On the use of constrained density-functional theory in determining model Hamiltonians. Test calculations for the H_2 molecule and a C chain

1998 J. Phys.: Condens. Matter 10 6953

(http://iopscience.iop.org/0953-8984/10/31/012)

J. Phys.: Condens. Matter **10** (1998) 6953–6971. Printed in the UK PII: S0953-8984(98)93144-7

On the use of constrained density-functional theory in determining model Hamiltonians. Test calculations for the H_2 molecule and a C chain

Heiko Meider and Michael Springborg

Fakultät für Chemie, Universität Konstanz, D-78457 Konstanz, Germany

Received 6 April 1998

Abstract. Constrained density-functional theory (CDFT) has been used by several authors for determining model Hamiltonian parameters of high-T_c superconductors and other transition metal compounds. These methods use the fact that the orbitals are well localized and can easily be split into atomic components. We generalize them to systems with relatively delocalized electrons, so that they can be used, for example, for determining parameters for a Hubbard model of conducting polymers (e.g., polyacetylene). We selected the H_2 molecule and a linear C chain as test cases. The results of the constrained density-functional calculations were fitted to Hartree–Fock model calculations of the Hubbard model. In the case of H_2 we also fitted the CDFT results to exact solutions of the Hubbard model. For H_2 we found a substantial discrepancy between the U-values from the exact and the Hartree–Fock model calculations, which shows that it is important to specify the approximations that are used in determining the model parameters. On the other hand, the CDFT calculations themselves lead to roughly uniquely defined model parameters.

1. Introduction

Parameter-free studies of the electronic properties of specific materials are becoming of increasing importance. Thereby larger and more complex systems as well as more properties have become accessible with *ab initio* methods based on the Hartree–Fock approximation and improvements thereof as well as with density-functional methods. This development has been mainly due to improvements in the computational schemes and their accuracy as well as in the computer technologies.

Nevertheless, not all properties and systems can be treated within these approaches, and in order to study theoretically those that are not directly accessible it is necessary to apply simplified models. These models are constructed such that only those properties of specific interest are treated accurately. For low-energy processes this amounts to studying only the orbitals closest to the Fermi level explicitly, whereas the remaining part of the total Hamiltonian is approximated as some simple form of the structure of the system.

To be more specific we study these properties by defining a model Hamiltonian for the system of interest as

$$\hat{H} = \hat{H}_{\mathrm{el}} + \hat{H}_{\mathrm{nuc}} \tag{1}$$

where $\hat{H}_{\mathrm{nuc}}$ depends only on the nuclear positions. In the present study we shall keep the nuclear positions fixed, and we can thus set $\hat{H}_{\mathrm{nuc}}$ equal to zero without loss of generality.

0953-8984/98/316953+19$19.50

As is common practice, we write that part of $\hat{H}$ that describes the electrons closest to the Fermi level as

$$\hat{H}_{\mathrm{el}} = \hat{H}_{\mathrm{sp}} + \hat{H}_{\mathrm{corr}} \tag{2}$$

with

$$\hat{H}_{\mathrm{sp}} = -\sum_{i,j,\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} \tag{3}$$

being the single-particle part, and

$$\hat{H}_{\mathrm{corr}} = \frac{1}{2} \sum_{i,j,k,l,\sigma,\sigma'} U_{ijkl} c^{\dagger}_{j\sigma'} c^{\dagger}_{i\sigma} c_{k\sigma} c_{l\sigma'} \tag{4}$$

being the many-body part.

Here i, j, k, and l are orbital indices, whereas σ and σ' are spin indices. $c^{\dagger}_{i\sigma}$ and $c_{i\sigma}$ are creation and annihilation operators which belong to some spin orbital $\varphi_{i\sigma}$. We denote the (position-space) orbital corresponding to $\varphi_{i\sigma}$ by φ_i. It is important to stress that the precise form of the orthonormal functions φ_i is not specified but that they are assumed to resemble atom-centred basis functions.

The Hamiltonian of equations (2)–(4) is the most general form. In the case in which only the elements $U_{ij} \equiv U_{ijij}$ are non-zero, we have a Pariser–Parr–Pople (PPP) Hamiltonian. If only the $U_{ii} \equiv U_{iiii}$ are non-zero, the Hubbard model results. Finally, if only those U_{ij} for which $i = j$ or i and j are nearest neighbours are non-zero, we have the extended Hubbard model.

Once the parameters t_{ij} and U_{ijkl} (and in the general case also the precise form of $\hat{H}_{\mathrm{nuc}}$) are known, one can attempt to study the properties of interest using this Hamiltonian. The advantage with these models is that, once a model for a given system has been defined, it can be used in studying the ground state, excited states, neutral systems, charged systems, and systems containing various kinds of symmetry reduction. One problem is, however, that it is far from trivial to determine the parameters. Most often one is guided by a combination of intuition and pieces of information from theoretical and experimental studies.

The approaches based on model Hamiltonians may be contrasted with parameter-free methods for calculating properties of materials. These are almost always based on expanding the single-particle eigenfunctions in some kinds of non-orthogonal and/or not-atom-centred basis function, such that a separation of the eigenfunctions into atomic and angular components becomes far from obvious. Furthermore, for a given system and structure, only a few states are accessible, and for extended (infinite) systems that are assumed periodic, it is not possible to treat charged systems exactly. This means that a direct comparison between the results of parameter-free calculations and model calculations is far from trivial, which in turn means that it is not obvious how to determine the parameters t_{ij} and U_{ij} of the model Hamiltonian from the results of parameter-free calculations.

In the present study we will present a general method that allows us to determine the parameters of the model Hamiltonian from parameter-free density-functional calculations. The method is based on the constrained density-functional (CDFT) method originally formulated by Dederichs *et al* [1]. This approach allows one to modify the charge density locally, whereupon the remaining part of the electron distribution will adjust to the new distribution. It has been applied mainly to systems in which the orbitals of interest are well localized and for which the density-functional method therefore allows for a direct separation of the electron density into atomic and angular components. Such studies included the original one of Dederichs *et al* [1] who studied Ce impurities in Ag and Pd. Later, the

CDFT method was used to study high-T_c superconductors [2, 3], the magnetic properties of magnetite [4], and MX chains [5].

The method that we will propose is similar to those, but applies also to cases in which the electrons are so delocalized that there is no natural separation into 'atomic parts', and for which the parameter-free method applies basis sets that are better than minimal. It is intimately related to the constrained density-functional method of Dederichs *et al* [1], which, for the sake of completeness, we briefly outline in section 2. The generalization is described in section 3, where we also describe our density-functional method. We test the method on the H_2 molecule and a linear C chain. The H_2 molecule is so small that both approximate and exact calculations for the model can be carried through, whereas only approximate studies can be performed for the C chain. The H_2 molecule allows therefore for an assessment of the consequences of incorporating various approximations in the model studies. The details of how we treat the model Hamiltonians are given in section 4.

In section 5 we give the results of the CDFT calculations for our test systems, study the variation of the model parameters when changing various internal parameters of the CDFT calculations, and compare the results for the exact and for the Hartree–Fock solution of the model Hamiltonians. Finally, section 6 contains our conclusions.

2. Constrained density-functional theory

Within the Born–Oppenheimer approximation and the density-functional formalism of Hohenberg and Kohn [6], the total electronic energy E of the system of interest is a functional of the electron density ρ:

$$E = E[\rho]. \tag{5}$$

Unfortunately, its precise form is not known. Kohn and Sham showed [7] that the task of minimizing $E[\rho]$ variationally may be recast into that of solving the eigenvalue equations

$$\hat{h}_{\mathrm{KS}}\psi_m(\boldsymbol{r}) = \varepsilon_m\psi_m(\boldsymbol{r}) \tag{6}$$

where (in Rydberg atomic units)

$$\hat{h}_{\mathrm{KS}} = -\Delta + V(\boldsymbol{r}). \tag{7}$$

Here $V(\boldsymbol{r})$ is the sum of the Coulomb potentials of the electrons, that of the nuclei, and a remainder. The latter is the so-called exchange–correlation potential V_{xc}. For V_{xc} we use the local density approximation (LDA) in the form given by von Barth and Hedin [8], but for the present discussion its precise form is not important.

The electron density is given by

$$\rho(\boldsymbol{r}) = \sum_{m=1}^{\mathrm{occ}} p_m|\psi_m(\boldsymbol{r})|^2 \tag{8}$$

where p_m is the occupation of ψ_m, and where the sum runs over all occupied orbitals.

To arrive at constrained density-functional theory we follow an approach close to that of Hybertsen *et al* [2]. We shall assume that we have somehow (to be specified in the following section) defined a set of orthonormal basis functions that can each be ascribed to a certain atom and (l, m) with the compound index I describing this classification. Denoting any of these functions by $\tilde{\varphi}_I$, this means that for any normalized function $f(\boldsymbol{r})$ we can consider

$$\hat{P}_I f(\boldsymbol{r}) = \tilde{\varphi}_I(\boldsymbol{r}) \int \tilde{\varphi}_I^*(\boldsymbol{r}')f(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \tag{9}$$

as that part of $f(\boldsymbol{r})$ that corresponds to the atom and angular dependence as specified by I. The operator $\hat{P}_I = |\tilde{\varphi}_I\rangle\langle\tilde{\varphi}_I|$ is a projection operator. We require the functions $\tilde{\varphi}_I$ to resemble those of our model Hamiltonian, but whereas we do not need to specify the precise spatial dependence for the latter, we have to do that here.

In order to arrive at the constrained density-functional method, we notice that the quantity

$$n_I = \sum_{m=1}^{\text{occ}} p_m \langle\psi_m|\hat{P}_I\psi_m\rangle \tag{10}$$

is that part of the total number of electrons (equation (8)) that is ascribed to a given atom and angular dependence as characterized by I.

When solving the Kohn–Sham equations (6) by expanding the function ψ_m in a set of atom-centred basis functions with a given angular dependence, it will be desirable for the functions $\tilde{\varphi}_I$ to resemble those. How this is achieved will be described in the following section.

The constrained density-functional approach corresponds now to minimizing the functional of equation (5), but with the additional constraints that for a certain subset of the Is the n_I have predefined values N_I. These constraints can be incorporated via Lagrange multipliers λ_I, and we thus seek the minimum

$$\tilde{E}(\{N_I\}) = \min\left\{E[\rho] + {\sum_I}' \lambda_I(n_I - N_I) + \mu\left(\int \rho(\boldsymbol{r})\ \mathrm{d}\boldsymbol{r} - N_{\text{el}}\right)\right\} \tag{11}$$

where the prime indicates that the constraints are only applied for a subset of the Is. In equation (11) the usual constraint that the total number of electrons N_{el} is conserved has been included too as the last term on the right-hand side. Here μ is the chemical potential.

Within the single-particle formulation of Kohn and Sham, the problem of solving equation (11) can be recast into that of solving the modified Kohn–Sham equations

$$\left(\hat{h}_{\text{KS}} + {\sum_I}' \lambda_I \hat{P}_I\right)\psi_m = \varepsilon_m\psi_m. \tag{12}$$

In principle, for a given set $\{N_I\}$, one arrives at a set $\{\lambda_I\}$. In any practical calculations, it is however easier to specify $\{\lambda_I\}$ and from the results of the calculation determine $\{N_I\}$.

3. Implementing constrained density-functional theory

In a typical calculation, the eigenfunctions of equation (6) or (12) are expanded in some set of basis functions. We shall here assume that they are atom centred (index $\boldsymbol{R}$), have a specific angular dependence (index $L \equiv (l, m)$), and that different functions with the same $\boldsymbol{R}$ and L are distinguished through the index κ. Thus, we write

$$\psi_i(\boldsymbol{r}) = \sum_{\boldsymbol{R},L,\kappa} c_{i\boldsymbol{R}L\kappa}\chi_{\boldsymbol{R}L\kappa}(\boldsymbol{r}). \tag{13}$$

In our density-functional method (described in detail in [9]), the basis functions are so-called linearized muffin-tin orbitals (LMTOs). They are defined as spherical Hankel functions multiplied by spherical harmonics:

$$\chi_{\boldsymbol{R}lm\kappa}(\boldsymbol{r}) = \frac{\mathrm{i}\kappa^{(l+1)}}{(2l-1)!!}h_l^{(1)}(\kappa|\boldsymbol{r}-\boldsymbol{R}|)Y_{lm}\left(\frac{\boldsymbol{r}-\boldsymbol{R}}{|\boldsymbol{r}-\boldsymbol{R}|}\right) \tag{14}$$

augmented continuously and differentiably inside non-overlapping atom-centred (muffin-tin) spheres with numerically given functions; these are in turn obtained from equation (6) by

replacing the potential with its spherically symmetric component. Thereby the functions become good approximations to the exact solutions of equation (6). In a typical calculation, the basis set consists of two subsets, each having s, p and d functions centred on all atomic sites. These two subsets differ in the decay constant κ of the spherical Hankel functions. We stress, however, that our general CDFT scheme can also be applied to other DFT methods using a (not necessarily minimal) basis set of atom-centred functions.

The radii of the muffin-tin spheres are free (internal) parameters. Experience has shown that calculations with different (reasonable) choices of these give similar results for geometry optimizations, electron densities, etc. Our aim is to construct a constrained density-functional method such that the results of these calculations become largely independent of the sphere radii.

Within the LMTO–ASA method as applied by e.g. Hybertsen *et al* [2] and by Zhang and Satpathy [4], one expands the muffin-tin spheres so that they become slightly overlapping but such that their total volume equals that of the crystal. Furthermore, the basis set contains only one function per atom and L. Thereby the basis functions χ become approximately orthonormal and can be chosen as the functions $\tilde{\varphi}$. However, for our basis set this approach cannot be followed because we have a considerable number of valence electrons outside the muffin-tin spheres (typically over 50%). Therefore it is desirable to have projection operators extending over all space. Each projector should in addition correspond to one of our basis functions (and therefore be centred at a certain atom), i.e., the function $\tilde{\varphi}$ should resemble the basis function χ as much as possible. This also means that requiring that the results are insensitive to the sphere radii is a very strong requirement.

It is well known [10] that Löwdin's symmetric-orthogonalization scheme [11] provides a set of orthonormal functions $\tilde{\varphi}_{\mathrm{I}}$ that resembles that of the original non-orthonormal functions χ_I as much as possible if the least-squares norm is taken as a measure of the difference. We shall therefore apply symmetric orthogonalization for defining orthogonal functions $\tilde{\varphi}_I$ and the corresponding projectors $\hat{P}_I$.

The procedure for symmetric orthogonalization is as follows. The overlap matrix $\mathbf{\Delta}$ of the basis functions $\{\chi_i\}$ is diagonalized:

$$\mathbf{\Delta} = \mathbf{USU}^{\dagger} \tag{15}$$

where $\mathbf{S}$ is a diagonal matrix containing the eigenvalues of $\mathbf{\Delta}$, and $\mathbf{U}$ is a unitary matrix whose columns are eigenvectors of $\mathbf{\Delta}$. The matrix $\mathbf{\Delta}^{-1/2}$ is obtained via

$$\mathbf{\Delta}^{-1/2} = \mathbf{US}^{-1/2}\mathbf{U}^{\dagger} \tag{16}$$

where $\mathbf{S}^{-1/2}$ is the diagonal matrix which is obtained by replacing each diagonal element S_I of $\mathbf{S}$ by $S_I^{-1/2}$. Letting $\boldsymbol{\chi}$ be the row vector of the original basis functions, the functions

$$\tilde{\boldsymbol{\varphi}} = \boldsymbol{\chi}\mathbf{\Delta}^{-1/2} \tag{17}$$

define a row vector $\tilde{\boldsymbol{\varphi}}$ of orthonormalized basis functions.

We shall use these functions in defining projection operators, and the constrained density-functional equations then take the matrix form

$$\left[\boldsymbol{h}_{\mathrm{KS}} + \sum_I{}' \lambda_I \mathbf{P}_I\right]\boldsymbol{c}_m = \varepsilon_m \mathbf{\Delta}\boldsymbol{c}_m \tag{18}$$

where $\boldsymbol{h}_{\mathrm{KS}}$ and $\mathbf{P}_I$ are the matrices of $\hat{h}_{\mathrm{KS}}$ and $\hat{P}_I$ in the original basis of the $\{\chi_I\}$, and $\boldsymbol{c}_m$ is a column vector containing the coefficients of the mth eigenfunction in this basis. The calculation of $\boldsymbol{h}_{\mathrm{KS}}$ and $\mathbf{\Delta}$ for our method can be found in reference [9]. The matrix elements for the projection operators are

$$\langle\chi_{I_1}|\hat{P}_I|\chi_{I_2}\rangle = \langle\chi_{I_1}|\tilde{\varphi}_I\rangle\langle\tilde{\varphi}_I|\chi_{I_2}\rangle \tag{19}$$

with

$$\langle\tilde{\varphi}_I|\chi_{I_i}\rangle = \sum_J (\mathbf{\Delta}^{1/2})_{JI_i}\langle\tilde{\varphi}_I|\tilde{\varphi}_J\rangle = (\mathbf{\Delta}^{1/2})_{II_i} \tag{20}$$

so we have

$$\langle\chi_{I_1}|\hat{P}_I|\chi_{I_2}\rangle = (\mathbf{\Delta}^{1/2})^*_{II_1}(\mathbf{\Delta}^{1/2})_{II_2}. \tag{21}$$

Thus, in order to include the constraints of equation (12), one does not actually have to carry out the transformation of equation (17)—the necessary matrix elements can be included via equation (21). It should be stressed in this context that in some cases the basis set appears to be nearly linear dependent, with the result that the calculation of $\mathbf{\Delta}^{-1/2}$ is numerically unstable. However, this is not the case for the calculation of $\mathbf{\Delta}^{1/2}$. We do, however, have to eliminate near-linear dependencies from the basis set, because otherwise the eigenvalue equation (18) becomes numerically unstable (this problem is not specific to CDFT). This can be done via the canonical transformation of Löwdin [12], where those linear combinations of the basis functions $\{\chi_I\}$ that correspond to the lowest eigenvalues S_I (i.e., those below a certain threshold) of the overlap matrix $\mathbf{\Delta}$ are excluded. As we found earlier for the unconstrained case, we found here that a reasonable variation of this threshold does not affect the results in the CDFT calculations either.

Once the projection operators are defined, we can also calculate the populations of equation (10). With our approach, they become the Löwdin populations

$$n_I = \sum_{m=1}^{\text{occ}} p_m|\langle\tilde{\varphi}_I|\psi_m\rangle|^2 = \sum_{m=1}^{\text{occ}} p_m\left|\sum_J c_{mJ}\langle\tilde{\varphi}_I|\chi_J\rangle\right|^2 = \sum_{m=1}^{\text{occ}} p_m\left|\sum_J c_{mJ}(\mathbf{\Delta}^{1/2})_{IJ}\right|^2. \tag{22}$$

With this approach, we have now defined a set of orthonormal basis functions. To each of those we can ascribe a certain atom as its site as well as an angular dependence. We can thus vary its population in a more or less controlled way, and by including the extra Hamilton matrix above with *chosen* λ_Is and by simultaneously calculating the populations of equation (22) together with the total energy (for the present method this is described in reference [9]), the function $\tilde{E}$ of equation (11) can be defined and compared with the model Hamiltonian, whereby the relevant model parameters can be determined.

It remains to be discussed how we take into account the fact that our LMTO basis set is better than minimal. Since this depends strongly on the system of interest, it is best discussed through a couple of examples.

For a H_2 molecule, we may want to change the occupation of one atom with respect to that of the other. If we place the molecule along the z-axis, the occupied orbital will contain contributions from the s, p_z, and d_{z^2} functions for the two different κs at each site. Thus, changing the population of one atom will require that we change the sum of all of those six populations. This is done by applying the same Lagrange multiplier to all six populations.

For planar organic molecules or conjugated polymers lying in the xz-plane, one may want to vary the number of π-electrons ascribed to a specific site. In this case the functions with π-symmetry contain contributions from p_y, d_{yz}, and d_{xy} basis functions. Since we have two κs for each L, there are six such functions for each atom, all of which will, in this case, be treated with the same Lagrange multiplier λ_I.

In both cases the populations of interest are the sums of the six $n_{\boldsymbol{R}L\kappa}$s.

Since one of our test systems is the H_2 molecule, we will discuss it in more detail, thereby emphasizing some further points. Our procedure corresponds to defining the

following two projection operators, one for each site (distinguished through their values of i):

$$\hat{P}_i = \sum_{\kappa=1}^{2} (\hat{P}_{\boldsymbol{R}_i \mathrm{s}\kappa} + \hat{P}_{\boldsymbol{R}_i \mathrm{p}_z\kappa} + \hat{P}_{\boldsymbol{R}_i \mathrm{d}_{z^2}\kappa}). \tag{23}$$

In principle, the application of more than one projector with corresponding Lagrange multipliers is not reasonable, since that corresponds to an overdetermination of the populations. However, if we apply both projectors, we get

$$(\hat{h}_{\mathrm{KS}} + \lambda_1 \hat{P}_1 + \lambda_2 \hat{P}_2)\psi = \varepsilon\psi \tag{24}$$

where ψ is the only occupied Kohn–Sham orbital (neglecting spin). This equation contains one constraint too many, since N_1 and N_2 cannot be varied independently, but we have $\hat{P}_1 + \hat{P}_2 = \hat{1}_\sigma$, where $\hat{1}_\sigma$ is the projector to the space of all eigenfunctions with σ-symmetry. Therefore equation (24) becomes

$$(\hat{h}_{\mathrm{KS}} + (\lambda_1 - \lambda_2)\hat{P}_1 + \lambda_2 \hat{1}_\sigma)\psi = \varepsilon\psi. \tag{25}$$

The term $\lambda_2 \hat{1}_\sigma$ gives only a constant shift of the eigenvalues, and does not affect the Kohn–Sham orbital ψ or the total energy (note that the total energy depends on ψ alone; by using ε—or, in the general case, the sum of eigenvalues of occupied orbitals—in the calculation of the total energy, one has to make a correction if extra Lagrange multipliers are included due to the extra constraints). Therefore the energy depends only on the difference $\lambda_1 - \lambda_2$. If we replace $\lambda_1 - \lambda_2$ by λ', equation (25) is equivalent to

$$(\hat{h}_{\mathrm{KS}} + \lambda' \hat{P}_1)\psi = \varepsilon\psi. \tag{26}$$

If we also apply Lagrange multipliers to our model Hamiltonian (this is the case whether we solve the model exactly or in the Hartree–Fock approximation; see below), the total energy also depends only on the difference of the Lagrange multipliers. On the other hand, if we were to define our projectors such that only the s orbitals would be involved (this would mean truncating the right-hand side of equation (23) after the terms $\hat{P}_{\boldsymbol{R}_i \mathrm{s}\kappa}$), the total energy would *not* depend on $\lambda_1 - \lambda_2$ alone. Thus, our choice of Lagrange multipliers is consistent with the model Hamiltonian.

4. Model calculations

4.1. The H_2 molecule

Due to its simplicity, the H_2 molecule is an excellent system on which to test new approaches, and therefore we will study it first. Neglecting spin, we construct a model based on two orthonormal basis functions φ_1 and φ_2, one per atom. The position-space part of the general two-electron wavefunction Ψ for a singlet state is then spanned by the three functions

$$\begin{aligned} \Phi_0(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \frac{1}{\sqrt{2}}(\varphi_1(\boldsymbol{r}_1)\varphi_2(\boldsymbol{r}_2) + \varphi_2(\boldsymbol{r}_1)\varphi_1(\boldsymbol{r}_2)) \\ \Phi_1(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \varphi_1(\boldsymbol{r}_1)\varphi_1(\boldsymbol{r}_2) \\ \Phi_2(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \varphi_2(\boldsymbol{r}_1)\varphi_2(\boldsymbol{r}_2) \end{aligned} \tag{27}$$

i.e.

$$\Psi = c_0\Phi_0 + c_1\Phi_1 + c_2\Phi_2. \tag{28}$$

c_0, c_1, and c_2 are determined by diagonalizing the following Hamilton matrix, which is the matrix representation of the Hamilton operator of equations (2)–(4) in the basis (Φ_0, Φ_1, Φ_2):

$$\mathbf{H} = \begin{pmatrix} -2t_{11} + U_{1212} + U_{1221} & -t_{12}\sqrt{2} + U_{1112}\sqrt{2} & -t_{12}\sqrt{2} + U_{1112}\sqrt{2} \\ -t_{12}\sqrt{2} + U_{1112}\sqrt{2} & -2t_{11} + U_{1111} & U_{1122} \\ -t\sqrt{2} + U_{1112}\sqrt{2} & U_{1122} & -2t_{11} + U_{1111} \end{pmatrix}. \tag{29}$$

We let the energy zero be at

$$-2t_{11} + U_{1212} + U_{1221} = 0 \tag{30}$$

and introduce

$$\begin{aligned} U &\equiv U_{1111} - U_{1212} - U_{1221} \\ t &\equiv t_{12}. \end{aligned} \tag{31}$$

Furthermore, we assume that

$$U_{1112} = U_{1122} = 0. \tag{32}$$

Then

$$\mathbf{H} = \begin{pmatrix} 0 & -t\sqrt{2} & -t\sqrt{2} \\ -t\sqrt{2} & U & 0 \\ -t\sqrt{2} & 0 & U \end{pmatrix}. \tag{33}$$

As discussed in section 3, the number of electrons that is ascribed to atom 1 can be controlled by adding the term $\lambda_1\hat{n}_1$ to the Hamilton operator, which produces the (constrained) Hamilton matrix

$$\mathbf{H}_c = \begin{pmatrix} \lambda_1 & -\sqrt{2}t & -\sqrt{2}t \\ -\sqrt{2}t & U + 2\lambda_1 & 0 \\ -\sqrt{2}t & 0 & U \end{pmatrix}. \tag{34}$$

The eigenvector $\boldsymbol{c} = (c_0, c_1, c_2)^{\mathrm{T}}$ for the lowest eigenvalue of $\mathbf{H}_c$ leads to the total energy

$$E_{\mathrm{tot}} = \boldsymbol{c}^{\dagger}\mathbf{H}\boldsymbol{c} = -2\sqrt{2}t\,\mathrm{Re}(c_0(c_1 + c_2)) + U(|c_1|^2 + |c_2|^2) \tag{35}$$

and the populations

$$\begin{aligned} n_1 &= |c_0|^2 + 2|c_1|^2 = 1 + |c_1|^2 - |c_2|^2 \\ n_2 &= |c_0|^2 + 2|c_2|^2 = 1 + |c_2|^2 - |c_1|^2 \end{aligned} \tag{36}$$

on using $|c_0|^2 + |c_1|^2 + |c_2|^2 = 1$.

So far, we have studied the exact solution to the model. Within the Hartree–Fock approximation, equation (28) is replaced by

$$\Psi = \psi(\boldsymbol{r}_1)\psi(\boldsymbol{r}_2) \tag{37}$$

with

$$\psi(\boldsymbol{r}) = a_1\phi_1(\boldsymbol{r}) + a_2\phi_2(\boldsymbol{r}). \tag{38}$$

This corresponds to

$$\begin{aligned} c_0 &= \sqrt{2}a_1a_2 \\ c_1 &= a_1^2 \\ c_2 &= a_2^2. \end{aligned} \tag{39}$$

Without loss of generality, we can chose c_0, c_1, and c_2 in equation (28) real. By normalization, we then have

$$c_0 = \pm\sqrt{1 - c_1^2 - c_2^2}. \tag{40}$$

The two signs correspond to two different wavefunctions that for $c_1 = -c_2$ become energetically degenerate.

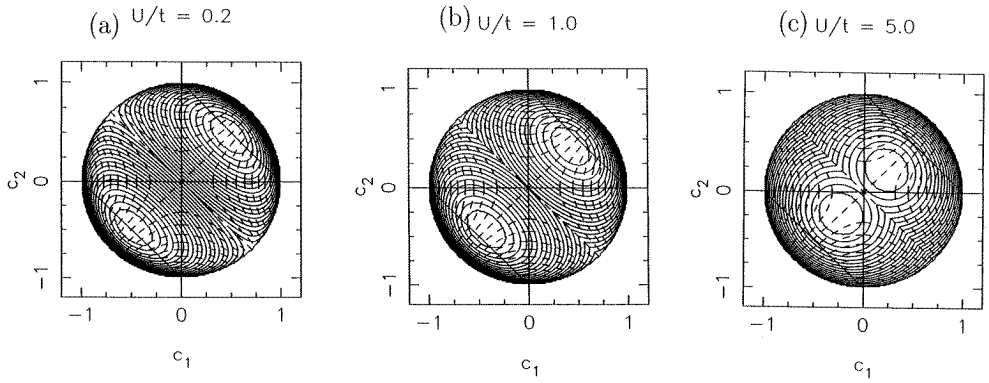

Figure 1. Contour plots of the lowest total energy for the H_2 molecule as a function of c_1, c_2 for different values of U/t, i.e. (a) $U/t = 0.2$, (b) $U/t = 1.0$, and (c) $U/t = 5.0$. The full curves are contour plots for the total energies at equidistant energies between the minimum ($-1.9025t$ in (a), $-1.5616t$ in (b), and $-0.7016t$ in (c); the two equivalent minima are placed along the $c_1 = c_2$ diagonal) and the maximum (i.e. U; this occurs for $c_1^2 + c_2^2 = 1$). The straight lines between (0, 1) and (1, 0) and between $(-1, 0)$ and $(0, -1)$ correspond to the Hartree–Fock wavefunctions. The dashed lines are contours for Δn of equation (41) for nine values equidistant between 0 and 1.

From equation (40), the lowest E_{tot} of equation (35) becomes a function of (c_1, c_2). Contour plots of this function are shown in figure 1, together with contour plots of the difference in occupation of the two atoms:

$$\Delta n = |c_1^2 - c_2^2|. \tag{41}$$

Within the Hartree–Fock approximation, it follows from equation (39) and from the normalization of ψ (i.e., $a_1^2 + a_2^2 = 1$) that the range of the constrained Hartree–Fock solution corresponds to $c_1 + c_2 = \pm 1$, with $0 \leqslant |c_1|, |c_2| \leqslant 1$.

The corresponding straight lines are depicted in figure 1. The Hartree–Fock approximation assumes that c_1 and c_2 are to lie on these lines and that the energy minimum is at $c_1 = c_2 = \pm 0.5$. From figure 1 we see that this is only realistic if $U/t \ll 1$.

4.2. Hartree–Fock calculations for linear C chains

Ultimately we intend to apply our method to conjugated polymers. These are quasi-one-dimensional materials which have a planar backbone of sp^2-bonded carbon atoms with the last valence electron per carbon atom occupying π-orbitals perpendicular to the plane of the backbone. The π-orbitals are the frontier orbitals, and are those of relevance when studying low-energy excitations. A Peierls distortion can be held responsible for the occurrence of alternating longer (double) and larger (single) bonds.

A linear carbon chain is the simplest possible model system resembling the conjugated polymers. It has two (and not one) π-orbitals per carbon atom and the σ-bonds are formed

by sp (and not sp^2) hybrids. But, as above, the frontier orbitals are defined by the π-orbitals, and the ground state for the neutral chain has alternating shorter (triple) and longer (single) bonds.

Assuming that there is no interaction between the two π-systems, the simplest many-body Hamiltonian for this system is the Hubbard Hamiltonian for each π-system:

$$\hat{H} = \sum_{i,\sigma} \varepsilon c_{i\sigma} c_{i\sigma}^{\dagger} - \sum_{i,\sigma} t_{i,i+1}(c_{i\sigma}^{\dagger} c_{i+1,\sigma} + c_{i+1,\sigma} c_i) + \sum_i U c_{i\uparrow}^{\dagger} c_{i\uparrow} c_{i\downarrow}^{\dagger} c_{i\downarrow} \quad (42)$$

where the hopping integrals $t_{i,i+1}$ alternate between t_1 and t_2 for the shorter and longer bonds, respectively.

We justify the omission of interactions between the two sets of π-electrons as follows. The simplest model corresponds to a constant repulsion $\tilde{U}$ between different spin orbitals of p type centred on a given atom. We thereby obtain (r discriminates between the p_x and the p_y orbital)

$$\sum_{\sigma\neq\sigma' \text{ or } r\neq r'} \frac{1}{2}\tilde{U} c_{i\sigma r}^{\dagger} c_{i\sigma r} c_{i\sigma' r'}^{\dagger} c_{i\sigma' r'} \approx \sum_{\sigma\neq\sigma' \text{ or } r\neq r'} \tilde{U}\langle c_{i\sigma r}^{\dagger} c_{i\sigma r}\rangle c_{i\sigma' r'}^{\dagger} c_{i\sigma' r'} = 3\sum_{\sigma' r'} \tilde{U} m_i c_{i\sigma' r'}^{\dagger} c_{i\sigma' r'} \quad (43)$$

with

$$m_i = \langle c_{i\sigma r}^{\dagger} c_{i\sigma r}\rangle \quad (44)$$

being independent of r and σ.

The result in equation (43) is three times the result that we would get if there was no repulsion between the p_x and p_y functions for the same atoms. Therefore the model with repulsion between the p_x and p_y orbitals can be approximated by a model with no repulsion and an effective $U = 3\tilde{U}$. The approximation that we have made is that we neglected the off-diagonal terms (like $\langle c_{i\sigma r}^{\dagger} c_{i\sigma r'}\rangle c_{i\sigma r'}^{\dagger} c_{i\sigma r}$, for $r \neq r'$). However, it should be stressed that our main goal is to demonstrate the performance of the CDFT approach and not to study the Hubbard model in detail. Therefore, we consider the question of the exact justification of our Hubbard model less relevant.

In the constrained density-functional calculations, we consider chains of periodically repeated units of four C atoms, and accordingly modify the populations of every fourth atom identically.

Applying Born–von Kármán boundary conditions to a dimerized chain with four atoms per unit cell and N unit cells, where $N \geqslant 3$, the Fock matrix (with Lagrange multipliers λ_i) can be written as a cyclic hypermatrix of dimension $N \times N$ of the form

$$\mathbf{F}_c = \begin{pmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{B}^{\dagger} \\ \mathbf{B}^{\dagger} & \mathbf{A} & \mathbf{B} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{\dagger} & \mathbf{A} & \mathbf{B} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{B}^{\dagger} & \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{B}^{\dagger} & \mathbf{A} \end{pmatrix} \quad (45)$$

with

$$\mathbf{A} = \begin{pmatrix} \varepsilon + \lambda_1 & -t_1 & 0 & 0 \\ -t_1 & \varepsilon + \lambda_2 & -t_2 & 0 \\ 0 & -t_2 & \varepsilon + \lambda_3 & -t_1 \\ 0 & 0 & -t_1 & \varepsilon + \lambda_4 \end{pmatrix} + \mathbf{G} \quad (46)$$

and

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -t_2 & 0 & 0 & 0 \end{pmatrix}. \tag{47}$$

$\mathbf{G}$ is a diagonal matrix, which stems from the two-electron interactions. The calculation of $\mathbf{G}$ is described below (equation (51)). The cyclic hypermatrix $\mathbf{F}_c$ can be block diagonalized with the help of the hypermatrix $\mathbf{U}$, given by

$$\mathbf{U}_{kl} = \frac{1}{\sqrt{N}} \mathrm{e}^{2\pi \mathrm{i}(k-1)(l-1)/N} \mathbf{I} \tag{48}$$

where $\mathbf{I}$ is the 4×4 unit matrix.

The kth diagonal block $\tilde{\mathbf{F}}_k$ of the matrix $\mathbf{U}^\dagger \mathbf{F}_c \mathbf{U}$ is then given by

$$\tilde{\mathbf{F}}_k = \mathbf{A} + \mathbf{B}\mathrm{e}^{2\pi \mathrm{i}(k-1)/N} + \mathbf{B}^\dagger \mathrm{e}^{2\pi \mathrm{i}(k-1)(N-1)/N} = \mathbf{A} + \mathbf{B}\mathrm{e}^{2\pi \mathrm{i}(k-1)/N} + \mathbf{B}^\dagger \mathrm{e}^{-2\pi \mathrm{i}(k-1)/N}. \tag{49}$$

In our calculations, we determine the eigenvectors $\boldsymbol{c}_{kj}$ of the matrices $\tilde{\mathbf{F}}_k$ as functions of k. Assuming that for each k there are n_{occ} occupied orbitals (in our case $n_{\mathrm{occ}} = 2$), we obtain the density matrices

$$\mathbf{R}_k = \sum_{j=1}^{n_{\mathrm{occ}}} \boldsymbol{c}_{kj} \boldsymbol{c}_{kj}^\dagger. \tag{50}$$

From these we can calculate the elements of the diagonal matrix $\mathbf{G}$ via

$$G_{ii} = \frac{U}{N} \sum_{k=1}^{N} R_{kii}. \tag{51}$$

We start with an initial guess for the density matrices $\mathbf{R}_k$, which allows us to calculate the block-diagonalized Fock matrices $\tilde{\mathbf{F}}_k$. This gives in turn a new set of density matrices. This procedure is repeated until self-consistency is achieved. To avoid oscillations we had to apply a feedback procedure for the density matrices.

For an undimerized C chain in the absence of Lagrange multipliers, the HOMO is degenerate with the LUMO for the first k-point. This may easily lead to oscillations. Although the degeneracy is lifted as Lagrange multipliers are included, the corresponding calculations may oscillate too. We therefore introduced a broadening for nearly degenerate orbitals, whereby equation (50) is replaced by

$$\mathbf{R}_k = \sum_{j=1}^{n_{\mathrm{occ}}-1} \boldsymbol{c}_{kj} \boldsymbol{c}_{kj}^\dagger + (1 - w(\Delta\varepsilon)) \boldsymbol{c}_{k,n_{\mathrm{occ}}} \boldsymbol{c}_{k,n_{\mathrm{occ}}}^\dagger + w(\Delta\varepsilon) \boldsymbol{c}_{k,n_{\mathrm{occ}}+1} \boldsymbol{c}_{k,n_{\mathrm{occ}}+1}^\dagger. \tag{52}$$

$\Delta\varepsilon$ is the difference in energy between the LUMO and the HOMO of the k-point in question. We choose w such that $w(0) = 0.5$; w is decreasing linearly up to the threshold value $\Delta\varepsilon_{\mathrm{trs}} = 0.001$ Ryd, for and above which $w(\Delta\varepsilon)$ is 0.

As soon as self-consistency is achieved, the population of the ith atom can be calculated as

$$N_i = \frac{4}{N} \sum_{k=1}^{N} R_{kii}. \tag{53}$$

The factor 4 arises because there are two perpendicular π-systems each having $2n_{\mathrm{occ}}$ electrons per unit cell. The total energy is calculated from the expectation value of the matrix $\mathbf{H}$, which is defined as the Fock matrix $\mathbf{F}_c$ of equation (45) except that the matrix

G of equation (46) is replaced by $\frac{1}{2}\mathbf{G}$ and that all λ_i are set to 0. Block diagonalization of **H** leads to diagonal blocks $\tilde{\mathbf{H}}_k$, and the total energy, normed to $4n_{\mathrm{occ}}$ electrons, is finally given by

$$E_{\mathrm{tot}} = \frac{4}{N}\sum_{k=1}^{N} \mathrm{Tr}\{\mathbf{R}_k\tilde{\mathbf{H}}_k\}. \tag{54}$$

In contrast to the case for H_2, we are not able to obtain exact solutions for the C chain. As we shall see, our results for H_2 raise some doubts about whether the Hartree–Fock approximation is justified. However, in the present work the focus is not on obtaining accurate eigensolutions to the model Hamiltonian, but on demonstrating that constrained density-functional calculations can provide parameter values that depend only weakly on internal parameters (e.g. sphere sizes for the muffin-tin orbitals) of the density-functional calculations, even for more delocalized electrons. We therefore consider our approach to be acceptable here.

5. Results

5.1. Results for the H_2 molecule

In the density-functional calculations for the H_2 molecule, we used a H–H bond length of 1.40 au and considered muffin-tin sphere radii of 0.5, 0.6, and 0.7 au. In the absence of Lagrange multipliers this corresponds to having only 0.11, 0.17, and 0.24 electrons, respectively, inside each of the two muffin-tin spheres. The Lagrange multipliers were set equal to 0, ± 0.05, ± 0.1, ± 0.15, ± 0.2, ± 0.25, and ± 0.30 Ryd.

The resulting E_{tot} as a function of Δn_1 was subsequently approximated through the first terms in a Taylor expansion:

$$E_{\mathrm{tot}} = E_0 + b\,\Delta n_1^2. \tag{55}$$

Least-squares fits for the three sets of calculations to equation (55) led to $b = 1.924$, 1.940, and 2.141 Ryd, respectively, for the three sphere radii. It is very important to note that b is fairly independent (within 6% from the mean value) of the muffin-tin sphere radii. This is a far from trivial finding, since a very large proportion of the electrons are in the interstitial region, so a separation into atomic components is not obvious.

In order to determine the parameters of the model, we performed model calculations with the same set of Lagrange multipliers. However, determining both U and t requires more information than the constrained density-functional calculations provide. We use therefore in addition the experimental value of the optical gap ($\Delta E_{\mathrm{gap}} = 0.8220$ Ryd [13]), which is the energy of the lowest $^1\Sigma_{\mathrm{u}}$ state relative to the ground state.

The optical gap for the model calculations is given by the difference of the two lowest eigenvalues of the Hamilton matrix **H** defined in equation (33), i.e.

$$\Delta E_{\mathrm{gap}} = U - \frac{U - \sqrt{U^2 + 16t^2}}{2} \tag{56}$$

or

$$U = \frac{\Delta E_{\mathrm{gap}}^2 - 4t^2}{\Delta E_{\mathrm{gap}}}. \tag{57}$$

Within the Hartree–Fock approximation, one may use $\Delta E_{\mathrm{gap}} = 2t$ by taking the difference between the Hartree–Fock eigenvalues corresponding to the excited and the lowest

one-electron orbitals. Note that this will lead to different values of t in the Hartree–Fock and the exact calculations. For the sake of comparison, we therefore performed calculations with the same values of t for the exact and the Hartree–Fock calculations.

We followed accordingly three different approaches for obtaining the model parameters from the results of the constrained density-functional calculations.

(i) Exact calculations varying both t and U according to equation (57).

(ii) Fixing $t = 0.133$ Ryd (which is the value from (i) for a muffin-tin radius of 0.6 au) and performing both exact and Hartree–Fock calculations.

(iii) Fixing $t = 0.411$ Ryd, which is half the optical gap (and thus the appropriate choice for Hartree–Fock calculations), and performing both exact and Hartree–Fock calculations.

The results of (i) are given in table 1, whereas the results of (ii) and (iii) are given in table 2. Figures 2 and 3 show the results of the constrained density-functional calculations at $r_{\mathrm{MTS}} = 0.6$ au, and the corresponding exact and Hartree–Fock calculations with $t = 0.133$ and $t = 0.411$ Ryd, respectively.

Table 1. Results of fitting the constrained density-functional calculations to exact solutions of the Hubbard model for the H_2 molecule. All quantities are given in Rydberg atomic units. r_{MTS} is the radius of the muffin-tin spheres in the constrained density-functional calculations, and 'Curvature' is the curvature of ΔE_{tot} versus Δn_1; cf. equation (55).

r_{MTS}:	0.5	0.6	0.7
Curvature	1.924	1.940	2.141
t	0.134	0.133	0.126
U	0.735	0.735	0.745

Table 2. The Hubbard U for the H_2 molecule obtained by fitting the curvatures of constrained density-functional calculations to model calculations at two values of the hopping integral t. U_{ex} and U_{HF} are the values of U for exact and Hartree–Fock calculations, respectively. All of the quantities are given in Rydberg atomic units. r_{MTS} is the radius of the muffin-tin spheres, and 'Curvature' is the curvature of ΔE_{tot} versus Δn_1; cf. equation (55)

r_{MTS}:	0.5		0.6		0.7	
Curvature:	1.924		1.940		2.141	
t	U_{ex}	U_{HF}	U_{ex}	U_{HF}	U_{ex}	U_{HF}
0.133	0.733	3.581	0.735	3.612	0.766	4.014
0.411	1.199	3.025	1.193	3.057	1.277	3.458

As can be seen from table 1, t and U for the exact calculations vary only little, indicating that our constrained density-functional approach is stable against variations in the computational details. From table 2 we learn also that fixing t may lead to other values of U, but that these are also stable. On the other hand, the Hartree–Fock calculations lead to markedly different values of U. This is, however, a problem of the Hartree–Fock approximation to the model and not of the constrained density-functional method, as can be seen in figure 1.

According to table 1, $U/t \approx 5$. Figure 1 shows that for the ground state, $c_1 = c_2 \approx 0.2$. Changing n_1 (adding the constraints) corresponds to moving away from $c_1 = c_2$. The system will respond to this change by finding that pair (c_1, c_2) that leads to the smallest

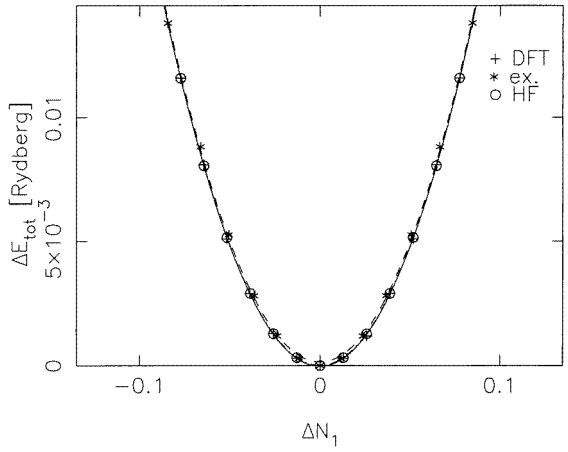

Figure 2. Results of constrained density-functional calculations for H_2 ($r_{MTS} = 0.6$ Bohr) and corresponding fitted parabolas for exact and Hartree–Fock calculations with $t = 0.133$ Ryd. Note that the dashed parabola is the least-squares fit to the exact calculations, whereas the curves for the constrained density-functional data and for the Hartree–Fock calculations coincide.

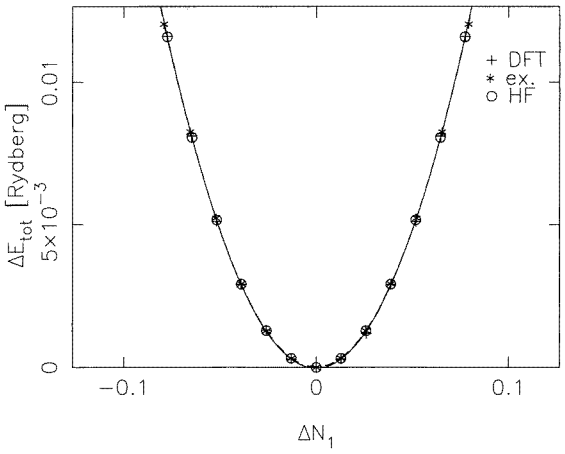

Figure 3. Results of constrained density-functional calculations for H_2 ($r_{MTS} = 0.6$ Bohr) and corresponding parabolas for exact and Hartree–Fock calculations with $t = 0.411$ Ryd.

change in total energy for a given change in n_1. From figure 1 we see that this is found upon shifting (c_1, c_2) away from $(0.2, 0.2)$ towards either $(1, 0)$ or $(0, 1)$.

The Hartree–Fock approximation assumes that the ground state occurs for $c_1 = c_2 = 0.5$, which is only valid for $U/t \ll 1$; cf. figure 1. This means that it is assumed that the changes in E_{tot} and n_1 follow the same patterns as that of figure 1(a). In that figure, E_{tot} changes more slowly on changing n_1 than in figure 1(c). Thus for a given set of $(\Delta E_{tot}, \Delta n_1)$, U has to be treated as much larger than it actually is, in order to describe the change correctly. This explains the difference between the exact and the Hartree–Fock results. It also demonstrates that the Hartree–Fock approximation may not be very good when U becomes comparable to or larger than t.

From figures 2 and 3, we see that the constrained density-functional calculations lead not only to approximately the same curvatures but also to sets of points lying on roughly the same curve. We see also that both the exact model and the Hartree–Fock approximation are able to fit the major points of the constrained density-functional results, except the exact calculations for the largest Δn_1.

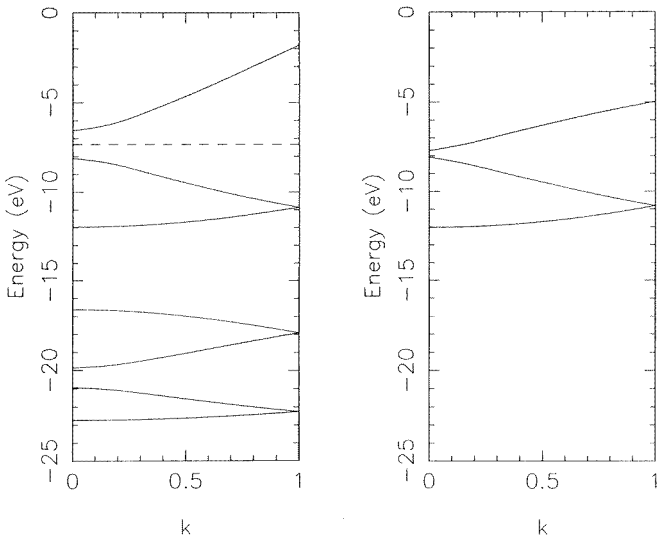

Figure 4. Band structures of a dimerized C chain from (left) the density-functional calculations and (right) the tight-binding fit. The dashed line marks the Fermi level in the DFT band structure.

5.2. Results for linear C chains

For the linear carbon chain we considered both undimerized and dimerized structures, i.e. chains with not-alternating and alternating bond lengths. The CDFT calculations for the dimerized C chain were carried through on systems containing repeated units of four atoms, and the bond lengths were set equal to 2.375 and 2.70 au [14], whereas we used the arithmetic mean (2.5375 au) for the undimerized chain.

As the radius for the muffin-tin spheres, we used $r_{\mathrm{MTS}} = 1.187$ au, which means that for the dimerized chain along the shorter bonds the muffin-tin spheres almost touch. We applied Lagrange multipliers to the π-functions of one atom (the second atom in the unit cell) in four, and their values were 0, ± 0.025, ± 0.05, ± 0.075, ± 0.100, and ± 0.125 Ryd. In our calculations we used seven k-points, the first and last being at the centre and the edge of the Brillouin zone, and the rest being equidistantly distributed. With Born–von Kármán boundary conditions, the seven k-points correspond to a cyclic system with $N = 12$ unit cells. We stress however that in our CDFT program we do not use Born–von Kármán boundary conditions, but consider an infinite chain. On the other hand, in our model calculations we used $N = 12$. We checked that enlarging N did not change the results for the dimerized C chain. For the undimerized C chain, the curvatures were 6% higher for $N = 288$ (which we consider converged) than for $N = 12$.

In order to analyse our results, we define a total population shift ΔN_{tot} via

$$\Delta N_{\mathrm{tot}} = \mathrm{sgn}(\Delta N_2)\sqrt{\sum_{i=1}^{4}\Delta N_i^2} \tag{58}$$

where ΔN_i is the population shift of the ith atom.

Our calculations showed that it was crucial that the Coulomb summations were carried through to much higher precision than is necessary in conventional (unconstrained) DFT calculations. If the range for the Coulomb summations is too small (in an actual calculation, a cut-off distance of 50 au turned out to be too small, whereas a distance above 60 au was sufficiently large), the plots of ΔE_{tot} versus ΔN_{tot} become asymmetric, and the lowest total energy is not at $\Delta N_{\mathrm{tot}} = 0$.

We studied the consequences of varying various internal parameters, such as the radii of the muffin-tin spheres, and the ranges of lattice summations other than the Coulomb summations. In all cases, our data points varied only little. The same was true when enlarging the range of Coulomb summations to over 70 au.

We first discuss the results for the dimerized chain. In order to perform HF model calculations, we fitted the band structures from the calculation with the Lagrange multipliers equal to 0 with those of a tight-binding model. This model is obtained from the HF model discussed in section 4.2 upon setting the matrix **G** of equation (51) to 0. The band structures of the density-functional calculation and of the tight-binding fit are shown in figure 4, and the fitted parameter values were $\varepsilon = -0.580$ Ryd, $t_1 = 0.159$ Ryd, $t_2 = 0.145$ Ryd.

The bands of the HF calculations in the absence of Lagrange multipliers are shifted by $U/2$ with respect to the bands of the tight-binding model. Therefore our tight-binding parameters remain valid except that ε has to be replaced by $\varepsilon - U/2$. This constant shift can be neglected, since it has no effect on ΔE_{tot} as a function of ΔN_{tot}.

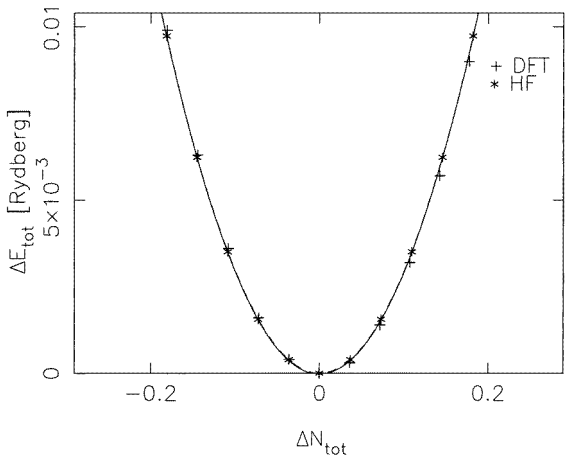

Figure 5. Results of constrained density-functional calculations for a dimerized C chain, and the corresponding parabola from the Hartree–Fock fit.

Our CDFT calculations result in the curve of figure 5. A least-squares fit of the form $\Delta E_{\mathrm{tot}} = a + b\,\Delta N_{\mathrm{tot}}^2$ leads to $b = 0.2953$ Ryd, from which we obtained $U = 1.840$ Ryd. This curve is also shown in figure 5. It can be seen that on the left-hand side of the plot, the HF results are very close to the CDFT data points, whereas on the right-hand

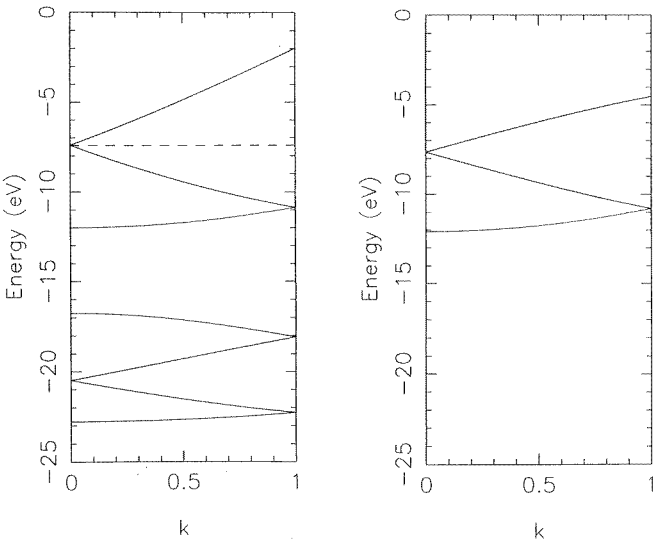

Figure 6. As figure 4, but for an undimerized C chain.

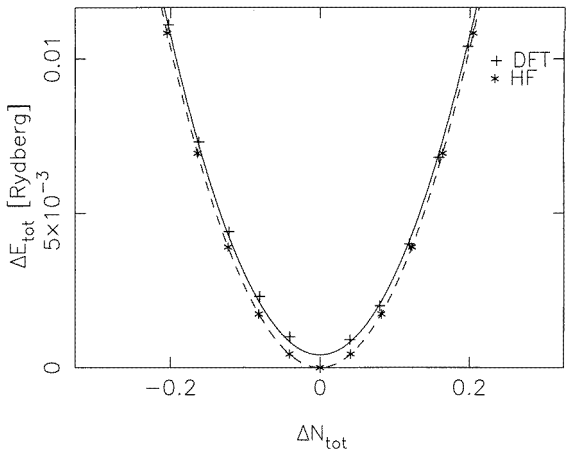

Figure 7. As figure 5, but for an undimerized C chain.

side, the data points differ, although the deviations are only small. For an undimerized C chain, the tight-binding fit of the band structures yielded $\varepsilon = -0.563$, and $t = 0.163$ Ryd (cf. figure 6). The curvature b of the $(\Delta N_{\mathrm{tot}}, \Delta E_{\mathrm{tot}})$ curve (figure 7) was 2.2582, which leads to $U = 1.671$ Ryd. In figure 7 we see that the fitted total-energy curve has a non-vanishing offset. This may be due to the fact that the undimerized C chain is metallic, so the Hubbard model with its implicit assumption of strong screening is an unrealistic starting point.

6. Conclusions

In the present work we have concentrated on presenting a generalization of the constrained density-functional methods. Such methods are applied in order to determine parameters for many-particle models from first-principles calculations, but have so far been applicable only for systems with strongly localized electrons and when the first-principles calculations have been performed with an essentially minimal basis set. Our approach allows for calculating these parameters also for systems with considerably less localized electrons and when the basis set is better than minimal.

The central point was that of transforming the original basis set of non-orthonormal functions into one of orthonormal functions that both resemble the original functions to a large extent and are so constructed that the final results depend only marginally on the calculational details of the density-functional calculations. To this end, we found that the symmetric-orthonormalization scheme of Löwdin provides the optimal methodology, whereby both (i) constraints fixing the number of electrons of specific atoms and (l, m) and (ii) the related populations were well defined.

A number of test calculations on the H_2 molecule showed that the results were stable against variations of the internal parameters of the density-functional calculations, whereby it turned out that the sizes of the muffin-tin spheres were the most critical quantities. The parameter values of the many-body model Hamiltonian can only be obtained by comparing the results of the constrained density-functional calculations with those of model calculations. For the H_2 molecule we could compare different ways of carrying the model calculations through, i.e., a so-called exact method and a Hartree–Fock method. We found that the results depended strongly on which method was used, and, in particular for large values of U, strong deviations were observed. This implies that whenever model parameters are determined with constrained density-functional calculations, it is important to specify how the accompanying model calculations have been performed and that the same approach is used in later applications of the model.

For the linear carbon chains, we studied the model only within the Hartree–Fock approximation. The fact that U was found to be large indicates that this approximation may not be fully justified. Our results for the H_2 molecule indicate that more exact model calculations would lead to a smaller value of U.

For the extended, quasi-one-dimensional systems, $U/4t$ (i.e., the strength of the many-body interactions over the total bandwidth) is of fundamental importance. For the carbon chains, we found values of around 2.5–3.2. (For $\tilde{U} = U/3$ of equation (43), these values are accordingly around 0.8–1.1.) We are not aware of any other estimates of this quantity for the carbon chains, but it is useful to compare the values with those that are considered realistic for the prototypical conjugated polymer, *trans*-polyacetylene. According to page 60 of Baeriswyl *et al* [15], a reasonable parameter range is $1.5t_0 < U < 3.5t_0$, where t_0 is something like a mean hopping integral. This results in $0.4 < U/4t_0 < 0.9$. However, in contrast to our study, that of Baeriswyl *et al* also considered the nearest-neighbour repulsion V, for which they regard as a reasonable range $0.5t_0 < V < 1.5t_0$. Therefore, for comparison with our values, we should use an effective $U_{\mathrm{eff}} = U - V$. This leads to $0 < U_{\mathrm{eff}} < 3t_0$, and $0 < U_{\mathrm{eff}}/4t_0 < 0.8$. Therefore our values for $U/4t$ and even for $\tilde{U}/4t$ for the C chains are rather high, which could be due to the Hartree–Fock approximation that we have used in determining U.

In conclusion, we stress that our aim has been to provide a scheme that allows for a first-principles determination of parameter values for many-body model Hamiltonians for relatively delocalized electrons, and not to provide exact results for our test systems. Our

results show that we have achieved this goal, and that our results are as independent of the computational details as one can hope for.

Acknowledgments

The authors thank R C Albers, J T Gammel, S Satpathy, and S Mazumdar for valuable discussions. HM was supported by the Deutsche Forschungsgemeinschaft (DFG) through project No II C1-Sp 439-2/1. Generous support by the Fond der Chemischen Industrie is gratefully acknowledged.

References

[1] Dederichs P H, Blügel S, Zeller R and Akai H 1984 *Phys. Rev. Lett.* **53** 2512
[2] Hybertsen M S, Schlüter M and Christensen N E 1989 *Phys. Rev.* B **39** 9028
[3] Hybertsen M S, Stechel E B, Schlüter M and Jennison D R 1990 *Phys. Rev.* B **41** 11 068
[4] Zhang Z and Satpathy S 1991 *Phys. Rev.* B **44** 13 319
[5] Alouani M, Albers R C and Willis J M 1993 *Synth. Met.* **55–57** 3358
[6] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[7] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[8] von Barth U and Hedin L 1972 *J. Phys. C: Solid State Phys.* **5** 1629
[9] Springborg M and Andersen O K 1987 *J. Chem. Phys.* **87** 7125
[10] Carlson B C and Keller J M 1957 *Phys. Rev.* **105** 102
[11] Löwdin P O 1950 *J. Chem. Phys.* **18** 365
[12] Löwdin P O 1956 *Adv. Phys.* **5** 1 (see in particular pp 40–56)
[13] Herzberg G 1950 *Spectra of Diatomic Molecules (Molecular Spectra and Molecular Structure I)* (New York: Van Nostrand Reinhold)
[14] Pohl A, Meider H and Springborg M 1994 *J. Mol. Struct. (Theochem)* **305** 165
[15] Baeriswyl D, Campbell D K and Mazumdar S 1992 *Conjugated Conducting Polymers* ed H G Kiess (New York: Springer)