke, M. Scheffler, F. Beeler, and P. Kelly for fruitful discussions.

¹P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

²See, e.g., C. S. Wang and B. M. Klein, Phys. Rev. B **24**, 3393 (1981).

³J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

⁴L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

⁵J. P. Perdew, R. G. Parr, M. Levy, and J. L. Baduz, Phys. Rev. Lett. **49**, 1691 (1982).

⁶U. von Barth, in *The Electronic Structure of Complex Systems*, edited by P. Phariseau and W. M. Temmermann (Plenum, New York, 1984); C. O. Almbladh and U. von Barth, in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985).

⁷Although the exponent $\frac{4}{3}$ and the coefficient C are somewhat arbitrary, Eq. (2) is found to give an accuracy typical for the "normal" LD approximation for the semiconductor model (positive errors in the range 5%–10%). For the Hubbard model the errors are larger and negative. Equation (2) also gives errors in the band gap of the type found in the "normal" LD approximation. Apart from a uniform shift, the value of C has a small influence on the eigenvalues and therefore on the band gap.

⁸E. H. Lieb and F. Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968).

⁹For the finite chain with $M = N = 9$ we also performed calculations with a finite staggered external magnetic field h , so that the ground state had a static antiferromagnetic ordering. As long as h is not very large a substantial discontinuity δ_{xc} in the XC potential remains. In the example shown ($U = 5$), $h = 1$ led to $\delta_{xc} = 1.27$.

¹⁰K. Schönhammer and O. Gunnarsson, to be published.

¹¹J. P. Perdew and D. C. Langreth, Solid State Commun. **17**, 1425 (1975); O. Gunnarsson and B. I. Lundquist, Phys. Rev. B **13**, 4274 (1976).

¹²G. Baraff and M. Schlüter, Phys. Rev. B **30**, 3460 (1984).

¹³Such an approach was suggested by U. von Barth, in Proceedings of the Conference on Total Energy Methods in Solid State Physics, Braunschweig, West Germany, 3–5 December 1984 (unpublished); see also G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **55**, 1327 (1985).

¹⁴The total energy can be calculated from

$$E_{\text{sciss corr}} = E_{\text{sciss}}(E_s) - E_s \partial E_{\text{sciss}}(E_s) / \partial E_s.$$

¹⁵Figure 1 does not show all such states, since the diagonalization method (Lanczos) used requires more work to give excited states which are not very low or very high in energy.