# Ab-initio Gutzwiller-DFT studies of diluted magnetic semiconductors

Dissertation
zur Erlangung des akademischen Grades
Dr. rer. nat.
der Fakultät Physik an der
Technischen Universität Dortmund

vorgelegt von

Thorben Linneweber
aus Dortmund

Dezember 2016

| | |
|---|---|
| Erster Gutachter | Priv.-Doz. Dr. Ute Löw |
| Zweiter Gutachter | Prof. Dr. Florian Gebhard |
| Tag der mündlichen Prüfung | 09.02.2017 |
| Version der Arbeit | 10.03.2017 |

Kontakt zum Autor thorben.linneweber@tu-dortmund.de

# Contents

# 1 Introduction

Strongly correlated electron systems, i.e., systems that cannot be fully described by a theory of free particles, play an important role in solid state physics. Their understanding and adequate description are mandatory for the explanation of emergent phenomena, such as high-temperature superconductivity or magnetism. Materials exhibiting these phenomena typically incorporate atoms with partly filled $d$- or $f$- shells in which electrons are subject to strong Coulomb interactions. Unfortunately, a standard tool of solid state physics, the Density Functional Theory (DFT), often fails to correctly predict certain properties of correlated systems. The failures are caused by intrinsic difficulties to account for strong electron-electron correlations within the DFT formalism. Therefore, a huge amount of research is devoted into curing these shortcomings by combining standard DFT methods with more elaborated many-particle descriptions, such as the dynamical mean field theory (DMFT) or the Gutzwiller variational method.

In this work we employ the latter method to study the magnetic properties of Diluted Magnetic Semiconductors (DMS). DMS are semiconductors doped with $3d$ transition-metal ions ("magnetic ions"). They combine two branches of solid state physics, namely semiconductor physics and magnetism. The most common DMS materials are compounds of the form $A_{1-x}M_xB$, i.e., a certain fraction $x$ of the cations A of a semiconductor AB is replaced by magnetic ions. Even though the doping concentration in these compounds is small, they display a variety of interesting magnetic and magneto-optical properties which are focus of recent studies far to extensive to summarize here. For a comprehensive review we refer to [1] and references therein.

DMS can be divided into two classes. In the first class, the magnetic ions isovalently replace semiconductor cations without introducing free charge carriers into the material. These DMS were subject of early experimental studies [2] which found a weak antiferromagnetic exchange between the magnetic ions dominantly caused by the superexchange mechanism. The studies concentrated mostly on II-VI based semiconductors in which magnetic ions can be easily incorporated in arbitrary concentrations. In contrast to II-VI semiconductors, the equilibrium solubility of transition-metal ions in III-V semiconductors is very low [3]. Advanced crystal growing methods like the Molecular Beam Epitaxy (MBE) are necessary for the production of samples with significant impurity concentrations x>1%. With the improvement of this method the focus of interest shifted towards the second class of DMS, the ferromagnetic semiconductors. Here, a hole-mediated ferromagnetic exchange between the magnetic impurities dominates. The most studied representative of this material class is $Ga_{1-x}Mn_xAs$, which exhibits experimentally observed Curie temperatures of up to 170 K [4]. The ultimate goal for this class of DMS is to find materials that exhibit ferromagnetism above room-temperature. Ferromagnetic DMS are considered as possible building blocks for electronic semiconductor devices which operate not only on charge but also incorporate spin degrees of freedom, so-called "spintronics".

It is believed [5] that spintronics are key to combine high data processing speeds with low energy consumption, significantly improving conventional semiconductor devices used in today's computing machines.

In this work we restrict our investigations to the study of manganese doped semiconductors. Existing semi-phenomenological models for manganese doped DMS are based on the assumption that the $d$-shells of the manganese impurities are in their atomic (half-filled) Hund's rule ground state and a description using effective $s = 5/2$ spins is valid. We employ density functional results in combination with the Gutzwiller variational method to review these assumptions. As an application we calculate effective exchange constants $J$ for the Heisenberg-type coupling between the Mn impurities embedded in various semiconductor hosts.

The thesis is structured as follows. In Chapter 2 we recapitulate the foundations of density functional theory and discuss some of the general problems that arise due to approximations in the exchange-correlation energy functional. We introduce DFT-based Wannier orbitals which are suitable to create ab-initio tight-binding models and discuss this procedure exemplary for the semiconductor CdTe. In Chapter 3 we introduce the multiband Gutzwiller variational method which we use to find approximative many-body ground-states of multiband Hubbard models. We outline the general idea behind the method and explain the numerical minimization procedure we employ. The interpretation of the obtained band structures as Landau quasi-particles closes this chapter.

In Chapter 4 the results from Chapter 2 and Chapter 3 are combined. We present the concept of large unit cells, so-called "supercells", to account for the small doping concentration in DMS. We introduce a linear-superposition scheme to set up effective tight-binding models for DMS supercells with arbitrary impurity positions, based on Wannier orbitals from DFT calculations. A local interaction term is added to the tight-binding models which accounts for the local correlations in the $d$-shells of the impurities. We discuss the problem of "double-counting" and present an algorithm to apply the Gutzwiller variational method to the final multiband Hubbard model for DMS, containing multiple impurities per unit cell. The question of the strength of the local interaction is partly addressed for in a constrained DFT calculation from which we extract local Coulomb parameters for some host-dopant combinations.

In Chapter 5 we apply our Gutzwiller-DFT scheme to manganese in various II-VI semiconductors. The exchange couplings are calculated and compared with exchange couplings obtained from pure DFT+U calculations. The coupling constants are used in a simple spin-cluster model to calculate magnetization curves for small doping concentrations. Finally, the result of a non-integer filling of the $d$-shell which seems to oppose the picture of localized $s = 5/2$ spins is resolved by the investigation of a toy model whose exact ground-state is determined using the Lanczos algorithm.

In Chapter 6 we apply the Gutzwiller-DFT method to manganese doped GaAs. The local $d$-shell spin- and occupation-distributions that are central results of the Gutzwiller treatment are compared with the results from II-VI semiconductors. As an application of our method we calculate exchange couplings using a large two-impurity supercell.

Conclusions and outlook end the thesis, some formal derivations are deferred to the appendix.

# 2 Density functional theory

> "The general theory of quantum mechanics is now almost complete, [...]. The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation."
> – Paul Dirac, 1929 [6]

There are two ways to approach the study of phenomena emerging from the many-body problem in solid-state physics. A first approach is "model-based" where a model Hamiltonian is built that captures the physical phenomena one is interested in, based on the understanding of the relevant mechanism present in the full problem. The description of magnetic materials by spin Hamiltonians is a good example: A huge part of the full problem is "downfolded" onto a more abstract and simpler spin Hamiltonian that can be studied for complex magnetic systems. The input parameters of the model Hamiltonian are usually derived from experiment or calculations following the second approach, the so-called "ab-initio" calculations, where quantum mechanical systems are studied without any adjustable parameters.

Unfortunately, solving such systems exactly, i.e., determining the full energy spectrum including all eigenfunctions, is only possible for small systems, e.g., certain atoms and very small molecules. Calculations for larger many-body systems are hitting the "exponential wall" [7], namely the computational effort grows exponentially with the number of particles in the system. The development of Density Functional Theory (DFT) which circumvents this problem for ground-state properties was therefore an immense achievement and led to the award of the nobel prize to W. Kohn and J. Pople in 1998. With the huge growth in computational power, present-day implementations of DFT allow the investigation of ground-state properties of systems with several hundred atoms, which, for example, helps to get insights into systems with complicated magnetic structures. In the last 20-30 years DFT indisputably became a main pillar of modern solid-state physics.

In this chapter we discuss the foundations of DFT as well as problems connected to approximations made within this formalism. Some of the implementation details of the DFT are discussed as far as they become important in the following chapters.

## 2.1 Fundamental solid-state Hamiltonian

We present the DFT in the framework of solid-state physics, even though it is also applicable to other fields. In the non-relativistic limit the universal microscopic Hamiltonian for a solid in first quantization is given by

$$\begin{aligned}\hat{H}^{1Q}_{\text{solid}} = &\sum_n \frac{\hat{\mathbf{P}}_n^2}{2M_n} + \frac{e^2}{2}\sum_{n\neq m}\frac{Z_n Z_m}{|\hat{\mathbf{R}}_n - \hat{\mathbf{R}}_m|} \\ &+ \sum_i \frac{\hat{\mathbf{p}}_i^2}{2m} + \frac{e^2}{2}\sum_{i\neq j}\frac{1}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|} - e^2\sum_n\sum_i \frac{Z_n}{|\hat{\mathbf{R}}_n - \hat{\mathbf{r}}_i|}\,.\end{aligned} \tag{2.1}$$

As usual, $\hat{\mathbf{P}}_n$ and $\hat{\mathbf{R}}_n$ denote momentum and position operators of the nuclei with atomic number $Z_n$, $\hat{\mathbf{p}}_i$ and $\hat{\mathbf{r}}_i$ are the momentum and position operators of a electrons and $e$ is the elemental charge. We denote the masses of the nuclei by $M_n$ and the mass of the electron by $m$.

We apply the so-called adiabatic or Born-Oppenheimer approximation [8] in which the dynamics of the nuclei and the electrons decouple in first order of an expansion using the expansion parameter $\gamma = (m/M)^{1/4}$. The remaining Hamiltonian describing the electron dynamics is given by

$$\hat{H}^{1Q}_{\text{elec}} = \sum_i \left(\frac{\hat{\mathbf{p}}_i^2}{2m} + V(\hat{\mathbf{r}})\right) + \frac{e^2}{2}\sum_{i\neq j}\frac{1}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|}\,, \tag{2.2}$$

where the external potential $\hat{V}(\mathbf{r})$ is given by superposition of static Coulomb potentials of point charges located at "sites" $\mathbf{R}_n$, which are the equilibrium positions of the nuclei,

$$V(\mathbf{r}) = -e^2\sum_n \frac{Z_n}{|\mathbf{R}_n - \mathbf{r}|}\,. \tag{2.3}$$

Introducing the field operators $\hat{\Psi}_s^\dagger(\mathbf{r})$ ($\hat{\Psi}_s(\mathbf{r})$) that create (annihilate) an electron at position $\mathbf{r}$ with spin $s \in \{\uparrow, \downarrow\}$, the electronic Hamilton operator in second quantization reads

$$\begin{aligned}\hat{H}_{\text{elec}} = &\sum_s \int \hat{\Psi}_s^\dagger(\mathbf{r})\left(-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}}^2 + V(\mathbf{r})\right)\hat{\Psi}_s(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ &+ \frac{1}{2}\sum_{s,s'}\int \hat{\Psi}_s^\dagger(\mathbf{r})\hat{\Psi}_{s'}^\dagger(\mathbf{r}')\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\hat{\Psi}_{s'}(\mathbf{r}')\hat{\Psi}_s(\mathbf{r})\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'.\end{aligned} \tag{2.4}$$

Although the Born-Oppenheimer approximation simplifies the problem vastly, the presence of the two-electron operator in (2.4) still renders the electronic problem generally unsolvable.

Please note that we use atomic units ($\hbar = m = e = 1/4\pi\epsilon_0 = 1$) from now on.

## 2.2 Foundations of density functional theory

### 2.2.1 Hohenberg-Kohn theorems

We follow the proofs of the Hohenberg-Kohn theorems as given in the original work [9]. The requirement of a non-degenerate ground state, that we assume from now on, can be relaxed in the "constrained search formulation" as shown by Levy [10].

**Theorem 1** *For any system of interacting particles in an external potential $v_{ext}(\mathbf{r})$, the external potential is uniquely determined by the ground-state density $n(\mathbf{r})$.*

We consider two systems $\hat{H}_i$ with external potential $\hat{V}_i$, ground-state $|\Psi_i\rangle$ and corresponding energy $E_i$, where $i \in \{1, 2\}$. Only the external potentials of both systems differ, i.e. $V_1(\mathbf{r}) \neq V_2(\mathbf{r}) + c$, where $c$ is an arbitrary constant. The Schrödinger equations read

$$\hat{H}_1|\Psi_1\rangle = (\hat{F} + \hat{V}_1)|\Psi_1\rangle = E_1|\Psi_1\rangle \,, \tag{2.5a}$$

$$\hat{H}_2|\Psi_2\rangle = (\hat{F} + \hat{V}_2)|\Psi_2\rangle = E_2|\Psi_2\rangle \,, \tag{2.5b}$$

where $\hat{F}$ represents the kinetic energy and the mutual interaction of the electrons. Obviously, $|\Psi_1\rangle \neq |\Psi_2\rangle$ holds, since both wave functions satisfy different Schrödinger equations.

To proof that the resulting ground-state densities

$$n_i(\mathbf{r}) = \sum_s \langle\Psi_i|\hat{\Psi}_s^\dagger(\mathbf{r})\hat{\Psi}_s(\mathbf{r})|\Psi_i\rangle \tag{2.6}$$

must be different, we assume the opposite, $n_1(\mathbf{r}) = n_2(\mathbf{r})$, and show that this assumption leads to a contradiction ("reductio ad absurdum"). Using the variational principle we get

$$\begin{aligned} E_1 &< \langle\Psi_2|\hat{H}_1|\Psi_2\rangle = \langle\Psi_2|\hat{H}_2 - \hat{V}_2 + \hat{V}_1|\Psi_2\rangle \\ &= E_2 + \langle\Psi_2|\hat{V}_1 - \hat{V}_2|\Psi_2\rangle = E_2 + \int n_2(\mathbf{r})\left[V_1(\mathbf{r}) - V_2(\mathbf{r})\right] \mathrm{d}\mathbf{r} \,. \end{aligned} \tag{2.7}$$

The same equation holds for interchanged indices

$$E_2 < E_1 - \int n_1(\mathbf{r})\left[V_1(\mathbf{r}) - V_2(\mathbf{r})\right] \mathrm{d}\mathbf{r} \,. \tag{2.8}$$

Under the assumption of $n_1(\mathbf{r}) = n_2(\mathbf{r})$, subtracting (2.7) from (2.8) leads to the immediate contradiction

$$E_1 + E_2 < E_1 + E_2 \,. \tag{2.9}$$

This proves **Theorem 1**.

**Theorem 2** *The exact ground-state density minimizes the energy functional*

$$E\left[n(\mathbf{r})\right] = \langle\Psi\left[n(\mathbf{r})\right]|\hat{H}|\Psi\left[n(\mathbf{r})\right]\rangle \,. \tag{2.10}$$

Theorem 1 states that the ground-state density determines the external potential. The external potential determines $\hat{H}$ and ultimately the ground-state $|\Psi\rangle$. Therefore, one can write the ground-state as a functional of the ground-state density, $|\Psi\rangle = |\Psi[n(\mathbf{r})]\rangle$. This implies that expectation values like the ground-state energy are also functionals of $n(\mathbf{r})$ (2.10). Given that $n(\mathbf{r})$ is the exact ground-state density of $\hat{H}$, the variational principle states

$$E[n(\mathbf{r})] < E[n'(\mathbf{r})], \tag{2.11}$$

which is the statement of the second theorem.

The Hohenberg-Kohn theorems are the foundation of density functional theory: All ground-state properties can be expressed as functionals of the electron density instead of a complicated many-body wave functions depending on $3N$ coordinates. This is an immense simplification of the problem but still does not lead to an immediate application.

### 2.2.2 Kohn-Sham formalism

The idea of the Kohn-Sham formalism is to find the potential of a non-interacting system $v_{\mathrm{KS}}(\mathbf{r})$ which leads to the same ground-state density as the interacting many-particle system. The question whether such a potential always exists in the mathematical sense is called the non-interacting $v$-representability problem.

We start by assuming that we found the exact expression of the energy functional which *solely* depends on the density $n(\mathbf{r})$

$$\begin{aligned} E[n(\mathbf{r})] &= \langle \Psi[n(\mathbf{r})] | \hat{H} | \Psi[n(\mathbf{r})] \rangle \\ &= \langle \Psi[n(\mathbf{r})] | \hat{F} | \Psi[n(\mathbf{r})] \rangle + \langle \Psi[n(\mathbf{r})] | \hat{V} | \Psi[n(\mathbf{r})] \rangle \\ &= F[n(\mathbf{r})] + V[n(\mathbf{r})]. \end{aligned} \tag{2.12}$$

According to Theorem 2, the charge density minimizing (2.12) is the charge density of the ground-state. In the Kohn-Sham formalism the (trial) ground-state density, with which the energy functional is evaluated, is the ground-state density of a fictitious non-interacting system. The degrees of freedom of the non-interacting system (the form of its external potential $v_{\mathrm{KS}}(\mathbf{r})$) are then used to minimize the energy functional $E[n]$. Strictly speaking, the non-interacting system does not have a direct physical meaning other than reproducing the ground-state density of the many-body problem.

The variational problem which has to be solved in order to minimize (2.12) has the form

$$\frac{\delta}{\delta n(\mathbf{r})} \left[ F[n(\mathbf{r})] + \int V(\mathbf{r}) n(\mathbf{r}) \, \mathrm{d}\mathbf{r} - \mu \left( \int n(\mathbf{r}) \, \mathrm{d}\mathbf{r} - N \right) \right] = 0, \tag{2.13}$$

where $\mu$ is a Lagrange parameter fixing the total number of particles. The first step is to write $F[n(\mathbf{r})]$ as the sum of three contributions

$$F[n(\mathbf{r})] = T_{\mathrm{S}}[n(\mathbf{r})] + E_{\mathrm{H}}[n(\mathbf{r})] + E_{\mathrm{xc}}[n(\mathbf{r})] \tag{2.14}$$

that are specified later. Assuming that the true ground-state density is $v$-representable, solving (2.13) is equivalent to finding the ground-state density

$$n(\mathbf{r}) = \sum_{\epsilon_i < E_\mathrm{F}} \Phi_i^*(\mathbf{r})\Phi_i(\mathbf{r}) \tag{2.15}$$

of the non interacting system

$$\left[-\frac{1}{2}\nabla^2 + v_\mathrm{KS}(\mathbf{r})\right]\Phi_i(\mathbf{r}) = \epsilon_i\Phi_i(\mathbf{r}) \tag{2.16}$$

with the Kohn-Sham potential

$$v_\mathrm{KS}(\mathbf{r}) = V(\mathbf{r}) + \frac{\delta E_\mathrm{H}[n(\mathbf{r})]}{\delta n(\mathbf{r})} + \frac{\delta E_\mathrm{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \ . \tag{2.17}$$

As usual, $E_\mathrm{F}$ denotes the Fermi-energy of the (non-interacting) system with $N$ particles. We define $T_\mathrm{S}\left[n(\mathbf{r})\right]$ as the kinetic energy of the non-interacting system

$$T_\mathrm{S}\left[n(\mathbf{r})\right] = -\frac{1}{2}\sum_{\epsilon_i < E_\mathrm{F}} \int \Phi_i^*(\mathbf{r})\nabla^2\Phi_i(\mathbf{r})\,\mathrm{d}\mathbf{r} \ , \tag{2.18}$$

and set $E_\mathrm{H}$ to be the classical electrostatic (Hartree) energy

$$E_\mathrm{H}\left[n(\mathbf{r})\right] = \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \Leftrightarrow \frac{\delta E_\mathrm{H}[n(\mathbf{r})]}{\delta n(\mathbf{r})} = \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}' \ . \tag{2.19}$$

The physical reason for this particular choice is to identify all major contributions to the energy functional $F$, where exact expressions are available. All other contributions are "hidden" in the exchange-correlation functional $E_\mathrm{xc}$. Since the exact functional $F$ is simply the sum of the kinetic energy $T$ and the electron-electron interaction $E_\mathrm{ee}$ of the interacting system, the exchange-correlation energy $E_\mathrm{xc}$ can be formally written as

$$E_\mathrm{xc}[n(\mathbf{r})] = T[n(\mathbf{r})] - T_\mathrm{S}[n(\mathbf{r})] + E_\mathrm{ee}[n(\mathbf{r})] - E_\mathrm{H}[n(\mathbf{r})] \ . \tag{2.20}$$

Note that no approximations have been made until this point. The possibility of mapping the many body ground-state problem to a non-linear set of non-interacting wave functions of a fictitious problem vastly simplifies the problem and the computational effort to solve such a problem numerically. However, all the complexity of evaluating the energy of a many-body wave function is hidden in $E_\mathrm{xc}[n(\mathbf{r})]$. It is obvious that there is no general exact expression for this functional and certain assumptions have to be made to find an appropriate approximation.

### 2.2.3 Approximations to the exchange-correlation energy

During the last decades many properties (sum rules, scaling properties, bounds, discontinuities in derivatives, etc.) were discovered which are fulfilled by the exact exchange correlation energy [11]. All approximations to $E_\mathrm{xc}$ lack at least one of the properties of

the exact functional. The most important, and also the most fundamental exchange correlation approximation, is the Local Density Approximation (LDA). The defining property of the LDA is that it can be written in the form

$$E_{\mathrm{xc}}^{\mathrm{LDA}}\left[n(\mathbf{r})\right] = \int n(\mathbf{r})\epsilon_{\mathrm{xc}}(n(\mathbf{r}))\,\mathrm{d}\mathbf{r}\ . \tag{2.21}$$

Here, $\epsilon_{\mathrm{xc}}(n(\mathbf{r}))$ is the exchange-correlation energy per particle of a Homogenous Electron Gas (HEG) of density $n(\mathbf{r})$. The energy functional is separated into a correlation and an exchange part

$$E_{\mathrm{xc}}^{\mathrm{LDA}}\left[n(\mathbf{r})\right] = E_{\mathrm{x}}^{\mathrm{LDA}}\left[n(\mathbf{r})\right] + E_{\mathrm{c}}^{\mathrm{LDA}}\left[n(\mathbf{r})\right]. \tag{2.22}$$

For the HEG the exchange energy can be evaluated exactly [12]

$$E_{\mathrm{x}}^{\mathrm{LDA}}[n(\mathbf{r})] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\int n(\mathbf{r})^{4/3}\mathrm{d}\mathbf{r}\ . \tag{2.23}$$

Since there is no analytic expression for the correlation energy, the functional dependence of the correlation energy on the density is usually obtained by fitting to accurate quantum Monte Carlo simulations [13]. Thus the common LDA functionals are mostly identical, differing only in their parametrization. The LDA is in principle exact for the HEG and a good approximation for systems with slowly varying density. Even though this requirement is not met in a real solid, the LDA produces satisfying results for a wide range of materials.

Semi-local approximations to the energy functional are the so-called Gradient-Expansion Approximations (GEA), where systematic gradient-corrections are added to the LDA functional. This approach almost never improves the LDA in practice [14, 15]. It was realized later that empiric functionals containing the gradient

$$E_{\mathrm{xc}}^{\mathrm{GGA}}\left[n(\mathbf{r})\right] = \int f(n(\mathbf{r}), \nabla n(\mathbf{r}))\,\mathrm{d}\mathbf{r} \tag{2.24}$$

are able to improve the LDA [16], leading to a whole class of so-called Generalized Gradient Approximation (GGA) functionals.

### 2.2.4 Extension to spin densities

In this subsection we summarize the extension of the DFT formalism to spin densities, following the work of von Barth and Hedin [17].

To describe the presence of an external magnetic field $\mathbf{B}(\mathbf{r})$ acting on the spins of the electrons, the Hamiltonian in (2.4) has to be extended. The external scalar potential in the one-electron operator is replaced by a spin dependent potential $w_{\alpha,\beta}(\mathbf{r})$. The Hamiltonian

reads

$$\begin{aligned}\hat{H}^{\mathrm{B}}_{\mathrm{elec}} = &-\frac{1}{2}\sum_{\alpha}\int \hat{\Psi}^{\dagger}_{\alpha}(\mathbf{r})\nabla^2\hat{\Psi}_{\alpha}(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ &+\frac{1}{2}\sum_{\alpha,\beta}\int \hat{\Psi}^{\dagger}_{\alpha}(\mathbf{r})\hat{\Psi}^{\dagger}_{\beta}(\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}\hat{\Psi}_{\beta}(\mathbf{r}')\hat{\Psi}_{\alpha}(\mathbf{r})\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \\ &+\sum_{\alpha,\beta}\int \hat{\Psi}^{\dagger}_{\alpha}(\mathbf{r})w_{\alpha,\beta}(\mathbf{r})\hat{\Psi}_{\beta}(\mathbf{r})\,\mathrm{d}\mathbf{r}\end{aligned} \tag{2.25}$$

with the $2\times 2$ matrix

$$\boldsymbol{w}(\mathbf{r}) = V(\mathbf{r})\mathbb{1}_2 + 2\mu_B\sum_i \mathbf{B}(\mathbf{r})\cdot\hat{\mathbf{s}}_i\ , \tag{2.26}$$

where $\mu_B$ is the Bohr magneton, $\hat{\mathbf{s}}_i$ the spin operator acting on the spin of the $i$-th electron.

In the spin-extended version of DFT, the fundamental role of the electron density $n(\mathbf{r})$ in the Hohenberg-Kohn theorems is adopted by the $2\times 2$ density matrix

$$\rho_{\alpha,\beta}(\mathbf{r}) = \langle\Psi|\hat{\Psi}^{\dagger}_{\beta}(\mathbf{r})\hat{\Psi}_{\alpha}(\mathbf{r})|\Psi\rangle\ . \tag{2.27}$$

The foundations of the Kohn-Sham equation, the Hohenberg-Kohn theorems, are still valid. There is a unique connection between the ground-state of the many-body system in an external magnetic field and the density matrix. The modified Kohn-Sham equations are given by

$$\sum_{\beta}\left(-\frac{1}{2}\nabla^2\delta_{\alpha,\beta} + \sum_{\gamma}\int\frac{\rho_{\gamma,\gamma}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}'\delta_{\alpha,\beta} + w_{\alpha,\beta}(\mathbf{r}) + \frac{\delta E_{\mathrm{xc}}}{\delta\rho_{\alpha,\beta}(\mathbf{r})}\right)\Phi_{i,\beta}(\mathbf{r}) = \epsilon_i\Phi_{i,\alpha}(\mathbf{r}) \tag{2.28}$$

and

$$\rho_{\alpha,\beta}(\mathbf{r}) = \sum_{\epsilon_i<E_{\mathrm{F}}}\Phi^{*}_{i,\alpha}(\mathbf{r})\Phi_{i,\beta}(\mathbf{r})\ . \tag{2.29}$$

Note that the Kohn-Sham potential is now also spin-dependent. For the extension of the LDA functional to the Local Spin Density Approximation (LSDA) functional the same approximations are made. Space is partitioned into small boxes and the density in each box is described by a (spin polarized) HEG with spin-up $\rho^{\uparrow}(\mathbf{r})$ and spin-down $\rho^{\downarrow}(\mathbf{r})$ densities, given by the eigenvalues of the density matrix $\rho_{\alpha,\beta}(\mathbf{r})$

$$E_{\mathrm{xc}}\left[\rho_{\alpha,\beta}(\mathbf{r})\right] = \int(\rho^{\uparrow}(\mathbf{r})+\rho^{\downarrow}(\mathbf{r}))\epsilon_{\mathrm{xc}}(\rho^{\uparrow}(\mathbf{r}),\rho^{\downarrow}(\mathbf{r}))\,\mathrm{d}\mathbf{r}\ . \tag{2.30}$$

The local rotation diagonalizing $\rho_{\alpha,\beta}(\mathbf{r})$ is applied. The exchange correlation potential appearing in (2.28) is then diagonal

$$v^{\alpha,\beta}_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}}{\delta\rho_{\alpha,\beta}(\mathbf{r})} \to v^{\alpha}_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}}{\delta\rho_{\alpha}(\mathbf{r})}\ , \tag{2.31}$$

which leads to Kohn-Sham equations with two (effective) spin densities, denoted in the following by $n^s(\mathbf{r})$, with $s \in \{\uparrow, \downarrow\}$, using the notation $\bar{\uparrow} = \downarrow$ and $\bar{\downarrow} = \uparrow$. The spin densities can also be expressed by the total spin density $n(\mathbf{r})$ and the magnetization density $m(\mathbf{r})$

$$n(\mathbf{r}) = n^{\downarrow}(\mathbf{r}) + n^{\uparrow}(\mathbf{r}) \,, \tag{2.32}$$

$$m(\mathbf{r}) = n^{\downarrow}(\mathbf{r}) - n^{\uparrow}(\mathbf{r}) \,. \tag{2.33}$$

The main reason for introducing a spin-dependent DFT is "the interest in systems which are intrinsically spin polarized, like transition metals and non-singlet atoms" [17]. Therefore, the external field $\mathbf{B}(\mathbf{r})$ introduced to break the symmetry between the spin-channels is set to zero. The intrinsic spin-polarization is then investigated by starting the self-consistent calculation of (2.28) and (2.29) using the scheme described in the upcoming Section 2.3.1 with a guess for $n^{\uparrow}(\mathbf{r})$ and $n^{\downarrow}(\mathbf{r})$ with $m(\mathbf{r}) \neq 0$ in the first iteration. Problems connected with the spin polarized DFT are briefly discussed in Section 2.4.1.

## 2.3 Practical implementation

In this section we discuss two of the main strategies that have been developed over the last decades to implement numerically the Kohn-Sham scheme introduced in Section 2.2.2, namely the "pseudopotential" and the "all-electron" approaches. Both approaches are connected to a certain choice of basis in which the Kohn-Sham equations are solved.

### 2.3.1 Self-consistent iteration

Equations (2.15), (2.16) and (2.17) are non-linearly coupled and solved iteratively:

(i) An initial guess for the electron density $n(\mathbf{r})$ is made, e.g., a Gaussian charge distributions around the nuclei.

(ii) The Kohn Sham potential $v_{\mathrm{KS}}(\mathbf{r})$ via (2.17) is determined from $n(\mathbf{r})$.

(iii) $v_{\mathrm{KS}}(\mathbf{r})$ is used to calculate the ground state via (2.16).

(iv) A new charge density $n'(\mathbf{r})$ is calculated and mixed with previous charge densities for better convergence. Go to step (ii) until convergence is achieved.

The algorithm terminates when the charge density is converged, i.e., a fix-point for the iteration scheme is found. For a numerical implementation, equations (2.15)-(2.17) are expressed in a basis of wave functions that lead to a set of algebraic equations. Two main formulations for the DFT are widely used that we discuss briefly.

### 2.3.2 Algebraic formulation

We aim to find solutions of the effective single-particle Hamiltonian in the Kohn-Sham equation given by (2.16). We assume that the material under investigation possesses

translational symmetry. Since the Kohn-Sham potential only depends on the external potential and on the electron density, it possesses the same periodicity. We can write

$$v_{\mathrm{KS}}(\mathbf{r}) = v_{\mathrm{KS}}(\mathbf{r} + \mathbf{R}_i) \quad \forall i \ , \tag{2.34}$$

where $\mathbf{R}_i$ denote the lattice vectors of the underlying Bravais lattice. For better readability we will drop the sub-index $i$ for the real space and the – yet to be introduced – reciprocal lattice vectors in the following. For a periodic potential, Bloch's theorem [18] states that the eigenfunctions can be written as

$$\Phi_{\mathbf{k},n}(\mathbf{r}) = u_{\mathbf{k},n}(\mathbf{r}) \exp(\mathrm{i}\mathbf{k}\mathbf{r}) \ , \quad \text{where} \quad u_{\mathbf{k},n}(\mathbf{r}) = u_{\mathbf{k},n}(\mathbf{r} + \mathbf{R}) \ . \tag{2.35}$$

Here, $n$ denotes the band-index and $\mathbf{k}$ the wave vector from the first Brillouin Zone (BZ). Expanding the periodic part of the wave functions $u_{\mathbf{k},n}(\mathbf{r})$ in a Fourier series results in

$$\Phi_{\mathbf{k},n}(\mathbf{r}) = \sum_{\mathbf{G}} c^n_{\mathbf{G}+\mathbf{k}} \ e^{\mathrm{i}(\mathbf{k}+\mathbf{G})\mathbf{r}} \ , \tag{2.36}$$

where we introduced the reciprocal lattice vectors $\mathbf{G}$, for which $e^{\mathrm{i}\mathbf{G}\mathbf{R}} = 1$ holds. Equation (2.36) is the plane wave (PW) expansion of the eigenfunction $\Phi_{\mathbf{k},n}(\mathbf{r})$. The wave function is periodic in $\mathbf{k}$-space, $\Phi_{\mathbf{k},n}(\mathbf{r}) = \Phi_{\mathbf{k}+\mathbf{G},n}(\mathbf{r})$, from which follows that all expectation values can be evaluated within the first BZ.

Since the energy of every plane wave in the infinite sum in (2.36) is proportional to $|\mathbf{G} + \mathbf{k}|^2$ the very high energy (or frequency) waves carry very small weight $c^n_{\mathbf{G}+\mathbf{k}}$ and are therefore neglected. This is done by introducing a material dependent cut-off parameter $K_{\mathrm{max}}$ and truncating the sum

$$\Phi_{\mathbf{k},n}(\mathbf{r}) = \sum_{|\mathbf{G}+\mathbf{k}|<K_{\mathrm{max}}} c^n_{\mathbf{G}+\mathbf{k}} \ e^{\mathrm{i}(\mathbf{k}+\mathbf{G})\mathbf{r}} \ . \tag{2.37}$$

As seen later in this chapter, plane waves are, despite of their simplicity, not generally the best choice for practical calculations. Often large values for $K_{\mathrm{max}}$ have to be chosen in order to describe a wave function correctly. Large values of $K_{\mathrm{max}}$ result in large basis sets which can not be handled numerically. Therefore, a trial wave function with not necessarily orthogonal basis functions $\chi_{\mathbf{G}+\mathbf{k}}(\mathbf{r})$ in the form of (2.37) is introduced

$$\Phi_{\mathbf{k},n}(\mathbf{r}) = \sum_{|\mathbf{G}+\mathbf{k}|<K_{\mathrm{max}}} c^n_{\mathbf{G}+\mathbf{k}} \ \chi_{\mathbf{G}+\mathbf{k}}(\mathbf{r}) \ . \tag{2.38}$$

The Rayleigh-Ritz principle is used to determine the expansion coefficients $c^n_{\mathbf{G}+\mathbf{k}}$ and the eigenvalues $\epsilon_{\mathbf{k},n}$, using

$$\sum_{\mathbf{G}'} \left[ H_{\mathbf{G},\mathbf{G}'}(\mathbf{k}) - \epsilon_{\mathbf{k},n} S_{\mathbf{G},\mathbf{G}'}(\mathbf{k}) \right] c^n_{\mathbf{G}'+\mathbf{k}} = 0 \ , \tag{2.39}$$

with

$$S_{\mathbf{G},\mathbf{G}'}(\mathbf{k}) = \int \chi^*_{\mathbf{G}+\mathbf{k}}(\mathbf{r}) \chi_{\mathbf{G}'+\mathbf{k}}(\mathbf{r}) \, \mathrm{d}\mathbf{r} \ , \tag{2.40}$$

$$H_{\mathbf{G},\mathbf{G}'}(\mathbf{k}) = \int \chi^*_{\mathbf{G}+\mathbf{k}}(\mathbf{r}) \left[ -\frac{1}{2}\nabla^2 + v_{\mathrm{KS}}(\mathbf{r}) \right] \chi_{\mathbf{G}'+\mathbf{k}}(\mathbf{r}) \, \mathrm{d}\mathbf{r} \ . \tag{2.41}$$

### 2.3.3 Pseudopotential (Plane Wave) Approach

The core states close to the nucleus of an atom remain relatively unaffected in a chemical environment. In solid-state physics most properties are determined by states in the vicinity of the Fermi energy. As we will see in the following chapters, one is often interested in the differences of ground-state energies, not in their absolute values. It seems therefore desirable to leave out the exact calculations of all electrons of an atom (which account for a huge fraction of the total energy) and focus on the valence states. This is realized in the "pseudopotential" (PP)-approach [19], in which the $1/\mathbf{r}$ potential of the nucleus gets replaced by an effective potential.

A first realization that comes to mind is an effective Coulomb potential screened by the core electrons. However, it was realized that it is not necessary to exactly reproduce the all-electron solution (not to be confused with a many-particle wave function) of the valence states in the core region, defined by $R < R_{\text{cut}}$, but to match the solution outside $R_{\text{cut}}$. The all-electron valence wave functions close to the nucleus are rapidly oscillating since they are orthogonal to all core states. Finding a suitable basis in which these oscillating states can be expressed is a challenging task. In the pseudopotential method this is avoided by constructing a potential that reproduces a node-less wave function in the core region, see Fig. 2.1, with the same energy eigenvalue as its all-electron counterpart.

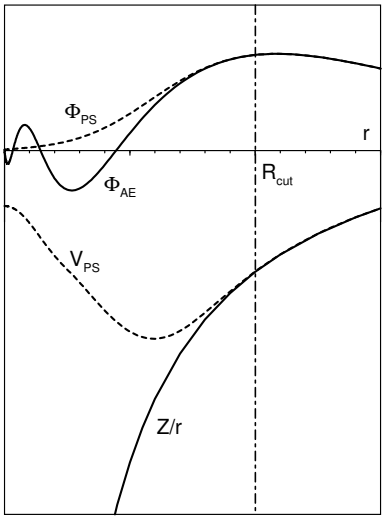


**Figure 2.1:** Schematic illustration of an atomic all-electron wave function (solid line) and the corresponding atomic pseudo wave function (dashed line) together with the respective external Coulomb potential and pseudopotential. Taken from [20].

The pseudopotentials are not unique. For a detailed discussion of the various classes (norm-conserving, fully non-local, and ultrasoft) we refer to [21]. The pseudopotentials are mostly constructed from all-electron atomic DFT calculations. For standard DFT codes, databases with precomputed pseudopotentials for various atoms are available.

The main advantage of the use of pseudopotential methods lies in the possibility to omit a huge part of the numerical effort which has to be carried out in all-electron calculations. Avoiding the description of the high frequency parts of the wave functions allows the expansion of wave function into Plane Waves (PWs), which are eigenfunctions of a particle in a constant potential. The Kohn-Sham equations are easily expressed in this basis and many mathematical operations can be implemented very efficiently in the PW basis.

The drawback of using pseudopotentials is that some a-priori knowledge of the system under consideration is sometimes needed. The atomic solution of the wave functions can be significantly different from the solutions of the atom embedded in a solid. To avoid this problem, the cut-off parameter $R_{\text{cut}}$ has to be sufficiently small which leads to a high cut-off parameter for the plane-wave basis set, increasing the basis set size. Another problem is that core and valence states are not well separated. An example connected to this work are the $3d$ levels of the transition metal atoms where the $3d$ level is not well separated from the $4s$ level and must be included into the valence states.

### 2.3.4 (Full Potential) Linear Augmented Plane Wave

Early on it was pointed out by J. C. Slater that the expansion of the all-electron wave function in terms of plane waves around the nuclei is not practicable for a numerical treatment [22]. Slater proposed the Augmented Plane Wave (APW) method in which space is partioned into spherical regions around each nucleus, the "muffin tin", and the interstitial region in between.

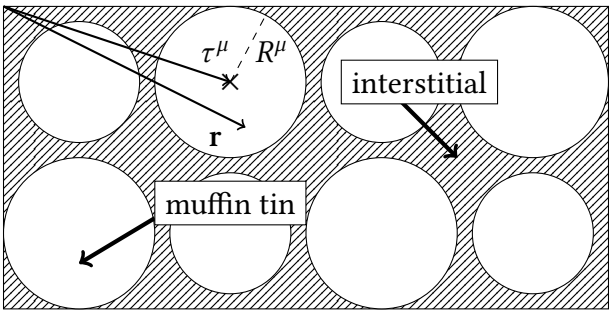


**Figure 2.2:** Schematic illustration of the space partitioning into "muffin tins" with radius $R^\mu$ centered around the atomic positions $\boldsymbol{\tau}^\mu$ and the interstitial region, after [23].

Wave functions within the muffin tin are expanded in terms of spherical harmonics, the interstitial region in plane waves,

$$\tilde{\chi}_{\mathbf{k},E}(\mathbf{r}) = \begin{cases} e^{\mathrm{i}\mathbf{k}\mathbf{r}} & \text{within the interstitial region ,} \\ \sum_{L=\{l,m\}} a_L^{\mu,\mathbf{k}} u_{l,E}^{\mu}(r^\mu) Y_L(\mathbf{r}^\mu) & \text{within muffin tin } \mu. \end{cases} \tag{2.42}$$

Here, we defined the relative coordinate $\mathbf{r}^\mu = \mathbf{r} - \boldsymbol{\tau}^\mu$ within the muffin tin with index $\mu$ centered at $\boldsymbol{\tau}^\mu$, compare Fig. 2.2. The wave function within the muffin tin is the solution of the Schrödinger equation with a radial symmetric potential $V(r)$. $Y_L(\mathbf{r})$ are spherical harmonics, $u_{l,E}(r)$ satisfy the radial Schrödinger equation

$$\left[ -\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{1}{2}\frac{l(l+1)}{r^2} + V(r) - E \right] r u_{l,E}(r) = 0 \ . \tag{2.43}$$

In order to have a well-defined kinetic energy, both functions in (2.42) must be matched at the muffin tin border. This uniquely determines the coefficients

$$a_L^{\mu,\mathbf{k}} = 4\pi \exp\left(\mathrm{i}\mathbf{k}\boldsymbol{\tau}^\mu\right) \mathrm{i}^l Y_L^*(\mathbf{k}) \frac{j_l(kR^\mu)}{u_{l,E}(R^\mu)} \ , \tag{2.44}$$

where $j_l(x)$ denotes the Bessel functions of integer order $l$, $R^\mu$ is the radius of the muffin tin $\mu$. It turns out [24, 25] that there is not enough variational freedom in the APW basis to describe accurately the eigenfunctions of $H$ for a general (fixed) value of $E$. To be able to describe the eigenstate $\Phi_{\mathbf{k},n}(\mathbf{r})$ one has to set $E = \epsilon_{\mathbf{k},n}$ in (2.42),

$$\Phi_{\mathbf{k},n}(\mathbf{r}) = \sum_{|\mathbf{G}+\mathbf{k}|<K_{\max}} c^n_{\mathbf{G}+\mathbf{k}}\ \tilde{\chi}_{\mathbf{G}+\mathbf{k},\epsilon_{\mathbf{k},n}}(\mathbf{r})\ . \tag{2.45}$$

The arising problem is difficult to solve: $\epsilon_{\mathbf{k},n}$ is determined by (2.39), using the basis set from (2.45), itself depending on $\epsilon_{\mathbf{k},n}$. This non-linear problem makes the APW method computationally expensive. In practice, the problem can be tackled by searching the $n$ roots of

$$\det\left[H_{\mathbf{G},\mathbf{G}'}(E,\mathbf{k}) - ES_{\mathbf{G},\mathbf{G}'}(E,\mathbf{k})\right] = 0\ , \tag{2.46}$$

using a sufficient fine energy grid. This has to be done for every $\mathbf{k}$-point.

The Linear Augmented Plane Wave (LAPW) method solves the main problem of the APW method, namely the energy dependence of the Hamiltonian. The idea is to expand the energy dependent radial part of the wave function in (2.42) in a Taylor series,

$$u_{l,E}(r) = u_{l,E_l}(r) + (E-E_l)\dot{u}_{l,E_l}(r) + O((E-E_l)^2)\ . \tag{2.47}$$

The basis-functions read

$$\bar{\chi}_{\mathbf{k},E_l}(\mathbf{r}) = \begin{cases} e^{i\mathbf{k}\mathbf{r}} & \text{in the interstitial region} \\ \sum_{L=\{l,m\}} \left(\bar{a}^{\mu,\mathbf{k}}_L u^\mu_{l,E_l}(r^\mu) + \bar{b}^{\mu,\mathbf{k}}_L \dot{u}^\mu_{l,E_l}(r^\mu)\right) Y_L(\mathbf{r}^\mu) & \text{in muffin tin } \mu, \end{cases} \tag{2.48}$$

where the radial solutions are defined as before, see (2.43).

The coefficients are determined by the condition to fit the plane waves at the muffin-tin boundary in value and slope [23],

$$\bar{a}^{\mu,\mathbf{k}}_L = e^{i\mathbf{k}\tau^\mu} 4\pi \frac{1}{W} i^l Y^*_L(\mathbf{k}) \left[\dot{u}_{l,E_l}(R^\mu) k j'_l(R^\mu k) - \dot{u}'_{l,E_l}(R^\mu) j_l(R^\mu k)\right]\ , \tag{2.49}$$

$$\bar{b}^{\mu,\mathbf{k}}_L = e^{i\mathbf{k}\tau^\mu} 4\pi \frac{1}{W} i^l Y^*_L(\mathbf{k}) \left[u'_{l,E_l}(R^\mu) j_l(R^\mu k) - u_{l,E_l}(R^\mu) k j'_l(R^\mu k)\right]\ , \tag{2.50}$$

$$\text{with} \quad W = \left[\dot{u}_{l,E_l}(R^\mu) u'_{l,E_l}(R^\mu) - u_{l,E_l}(R^\mu) \dot{u}'_{l,E_l}(R^\mu)\right]\ , \tag{2.51}$$

where $\dot{u}$ denotes the energy derivative and $j'$ and $u'$ spatial derivates, e.g. $\dot{u}_{l,E_l} = \partial u_{l,E_l}/\partial E$ and $u'_{l,E_l} = \partial u_{l,E_l}/\partial r$. The additional variational freedom in the basis allows the use of a fixed set of energies $E_l$ for the radial solution of the Schrödinger equation within the muffin tins. The energies are fixed for each iteration in the self-consistent scheme introduced in Section 2.3.1. They are set to minimize the error due to the linearization based on the knowledge of the $\ell$-like bands of the previous iteration; an explicit expression can be found for example in [23]. $H$ and $S$ in (2.39) no longer depend on the exact eigenenergies, thus solving (2.39) reduces from a problem of root finding to a standard problem in linear algebra.

The Full Potential Linear Augmented Plane Wave (FLAPW) method combines the particular choice of the LAPW basis with the treatment of the exact core potential. That means that the LAPW basis is used to evaluate the Hamilton matrix in (2.41) without approximating $v_{\mathrm{KS}}(\mathbf{r})$. For a numerical exact representation of the (Kohn-Sham) potential, $v_{\mathrm{KS}}(\mathbf{r})$ is itself expanded into spherical harmonics and plane waves

$$v_{\mathrm{KS}}(\mathbf{r}) = \begin{cases} \sum_{\mathrm{G}} V_{\mathrm{I}}^{G} e^{i\mathrm{G}\mathbf{r}} & \text{within the interstitial region ,} \\ \sum_{L} V_{\mathrm{MT}}^{\mu,L}(r) Y_L(\mathbf{r}^\mu) & \text{within muffin tin } \mu. \end{cases} \tag{2.52}$$

The numerical exact evaluation of the integrals appearing for example in the calculation of the potential energy pose a difficult technical challenge, which is owed to the complexity of the basis. FLAPW codes are therefore more complicated but also better than other implementations of the Kohn-Sham scheme available today.

### 2.3.5 $\ell$-like charge in the LAPW basis

Using (2.38), the charge density can be written as

$$n(\mathbf{r}) = \sum_{\mathbf{k},n} f_{\mathbf{k},n} |\bar{\Phi}_{\mathbf{k},n}(\mathbf{r})|^2 \ , \tag{2.53}$$

where

$$\bar{\Phi}_{\mathbf{k},n}(\mathbf{r}) = \sum_{|\mathrm{G}+\mathbf{k}|<K_{\max}} c^n_{\mathrm{G}+\mathbf{k}} \ \bar{\chi}_{\mathrm{G}+\mathbf{k}}(\mathbf{r}) \tag{2.54}$$

is the expansion of the eigenfunctions in the LAPW basis and $f_{\mathbf{k},n}$ is the corresponding occupation number. With (2.48) the charge density within each muffin-tin sphere $\mu$ has the explicit form

$$\begin{aligned} n^\mu(\mathbf{r}^\mu) = \sum_{\mathbf{k},n} f_{\mathbf{k},n} \sum_{\mathrm{G},\mathrm{G}'} c^n_{\mathrm{G}+\mathbf{k}} c^{n*}_{\mathrm{G}'+\mathbf{k}} \sum_{L,L'} \left( \bar{a}^{\mu,\mathbf{k}}_L u^\mu_{l,E_l}(r^\mu) + \bar{b}^{\mu,\mathbf{k}}_L \dot{u}^\mu_{l,E_l}(r^\mu) \right) \times \\ \left( \bar{a}^{\mu,\mathbf{k}}_{L'} u^\mu_{l',E_l}(r^\mu) + \bar{b}^{\mu,\mathbf{k}}_{L'} \dot{u}^\mu_{l',E_l}(r^\mu) \right)^* Y_L(\mathbf{r}^\mu) Y^*_{L'}(\mathbf{r}^\mu) \ . \end{aligned} \tag{2.55}$$

The LAPW basis set allows for a straightforward identification of a local density matrix by projecting the wave function onto the $Y_L$ subspace and integrating over the spatial coordinates [26],

$$\begin{aligned} n^\mu_{L,L'} = \sum_{\mathbf{k},n} f_{\mathbf{k},n} \sum_{\mathrm{G},\mathrm{G}'} c^n_{\mathrm{G}+\mathbf{k}} c^{n*}_{\mathrm{G}'+\mathbf{k}} \times \\ \int r^2 \left( \bar{a}^{\mu,\mathbf{k}}_L u^\mu_{l,E_l}(r) + \bar{b}^{\mu,\mathbf{k}}_L \dot{u}^\mu_{l,E_l}(r) \right) \left( \bar{a}^{\mu,\mathbf{k}}_{L'} u^\mu_{l',E_l}(r) + \bar{b}^{\mu,\mathbf{k}}_{L'} \dot{u}^\mu_{l',E_l}(r) \right)^* \mathrm{d}r \ . \end{aligned} \tag{2.56}$$

In the atomic limit (an isolated atom), this matrix is block-diagonal with respect to the angular quantum number $l$. In a solid the off-diagonal elements are expected to be small and are therefore neglected. Setting $l = l'$ we get

$$n^{\mu,l}_{m,m'} = \sum_{\mathbf{k},n} f_{\mathbf{k},n} \sum_{\mathrm{G},\mathrm{G}'} c^n_{\mathrm{G}+\mathbf{k}} c^{n*}_{\mathrm{G}'+\mathbf{k}} \left( \bar{a}^{\mu,\mathbf{k}*}_{l,m'} \bar{a}^{\mu,\mathbf{k}}_{l,m} + \bar{b}^{\mu,\mathbf{k}*}_{l,m'} \bar{b}^{\mu,\mathbf{k}}_{l,m} \langle \dot{u}_{l,E_l} | \dot{u}_{l,E_l} \rangle \right) \ , \tag{2.57}$$

where we used that $\langle u_{l,E_l} | \dot{u}_{l,E_l} \rangle = 0$ and $\langle u_{l,E_l} | u_{l,E_l} \rangle = 1$. This density matrix is used in the LDA+U formalism, see Section 2.5.

The "$\ell$-like charges" are retrieved by summing over the diagonal entries

$$n^{\mu,l} = \sum_m n^{\mu,l}_{m,m} \ . \tag{2.58}$$

This quantity allows to calculate the formation of local magnetic moments in the muffin tins, $m^{\mu,l} = n^{\mu,s,l} - n^{\mu,\bar{s},l}$, where we reintroduced the spin index $s$. The choice of the muffin-tin radius has a noticeable influence on the matrix elements $n^{\mu,l}_{m,m'}$ and also on the $\ell$-like charges $n^{\mu,l}$, as shown in Section 4.4. Note, however, that the local magnetic moments $m^{\mu,l}$ are largely independent of the muffin-tin radius (neutral charges just cancel out) if the moment is well localized within the muffin-tin sphere.

## 2.4 Shortcomings of density functional theory

Since DFT is exact but in practice relies on approximations of the exchange correlation functional, all shortcomings stem from errors in these approximations. In this section we briefly discuss some of the main issues of the DFT connected to this work.

### 2.4.1 Spin contamination

One of the fundamental problems of DFT is the impact of approximative forms of the exchange-correlation functional on spin-polarized calculations. As discussed in Section 2.2.4, systems with intrinsic magnetization are investigated using a spin-dependent version of the DFT. The total electron density $n(\mathbf{r})$ of the spin-polarized calculation (2.32) should match the results from the paramagnetic calculation (2.15). Since both formulations are exact, given the exact exchange-correlation functionals, also the total ground-state energies should be equal.

In practice, however, the spin-polarized calculation only matches the paramagnetic calculation if the converged magnetization density of the spin-polarized calculation is zero. In all other cases the spin-polarized calculation lowers the total energy. It turns out that the symmetry-broken exchange correlation functional for the non-interacting system is often a better approximation than the paramagnetic functional [27]. This can lead to improved results for the ground-state energy by introducing a magnetization which is absent in the real system. This problem is called "spin contamination" [28].

### 2.4.2 Band-gap problem

A well known shortcoming of DFT is the band-gap problem: The band gap in semiconductors or insulators derived from the Kohn-Sham equations, which we will further refer to as "KS band gap", does not match the fundamental band gap of the real solid. It is systematically underestimated, typical deviations are of the order of 40% and more [29]. The band-gap problem is an example of the subtleties hidden in the DFT formalism. The

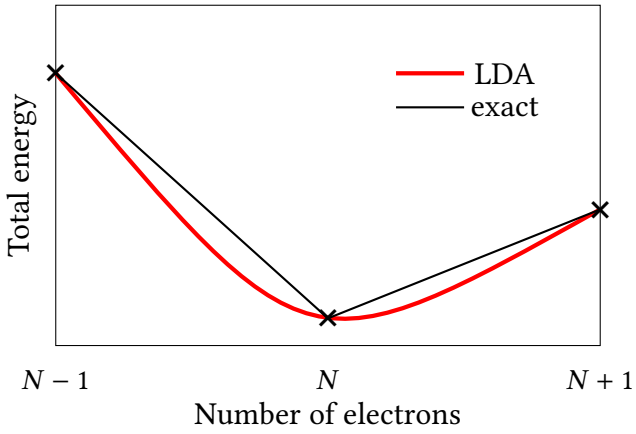


**Figure 2.3:** Illustration of the derivative discontinuity in the exact energy functional at integer electron number $N$.

assertion that DFT is a ground-state theory and, therefore, cannot predict observables of excited states is mathematically not satisfactory.

A strictly mathematical statement is made by Janaks' theorem [30], connecting the (differential) change in total energy with respect to a change in occupation number of a Kohn-Sham orbital to its corresponding eigenvalue

$$\frac{\partial E}{\partial n_i} = \epsilon_i \ , \tag{2.59}$$

where we assume in the following that the orbital energies are ordered

$$\epsilon_1 \leq \epsilon_2 \leq \ \ldots \ . \tag{2.60}$$

Perdew et al. have shown that the energy of a system with fractional electron number[1] is given by linear interpolation of values at integer electron numbers $N$ [31]

$$E(N+\delta) = (1-\delta)E(N) + \delta E(N+1) \ . \tag{2.61}$$

The problem of the unknown derivative discontinuities is illustrated in Fig. 2.3. Following [29] we are able to express the ionization energy of a system with $N$ electrons in terms of Kohn-Sham eigenenergies using (2.59) and (2.61)

$$E(N) - E(N-1) = \int_0^1 \mathrm{d}n \, \epsilon_N(N-1+n) = \epsilon_N(N) \ . \tag{2.62}$$

The numbers in brackets denote the number of electrons in the system. For example the KS band gap $\epsilon_\mathrm{g}$ in a system with $N$ electrons is given by

$$\epsilon_\mathrm{g} = \epsilon_{N+1}(N) - \epsilon_N(N) \ . \tag{2.63}$$

[1] The appearance of non-integer electron numbers is justified in the density functional theory of open systems [31].

Equation (2.62) is called Koopmans' theorem [32]. It states that the negative ionization energy is equal to the highest occupied eigenvalue. We also get an expression in term of Kohn-Sham eigenvalues for the electron affinity

$$E(N+1) - E(N) = \epsilon_{N+1}(N+1). \tag{2.64}$$

The fundamental gap $E_\mathrm{g}$ is defined via

$$E_\mathrm{g} = (E(N+1) - E(N)) - (E(N) - E(N-1)) \tag{2.65}$$
$$= \epsilon_{N+1}(N+1) - \epsilon_N(N)\ . \tag{2.66}$$

Thus we get a deviation $\Delta$ for the fundamental gap and the KS band gap

$$E_\mathrm{g} = \epsilon_\mathrm{g} + [\epsilon_{N+1}(N+1) - \epsilon_{N+1}(N)] = \epsilon_\mathrm{g} + \Delta\ . \tag{2.67}$$

The correction $\Delta$ is the derivative discontinuity appearing in the exact expression for the exchange-energy

$$\left.\frac{\delta E_\mathrm{ex}}{\delta n(\mathbf{r})}\right|_{N+\delta} - \left.\frac{\delta E_\mathrm{ex}}{\delta n(\mathbf{r})}\right|_{N-\delta} = \Delta\ . \tag{2.68}$$

For insulators, $\Delta$ is in the order of one eV, and cannot be neglected. Identifying the KS band gap with the fundamental gap, which means to interpret the Kohn-Sham band structure as a quasi-particle system, is therefore flawed and explains why the development of "better" exchange-functionals did have very little effect on the improvement of the gap [33]. One could think about using (2.65) directly and calculate the energies for a system with $N$ and $N \pm 1$ electrons. Adding/Removing an electron to a solid changes the total electron density slightly, of the order of $1/N$. However, since all standard functionals, such as the LDA, do not posses the property of a derivative discontinuity, the results for the band gap would not improve.

### 2.4.3 Strongly correlated systems

By definition, systems with strong correlations cannot be described by a theory of independent particles. Instead a many-body description/theory is mandatory. The DFT does include all the many-body effects if the exact exchange-correlation potential was known. Due to intrinsic difficulties in transferring the many-body features into a functional depending on the electron density, all standard approximations to $E_\mathrm{xc}$ are bound to fail to describe such systems.

Typically, strongly correlated systems are materials that incorporate atoms with incompletely filled $d$- or $f$-shells. The electrons in these shells are under the influence of a strong Coulomb repulsion which leads to a wide range of interesting phenomena, such as magnetism, metal-to-insulator transitions, and high-temperature superconductivity. Much effort is put into correctly describing such systems by extending the DFT, most of them come with a loss of the "ab-inito" character of the method. A simple attempt to better account for correlations is made with the LDA+U method, described in the next section.

## 2.5 LDA+U method

The standard exchange-correlation functionals have a tendency to over-delocalize valence electrons which leads to a stabilization of metallic phases [34]. A typical example for this is NiO which is a prototype charge-transfer insulator [35]. For NiO the LDA predicts a metal whereas the LSDA correctly predicts an insulator, but with a much too small band gap and magnetic moment. In the LDA+U method [36, 37, 38] the tendency to over-delocalize electrons is corrected by identifying "strongly correlated" orbitals. Their electron-electron interaction is described by a Hubbard-like interaction term on a mean-field level.

In the presentation of the LDA+U formalism we follow the notation in [34]. We set $n(\mathbf{r}) := (n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r}))$ in this section and use the abbreviation "LDA" also for the spin polarized LSDA functional. The LDA+U energy functional is given by

$$E_{\mathrm{LDA+U}} = E_{\mathrm{LDA}}[n(\mathbf{r})] + E_{\mathrm{Hub}}[\left\{n^{\mu,s}_{m,m'}\right\}] - E_{\mathrm{dc}}[\left\{n^{\mu,s}_{m,m'}\right\}] \,, \tag{2.69}$$

where $E_{\mathrm{LDA}}[n(\mathbf{r})]$ is the total energy functional,

$$E_{\mathrm{LDA}}[n(\mathbf{r})] = T_S\left[n(\mathbf{r})\right] + V[n(\mathbf{r})] + E_{\mathrm{H}}\left[n(\mathbf{r})\right] + E^{\mathrm{LDA}}_{\mathrm{xc}}\left[n(\mathbf{r})\right] . \tag{2.70}$$

The values $\left\{n^{\mu,s}_{m,m'}\right\}$ represent the entries of the local density matrix of the orbitals "identified" as correlated ($m$ and $m'$ are the magnetic quantum numbers, $\mu$ the atomic site and $s$ the spin index). This identification is not completely unambiguous. Usually the Kohn-Sham eigenvalues get projected onto a set of predefined atomic $d$-levels ($l = 2$) or $f$-levels ($l = 3$) under the assumption that the atomic character of the corresponding wave function is mostly retained in the solid. Due to the strong localization of the correlated orbitals this assumption is reasonable. A derivation of the local density matrix for the LAPW basis is given in Section 2.3.5.

The term $E_{\mathrm{Hub}}[\left\{n^{\mu,s}_{m,m'}\right\}]$ in (2.69) reads in the rotationally invariant formulation [39]

$$\begin{aligned} E_{\mathrm{Hub}}[\left\{n^{\mu,s}_{m,m'}\right\}] = \frac{1}{2}\sum_{m,s,\mu} \Big\{ &\langle m, m''|V_{ee}|m', m'''\rangle n^{\mu,s}_{m,m'} n^{\mu,\bar{s}}_{m'',m'''} + \\ &\left(\langle m, m''|V_{ee}|m', m'''\rangle - \langle m, m''|V_{ee}|m''', m'\rangle\right) n^{\mu,s}_{m,m'} n^{\mu,s}_{m'',m'''} \Big\} \,. \end{aligned} \tag{2.71}$$

Obviously, only density-density terms of the Hubbard interaction Hamiltonian (3.4) enter the functional. The matrix elements of the Coulomb interaction

$$\langle m, m''|V_{ee}|m', m'''\rangle = \int \Psi^*_{l,m}(\mathbf{r})\Psi_{l,m'}(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\Psi^*_{l,m''}(\mathbf{r}')\Psi_{l,m'''}(\mathbf{r}')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \tag{2.72}$$

are written as a sum of Slater integrals $F^k$ and geometrical factors $a_k$

$$\langle m, m''|V_{ee}|m', m'''\rangle = \sum_{k=0}^{2l} a_k(m, m', m'', m''')F^k \,. \tag{2.73}$$

The geometrical factors read

$$a_k(m, m', m'', m''') = \frac{4\pi}{2k+1} \sum_{q=-k}^{k} \langle lm|Y_{k,q}|lm'\rangle\langle lm''|Y^*_{k,q}|lm'''\rangle \,, \tag{2.74}$$

$$\langle l, m|Y_{k,q}|l, m'\rangle = \int Y^*_{l,m}(\mathbf{r})Y_{k,q}(\mathbf{r})Y_{l,m'}(\mathbf{r})\,\mathrm{d}\Omega_r \,. \tag{2.75}$$

The explicit form of the Slater integrals is given by

$$F^k = \int r^2 r'^2 R^2(r) \frac{r^k_<}{r^{k+1}_>} R^2(r')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \,. \tag{2.76}$$

The symbol $r_<$ indicates the smaller value of $r$ and $r'$, $r_>$ indicates the larger value. $R(r)$ is the radial part of the wave function $\Psi_{l,m}(\mathbf{r})$ in (2.72), $Y_{l,m}(\mathbf{r})$ the angular part given by the usual spherical harmonics.

Note that the Slater integrals are not explicitly evaluated in the LDA+U method. For example for $l = 2$, only $F^0, F^2$ and $F^4$ are needed to evaluate (2.73). They are often expressed under further assumptions by only two parameters

$$U^{\mathrm{avg}} = \frac{1}{(2l+1)^2} \sum_{m,m'} \langle m, m'|V_{ee}|m, m'\rangle = F^0 \,, \tag{2.77}$$

$$J^{\mathrm{avg}} = \frac{1}{2l(2l+1)} \sum_{m\neq m'} \langle m, m'|V_{ee}|m', m\rangle = \frac{F^2 + F^4}{14} \,. \tag{2.78}$$

$U^{\mathrm{avg}}$ and $J^{\mathrm{avg}}$ enter the LDA+U method as free parameters and determine the Coulomb matrix elements in (2.71).

By adding (2.71) to the LDA Hamiltonian (2.70), the electron-electron interaction is accounted twice. This is the famous "double-counting" problem affecting all approaches that combine DFT calculations with model-based methods. The double-counting correction term $E_{\mathrm{dc}}[\{n^{\mu,s}_{m,m'}\}]$ subtracts the interaction energy which the LDA functional adds to the total energy for the correlated electrons. Due to the very nature of the density functional theory there is no rigorous derivation for the double-counting term. Thus, all double-counting corrections are at least semi-empiric. The double-counting problem still poses a difficult problem and is focus of current research.

A frequently used variant is the double-counting correction in the Fully Localized Limit (FLL) formulation [38]

$$E_{\mathrm{dc}}[\{n^{\mu,s}_{m,m'}\}] = \sum_{\mu} \left\{ \frac{U^{\mathrm{avg}}}{2} n^{\mu}(n^{\mu} - 1) - \frac{J^{\mathrm{avg}}}{2} \left[ n^{\mu,\uparrow}(n^{\mu,\uparrow} - 1) + n^{\mu,\downarrow}(n^{\mu,\downarrow} - 1) \right] \right\} \tag{2.79}$$

with the occupation numbers $n^{\mu,s} = \sum_m n^{\mu,s}_{m,m}$ and $n^{\mu} = n^{\mu,\uparrow} + n^{\mu,\downarrow}$. To get an intuitive understanding of this expression it is useful to be aware that terms of the form $n(n-1)/2$ in (2.79) simply count the number of pairs which can formed in a system of $n$ electrons. The interaction parameter $U$ is the screened Coulomb interaction and typically in the range of 2-10 eV, the exchange parameter $J$ lies in between 0 and 3 eV. For a half-filled

$d$-shell ($n = 5$) the double-counting correction energy can therefore reach up to 100 eV. The double-counting correction is discussed in more detail in Section 4.5.

The functional (2.69) is minimized by solving the Kohn-Sham equations with an additional term in the Kohn-Sham potential due to the Hubbard and the double-counting term. For $U^{\text{avg}} = J^{\text{avg}} = 0$, the original LDA functional is recovered. Often, the LDA functional is replaced by other semi-local functionals like the GGA leading to a more general class of "DFT+U" methods. The advantage of these methods lies in the fact that their computational cost does not increase much when the U-term is added. However the mean-field treatment of the correlations indicates that complex many-particle effects are not captured entirely.

## 2.6 Wannier functions: ab-initio tight-binding

In this section we introduce two methods to obtain Wannier functions from density functional theory calculations, first, the projection method, and second, the Marzari-Vanderbilt scheme for the construction of Maximally Localized Wannier Function (MLWF). The Wannier functions are used in an effective tight-binding description of the physical relevant bands of the underlying DFT Hamiltonian. As an example, in the scope of this work, we discuss results for CdTe.

### 2.6.1 Definition

Wannier Functions (WFs) were originally introduced by Wannier as Fourier-transformed Bloch functions [40]. For a single band in $\mathbf{k}$-space the Wannier function is written as

$$|\bar{W}_{\mathbf{R}}\rangle = \sqrt{\frac{1}{N}} \sum_{\mathbf{k}} e^{-\mathrm{i}\mathbf{k}\mathbf{R}} |\Phi_{\mathbf{k}}\rangle \,, \tag{2.80}$$

where $N$ is the number of primitive cells in the crystal, $|\Phi_{\mathbf{k}}\rangle$ a Bloch state. However, since the Bloch functions are only defined up to a phase factor, a more general definition is given by [41]

$$|W_{\mathbf{R}}\rangle = \sqrt{\frac{1}{N}} \sum_{\mathbf{k}} e^{-\mathrm{i}\mathbf{k}\mathbf{R}} e^{\mathrm{i}\phi(\mathbf{k})} |\Phi_{\mathbf{k}}\rangle \,, \tag{2.81}$$

with the $\mathbf{k}$-dependent real phase factor $\phi(\mathbf{k})$. In the presence of several bands there is no simple one-to-one mapping between a single band and a Wannier function. The general expression for the Wannier function in a multiband system is therefore [42]

$$|W_{\mathbf{R},n}\rangle = \sqrt{\frac{1}{N}} \sum_{\mathbf{k}} e^{-\mathrm{i}\mathbf{k}\mathbf{R}} \sum_{m} U^{(\mathbf{k})}_{m,n} |\Phi_{\mathbf{k},m}\rangle \,, \tag{2.82}$$

where $U^{(\mathbf{k})}_{m,n}$ is an arbitrary unitary matrix referred to as "gauge freedom" in the definition of Wannier functions. The such defined functions carry an index for the orbital $n$ and for

the real-space vector **R**. From a mathematical point of view, Bloch functions and Wannier functions are equivalent because they are connected by a unitary transformation. The gauge freedom in (2.82) can be chosen in such a way that the Wannier functions are well localized in real space. This localization is quantified by their spread, a property which we will define later in this section.

The Wannier functions form a set of orthogonal functions, so that "Wannier function" and "Wannier orbital" are used synonymously. In this work we will later use Wannier functions obtained from DFT calculations as a suitable basis for a Hubbard model (Section 3.1). The localization feature of the Wannier functions is then used to model a strong local Coulomb interaction present in $d$-type or $f$-type orbitals. Since this combination of standard band theory with more elaborated methods is an active field of research, Wannier functions gained a recent boost in practical importance [43]. In their seminal paper Marzari and Vanderbilt showed that the gauge freedom in the definition of Wannier functions can be completely eliminated by the requirement of a minimal spread [44]. This maximally MLWFs are not only used as "glue" between DFT and model based methods, MLWFs also provide insight in chemical bonding, for example by their visualization or by calculation of their centers and spreads [45].

### 2.6.2 Wannier functions by projection

The projection method, presented for example in [46], is a simple approach to construct Wannier functions. Trial orbitals $|g_n\rangle$ get projected onto the Bloch functions via

$$|\Phi'_{\mathbf{k},n}\rangle = \sum_m |\Phi_{\mathbf{k},m}\rangle\langle\Phi_{\mathbf{k},m}|g_n\rangle = \sum_m A^{(\mathbf{k})}_{m,n}|\Phi_{\mathbf{k},m}\rangle\,. \tag{2.83}$$

For these states, the overlap matrix $S'^{(\mathbf{k})}_{m,n} = \langle\Phi'_{\mathbf{k},m}|\Phi'_{\mathbf{k},n}\rangle$ is calculated. Using a standard Löwdin-orthogonalization [47] a set of orthogonal wave functions is constructed

$$|\tilde{\Phi}_{\mathbf{k},n}\rangle = \sum_m \left((S'^{(\mathbf{k})})^{-1/2}\right)_{m,n} |\Phi'_{\mathbf{k},m}\rangle\,, \tag{2.84}$$

from which the Projected Wannier Functions (PWFs) are calculated,

$$|W_{\mathbf{R},n}\rangle = \sqrt{\frac{1}{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{R}}|\tilde{\Phi}_{\mathbf{k},n}\rangle\,. \tag{2.85}$$

The projection-based Wannier orbitals depend on the choice of the trial orbitals $|g_n\rangle$. In many cases the Wannier orbitals retrieved by the projection method agree well with the maximally localized orbitals. This is especially true for localized $d$-orbitals and $f$-orbitals. The Löwdin orthogonalization conserves the symmetries of the trial wave functions as shown in Appendix A.2.

In this thesis we use Wannier orbitals from the LAPW basis. An obvious choice for the trial orbitals in this basis is [46]

$$g_n(\mathbf{r}^\mu) = \sum_L c_{n,L} u^\mu_{l,E_l}(\mathbf{r}^\mu) Y_L(\mathbf{r}^\mu)\,, \tag{2.86}$$

where $u_{l,E_l}(\mathbf{r}^\mu)$ is the solution of the radial Schrödinger equation in muffin tin $\mu$. The factors $c_{n,L}$ determine the symmetry of the trial orbital, e.g., to construct a trial orbital with $sp^3$ character we set $c_{n,(l,m)} = 1/2$ for $l \leq 1$, else the factor is zero. Using (2.54) the matrix elements $A^{(\mathbf{k})}_{m,n}$ in (2.83) take the simple form

$$A^{(\mathbf{k})}_{m,n} = \langle \Phi_{\mathbf{k},m} | g_n \rangle = \sum_L c_{n,L} \left[ \bar{a}^{\mu,\mathbf{k}}_L \right]^* . \tag{2.87}$$

### 2.6.3 Maximally localized Wannier functions

Maximally localized Wannier functions $|W_{\mathbf{R},n}\rangle$ minimize the total spread, given by [44]

$$\begin{aligned} \Omega[\left\{U^{(\mathbf{k})}_{m,n}\right\}] &= \sum_n \left[ \langle \hat{\mathbf{r}}^2 \rangle_n - \langle \hat{\mathbf{r}} \rangle^2_n \right] \\ &= \sum_n \left[ \langle W_{\mathbf{0},n} | \hat{\mathbf{r}}^2 | W_{\mathbf{0},n} \rangle - \langle W_{\mathbf{0},n} | \hat{\mathbf{r}} | W_{\mathbf{0},n} \rangle^2 \right] . \end{aligned} \tag{2.88}$$

The spread functional and its derivative with respect to the unitary matrix $U^{(\mathbf{k})}_{m,n}$ (for the implementation of a minimization algorithm) solely depend on the overlap matrix elements between the lattice periodic parts $u_{\mathbf{k},m}$ of the Bloch functions at nearest-neighbor $\mathbf{k}$-points $\mathbf{k}$ and $\mathbf{k}+\mathbf{b}$, $M^{\mathbf{k},\mathbf{b}}_{m,n} = \langle u_{\mathbf{k},m} | u_{\mathbf{k}+\mathbf{b},n} \rangle$. The ingenious derivation of the corresponding analytical expression is given in [44].

### 2.6.4 Wannier representation of the DFT Hamiltonian

The formulation of the DFT Hamiltonian in terms of the constructed Wannier orbitals is useful for the study of correlation effects using many-body techniques. The DFT Hamiltonian in spectral representation is given by

$$\hat{H} = \sum_{\mathbf{k},n} \epsilon_{\mathbf{k},n} | \Phi_{\mathbf{k},n} \rangle \langle \Phi_{\mathbf{k},n} | . \tag{2.89}$$

If the Bloch functions $|\Phi_{\mathbf{k},n}\rangle$ in (2.89) span the same subspace as the Wannier functions $|W_{\mathbf{R},n}\rangle$, $\hat{H}$ can be expressed in the Wannier basis

$$\hat{H} = \sum_{\mathbf{R},m} \sum_{\mathbf{R}',m'} H^{\mathbf{R}-\mathbf{R}'}_{m,m'} | W_{\mathbf{R},m} \rangle \langle W_{\mathbf{R}',m'} | , \tag{2.90}$$

where $H^{\mathbf{R}-\mathbf{R}'}_{m,m'}$ is the hopping integral between the Wannier orbitals $|W_{\mathbf{R},m}\rangle$ and $|W_{\mathbf{R}',m'}\rangle$,

$$H^{\mathbf{R}-\mathbf{R}'}_{m,m'} = \sum_{\mathbf{k},n} \epsilon_{\mathbf{k},n} \langle W_{\mathbf{R},m} | \Phi_{\mathbf{k},n} \rangle \langle \Phi_{\mathbf{k},n} | W_{\mathbf{R}',m'} \rangle . \tag{2.91}$$

Using operators $\hat{c}^\dagger_{\mathbf{R},m}$ ($\hat{c}_{\mathbf{R},m}$) creating (annihilating) an electron in the Wannier orbital with orbital index $m$ and real space index $\mathbf{R}$, the Hamiltonian can be written in the well-known tight-binding form

$$\hat{H} = \sum_{\mathbf{R},m} \sum_{\mathbf{R}',m'} \left( H^{\mathbf{R}-\mathbf{R}'}_{m,m'} \right) \hat{c}^\dagger_{\mathbf{R},m} \hat{c}_{\mathbf{R}',m'} . \tag{2.92}$$

Typically, a cut-off parameter is introduced, discarding all terms in the sum with $|\mathbf{R}-\mathbf{R}'| > R_{\max}$. The diagonalization of the tight-binding Hamiltonian reproduces the part of the DFT band structure from which the Wannier functions were constructed.

Since only a few bands (typically bands around the Fermi energy) of the full DFT Hamiltonian are reproduced, this approach is called "downfolding". In principle, the downfolded band structure is independent of the choice of $U_{mn}^{(\mathbf{k})}$. However, the cut-off parameter $R_{\max}$ needs to be larger for WF with large spreads. For too small $R_{\max}$, the resulting band structure does not reproduce the DFT band structure.

### 2.6.5 Entangled bands

In [48], the method of Marzari and Vanderbilt was extended to systems with entangled energy bands. It was realized that it is useful to decompose the spread functional (2.88) into two terms,

$$\Omega = \Omega_{\mathrm{I}} + \tilde{\Omega} \,, \tag{2.93}$$

where

$$\Omega_{\mathrm{I}} = \sum_{n} \left[ \langle \hat{\mathbf{r}}^2 \rangle_n - \sum_{\mathbf{R},m} |\langle W_{\mathbf{R},m} | \hat{\mathbf{r}} | W_{\mathbf{0},\mathbf{n}} \rangle|^2 \right] \tag{2.94}$$

and

$$\tilde{\Omega} = \sum_{n} \sum_{(\mathbf{R},m)\neq(\mathbf{0},n)} |\langle W_{\mathbf{R},m} | \hat{\mathbf{r}} | W_{\mathbf{0},n} \rangle|^2 \,. \tag{2.95}$$

The term (2.95) is gauge invariant if the Bloch functions from which the MLWF are constructed are not entangled with other bands [44].

An example for non-entangled bands are the valence bands of a semiconductor. The band gap separates the Bloch functions of the valence bands from the Bloch functions of the conduction bands. A well-defined set of $n$ Bloch functions can be specified from which $n$ Wannier functions are constructed.

However, if we want to construct Wannier functions for the conduction bands, we face the problem that we can not specify an energy window that includes a distinct set of $n$ bands, because several bands with different character are present, the so-called "entangled bands". A strategy to cope with the problem is to search for an $n$-dimensional "optimally connected" subspace in an energy window that fully includes the target bands. "Optimally connected" means that the dispersion of the bands in the subspace is minimal across the BZ [48]. In the corresponding reference it is shown that this is archived by minimizing the functional $\Omega_{\mathrm{I}}$ (the so called "disentanglement" procedure) before starting with the minimization of $\tilde{\Omega}$. An example for the choice of the energy window and the resulting downfolded band structure for the conduction bands above the Fermi energy is given for CdTe in the next subsection.

### 2.6.6 Wannier functions for CdTe

As an example, we discuss the construction of a tight-binding model for the semiconductor CdTe based on an FLAPW DFT calculation. CdTe crystallizes in the zincblende structure and can be described by an fcc lattice with a two atomic basis. For the DFT calculations we use the Jülich FLEUR program package [49] in combination with the Perdew–Burke–Enzernhof (PBE) exchange functional. The calculated band structure of the Kohn-Sham Hamiltonian is shown in Fig. 2.4. The band structure is gapped at the Fermi energy, as expected. The size of the gap at the Γ-point is $\epsilon_g = 0.76\,\text{eV}$, in agreement with previous calculations [50]. As described in Section 2.4.2, the DFT underestimates the single-particle gap for insulators. The (exciton) gap, which is a lower bound on the single-particle gap, is experimentally found at $\epsilon_{\text{ex}} = 1.5\,\text{eV}$ [51].

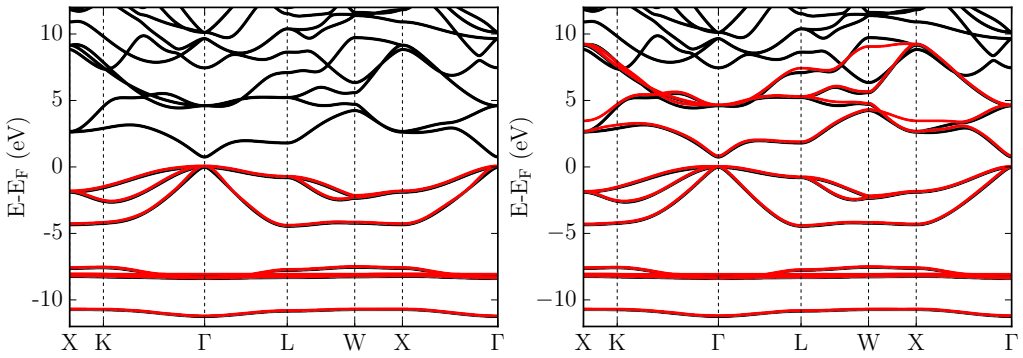


**Figure 2.4:** Black: DFT (GGA) band structure for CdTe. Red: Band structure of the effective downfolded tight-binding model using Wannier orbitals obtained from the projection method (b), see text. Left: downfolded band structure using Te-centered $sp^3$-orbitals and Cd-centered $d$-orbitals. Right: Additional Cd-centered $sp^3$-orbitals are used within the disentanglement procedure to obtain a tight-binding model, that includes the four lowest-lying conduction bands. For the disentanglement procedure an upper energy window of 10 eV was used.

In the atomic ground state, the electronic configuration of Cd is $[\text{Kr}]5\text{s}^2 4\text{d}^{10}$, the configuration of Te is $[\text{Kr}]5\text{s}^2 4\text{d}^{10} 5\text{p}^4$. Therefore, we expect the valence bands with $s$-type symmetry, located at about −11 eV, and $p$-type symmetry, located right below the Fermi energy, to have predominantly Te $s$-character and Te $p$-character. Hence, we chose atomic Te $s$-, $p_x$-, $p_y$-, and $p_z$-orbitals as trial orbitals in the Wannier procedure based on the projection method described in Section 2.6.2. The Cd 4d bands at around −8 eV are described using trial orbitals with $t_{2g}$ and $e_g$ character, see Section 4.3 for their definition. We have a basis of nine orbitals to describe the valence bands obtained in the DFT calculation. We refer to this particular choice of orbitals for the calculation using projected Wannier orbitals as method (a). For comparison, we repeat the calculations with $\text{sp}^3$-orbitals located at the Te-sites in (b). In a third calculation (c) we use the Wannier90 program code [52] to obtain maximally localized Wannier functions. The results for the spreads and centers of the orbitals of all three calculations are listed in Table 2.1.

| | | PWF (a) | | | PWF (b) | | | MLWF (c) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| WF | site | proj | $\Omega_n$ [a.u.$^2$] | $\langle\hat{\mathbf{r}}\rangle_n$ [a.u.] | proj | $\Omega_n$ [a.u.$^2$] | $\langle\hat{\mathbf{r}}\rangle_n$ [a.u.] | proj | $\Omega_n$ [a.u.$^2$] | $\langle\hat{\mathbf{r}}\rangle_n$ [a.u.] |
| 1 | Cd | dz2 | 2.134 | 1.53, 1.53, 1.53 | dz2 | 2.134 | 1.53, 1.53, 1.53 | dz2 | 2.094 | 1.53, 1.53, 1.53 |
| 2 | Cd | dxz | 2.187 | 1.53, 1.53, 1.53 | dxz | 2.187 | 1.53, 1.53, 1.53 | dxz | 2.229 | 1.53, 1.53, 1.53 |
| 3 | Cd | dyz | 2.187 | 1.53, 1.53, 1.53 | dyz | 2.187 | 1.53, 1.53, 1.53 | dyz | 2.229 | 1.53, 1.53, 1.53 |
| 4 | Cd | dx2-y2 | 2.134 | 1.53, 1.53, 1.53 | dx2-y2 | 2.134 | 1.53, 1.53, 1.53 | dx2-y2 | 2.094 | 1.53, 1.53, 1.53 |
| 5 | Cd | dxy | 2.187 | 1.53, 1.53, 1.53 | dxy | 2.187 | 1.53, 1.53, 1.53 | dxy | 2.229 | 1.53, 1.53, 1.53 |
| 6 | Te | s | 5.764 | −1.53, −1.53, −1.53 | sp3-1 | 10.53 | −1.95, −1.95, −1.95 | sp3-1 | 8.312 | −0.60, −0.60, −0.60 |
| 7 | Te | pz | 12.76 | −1.53, −1.53, −1.53 | sp3-2 | 10.53 | −1.95, −1.11, −1.11 | sp3-2 | 8.312 | −0.60, −2.46, −2.46 |
| 8 | Te | px | 12.76 | −1.53, −1.53, −1.53 | sp3-3 | 10.53 | −1.11, −1.95, −1.11 | sp3-3 | 8.312 | −2.46, −0.60, −2.46 |
| 9 | Te | py | 12.76 | −1.53, −1.53, −1.53 | sp3-4 | 10.53 | −1.11, −1.11, −1.95 | sp3-4 | 8.312 | −2.46, −2.46, −0.60 |
| $\sum \Omega_n$[a.u.$^2$] | | 54.89 | | | 52.93 | | | 44.13 | | |

**Table 2.1:** Numerical results for three methods to construct Wannier functions from DFT results for the valence bands of CdTe. In method (a) and (b) the Wannier orbitals are obtained by the projection method described in Section 2.6.2. The symmetry of the projection is given in the "proj"-column. The *d*-type wave functions are projected onto the Cd-site, all other onto the Te-site. Method (c) minimizes the spread functional Ω. The table shows the center and spreads of each Wannier function $\langle\hat{r}\rangle_n$. The positions of the two atoms in the unit cell of are (1.53, 1.53, 1.53) (a.u.) for the Cd atom, and -(1.53, 1.53, 1.53) a.u. for the Te atom. WF 9 is visualized for each of the three methods in Fig. 2.5.

In all three methods the center of the $d$-orbitals is located exactly at the Cd site. The spreads of the $d$-orbitals are small (about 2 a.u.$^2$), compared to the spreads of the $s$-type and $p$-type orbitals (10-12 a.u.$^2$). We also find a slight deviation between the spreads of the $e_g$ and $t_{2g}$ orbitals. The total spread in method (b) is smaller than in method (a). The optimal choice of orbitals, in the sense of minimal overlap (or maximally localization), is given by the orbitals obtained in method (c). The total spread is significantly reduced (by about 20%) for these orbitals. The gain in localization stems predominantly from a better localization of the $s$-type and $p$-type orbitals.

We plot the isosurface of the square of one of the Wannier functions of each method in Fig. 2.5. From the isosurface plot we see that the MLWF obtained in method (c) are shifted towards the Cd atom. The orbitals from the MLWF method matches the chemical picture of four bonding orbitals between a Te atom and its four nearest neighbor Cd atoms.

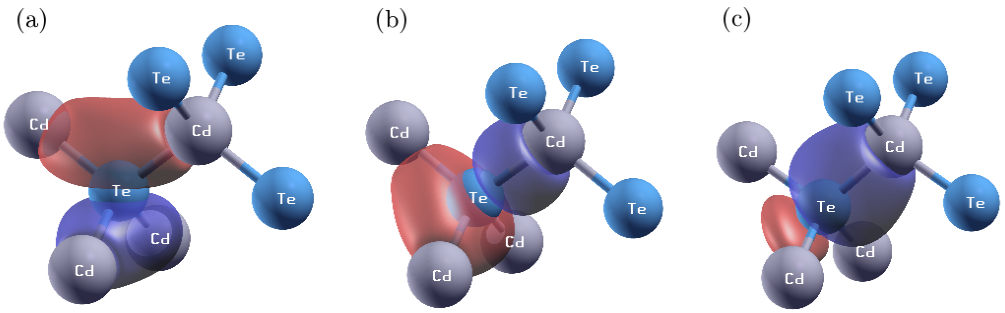


**Figure 2.5:** Isosurface= ±0.03 plot of the square of the Wannier orbitals constructed from DFT calculations for the valence bands of CdTe. (a) the $p$-type Wannier orbital WF9 using the projection method with one $s$- and three $p$-orbitals, (b) the $sp^3$-type Wannier orbital WF9 using the projection method, and (c) the maximally localized Wannier orbital WF9.

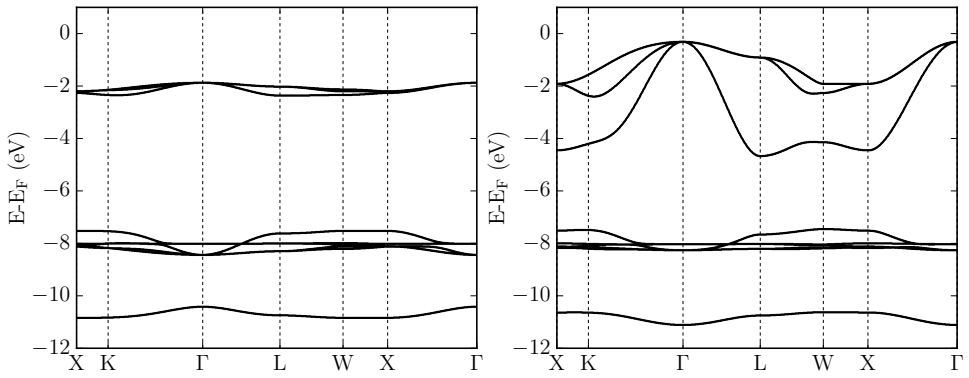


**Figure 2.6:** Band structure of the tight-binding model calculated with method (b) for different values of the cut-off parameter $R_{\max}$. Left: $R_{\max}$ is chosen to include only nearest neighbor hopping matrix elements. Right: Next-nearest neighbor hopping reproduces the main characteristics of the DFT band structure.

As seen from Fig. 2.6 the inclusion of at least next-nearest neighbor hoppings is required to reproduce the main features of the DFT band structure. To also describe the lowest lying conduction bands we have to employ the disentanglement procedure described in Section 2.6.5. We chose four additional $sp^3$ orbitals located at the Cd site as anti-bonding orbitals in the projection method (b). The upper limit for the energy window is chosen to be 10 eV. Using the disentanglement procedure we are now able to recover also the band structure of the conduction bands, as shown in Fig. 2.4. The orbital character of the band structure using the projection method is shown in Fig. 2.7. As expected, the valence bands have mostly Te character, whereas the conduction bands have Cd character. The narrow Cd 4*d*-bands do not significantly hybridize. Therefore, these bands are left out in our effective tight-binding description in Section 4.2.

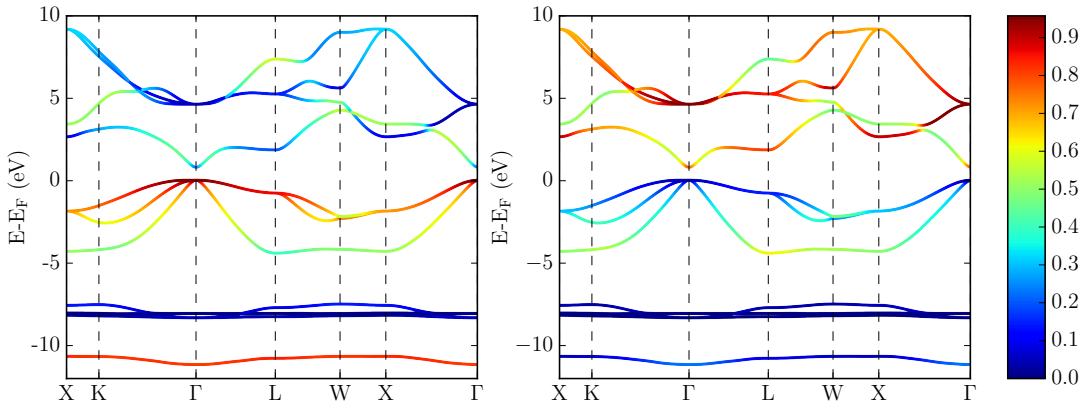


**Figure 2.7:** Band structure of the tight-binding model for CdTe using $sp^3$ orbitals at the Te and $sp^3$ and $d$ orbitals on the Cd sites. The color coding shows the square of the overlap of the band structure with the Te-$sp^3$ orbitals (left) and the overlap with the Cd-$sp^3$ orbitals (right).

In this work we use the DFT Wannier basis as a starting point for a more elaborated many-body calculations using the Gutzwiller method that we introduce in the next chapter. We expect that the particular choice of the Wannier basis is not crucial for the outcome of our calculations because the $d$-orbitals which are treated as correlated in the Gutzwiller method are strongly localized and vary only slightly across the three methods presented here. Therefore, we use orbitals from the simple projection method (b) in all following Wannier calculations. The influence of the choice of the Wannier basis is discussed in the literature on the Dynamical Mean Field Theory (DMFT) [53].

# 3 Gutzwiller variational method

The Gutzwiller approach to the many-body problem is a variational method originally introduced by M. Gutzwiller in a series of seminal papers [54, 55, 56] to study the effect of Coulomb interactions in a ferromagnetic single-band Hubbard model[1]. The generalization of the Gutzwiller method to multiband models [58] provides a powerful tool for the study of strong correlations in transition metals and their compounds.

In this chapter we recapitulate the derivation of the multiband Hubbard model from the fundamental electronic Hamiltonian, and present the Gutzwiller formalism as a suitable method for investigating its ground-state properties. We summarize the numerical procedure which is used in this work. An intepretation of obtained band structures as energies of a quasi-particle system closes this chapter.

## 3.1 Hubbard model

To motivate the Hubbard model we express the electronic Hamiltonian (2.2) in a basis of single-particle wave functions. We start by solving the atomic problem (one electron in a static external Coulomb potential)

$$\left(-\frac{1}{2}\nabla_{\mathbf{r}}^2 + \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}|}\right)\varphi_{i,\sigma}(\mathbf{r}) = \epsilon_{i,\sigma}(\mathbf{r})\varphi_{i,\sigma}(\mathbf{r}) \tag{3.1}$$

for each site $i$ located at $\mathbf{R}_i$. The combined spin-orbital index is denoted by $\sigma$. The wave functions at each site are orthogonal but have a non-vanishing overlap with wave functions on neighboring sites. We introduce the combined index $\tilde{\sigma} = (i, \sigma)$ and calculate the overlap matrix

$$S_{\tilde{\sigma},\tilde{\sigma}'} = \langle\varphi_{\tilde{\sigma}}|\varphi_{\tilde{\sigma}'}\rangle = \int \mathrm{d}\mathbf{r}\, \varphi_{\tilde{\sigma}}^*(\mathbf{r})\varphi_{\tilde{\sigma}'}(\mathbf{r}) \ . \tag{3.2}$$

Using a Löwdin orthogonalization [47]

$$|\phi_{\tilde{\sigma}}\rangle = \sum_{\tilde{\sigma}'}(\boldsymbol{S}^{-1/2})_{\tilde{\sigma}',\tilde{\sigma}}|\varphi_{\tilde{\sigma}'}\rangle \ , \tag{3.3}$$

we construct a new set of orthogonal orbitals $|\phi_{\tilde{\sigma}}\rangle$. In the limit of an infinite lattice constant, where the overlap of neighboring sites vanishes, these new orbitals resemble the original orbitals $|\varphi_{\tilde{\sigma}}\rangle$. Bringing neighboring sites closer together changes the character of the orthogonal orbitals, which fulfil all properties of Wannier orbitals. If the overlap matrix

[1] The Hubbard model itself was independently introduced by Gutzwiller, Hubbard, and Kanamori in 1963 [57].

elements is not too large, these Wannier orbitals retain most of their atomic character, and the index $i$ remains meaningful in the sense that the new orbitals are well localized at each site.

We continue by expressing the electronic Hamiltonian (2.4) in this new orthogonal basis and drop all terms of the Coulomb interaction that include different sites ("Hubbard approximation"). One arrives at the well-known form of the multiband Hubbard model

$$\begin{aligned}\hat{H}_{\mathrm{Hub}} =& \hat{H}_{\mathrm{kin}} + \hat{H}_{\mathrm{int}} \\ =& \sum_{i,j} \sum_{\sigma_1,\sigma_2} t_{i,j}^{\sigma_1,\sigma_2} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{j,\sigma_2} + \frac{1}{2} \sum_{i} \sum_{\sigma_1,\sigma_2,\sigma_3,\sigma_4} U_i^{\sigma_1,\sigma_2,\sigma_3,\sigma_4} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{i,\sigma_2}^{\dagger} \hat{c}_{i,\sigma_3} \hat{c}_{i,\sigma_4}\end{aligned} \tag{3.4}$$

with

$$t_{i,j}^{\sigma_1,\sigma_2} = \delta_{s_1,s_2} \int \mathrm{d}\mathbf{r}\, \phi_{i,\sigma_1}^{*}(\mathbf{r}) \left( -\frac{1}{2}\nabla_{\mathbf{r}}^2 + V(\mathbf{r}) \right) \phi_{j,\sigma_2}(\mathbf{r}) \,, \tag{3.5}$$

$$U_i^{\sigma_1,\sigma_2,\sigma_3,\sigma_4} = \delta_{s_1,s_4}\delta_{s_2,s_3} \int \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{r}'\, \phi_{i,\sigma_1}^{*}(\mathbf{r}) \phi_{i,\sigma_2}^{*}(\mathbf{r}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \phi_{i,\sigma_3}(\mathbf{r}') \phi_{i,\sigma_4}(\mathbf{r}) \,, \tag{3.6}$$

where $s_i$ denotes the spin component of $\sigma_i$.

Following the procedure described here, we end up with operators $\hat{c}_{i,\sigma}^{\dagger}$ ($\hat{c}_{i,\sigma}$) that create (annihilate) electrons in orbitals which generally do not possess the proper site symmetry, i.e., the corresponding wave functions can not be classified by the representation of the point-symmetry group of the lattice. This property can be recovered by initially choosing the eigenfunctions $\varphi_{i,\sigma}(\mathbf{r})$ in (3.1) to respect the symmetry of the lattice (using a proper unitary transformation of the basis functions spanning degenerate subspaces), c.f., $d$-orbitals in a cubic environment (4.7).

We point out the similarity between the construction of Wannier orbitals from DFT calculations using the projection method in Section 2.6.2 and the procedure described here. In this work we use the Hubbard model with basis functions derived from DFT as a starting point for Gutzwiller calculations.

## 3.2 Hartree-Fock method

We derive the Hartree-Fock equations using a variational method, following [59]. The Hartree-Fock method yields an upper bound for the ground-state energy of the exact solution and for the solution of the Gutzwiller method introduced in the next section. The solution of the Hartree-Fock equations leads to the optimal Slater determinant for an interacting system, in the sense of a minimal expectation value for the energy.

A general Slater determinant in second quantization is written as

$$|\Phi_0\rangle = \prod_{\epsilon_\gamma < E_{\mathrm{F}}} h_\gamma^{\dagger} |\mathrm{vac}\rangle \,, \tag{3.7}$$

where all one-particle states $|\gamma\rangle$ with energy $\epsilon_\gamma$ below the Fermi-energy $E_{\mathrm{F}}$ are occupied. Our goal is to determine a state of the form (3.7) that minimizes the energy of the Hubbard

Hamiltonian (3.4). The operators $\hat{h}^\dagger_\gamma$ and $\hat{c}^\dagger_\sigma$ are connected by the unitary transformation $\boldsymbol{u}$,

$$\hat{h}^\dagger_\gamma = \sum_{\tilde{\sigma}} u_{\gamma,\tilde{\sigma}} \hat{c}^\dagger_{\tilde{\sigma}} , \quad \hat{c}^\dagger_{\tilde{\sigma}} = \sum_{\gamma} u^*_{\gamma,\tilde{\sigma}} \hat{h}^\dagger_\gamma \quad \text{with} \quad \sum_{\tilde{\sigma}} u^*_{\gamma,\tilde{\sigma}} u_{\gamma',\tilde{\sigma}} = \delta_{\gamma,\gamma'} . \tag{3.8}$$

Again, we use the index $\tilde{\sigma} = (\sigma, i)$ that combines the site and the spin-orbital index. For expectation values of an operator $\hat{A}$ with respect to a single-particle wave function we introduce the notation

$$\langle \hat{A} \rangle_{\Phi_0} = \langle \Phi_0 | \hat{A} | \Phi_0 \rangle . \tag{3.9}$$

The expectation value of the energy is given by

$$\langle \hat{H}_{\text{Hub}} \rangle_{\Phi_0} = \sum_{\tilde{\sigma}_1,\tilde{\sigma}_2} t_{\tilde{\sigma}_1,\tilde{\sigma}_2} \langle \hat{c}^\dagger_{\tilde{\sigma}_1} \hat{c}_{\tilde{\sigma}_2} \rangle_{\Phi_0} + \frac{1}{2} \sum_{\tilde{\sigma}_1,\tilde{\sigma}_2,\tilde{\sigma}_3,\tilde{\sigma}_4} U_{\tilde{\sigma}_1,\tilde{\sigma}_2,\tilde{\sigma}_3,\tilde{\sigma}_4} \langle \hat{c}^\dagger_{\tilde{\sigma}_1} \hat{c}^\dagger_{\tilde{\sigma}_2} \hat{c}_{\tilde{\sigma}_3} \hat{c}_{\tilde{\sigma}_4} \rangle_{\Phi_0} \tag{3.10}$$

with the obvious condition that $U_{\tilde{\sigma}_1,\ldots,\tilde{\sigma}_4}$ is zero for indices containing different sites, c.f. (3.4). Applying Wick's theorem to the interaction part yields

$$\langle \hat{c}^\dagger_{\tilde{\sigma}_1} \hat{c}^\dagger_{\tilde{\sigma}_2} \hat{c}_{\tilde{\sigma}_3} \hat{c}_{\tilde{\sigma}_4} \rangle_{\Phi_0} = \langle \hat{c}^\dagger_{\tilde{\sigma}_1} \hat{c}_{\tilde{\sigma}_4} \rangle_{\Phi_0} \langle \hat{c}^\dagger_{\tilde{\sigma}_2} \hat{c}_{\tilde{\sigma}_3} \rangle_{\Phi_0} - \langle \hat{c}^\dagger_{\tilde{\sigma}_1} \hat{c}_{\tilde{\sigma}_3} \rangle_{\Phi_0} \langle \hat{c}^\dagger_{\tilde{\sigma}_2} \hat{c}_{\tilde{\sigma}_4} \rangle_{\Phi_0} . \tag{3.11}$$

Therefore, the expression that has to be minimized with respect to the entries of the transformation matrix $\boldsymbol{u}$ is given by

$$\begin{aligned} \langle \hat{H}_{\text{Hub}} \rangle_{\Phi_0} = & \sum_{\tilde{\sigma}_1,\tilde{\sigma}_2} t_{\tilde{\sigma}_1,\tilde{\sigma}_2} \sum_{\epsilon_\gamma < E_\text{F}} u^*_{\gamma,\tilde{\sigma}_1} u_{\gamma,\tilde{\sigma}_2} + \\ & \sum_{\tilde{\sigma}_1,\tilde{\sigma}_2,\tilde{\sigma}_3,\tilde{\sigma}_4} \frac{U_{\tilde{\sigma}_1,\tilde{\sigma}_2,\tilde{\sigma}_3,\tilde{\sigma}_4}}{2} \sum_{\epsilon_\gamma,\epsilon_{\gamma'} < E_\text{F}} \left( u^*_{\gamma,\tilde{\sigma}_1} u_{\gamma,\tilde{\sigma}_4} u^*_{\gamma',\tilde{\sigma}_2} u_{\gamma',\tilde{\sigma}_3} - u^*_{\gamma,\tilde{\sigma}_1} u_{\gamma,\tilde{\sigma}_3} u^*_{\gamma',\tilde{\sigma}_2} u_{\gamma',\tilde{\sigma}_4} \right) . \end{aligned} \tag{3.12}$$

Minimizing (3.12) and accounting for the property of unitarity of $\boldsymbol{u}$ (3.8) by introducing the Lagrange parameters $\epsilon_{\gamma'}$ and $s_{\gamma'\gamma''}$ yields the equation

$$\frac{\partial}{\partial u^*_{\gamma,\tilde{\sigma}}} \left[ \langle \hat{H}_{\text{Hub}} \rangle_{\Phi_0} - \sum_{\gamma'} \epsilon_{\gamma'} \left( \sum_{\tilde{\sigma}'} |u_{\gamma',\tilde{\sigma}'}|^2 - 1 \right) - \sum_{\gamma' \neq \gamma''} s_{\gamma'\gamma''} \sum_{\tilde{\sigma}'} u_{\gamma'\tilde{\sigma}'} \left( u_{\gamma'',\tilde{\sigma}'} \right)^* \right] = 0 . \tag{3.13}$$

Equation (3.13) is solved by the effective one-particle Schrödinger equation

$$\sum_{\tilde{\sigma}'} \left( t_{\tilde{\sigma},\tilde{\sigma}'} + \Sigma_{\tilde{\sigma},\tilde{\sigma}'} \right) u_{\gamma,\tilde{\sigma}'} = \epsilon_\gamma u_{\gamma,\tilde{\sigma}} \tag{3.14}$$

with the Hartree-Fock "self-energy"

$$\Sigma_{\tilde{\sigma},\tilde{\sigma}'} = \sum_{\tilde{\sigma}_1,\tilde{\sigma}_2} \left( U_{\tilde{\sigma},\tilde{\sigma}_1,\tilde{\sigma}_2,\tilde{\sigma}'} - U_{\tilde{\sigma},\tilde{\sigma}_1,\tilde{\sigma}',\tilde{\sigma}_2} \right) \sum_{\epsilon_\gamma < E_\text{F}} u^*_{\gamma,\tilde{\sigma}_1} u_{\gamma,\tilde{\sigma}_2} , \tag{3.15}$$

where we have used the symmetry relation $U_{\tilde{\sigma}_1,\tilde{\sigma}_2,\tilde{\sigma}_3,\tilde{\sigma}_4} = U_{\tilde{\sigma}_2,\tilde{\sigma}_1,\tilde{\sigma}_4,\tilde{\sigma}_3}$. The first term in (3.15) is known as the "Hartree" contribution, the second term as the "Fock" contribution. Equations (3.14) and (3.15) need to be solved self-consistently. The expectation value for the ground-state energy is then given by

$$\langle \hat{H}_{\text{Hub}} \rangle_{\Phi_0} = \sum_{\epsilon_\gamma < E_\text{F}} \epsilon_\gamma - \langle \hat{H}_{\text{int}} \rangle_{\Phi_0} . \tag{3.16}$$

## 3.3 Gutzwiller wave function and approximation

### 3.3.1 Single-band case

Historically, M. Gutzwiller introduced the Gutzwiller wave function to study the effect of strong correlations in a single-band Hubbard model

$$\hat{H}^{\mathrm{s}}_{\mathrm{Hub}} = -t \sum_{i,j,s} \hat{c}^{\dagger}_{i,s} \hat{c}_{j,s} + \frac{U}{2} \sum_{i} \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow} \,. \tag{3.17}$$

For the purpose of illustrating the fundamental ideas in the Gutzwiller method, we start our considerations using this model.

Typically, one has $t > 0$ and a repulsive potential $U > 0$. For the half-filled model, with one electron per orbital, it is obvious that the former term favours delocalized electrons, whereas the latter term favours localization of the electrons at the sites. As we will see, the Gutzwiller wave function accommodates for both of these "limiting cases" ($t \to \infty$ or $U \to \infty$).

An isolated orbital can be unoccupied, occupied by two electrons in different spin states, or occupied by one electron carrying spin $s$. We denote the expectation values of the four eigenstates of the interaction term by $m^0_{\varnothing}$ (no occupation), $m^0_s$ (single occupied by an electron with spin $s$) and $m^0_d$ (doubly occupied). Evaluating this expectation values with respect to any Slater determinant (for example obtained from the Hartree-Fock method) yields

$$m^0_d = n^0_{\uparrow} n^0_{\downarrow} \,, \tag{3.18a}$$

$$m^0_s = n^0_s (1 - n^0_{\bar{s}}) \,, \tag{3.18b}$$

$$m^0_{\varnothing} = (1 - n^0_{\uparrow})(1 - n^0_{\downarrow}) \,, \tag{3.18c}$$

where $n^0_s$ are the expectation values for finding an electron with spin $s$ on a specific site, e.g., for a half-filled Hubbard model $n^0_{\uparrow} + n^0_{\downarrow} = 1$ holds. The expectation values for the four eigenstates of the isolated orbital are given by simple statistical considerations. For example in the case of no spin-polarization and half-filling the double occupation is always given by $m_d = 1/4$.

It is readily understood that this limitation of the single-particle wave function is problematic if we want to account for correlations. At least for the half-filled case with $U \to \infty$ one would expect that the system becomes insulating: charge fluctuations are completely suppressed, resulting in a expectation value of the double occupation of $m_d = 0$. As shown in Appendix A.3, the Hartree-Fock solution displays unphysical spin-polarization already for moderate values of $U$, because breaking the $s_z$ symmetry is the only mechanism to avoid costly double occupations, compare (3.18).

Gutzwiller introduced a wave function, that includes the variational freedom to suppress double occupancies in a Slater determinant $|\Phi_0\rangle$ by applying the Gutzwiller correlator $\hat{P}^s_{\mathrm{G}}$, yielding the Gutzwiller wave function

$$|\Psi^s_{\mathrm{G}}\rangle = \hat{P}^s_{\mathrm{G}} |\Phi_0\rangle = \prod_i \hat{P}^s_i |\Phi_0\rangle \quad \text{with} \quad \hat{P}^s_i = 1 - (1 - g)\hat{d}_i \,, \tag{3.19}$$

where $\hat{P}_i^s$ is the local Gutzwiller correlator acting on site $i$, $\hat{d}_i = \hat{n}_{i,\uparrow}\hat{n}_{i,\downarrow}$ the double occupation operator, and $g \in [0,1]$, a variational parameter. For $g = 1$ the Gutzwiller wave function is equal to the Slater determinant, for $g = 0$ the double occupied sites are completely projected out. Since the Slater determinant is the exact solution for the case $U = 0$, $g = 1$ is called the "uncorrelated" limit, whereas $U \to \infty$, leading to $g = 0$, is the "atomic" limit. The Gutzwiller method is a variational method, i.e., the variational principle provides an upper bound on the ground-state energy, by minimizing $|\Psi_{\mathrm{G}}^{\mathrm{s}}\rangle$ with respect to the single-particle state $|\Phi_0\rangle$ and the variational parameter $g$. Since the variational space includes the Hartree-Fock solution, the Gutzwiller method results in equal or lower energies.

It is important to note that the single-particle state $|\Phi_0\rangle$ (accounting for the kinetic term) is a delocalized wave function "living" in $\mathbf{k}$-space, whereas the Gutzwiller correlator (accounting for the correlations) is a local operator acting on each site in real space. This complicates the evaluation of expectation values with respect to the many-particle wave function $|\Psi_{\mathrm{G}}^{\mathrm{s}}\rangle$. Even for the one-band model discussed here, the exact expressions for the energy are only available for the case of one dimension [60, 61] or infinite dimensions [58]. Gutzwiller used an approximation of the ground-state energy which later turned out to be the exact solution in infinite spatial dimensions. The corresponding Gutzwiller solution for the one-band model can be readily derived from the expressions for the general multiband case introduced in the next subsection.

### 3.3.2 Multiband case

It is convenient to separate the multiband Hubbard model (3.4) into a non-local and a local part

$$\hat{H}_{\mathrm{Hub}} = \sum_{i \neq j} \sum_{\sigma,\sigma'} t_{i,j}^{\sigma,\sigma'} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{j,\sigma'} + \sum_i \hat{H}_{i,\mathrm{loc}} \tag{3.20}$$

with

$$\hat{H}_{i,\mathrm{loc}} = \sum_{\sigma_1,\sigma_2} t_{i,i}^{\sigma_1,\sigma_2} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{i,\sigma_2}^{\dagger} + \sum_{\sigma_1,\sigma_2,\sigma_3,\sigma_4} U_i^{\sigma_1,\sigma_2,\sigma_3,\sigma_4} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{i,\sigma_2}^{\dagger} \hat{c}_{i,\sigma_3} \hat{c}_{i,\sigma_4} \ . \tag{3.21}$$

The Gutzwiller wave function

$$|\Psi_{\mathrm{G}}\rangle = \hat{P}_{\mathrm{G}} |\Phi_0\rangle = \prod_i \hat{P}_i |\Phi_0\rangle \tag{3.22}$$

is generalized to the multiband case by introducing the local Gutzwiller correlator

$$\hat{P}_i = \sum_{\Gamma_i \Gamma_i'} \lambda_{i;\Gamma_i,\Gamma_i'} \hat{m}_{i;\Gamma_i,\Gamma_i'} \quad \text{with} \quad \hat{m}_{i;\Gamma_i,\Gamma_i'} = |\Gamma\rangle_{ii}\langle\Gamma'| \ . \tag{3.23}$$

The variational parameters are denoted by $\lambda_{i;\Gamma_i,\Gamma_i'}$ where $|\Gamma\rangle_i$ are the eigenstates of the local Hamiltonian (3.21)

$$\hat{H}_{i,\mathrm{loc}} = \sum_{\Gamma} E_{i;\Gamma} \hat{m}_{i;\Gamma_i,\Gamma_i} \ . \tag{3.24}$$

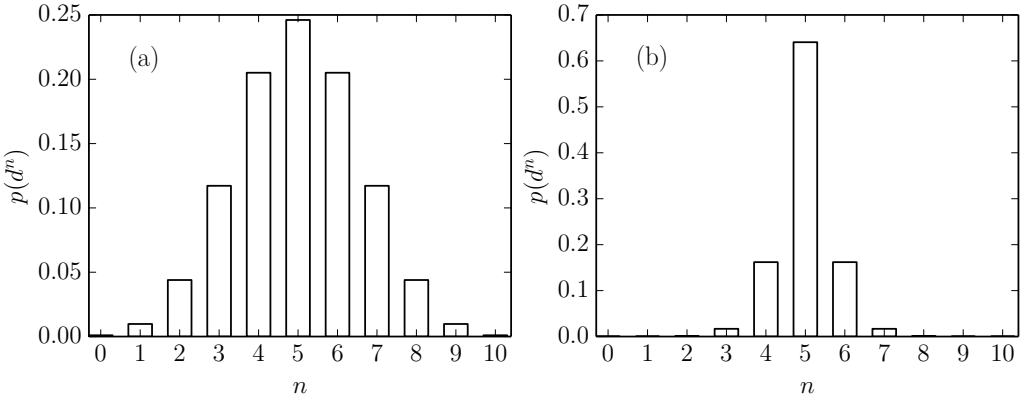


**Figure 3.1:** Local Hartree-Fock expectation values for a $d^n$-occupation of the half-filled $d$-shell with 50% spin-polarization (left) and 95% spin polarization (right).

The statistical considerations made for the four eigenstates of the interaction Hamiltonian of the single-band model can also be generalized to the multiband case. The diagonalization of the local interaction Hamiltonian in a multiband Hubbard model, accounting for the correlations in $d$-orbitals (see Section 4.3), yields a set of $2^{10}$ eigenstates. All 1024 corresponding diagonal expectation values, which are evaluated with respect to a single-particle wave function, can be expressed by the ratio of the two values $n_\uparrow^0$ and $n_\downarrow^0$ (assuming degeneracy of the orbitals). To illustrate this we plot the Hartree-Fock expectation values for a $d^n$ occupation ($n = 0, \ldots, 10$) of a half-filled model in Fig. 3.1 for two values of the spin-polarization. Since $d^n$ eigenstates are energetically more favourable then $d^{n+1}$ states one expects, for strong local correlations, a much more narrow distribution around the expectation value of five electrons than given by the paramagnetic Hartree-Fock solution, Fig. 3.1(a). As discussed for the one-band model in the previous subsection, a spin-polarized Hartree-Fock wave function also typically lowers the energy for half-filled shells, compare with the "improved" distribution for the spin-polarized case shown in Fig. 3.1(b). Note, however, that such spin-polarization is not observed in most real systems. From these considerations it becomes clear that the generalization of the Gutzwiller wave function (3.19) to multiband models greatly improves the ground-state description of strongly correlated multiband systems.

### 3.3.3 Gutzwiller approximation and energy functional

The evaluation of expectation values with respect to the Gutzwiller wave function without further approximations poses a generally unsolvable many-body problem. However, it is possible to evaluate expectation values in the limit of infinite lattice coordination number. Applying these results to models in finite dimensions is the so-called "Gutzwiller approximation". In this subsection we do not derive but only summarize the results of the diagrammatic evaluation in infinite-dimensions obtained by Bünemann et al. [59].

We are interested in evaluating expectation values of the form

$$\langle \hat{O}_i \rangle_{\Psi_G} = \frac{\langle \Phi_0 | \left[ \prod_{m \neq i} \hat{P}_m^\dagger \hat{P}_m \right] \hat{P}_i^\dagger \hat{O}_i \hat{P}_i | \Phi_0 \rangle}{\langle \Phi_0 | \left[ \prod_m \hat{P}_m^\dagger \hat{P}_m \right] | \Phi_0 \rangle} \tag{3.25}$$

$$\langle \hat{O}_{i,j} \rangle_{\Psi_G} = \frac{\langle \Phi_0 | \left[ \prod_{m \neq i,j} \hat{P}_m^\dagger \hat{P}_m \right] \hat{P}_j^\dagger \hat{P}_i^\dagger \hat{O}_{i,j} \hat{P}_i \hat{P}_j | \Phi_0 \rangle}{\langle \Phi_0 | \left[ \prod_m \hat{P}_m^\dagger \hat{P}_m \right] | \Phi_0 \rangle}, \tag{3.26}$$

where $\hat{O}_i$ is any operator acting only on site $i$ and $\hat{O}_{i,j} = \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma'}$ for $i \neq j$. Imposing the constraints

$$\langle \hat{P}_i^\dagger \hat{P}_i \rangle_{\Phi_0} = 1 \,, \tag{3.27}$$

$$\langle \hat{c}_{i,\sigma}^\dagger \hat{P}_i^\dagger \hat{P}_i \, \hat{c}_{i,\sigma'} \rangle_{\Phi_0} = \langle \hat{c}_{i,\sigma}^\dagger \hat{c}_{i,\sigma'} \rangle_{\Phi_0} \,, \tag{3.28}$$

the expectation values in infinite dimensions are given by

$$\langle \hat{O}_i \rangle_{\Psi_G} = \langle \Phi_0 | \hat{P}_i^\dagger \hat{O}_i \hat{P}_i | \Phi_0 \rangle \,, \tag{3.29}$$

$$\langle \hat{O}_{i,j} \rangle_{\Psi_G} = \langle \Phi_0 | \hat{P}_j^\dagger \hat{P}_i^\dagger \hat{O}_{i,j} \hat{P}_i \hat{P}_j | \Phi_0 \rangle \,. \tag{3.30}$$

Using these results, we are able to derive the energy functional in the Gutzwiller approximation. We assume translational invariance from now on. The site index "i" is suppressed, whenever this does not create ambiguities. A local basis can be chosen such that the local density matrix is diagonal

$$C_{\sigma,\sigma'} = \langle \hat{c}_{i,\sigma}^\dagger \hat{c}_{i,\sigma'} \rangle_{\Phi_0} = \delta_{\sigma,\sigma'} n_\sigma \,. \tag{3.31}$$

The constraints (3.27) and (3.28) are expressed in terms of the variational parameters using (3.23),

$$\sum_{\Gamma,\Gamma_1,\Gamma_2} \lambda_{\Gamma,\Gamma_1}^* \lambda_{\Gamma,\Gamma_2} m_{\Gamma_1,\Gamma_2}^0 = 1 \,, \tag{3.32}$$

$$\sum_{\Gamma,\Gamma_1,\Gamma_2} \lambda_{\Gamma,\Gamma_1}^* \lambda_{\Gamma,\Gamma_2} m_{\Gamma_1 \cup \sigma, \Gamma_2 \cup \sigma'}^0 = \delta_{\sigma,\sigma'} n_\sigma \tag{3.33}$$

with $|\Gamma \cup \sigma\rangle := \hat{c}_\sigma^\dagger |\Gamma\rangle$ and $m_{\Gamma,\Gamma'}^0 = \langle \hat{m}_{\Gamma,\Gamma'} \rangle_{\Phi_0}$. Each local operator $\hat{O}_i$, e.g., the local Hamiltonian (3.21), can be written as

$$\hat{O}_i = \sum_{\Gamma,\Gamma'} O_{\Gamma,\Gamma'} \hat{m}_{i;\Gamma,\Gamma'} \,. \tag{3.34}$$

Using (3.23) and (3.29) its expectation value with respect to the Gutzwiller wave function is then given by

$$\langle \hat{O} \rangle_{\Psi_G} = \sum_{\Gamma_1,\Gamma_2,\Gamma_3,\Gamma_4} O_{\Gamma_2,\Gamma_3} \lambda_{\Gamma_2,\Gamma_1}^* \lambda_{\Gamma_3,\Gamma_4} m_{\Gamma_1,\Gamma_4}^0 \,. \tag{3.35}$$

For the local Hamiltonian (3.24) we therefore obtain

$$\langle \hat{H}_{i,\mathrm{loc}} \rangle_{\Psi_\mathrm{G}} = \sum_{\Gamma,\Gamma_1,\Gamma_2} E_\Gamma \lambda^*_{\Gamma,\Gamma_1} \lambda_{\Gamma,\Gamma_2} m^0_{\Gamma_1,\Gamma_2} \ . \tag{3.36}$$

Expectations values for the kinetic term are evaluated using (3.29)

$$\langle \hat{c}^\dagger_{i,\sigma_1} \hat{c}_{j,\sigma_2} \rangle_{\Psi_G} = \left\langle \left( \hat{P}^\dagger_i \hat{c}^\dagger_{i,\sigma_1} \hat{P}_i \right) \left( \hat{P}^\dagger_j \hat{c}^\dagger_{j,\sigma_1} \hat{P}_j \right) \right\rangle_{\Phi_0} \ . \tag{3.37}$$

In infinite dimensions the expectations values take the form

$$\langle \hat{c}^\dagger_{i,\sigma_1} \hat{c}_{j,\sigma_2} \rangle_{\Psi_G} = \sum_{\sigma_1',\sigma_2'} q^{\sigma_1'}_{\sigma_1} \left( q^{\sigma_2'}_{\sigma_2} \right)^* \langle \hat{c}^\dagger_{i,\sigma_1} \hat{c}_{j,\sigma_2} \rangle_{\Psi_G} \tag{3.38}$$

with the "hopping renormalization" factors

$$q^{\sigma'}_{\sigma} = \frac{1}{n_{\sigma'}} \sum_{\Gamma_1,\ldots,\Gamma_4} \lambda^*_{\Gamma_2,\Gamma_1} \lambda_{\Gamma_3,\Gamma_4} \langle \Gamma_2 | \hat{c}^\dagger_\sigma | \Gamma_3 \rangle \left\langle \left( |\Gamma_1\rangle\langle\Gamma_4| \hat{c}_{\sigma'} \right) \right\rangle_{\Phi_0} \ . \tag{3.39}$$

With expectation values for the local and the kinetic part of the Hamiltonian (3.20) at hand we retrieve the following variational energy functional (per lattice site)

$$E_\mathrm{G}\left( \lambda_{\Gamma,\Gamma'}, |\Phi_0\rangle \right) = \sum_{\substack{\sigma_1,\sigma_2 \\ \sigma_1',\sigma_2'}} q^{\sigma_1'}_{\sigma_1} \left( q^{\sigma_2'}_{\sigma_2} \right)^* E_{\sigma_1,\sigma_2,\sigma_1',\sigma_2'} + \sum_{\Gamma,\Gamma_1,\Gamma_2} E_\Gamma \lambda^*_{\Gamma,\Gamma_1} \lambda_{\Gamma,\Gamma_2} m^0_{\Gamma_1,\Gamma_2} \tag{3.40}$$

with the energy tensor

$$E_{\sigma_1,\sigma_2,\sigma_1',\sigma_2'} = \frac{1}{L} \sum_{i \neq j} t^{\sigma_1,\sigma_2}_{i,j} \left\langle \hat{c}^\dagger_{i,\sigma_1'} \hat{c}_{j,\sigma_2'} \right\rangle_{\Phi_0} \tag{3.41}$$

As usual, $L$ denotes the number of sites in the crystal. Equation (3.40) has to be minimized, fulfilling the constraints (3.27), (3.28).

## 3.4 Minimizing the Gutzwiller energy functional

The direct minimization of (3.40) poses a difficult (numerical) problem. For a local Hamiltonian with $\ell$-like orbitals, the dimension of the Hilbert space of the atomic problem is $2^{2(2l+1)}$. Even if we only take into account diagonal entries $\lambda_{\Gamma,\Gamma}$ of the full parameter matrix $\lambda_{\Gamma,\Gamma'}$ we end up with 1024 parameters for $d$-type ($\ell = 2$) orbitals. The constraints (3.27), (3.28) introduce dependencies between the variational parameters that complicate the numerical procedure.

The minimization of the energy functional is achieved by subsequent inner and outer minimization cycles. With an "inner minimization" we refer to the optimization with respect to the variational parameters, with "outer minimization" to the optimization with respect to the single-particle state $|\Phi_0\rangle$. In this section we outline the general strategy for the numerical scheme used in this work, which is based on [62, 63].

### 3.4.1 Inner minimization

The single-particle state enters the energy functional (3.40) through its density matrix

$$\rho_{(i\sigma),(j,\sigma')} = \langle \hat{c}^{\dagger}_{j,\sigma'} \hat{c}_{i,\sigma} \rangle_{\Phi_0} \quad \text{with} \quad \boldsymbol{\rho} = \boldsymbol{\rho}^2 \ . \tag{3.42}$$

Since we perform an inner minimization we only minimize with respect to the variational parameters. For a fixed density matrix $\boldsymbol{\rho}$, the energy functional can be cast into the form

$$E_{\mathrm{G}}(\mathbf{v}) = \sum_{\sigma_1,\sigma_2,\sigma_1',\sigma_2'} q^{\sigma_1'}_{\sigma_1}(\mathbf{v}) q^{\sigma_2'}_{\sigma_2}(\mathbf{v}) E_{\sigma_1,\sigma_2,\sigma_1',\sigma_2'} + \sum_{Z,Z'} U(Z,Z') v_Z v_{Z'} \ , \tag{3.43}$$

where we used the abbreviation $v_Z$ for the variational parameters

$$v_Z = \frac{\lambda_{\Gamma,\Gamma'}}{\sqrt{m^0_{\Gamma,\Gamma} m^0_{\Gamma',\Gamma'}}} \ , \tag{3.44}$$

which are the elements of a vector $\mathbf{v}$. Since the $q$-factors depend quadratically on the variational parameters we can write them as

$$q^{\sigma'}_{\sigma}(\mathbf{v}) = \sum_{Z,Z'} S^{\sigma'}_{\sigma}(Z,Z') v_Z v_{Z'} \ . \tag{3.45}$$

The same holds for the $n_{\mathrm{c}}$ constraints

$$g_l(\mathbf{v}) = \sum_{Z,Z'} f_l(Z,Z') v_Z v_{Z'} - g^0_l = 0 \qquad (l = 1, \ldots, n_{\mathrm{c}}) \ . \tag{3.46}$$

For each inner minimization cycle, the coefficients $C_{Z,Z'} = \left\{ S^{\sigma'}_{\sigma}(Z,Z'), f_l(Z,Z'), U(Z,Z') \right\}$ only have to be calculated once. Differentiating the energy functional (3.43) and the constraints (3.46) with respect to the variational parameters $v_Z$ is done analytically. The analytic expressions are readily evaluated numerically and used in a constrained gradient descent scheme. For the detailed technical description of this scheme, we refer to [62].

### 3.4.2 Outer minimization

The inner minimization yields an optimum set of variational parameters $\mathbf{v}^{\min}$. For the outer minimization we have to optimize the functional

$$E_{\mathrm{G}}(\boldsymbol{\rho}) = \sum_{i \neq j} \sum_{\sigma,\sigma'} \bar{t}^{\sigma,\sigma'}_{i,j}(\boldsymbol{\rho}) \rho_{(j\sigma'),(i\sigma)} + L \sum_{Z,Z'} U(Z,Z')(\boldsymbol{\rho}) v^{\min}_Z v^{\min}_{Z'}$$

with respect to $\boldsymbol{\rho}$. Here we introduced the renormalised hopping parameters

$$\bar{t}^{\sigma_1,\sigma_2}_{i,j}(\boldsymbol{\rho}) = \sum_{\sigma_1',\sigma_2'} q^{\sigma_1}_{\sigma_1'}(\boldsymbol{\rho}) q^{\sigma_2}_{\sigma_2'}(\boldsymbol{\rho}) t^{\sigma_1',\sigma_2'}_{i,j} \tag{3.47}$$

with

$$q^{\sigma'}_{\sigma}(\boldsymbol{\rho}) = \sum_{Z,Z'} S^{\sigma'}_{\sigma}(Z,Z';\boldsymbol{\rho}) v^{\min}_Z v^{\min}_{Z'} \ . \tag{3.48}$$

In addition, the (independent) constraints

$$g_l(\boldsymbol{\rho}) = \sum_{Z,Z'} f_l(Z,Z',\boldsymbol{\rho}) v_Z^{\min} v_{Z'}^{\min} - g_l^0 = 0$$

need to be obeyed. In the following, we describe two strategies to perform the outer minimization. We denote them as "Fixed local density matrix" and "Unrestricted minimization".

#### Fixed local density matrix

Only the local density matrix

$$C_{\sigma,\sigma'} = \rho_{(i,\sigma'),(i,\sigma)} \tag{3.49}$$

enters the coefficients $S_\sigma^{\sigma'}(Z,Z')$, $f_l(Z,Z')$ and $U(Z,Z')$. When we keep the local density matrix fixed we have to minimize the functional

$$E_{\mathrm{G},0}(\boldsymbol{\rho}) \equiv \sum_{i \neq j} \sum_{\sigma,\sigma'} \bar{t}_{i,j}^{\sigma,\sigma'} \rho_{(j\sigma'),(i\sigma)} \ . \tag{3.50}$$

We obey the constraints using the Lagrange functional

$$L_{\mathrm{G}} := E_{\mathrm{G},0}(\boldsymbol{\rho}) - \sum_{\sigma,\sigma'} \eta_{\sigma,\sigma'} \sum_i (C_{\sigma,\sigma'} - \rho_{(i\sigma'),(i\sigma)}) - \sum_{i,j} \sum_{\sigma,\sigma'} \Omega_{(i\sigma),(j\sigma')} [\boldsymbol{\rho}^2 - \boldsymbol{\rho}]_{(j\sigma'),(i\sigma)} \ , \tag{3.51}$$

where $\eta_{\sigma,\sigma'}$ are Lagrange parameters for fixing the local density matrix, and $\Omega_{(i\sigma),(j\sigma')}$ are Lagrange parameters for ensuring that the density matrix corresponds to the density matrix of a single-particle wave function. Minimizing the Lagrange functional is equal to finding the ground state of the effective single-particle Hamiltonian

$$\hat{H}_0^{\mathrm{eff}} = \sum_{i,j} \sum_{\sigma,\sigma'} ((1-\delta_{i,j}) \bar{t}_{i,j}^{\sigma,\sigma'} + \delta_{i,j} \eta_{\sigma,\sigma'}) \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma'} \ , \tag{3.52}$$

that then yields the optimal density matrix.

Up to this point we developed a working strategy to minimize the Gutzwiller energy functional: By a single inner and a single outer minimization step we find the optimal density matrix $\boldsymbol{\rho}$ and the optimal variational parameters $\mathbf{v}$ for a given (fixed) local density matrix $\boldsymbol{C}$. All is left is to "scan" for the optimal local density matrix. We can simplify this approach by directly choosing a fixed set of effective on-site elements $\eta_{\sigma,\sigma'}$, our Lagrange parameters in (3.51). Next we apply several inner and outer minimization steps until convergence with respect to the local density matrix is achieved. We then found a density matrix $\boldsymbol{\rho}$ that is optimal for a set of effective on-site energies $\eta_{\sigma,\sigma'}$ that indirectly specify the local density matrix. If we take symmetries into account, the effective on-site energies can often be reduced to a very small set of independent parameters so that a simple scan of the remaining parameters becomes numerically feasible.

#### Unrestricted minimization

The optimal set of effective on-site energies in the scheme of subsequent inner and outer minimization steps described in the previous section can also be determined directly. We drop the requirement of a fixed local density matrix and explicitly fulfil the Gutzwiller constraints using the Lagrange parameters $\Lambda_l$, yielding the Lagrange functional for the unrestricted minimization

$$L_{\mathrm{G}} = E_{\mathrm{G}}(\boldsymbol{\rho}) - \sum_{l} \Lambda_l g_l(\boldsymbol{\rho}) - \sum_{i,j} \sum_{\sigma,\sigma'} \Omega_{(i\sigma),(j\sigma')} [\boldsymbol{\rho}^2 - \boldsymbol{\rho}]_{(j\sigma'),(i\sigma)} \ . \tag{3.53}$$

Again, the minimization yields an effective single-particle Hamiltonian (3.52) where the effective on-site energies are now given by

$$\eta_{\sigma,\sigma'} = \frac{\partial}{\partial C_{\sigma,\sigma'}} E_{\mathrm{G}}(\boldsymbol{\rho}) - \sum_{l} \Lambda_l \frac{\partial}{\partial C_{\sigma,\sigma'}} g_l(\boldsymbol{\rho}) \ . \tag{3.54}$$

Please note that even though the local density matrix is diagonal (which always can be archived by a unitary transformation) derivatives with respect to off-diagonal elements do not necessarily have to be zero. The remaining Lagrange parameter $\Lambda_l$ can be obtained by exploiting the fact that, for the variational ground state, the Lagrange functional (3.53) is also minimal with the respect of the parameters $v_Z$, yielding the equation

$$\frac{\partial}{\partial v_Z} E_{\mathrm{G}}(\boldsymbol{\rho}, v)\Big|_{v=v^{\min}} - \sum_{l} \Lambda_l \frac{\partial}{\partial v_Z} g_l(\boldsymbol{\rho}, v)\Big|_{v=v^{\min}} = 0. \tag{3.55}$$

From this equation we can determine $\Lambda_l$ and finally calculate the effective on-site energy $\eta_{\sigma,\sigma'}$.

The unrestricted minimization scheme is summarized by the following steps

(i) Start from the uncorrelated limit, i.e. set $q_\sigma^{\sigma'} = \delta_{\sigma,\sigma'}$ and $\eta_{\sigma,\sigma'} = t_{i,i}^{\sigma,\sigma'}$.

(ii) Determine the ground state $|\Phi_0\rangle$ of the Hamiltonian (3.52) and calculate $C_{\sigma,\sigma'}$.
If $C_{\sigma,\sigma'}$ is not diagonal, change to a basis where it is.

(iii) Perform an inner minimization.
Calculate a new set of hopping reduction factors $q_{\sigma'}^{\sigma}$.

(iv) Determine $\Lambda_l$ and calculate a new set of effective on-site energies $\eta_{\sigma,\sigma'}$.
Go to step (ii).

As usual, a simple linear mixing scheme for the $q$-factors and the $\eta$- parameters is applied. The algorithm terminates as soon as a fixed point is reached.

## 3.5 Landau-Gutzwiller quasi particles

It is shown in [64] that the eigenenergies of the effective single-particle Hamiltonian (3.52) can be interpreted as quasi-particle bands which are suitable for a comparison with real, three-dimensional Fermi liquids. We give a short summary of the publication in this section.

The Gutzwiller correlator acts on the Fermi-gas ground state $|\Phi_0\rangle$ and gradually reduces the energetically unfavourable configurations. This gradual reduction is in the spirit of the Landau Fermi-liquid theory which assumes that the Fermi gas picture remains qualitatively valid when electron-electron interactions are switched on. The eigenstates of the uncorrelated Fermi-gas adiabatically transform into those of the Fermi liquid.

To obtain the eigenstates of (3.52) we transform the Hamiltonian to momentum space,

$$H_0^{\text{eff}} = \sum_{\mathbf{k},\sigma,\sigma'} \left(\bar{\epsilon}_{\mathbf{k},\sigma,\sigma'} - \eta_{\sigma,\sigma'}\right) \hat{c}^\dagger_{\mathbf{k},\sigma} \hat{c}_{\mathbf{k},\sigma'} \tag{3.56}$$

with

$$\bar{\epsilon}_{\mathbf{k},\sigma,\sigma'} = \frac{1}{L} \sum_{i \neq j} \bar{t}^{\sigma\sigma'}_{i,j} e^{\mathrm{i}\mathbf{k}(\mathbf{R}_i - \mathbf{R}_j)} \quad \text{and} \quad \hat{c}^{(\dagger)}_{\mathbf{k}} = \frac{1}{\sqrt{L}} \sum_i e^{(\frac{+}{-})\mathrm{i}\mathbf{k}\mathbf{R}_i} \hat{c}^{(\dagger)}_{i,\sigma} \,, \tag{3.57}$$

where $L$ is the number of sites. Using another unitary transformation

$$\hat{h}^\dagger_{\mathbf{k},\gamma} = \sum_\sigma u^{(\mathbf{k})}_{\gamma,\sigma} \hat{c}^\dagger_{\mathbf{k},\sigma} \tag{3.58}$$

we eventually arrive at the diagonal form

$$H_0^{\text{eff}} = \sum_{\mathbf{k},\gamma} E_{\mathbf{k},\gamma} \hat{h}^\dagger_{\mathbf{k},\gamma} \hat{h}_{\mathbf{k},\gamma} \tag{3.59}$$

with the dispersion $E_{\mathbf{k},\gamma}$. The ground state of (3.59) is then assumed to be given by

$$|\Phi_0\rangle = \prod_{E_{\mathbf{k},\gamma} < E_\mathrm{F}} \hat{h}^\dagger_{\mathbf{k},\gamma} |\mathrm{vac}\rangle \,, \tag{3.60}$$

where the Fermi energy is indirectly determined by the number of particles in the system. We define the quasi-particle and quasi-hole creation operators

$$\hat{e}^\dagger_{\mathbf{k},\gamma} = \hat{P}_\mathrm{G} h^\dagger_{\mathbf{k},\gamma} (\hat{P}_\mathrm{G})^{-1} \,, \tag{3.61}$$

$$\hat{v}_{\mathbf{k},\gamma} = \hat{P}_\mathrm{G} h_{\mathbf{k},\gamma} (\hat{P}_\mathrm{G})^{-1} \,, \tag{3.62}$$

which fulfil Fermi anti-commutation relations. These operators create quasi-particles/quasi-holes in the Gutzwiller ground state

$$\langle \Psi_\mathrm{G} | \hat{e}^\dagger_{\mathbf{k},\gamma} \hat{v}_{\mathbf{k},\gamma} | \Psi_\mathrm{G} \rangle = \Theta(E_\mathrm{F} - E_{\mathbf{k},\gamma}) \,, \tag{3.63}$$

$$\langle \Psi_\mathrm{G} | \hat{v}^\dagger_{\mathbf{k},\gamma} \hat{e}_{\mathbf{k},\gamma} | \Psi_\mathrm{G} \rangle = \Theta(E_{\mathbf{k},\gamma} - E_\mathrm{F}) \,. \tag{3.64}$$

As seen from the expressions for the corresponding states,

$$|\Psi_{\mathrm{G},+}^{\mathbf{k},\gamma}\rangle = \hat{e}_{\mathbf{k},\gamma}^{\dagger}|\Psi_{\mathrm{G}}\rangle = P_{\mathrm{G}}\hat{h}_{\mathbf{k},\gamma}^{\dagger}|\Phi_0\rangle\,, \tag{3.65}$$

$$|\Psi_{\mathrm{G},-}^{\mathbf{k},\gamma}\rangle = \hat{v}_{\mathbf{k},\gamma}|\Psi_{\mathrm{G}}\rangle = P_{\mathrm{G}}\hat{h}_{\mathbf{k},\gamma}|\Phi_0\rangle\,, \tag{3.66}$$

the excitations of the correlated system are uniquely connected to the excitations of the uncorrelated Fermi gas $|\Phi_0\rangle$. The energies of the excitations are given by

$$E_{\mathbf{k},\gamma}^{\mathrm{qp}} = \pm\left(E_{\pm,\mathbf{k},\gamma}^{\mathrm{var}} - E_0^{\mathrm{var}}\right), \tag{3.67}$$

where $E_0^{\mathrm{var}}$ is the variational ground state energy, and

$$E_{\pm,\mathbf{k},\gamma}^{\mathrm{var}} = \frac{\langle\Psi_{\mathrm{G},\pm}^{\mathbf{k},\gamma}|H_{\mathrm{Hub}}|\Psi_{\mathrm{G},\pm}^{\mathbf{k},\gamma}\rangle}{\langle\Psi_{\mathrm{G},\pm}^{\mathbf{k},\gamma}|\Psi_{\mathrm{G},\pm}^{\mathbf{k},\gamma}\rangle} \tag{3.68}$$

is the expectation value of the energy with respect to the excited states. It can be shown [59] that the quasi-particle energies are directly connected to the dispersion of the effective Hamiltonian $H_0^{\mathrm{eff}}$ in (3.59)

$$E_{\mathbf{k},\gamma}^{\mathrm{qp}} = E_{\mathbf{k},\gamma} - E_{\mathrm{F}}\,. \tag{3.69}$$

Hence, the eigenenergies $E_{\mathbf{k},\gamma}$ of $H_0^{\mathrm{eff}}$ have direct physical meaning and can be compared to band structures obtained from Angle-Resolved Photoemission Spectroscopy (ARPES) experiments.

The effective on-site energies $\eta$ entering the Hamiltonian (3.52) change the band-positions, the $q$-factors renormalize the band width. The square of the renormalization factor is related to the effective mass $m^*$ of the quasi-particle, $m^* \propto 1/q^2$. A renormalization of the hopping reduction factor to zero can therefore be interpreted as a metal-to-insulator transition. Such a transition, the so-called "Brinkmann-Rice" transition [65], is shown to occur for the one-band model for a finite value of the interaction $U$. It was shown that the occurrence of the Brinkmann-Rice transition in the one-band Hubbard model is an artefact of the Gutzwiller approximation [66]. Results with very small renormalization factors or the observation of metal-insulator transitions must therefore be taken with care in the variational method described here [59]. We shall not encounter vanishingly small $q$-factors in this work.

# 4 Gutzwiller DFT for diluted magnetic semiconductors

The study of Diluted Magnetic Semiconductors (DMS) requires to solve two main problems: (i) an adequate treatment of electron correlations, due to the presence of a strong Coulomb interaction at the impurity sites; (ii) an adequate modelling of the low doping concentrations that break translational and rotational symmetries of the host semiconductor. Problem (i) is tackled by employing the Gutzwiller variational method, problem (ii) by using large unit cells, so called "supercells".

In this chapter we introduce the assumptions and approximations that we make in order to investigate numerically the magnetic properties of DMS materials. We do not only aim for a qualitative model description, but are interested in quantitative results. For this purpose we utilize accurate DFT calculations as input for our multiband Hubbard model Hamiltonian. We study different host-dopant combinations which allow us to draw conclusion for the general coupling mechanism between magnetic ions in DMS materials.

## 4.1 Crystal structure

The two most important crystal structures for semiconductors are the hexagonal wurtzite structure and the cubic zincblende structure ($\beta$-ZnS). In this work we will solely focus on the more common zincblende structure. We investigate the magnetic properties of several II-VI semiconductors doped with manganese, as well as GaAs doped with manganese. The fcc sub-lattice of the cations is shifted against the fcc sub-lattice of the anions by $a/4$ along the diagonal of the cubic unit cell, see Fig. 4.1. Values for the experimentally determined lattice constant $a$ for the various semiconductors are collected in Table 4.1.

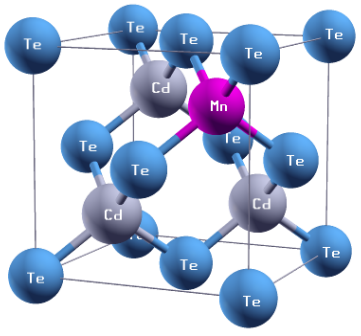


**Figure 4.1:** Unit cell of the prototype II-VI semiconductor $Cd_3MnTe_4$. One of the four Cd-atoms within the unit cell is replaced by a magnetic impurtiy, leading to an impurity concentration of $x = 25\%$. This configuration conserves the cubic symmetry at the Mn-site.

We assume that the magnetic impurity substitutes the semiconductor cation without inducing any deformation of the original lattice. We restrain our investigations to materials without impurity-induced defects (occupations by magnetic impurities at interstitial

| ABC | CdMnTe | ZnMnTe | ZnMnSe | ZnMnS | GaMnAs | CdCrTe |
|---|---|---|---|---|---|---|
| $A_3B_1C_4$ | 12.39 | 11.76 | 10.88 | 10.37 | 10.83 | 12.37 |
| $A_4B_0C_4$ | 12.52 | 11.70 | 10.81 | 10.30 | 10.86 | 12.52 |
| $A_4B_0C_4$ (exp) | 12.25 | 11.53 | 10.71 | 10.22 | 10.68 | 12.25 |

**Table 4.1:** Lattice constants determined by a non-relativistic DFT calculation in the FLAPW formulation using the PBE functional. Experimental values are taken from [67]. All values are given in atomic units.

sites, vacancies, and so on). Moreover, we assume a random distribution of the magnetic ions without a bias towards clustering, i.e., whether a site is occupied by an impurity or a semiconductor cation is completely random. The validity of this assumption is supported by the experimental analysis of crystals by x-ray diffraction for the cases where the dopant iso-electronically replaces the cation, for example, in a rather crude chemical picture, the Mn $4s^2$ valence electrons overtake the function of the $5s^2$ valence electrons of Cd in the bonding in $Cd_{1-x}Mn_xTe$. However, for the case of manganese doped GaAs it is known from experiment that the crystals can have defects which largely influence their physical properties. For example, the Curie temperature of $Ga_{1-x}Mn_xAs$ can be influenced by a post-growth annealing process [68] which is believed to cure certain lattice defects. The study of effects caused by lattice defects is not within the scope of this thesis.

We calculate the lattice constant of the undoped semiconductor and of a prototype material with eight atoms per unit cell where we replaced one atom by a magnetic impurity, corresponding to a dopant concentration of $x = 25\%$ as exemplary shown for $Cd_{1-x}Mn_xTe$ in Fig. 4.1. The calculation is performed using a non-relativistic spin-polarized DFT calculation where we used the GGA with the PBE functional. The lattice constant is determined from the minimum of a quadratic fit to the ground-state energy, see Fig. 4.2. Results for various materials are listed in Table 4.1. The density functional theory systematically overestimates the lattice parameters by about 2%. The lattice parameter of the unit cell containing the impurity differs only about 1% from the non-doped material. Therefore, the assumption that the lattice distortion around a single impurity can be neglected in our studies seems reasonable. For all further calculations we use the experimental lattice constant of the host semiconductor.

## 4.2 Supercells

Performing a DFT calculation for a zincblende semiconductor material is computationally inexpensive. A fcc lattice with a two-atomic basis is used. The evaluation of integrals within the fcc BZ on a discrete set of **k**-points ("**k**-mesh") is vastly simplified by exploiting the 24 point-group symmetries present, cutting down the computational effort by about the same factor. Additionally, the **k**-mesh for non-metallic systems can be kept coarse, since it is not necessary to sample accurately the two-dimensional energy surface at the Fermi level. By exploiting all symmetry properties, a typical DFT calculation for such a system converges on today's computing machines within seconds, even when the expensive FLAPW formalism is used.

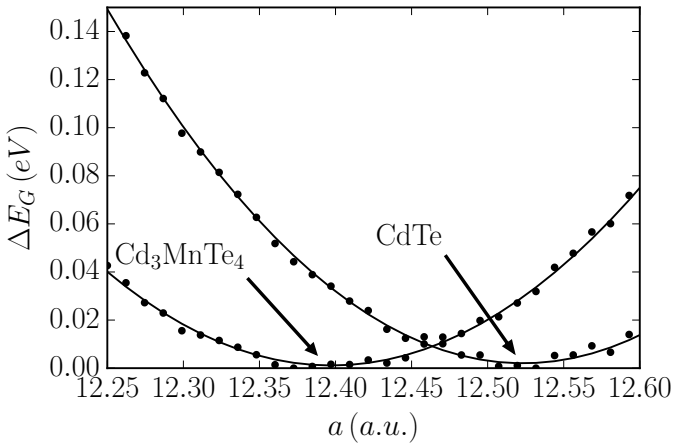


**Figure 4.2:** Normalized ground-state energy for CdTe and $Cd_3MnTe_4$, as a function of the lattice constant $a$, determined by a non-relativistic DFT calculation in the FLAPW formulation using the PBE functional (points) together with a quadratic fit.

To be able to describe the physics of doped crystals adequately we employ the supercell method. A large unit cell of several atoms with a certain concentration of magnetic impurities mimics the random nature of the impurity distribution over the whole crystal. The material then still possesses translational symmetry which allows us to employ band structure methods for infinite large systems. In comparison to the two-atomic calculation these supercell calculations are computationally very demanding. The unit cell of $Cd_3MnTe_4$, see Fig. 4.1, is a realization of such a supercell but obviously too small to study highly diluted systems. The Mn atoms form a cubic sub-lattice in the crystal, so that all symmetries are conserved. Even though there are only three distinct atoms (atoms, that cannot be mapped onto each other by symmetry operations of the lattice), the computational time goes up from seconds to minutes. The time complexity of FLAPW-methods is typically given by $O(N_A^3)$ (ultimately determined by matrix diagonalization), where $N_A$ is the number of atoms in the unit cell [23]. In some of our calculations we study isolated impurities in a host semiconductor. In this context "isolated" means that the supercell containing these atoms is large enough so that the ions are not influenced by their periodic counterparts in neighboring cells. To meet this requirements, even for the simple case of two isolated ions in a II-VI semiconductor, unit cells with sizes up to 128 atoms are necessary. All DFT results in this work for unit cells with more than 16 atoms were performed on high-performance computing systems [69].

### 4.2.1 Linear interpolation scheme

In Section 2.6 we introduced a method to obtain a tight-binding model derived from DFT calculations based on Wannier orbitals. The combination of density functional theory and the Gutzwiller method appears straightforward. (i) a DFT supercell calculation is downfolded using Wannier orbitals; (ii) the resulting tight-binding model is extended to include local interaction terms for the impurity sites; (iii) the final Hubbard model is investigated using the Gutzwiller variational method. However, this approach fails due

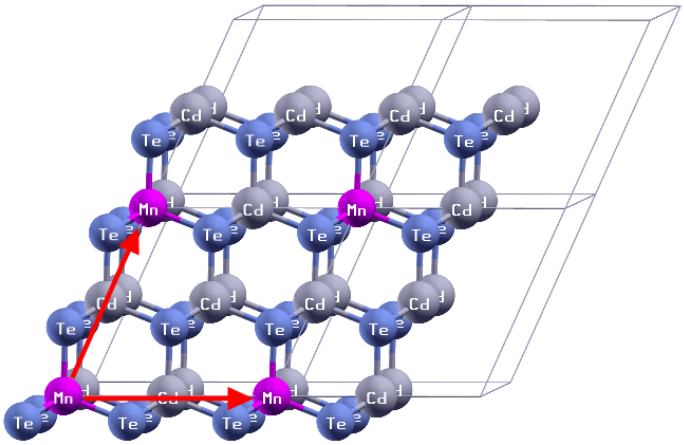


**Figure 4.3:** Crystal structure (exemplary for the case of Mn doped CdTe) based on the 16 atom unit cell with one impurity continued $2 \times 2 \times 1$ times. Nearest-neighbor distances between two impurities are indicated by red arrows.

to the complexity of downfolding the DFT band structures of large supercells. As an instructive example we consider a DFT calculation with 128 atoms/supercell. A unit cell of this size requires a downfolded basis of 512 Wannier orbitals (four orbitals per atom to describe the valence and conduction bands, c.f. Section 2.6.6). The number of matrix elements which needs to be computed just within this unit cell is $512^2$. We also need to compute hopping matrix elements into neighboring cells, so that an efficient algorithm is required which implements strategies for neglecting small and storing large matrix elements. We also find that numerical inaccuracies build up in the downfolding procedure for large unit cells. This renders the downfolding procedure impracticable for supercells with much more than 16 atoms.

In order to cope with these problems we use a different approach. The main idea is that an impurity placed in the host semiconductor only has an influence on neighboring hopping matrix elements. Matrix elements between orbitals that are "further away" stay unaffected by the impurity. We exploit the fact that Wannier orbitals are defined in real space to "extract" the influence of the impurity on its surrounding. This is done by performing the following steps.

(i) Carry out a non-spin polarized DFT calculation with a 16-atom unit cell for the host semiconductor and a subsequent downfolding of the DFT band structure onto a tight-binding model with a basis of $16 \times 4 = 64$ Wannier orbitals.

(ii) Carry out a non-spin polarized DFT calculation with a 16-atom unit cell for the host semiconductor in which one cation is replaced by a magnetic impurity and a subsequent downfolding of the DFT band structure onto a tight-binding model with a basis of $16 \times 4 + 5 = 69$ Wannier orbitals. Note that there are five additional $d$-type orbitals from the impurity.

Procedure (i) yields a set of orbitals and matrix elements describing valence and conduction bands around the Fermi energy. These bands have $s$-type and $p$-type character. In procedure (ii) we obtain additional matrix elements due to the impurity which includes

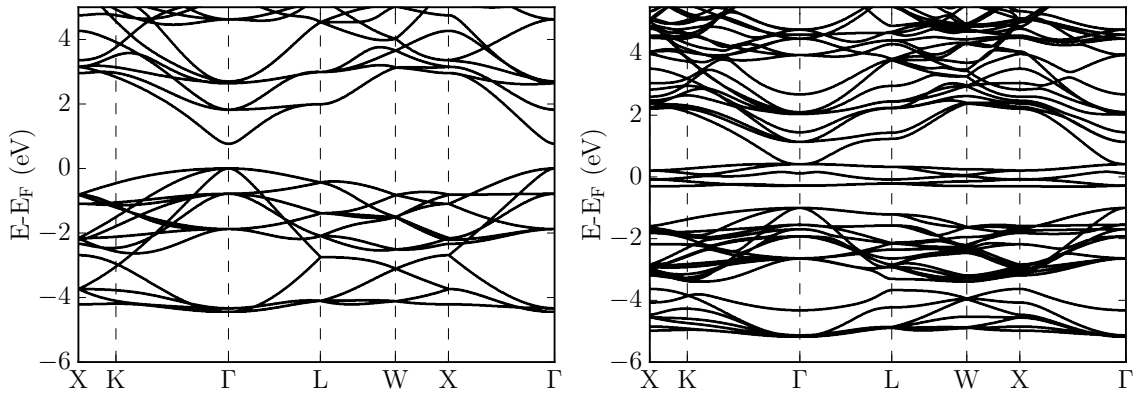


**Figure 4.4:** Band structures of $Cd_8Te_8$ (left) and $Cd_7MnTe_8$ (right) for the 16-atom unit cell shown in Fig. 4.3 obtained from paramagnetic DFT calculations utilizing the PBE functional.

matrix elements from or onto $d$-orbitals. In this subsection we use the indices $\sigma_{1,2}$ for $s$- and $p$-type orbitals, and $\tilde{\sigma}_{1,2}$ for $d$-type orbitals.

With this notation the tight-binding model obtained from step (i) reads

$$\hat{H}^{\mathrm{AB}} = \sum_{i,j,\sigma_1,\sigma_2} t_{i,j}^{\sigma_1,\sigma_2} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{j,\sigma_2} \ . \tag{4.1}$$

We remind the reader that $i$ and $j$ indicates atoms over the whole crystal, they are not restricted to the 16 atoms of one unit cell. Because of the translational symmetry of the crystal, all hopping values occur infinitely often. The set of distinct hopping values obtained from the downfolding procedure is only finite because a cut-off value is defined, neglecting matrix elements of distant orbitals, as explained in Section 2.6.4. The total number of distinct matrix elements obtained from procedure (i) and (ii) are of the order of one million.

Similarly, we obtain a tight-binding Hamiltonian from procedure (ii),

$$\begin{aligned} \hat{H}^{\mathrm{AMB}} = &\sum_{i,j,\sigma_1,\sigma_2} \bar{t}_{i,j}^{\sigma_1,\sigma_2} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{j,\sigma_2} + \sum_{i,j,\tilde{\sigma}_1,\sigma_2} \left\{ \bar{t}_{i,j}^{\tilde{\sigma}_1,\sigma_2} \hat{c}_{i,\tilde{\sigma_1}}^{\dagger} \hat{c}_{j,\sigma_2} + \mathrm{h.\,c.} \right\} \\ &+ \sum_{i,j,\tilde{\sigma}_1,\tilde{\sigma}_2} \bar{t}_{i,j}^{\tilde{\sigma}_1,\tilde{\sigma}_2} \hat{c}_{i,\tilde{\sigma}_1}^{\dagger} \hat{c}_{j,\tilde{\sigma}_2} \ . \end{aligned} \tag{4.2}$$

Again, a lot of complexity is hidden within the notation. To give an illustrative example we pick an atom right next to a magnetic impurity. The orbitals of this atom/site not only have a finite overlap with the orbitals of the closest impurity, but also with all periodically reoccurring impurity orbitals from all over the crystal that are within the range defined by the cut-off value $R_{\mathrm{max}}$.

We assume that the values $t_{i,j}^{\sigma_1,\sigma_2}$ obtained from procedure (i) correspond to values $\bar{t}_{i,j}^{\sigma_1,\sigma_2}$ obtained from procedure (ii), i.e., they are both matrix elements between a specific orbital located at site $a$ and a specific orbital located at site $b$, regardless of the type of the crystal

(pure or doped). Here, an additional difficulty arises. In principle, the Wannier orbitals obtained from the downfolding-procedure are only known up to a phase-factor. This phase-factors can be chosen in such a way that all hopping matrix elements become real [44]. In order to get meaningful results in the linear position scheme described in the following, we only have to ensure that there are no sign changes between the definition of the orbitals from procedure (i) and (ii).

Using this notation we are finally able to express the influence of *one* impurity on the matrix elements. We pick a lattice site $f$ occupied by an impurity and calculate all differences in matrix elements in the proximity given by the cut-off values $C_1$ and $C_2$.

$$\hat{H}_f^M = \sum_{\substack{i,j \\ \min(d_i^f, d_j^f) < C_1 \\ \max(d_i^f, d_j^f) < C_2}} \left\{ \sum_{\sigma_1,\sigma_2} \left( \bar{t}_{i,j}^{\sigma_1,\sigma_2} - t_{i,j}^{\sigma_1,\sigma_2} \right) \hat{c}_{i,\sigma_1}^\dagger \hat{c}_{j,\sigma_2} + \sum_{\tilde{\sigma}_1,\sigma_2} \left\{ \bar{t}_{i,j}^{\tilde{\sigma}_1,\sigma_2} \hat{c}_{i,\tilde{\sigma}_1}^\dagger \hat{c}_{j,\sigma_2} + \text{h. c.} \right\} + \sum_{\tilde{\sigma}_1,\tilde{\sigma}_2} \bar{t}_{i,j}^{\tilde{\sigma}_1,\tilde{\sigma}_2} \hat{c}_{i,\tilde{\sigma}_1}^\dagger \hat{c}_{j,\tilde{\sigma}_2} \right\}, \tag{4.3}$$

where we used the distance to the impurity at site $f$

$$d_i^f = |\mathbf{R}_i - \mathbf{R}_f|_2 \ . \tag{4.4}$$

Here $\mathbf{R}_i$ denote the atomic positions. We always have $C_2 > C_1$ for the choice of the cut-off parameters. In our model we set $C_1$ to include nearest neighbors, $C_2$ to include next-nearest neighbors of the impurity.

We can simply shift $\hat{H}_f^M$ to obtain the Hamiltonian for an impurity at a different site $f'$. With this at hand we can finally use a linear interpolation scheme to obtain a tight-binding model of arbitrarily placed impurities within the host semiconductor

$$\hat{H}^{\text{TB}} = \hat{H}^{AB} + \sum_{f \in \text{occ.}} \hat{H}_f^M, \tag{4.5}$$

where $f$ runs over sites occupied by magnetic ions. Note that in this model a hopping matrix element between two semiconductor orbitals in the proximity of two impurities "feels" the combined influence of both impurities, due to the linear superposition scheme. Of course, our model works best if the impurities are not too close together. Fortunately, the magnetic ions can only occupy cation sites, so that they are never nearest neighbors on the zincblende lattice.

Using this approach, we are able to construct large supercells with up to about 1000 atoms with arbitrary impurity placements on the fcc cation sub-lattice. We use sparse matrices to efficiently store non-vanishing matrix elements. The huge advantage of this method is that the computationally expensive DFT calculations only have to be performed once.

In order to test our method we compare an actual calculation using DFT for $Cd_3Mn_1Te_4$, see Fig. 4.1, and $Cd_2Mn_2Te_4$ with a result from our linear interpolation scheme. Note that in the latter case we have a quasi one-dimensional nearest neighbor chain of Mn ions on the fcc cation-sublattice of the crystal. A comparison of the band structures obtained is shown

in Fig. 4.5. Even for the case of impurities at nearest neighbors sites on the cation lattice, the band structures of our method match the band structures from actual calculations remarkably well. Thus, we are confident that our method is suitable to approximate accurately the DFT band structure for very large supercells.

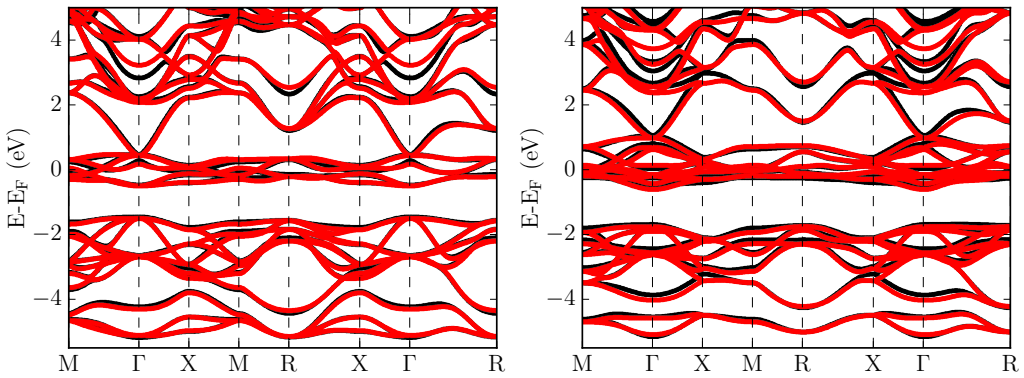


**Figure 4.5:** Band structures for $Cd_3Mn_1Te_4$ (left) (compare Fig. 4.1) and $Cd_2Mn_2Te_4$ (right) obtained from a paramagnetic DFT calculation (black lines) and our linear superposition scheme (red lines).

## 4.3 Atomic problem

Using the method introduced in the last section we are now able to derive a tight-binding model for DMS. Our final Hamiltonian is obtained by adding local interaction terms for all $d$-shells occupied by an impurity, and subtract a double-counting term, similar to the LDA+U method in Section 2.5. The final multiband Hubbard Hamiltonian reads

$$\hat{H} = \hat{H}^{\mathrm{TB}} + \sum_{f \in \mathrm{occ.}} \left[ \hat{H}_f^{\mathrm{AT}} - \hat{H}_f^{\mathrm{DC}} \right], \tag{4.6}$$

where $\hat{H}^{\mathrm{TB}}$ is given in (4.5) and the sum runs again over all sites occupied by impurities. In this section we will focus on the explicit form of $\hat{H}_f^{\mathrm{AT}}$; $\hat{H}_f^{\mathrm{DC}}$ is discussed in Section 4.5.

Since the interaction Hamiltonian is the same for every correlated site, we drop the site index in the following. We use results from crystal field theory (in chemistry: ligand field theory) to employ a more sophisticated description of the local $d$-type orbitals as done in the LDA+U method. In cubic symmetry the five $d$-orbitals split up in $e_g$ ($\varphi_u$, $\varphi_v$) and $t_{2g}$ ($\varphi_\xi$, $\varphi_\eta$, $\varphi_\zeta$) orbitals as visualized in Fig. 4.6.

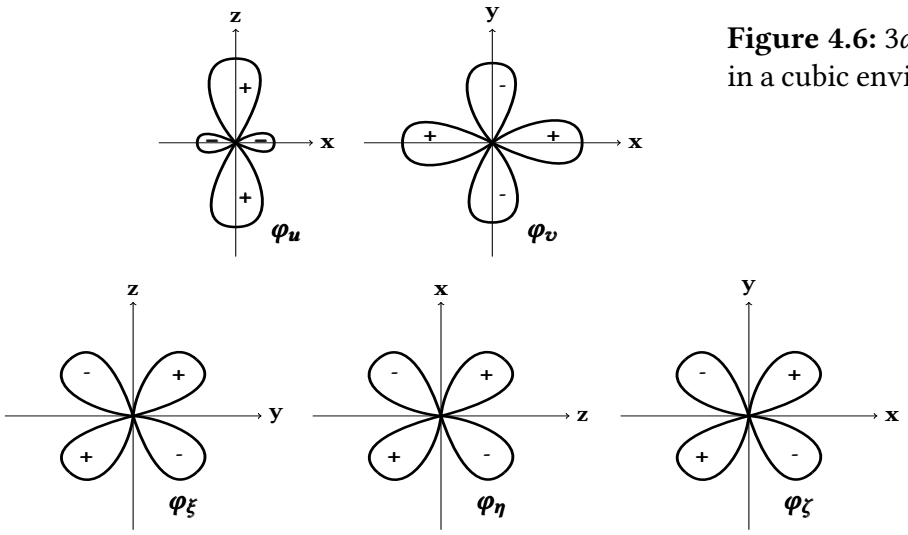


**Figure 4.6:** 3*d*-orbitals in a cubic environment.

The wave functions are defined as:

$$\varphi_\xi(r,\phi,\Theta) = \frac{\mathrm{i}}{\sqrt{2}} R_{t2_g}(r)\left[Y_{2,1}(\phi,\Theta) + Y_{2,-1}(\phi,\Theta)\right] \propto yz\,, \tag{4.7a}$$

$$\varphi_\eta(r,\phi,\Theta) = -\frac{1}{\sqrt{2}} R_{t2_g}(r)\left[Y_{2,1}(\phi,\Theta) - Y_{2,-1}(\phi,\Theta)\right] \propto xz\,, \tag{4.7b}$$

$$\varphi_\zeta(r,\phi,\Theta) = -\frac{\mathrm{i}}{\sqrt{2}} R_{t2_g}(r)\left[Y_{2,2}(\phi,\Theta) - Y_{2,-2}(\phi,\Theta)\right] \propto xy\,, \tag{4.7c}$$

$$\varphi_u(r,\phi,\Theta) = R_{e_g}(r) Y_{2,0}(\phi,\Theta) \propto (3z^2 - r^2)\,, \tag{4.7d}$$

$$\varphi_v(r,\phi,\Theta) = \frac{1}{\sqrt{2}} R_{e_g}(r)\left[Y_{2,2}(\phi,\Theta) + Y_{2,-2}(\phi,\Theta)\right] \propto (x^2 - y^2)\,. \tag{4.7e}$$

We use the same definition to obtain the trial orbitals in the construction of the projected Wannier functions in (2.86). Since their symmetry is conserved (see Appendix A.2) it is meaningful to identify the Wannier orbitals from the tight-binding model with the orbitals we use to describe the local interaction term.

The full interaction Hamiltonian for *d*-orbitals is given by [70, 59]

$$\begin{aligned}
\hat{H}^{\mathrm{AT}} = &\sum_{c,s} U(c,c)\hat{n}_{c,s}\hat{n}_{c',\bar{s}} + \sum_{c(\neq)c'}\sum_{s,s'}\left(U(c,c') - \delta_{s,s'}J(c,c')\right)\hat{n}_{c,s}\hat{n}_{c',s'} \\
&+ \sum_{c(\neq)c'} J(c,c')\left(\hat{c}^\dagger_{c,\uparrow}\hat{c}^\dagger_{c,\downarrow}\hat{c}_{c',\downarrow}\hat{c}_{c',\uparrow} + \text{h. c.}\right) + \sum_{c(\neq)c';s} J(c,c')\hat{c}^\dagger_{c,s}\hat{c}^\dagger_{c',\bar{s}}\hat{c}_{c,\bar{s}}\hat{c}_{c',s} \\
&+ \Bigg[\sum_{t;s,s'}(T(t) - \delta_{s,s'}A(t))\hat{n}_{t,s}\hat{c}^\dagger_{u,s'}\hat{c}_{v,s'} \quad + \sum_{t,s} A(t)\left(\hat{c}^\dagger_{t,s}\hat{c}^\dagger_{t,\bar{s}}\hat{c}_{u,\bar{s}}\hat{c}_{v,s} + \hat{c}^\dagger_{t,s}\hat{c}^\dagger_{u,\bar{s}}\hat{c}_{t,\bar{s}}\hat{c}_{v,s}\right) \\
&\quad + \sum_{t(\neq)t'(\neq)t''}\sum_{e,s,s'} S(t,t';t'',e)\hat{c}^\dagger_{t,s}\hat{c}^\dagger_{t',s'}\hat{c}_{t'',s'}\hat{c}_{e,s} + \text{h. c.}\Bigg]\,.
\end{aligned} \tag{4.8}$$

As usual, the operators $\hat{c}^\dagger_{c,s}$ ($\hat{c}_{c,s}$) create (annihilate) an electron with spin $s$ in orbital $c$, where $c \in \{\xi, \eta, \zeta, u, v\}$. We use the indices $e$ and $t$ to indicate that sums run over $e_g$ and

$t_{2g}$ orbitals, respectively. In cubic symmetry we can express all matrix elements in (4.8) in terms of ten independent parameters. When we also apply the "spherical approximation" ($R_{t_{2g}}(r) = R_{e_g}(r)$), all of the matrix elements can be expressed in terms of the three Racah parameters $A$, $B$ and $C$, see Appendix A.1. We further use $C = 4B$, which is a reasonable assumption for transition metal ions [70].

We need to be aware of the fact that in our calculations the cubic symmetry at an impurity site can be broken due to the presence of other impurities, e.g., a nearest-neighbor pair located on the fcc sub-lattice of the cations. As seen in later calculations, it is a good approximation to always assume cubic symmetry at the impurity sites. This assumption also vastly simplifies our Gutzwiller calculations.

## 4.4 Hubbard U from constrained DFT

It is desirable to calculate the interaction parameters for the atomic (model) Hamiltonian (4.8) from first principles. This would allow us to perform calculations without further input. However, it turns out that this is a difficult task even when we are only interested in the averaged values $U^{\text{avg}}$ (2.78) and $J^{\text{avg}}$ (2.77). One of the main problems is that the definition of the orbitals in a solid is not unique and an estimation of $U$ is directly linked to the definition of the interacting orbitals. Literature values for $U$ and $J$ for specific materials must therefore be interpreted as reference values and can deviate significantly from values appropriate for the model under consideration.

One strategy to obtain the screened Coulomb interaction parameters is to perform a constrained DFT calculation [71]. A Lagrange parameter $\nu$ is introduced that fixes the total occupation of $d$-type orbitals of the magnetic impurity. The resulting functional

$$E[n(\mathbf{r}](N_d) = E_{\text{LDA/GGA}}\left[n(\mathbf{r})\right] + \nu\left(n_d\left[n(\mathbf{r})\right] - N_d\right) \tag{4.9}$$

has to be minimized with respect to the charge density $n(\mathbf{r})$ for a fixed value of $N_d$. The $d$-type occupation number is denoted by $n_d$. An explicit expression for $n_d$ in the LAPW basis is given by (2.58). The minimization of (4.9) leads to a modified set of Kohn-Sham equations, similar to the LDA+U method, described in Section 2.5. We implemented the constrained DFT method into the FLEUR program package.

It is now possible to calculate the Hubbard interaction parameter from the change in the ground-state energy with respect to a variation in the $d$-shell occupation. We choose to be consistent with the double-counting correction formula (2.79). For a paramagnetic calculation the following relation holds,

$$U^{\text{avg}} - \frac{J^{\text{avg}}}{2} = \frac{\delta^2 E}{\delta N_d^2}\ . \tag{4.10}$$

We performed paramagnetic GGA calculations using a unit cell with eight atoms where one of the semiconductor cations is replaced by an impurity. The dependence of the Lagrange parameter and of the ground-state energy on $N_d$ for $Cd_3Mn_1Te_4$ are shown in Fig. 4.7. For $\nu = 0$, which corresponds to an unconstrained calculation, the energy is minimal, as expected. Driving the system away from the optimal $d$-shell occupation

results in a quadratic increase in energy. Since the Lagrange parameter $\nu$ is nothing but an external potential, the total occupation depends approximatively linearly on $\nu$. We use a quadratic fit to the energy to obtain $\delta^2 E/N_d^2$. This procedure is repeated for various host-dopant combinations. In order to show the dramatic effect of the choice of the muffin tin radius in this type of methods, we perform all calculations for two different values. A small value of the muffin tin radius yields large values for the interaction strength. The resulting interaction parameters are listed in Table 4.2.

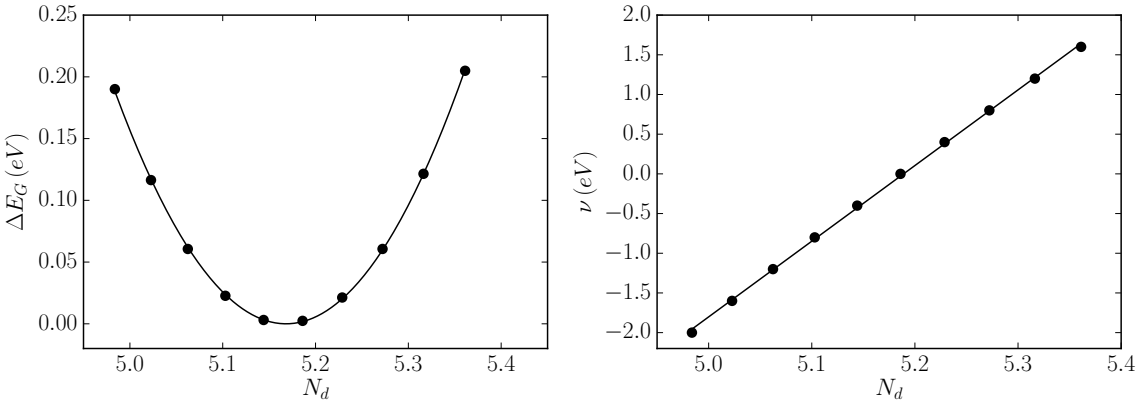


**Figure 4.7:** Results of a constrained DFT calculation for $Cd_3Mn_1Te_4$ using a muffin tin radius of 2.4 a.u.. Left: Normalized ground-state energy obtained for different values of the $d$-occupation $N_d$ of the manganese atom (dots) together with a quadratic fit (line). Right: Dependence of the Lagrange parameter $\nu$ on the filling $N_d$ (dots) together with a linear fit (line).

| $R_{\mathrm{MT}}$ $(a.u.)$ | $Cd_4MnTe_3$ | $Zn_4MnTe_3$ | $Zn_4MnSe_3$ | $Zn_4MnS_3$ | $Ga_4MnAs_3$ | $Cd_4CrTe_3$ |
|---|---|---|---|---|---|---|
| 2.4 | 11.1 | 10.5 | 11.5 | 12.1 | 9.9 | 10.7 |
| 2.0 | 16.5 | 15.7 | 16.7 | 17.3 | 14.9 | 15.1 |

**Table 4.2:** Values for $U^{\mathrm{avg}} - J^{\mathrm{avg}}/2$ obtained from constrained DFT calculations for various DMS in eV. All values were obtained for two different muffin tin radii $R_{\mathrm{MT}}$ of the magnetic impurity.

The constrained DFT method invites further research, such as the investigation of the dependence of the interaction on the muffin tin radius or the analysis of the charge flow. However, these considerations are not within the scope of this thesis. The values in Table 4.2 clearly show that the on-site interaction at the impurity sites are of comparable magnitude for the II-VI semiconductors. Even though it is believed that the constrained DFT method properly includes screening effects, the obtained values are too large. Using the same method, a comparable overestimation was found for $d$-orbitals in copper [72]. For our further calculations, Gutzwiller as well as LDA+U, considerably smaller values for the Coulomb interaction are needed to recover experimental results.

## 4.5 Double-counting correction

The double-counting correction is one of the fundamental problems in combining band-theory with a model-based (many-particle) description of correlations. Methods such as LDA+U, DMFT, and the Gutzwiller-DFT depend on this correction. As already mentioned in Section 2.5, the double-counting term accounts for the interaction which is already included in a mean-field way in the DFT band structure. Since we handle the interaction explicitly by adding a suitable term to the DFT tight-binding model, we are interested in removing the mean-field contribution. This poses a difficult problem due to the very nature of density functional theory, which does not allow us to identify such contributions in a straightforward manner, unlike for example, in the Hartree-Fock method. If we add an interaction term to the DFT tight-binding Hamiltonian and perform a Gutzwiller calculation without accounting for the correction we get an unphysical charge-flow from the correlated $d$-orbitals into the uncorrelated orbitals.

There exist several formulations for the double-counting correction that are all semi-empirical. The most famous formulations are the Around Mean Field (AMF) and the Fully Localized Limit (FLL) method. The choice of the double-counting correction can strongly influence the outcome of a calculation. For a comparison of the methods, see for example [73]. We will briefly describe the double-counting correction we use in our calculation scheme.

Our calculation scheme for DMS presented in this chapter can be strictly separated into two parts. (i) The non-magnetic DFT calculation serves as a basis for a realistic tight-binding model, (ii) The Gutzwiller calculation treats the Hubbard Hamiltonian obtained by adding a local interaction to the tight-binding model. The formation of magnetic moments on the impurity sites in (ii) is a result of the interaction term, our DFT calculations in (i) are not spin-polarized. It is important to note that we are not using the Gutzwiller method to extend density functional theory, which would mean that we incorporate the Gutzwiller method into the Kohn-Sham self-consistency cycle, feeding back a possible charge flow into the DFT calculation until self-consistency is achieved. We are solely interested in obtaining a reasonable non-spin-polarized tight-binding model as a starting point for our calculations; this has to be considered in our formulation of the double-counting correction.

In a fully self-consistent method, such as the LDA+U method, the double-counting correction in the FLL formulation is given by [38]

$$E_{\mathrm{dc}}^{\mathrm{FLL}} = \frac{U^{\mathrm{avg}}}{2} n(n-1) - \frac{J^{\mathrm{avg}}}{2}\left[n^{\uparrow}(n^{\uparrow}-1) + n^{\downarrow}(n^{\downarrow}-1)\right], \tag{4.11}$$

where $n^{\uparrow}$, $n^{\downarrow}$ and $n = n^{\uparrow} + n^{\downarrow}$ are the occupations of the correlated orbitals. This energy correction enters the KS potential by an occupation-dependent potential

$$\tau^{s} = \frac{\partial E_{\mathrm{dc}}}{\partial n^{s}} = U^{\mathrm{avg}}\left(n - \frac{1}{2}\right) - J^{\mathrm{avg}}\left(n^{s} - \frac{1}{2}\right). \tag{4.12}$$

In each iteration of the self-consistency cycle, the value $\tau^{s}$ changes, correcting for the LDA mean-field contribution from the last iteration, until self-consistency is achieved.

In our calculation scheme, however, $\tau^s$ is fixed and given by the DFT result we are starting from. We allow a charge flow within our Gutzwiller method but do *not* update $\tau^s$ in part (ii) of our calculations. Since our DFT calculation is paramagnetic, $\hat{H}^{\mathrm{DC}}$ in (4.6) is simply given by

$$\hat{H}^{\mathrm{DC}} = \tau \sum_{c,s} \hat{c}^{\dagger}_{c,s} \hat{c}_{c,s}. \tag{4.13}$$

Using (4.12) and

$$U^{\mathrm{avg}} = F^0 = A + \frac{7}{5} C\,, \tag{4.14}$$

$$J^{\mathrm{avg}} = \frac{F^2 + F^4}{14} = \frac{7}{2} B + \frac{7}{5} C\,, \tag{4.15}$$

we are able to obtain a rough estimate of the interaction energy present in the LDA calculation.

The exact meaning of the double-counting correction in the self-consistent calculations is a difficult topic for itself. We just mention a few complications connected with the double-counting term. Strictly speaking, the interaction present in the LDA calculation should not depend on our choice of $U$ and $J$. The double-counting correction is therefore often tacitly interpreted as a mean-field value of the additional interaction term. We do not correct for the LDA interaction energy, but for the "mean-field" value of the interaction term we add to the calculation. This helps to understand why correlated orbitals in LDA+U calculations are still occupied, even if we consider the $U \to \infty$ limit. On the other hand, exact expressions for the mean-field value of the interaction term are known and the question arises why the exact form is not chosen as the proper double-counting correction. In our numerical experiments it turned out that this choice suppresses important physical effects, e.g. spin-polarization.

To avoid all these complications we treat $\tau$ as a free parameter adjusting the total occupation of our correlated orbitals. It turns out that the value for $\tau$ reproducing experimental values is very close to the estimate from (4.12).

## 4.6 Minimization scheme for several impurities

We use the Gutzwiller method to account for several, different correlated sites per unit cell. This complicates the minimization procedure described in Section 3.4. To keep the notation as simple as possible we use from now on $i$ and $j$ as index for the unit cell and $\sigma$ to indicate all orbitals within the unit cell.

We denote the $z$ impurities within the unit cell by $I_n$, with $n = 1, 2, \ldots, z$. We define the set $\mathbf{I}_n$ to include all ten spin-orbital indices of the five correlated ($d$-type) orbitals belonging to impurity $I_n$

$$\mathbf{I_n} = \left\{ \sigma^n_{\zeta,\uparrow}, \sigma^n_{\zeta,\downarrow}, \sigma^n_{\eta,\uparrow}, \sigma^n_{\eta,\downarrow}, \sigma^n_{\xi,\uparrow}, \sigma^n_{\xi,\downarrow}, \sigma^n_{u,\uparrow}, \sigma^n_{u,\downarrow}, \sigma^n_{v,\uparrow}, \sigma^n_{v,\downarrow} \right\}. \tag{4.16}$$

The set of all correlated orbital indices is then given by

$$\mathbf{C} = \mathbf{I_1} \cup \mathbf{I_2} \cup \ldots \mathbf{I_z} \ . \tag{4.17}$$

The multiband Hubbard Hamiltonian (4.6) is written as

$$\hat{H} = \sum_{(i,\sigma_1)\neq(j,\sigma_2)} t_{i,j}^{\sigma_1,\sigma_2} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{j,\sigma_2} + \sum_{i,\sigma\notin\mathbf{C}} t_{i,i}^{\sigma,\sigma} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{i,\sigma} + \sum_{i,n} \hat{H}_{i,n;\mathrm{loc}} \ , \tag{4.18}$$

where the local correlated part of the Hamiltonian for impurity $n$ in unit cell $i$ is given by

$$\begin{aligned} \hat{H}_{i,n;\mathrm{loc}} = & \sum_{(\sigma_1,\sigma_2)\in\mathbf{I}_n} \left( t_{i,i}^{\sigma_1,\sigma_2} - \delta_{\sigma_1,\sigma_2}\tau_{i,n} \right) \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{i,\sigma_2}^{\dagger} + \\ & \sum_{(\sigma_1,\sigma_2,\sigma_3,\sigma_4)\in\mathbf{I}_n} U_{i,n}^{\sigma_1,\sigma_2,\sigma_3,\sigma_4} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{i,\sigma_2}^{\dagger} \hat{c}_{i,\sigma_3} \hat{c}_{i,\sigma_4} \end{aligned} \tag{4.19}$$

with the double-counting correction $\tau_{i,n}$. We apply the Gutzwiller correlator to each correlated site so that we can formally write

$$|\Psi_{\mathrm{G}}\rangle = \prod_{n=1}^{z} \hat{P}_n |\Phi_0\rangle \ . \tag{4.20}$$

Due to cubic symmetry at the impurity site, we only have to include diagonal variational parameters in the single-site Gutzwiller correlator,

$$\hat{P}_n = \sum_{\Gamma} \lambda_{\Gamma}^{n} |\Gamma\rangle_{nn}\langle\Gamma| \ , \tag{4.21}$$

where $|\Gamma\rangle_n$ are the eigenstates of (4.19)

$$\hat{H}_{i,n,\mathrm{loc}} = \sum_{\Gamma} E_{\Gamma}^{n} |\Gamma\rangle_{nn}\langle\Gamma| \ . \tag{4.22}$$

The Gutzwiller wave-function has to be minimized with respect to an optimal single-particle wave function $|\Phi_0\rangle$ (Slater determinant) and the $z$ sets of variational parameters $\lambda_{\Gamma}^{n}$. Since the local Hamiltonian describes the $d$-shell of the impurity we have $2^{10}$ variational parameters for each impurity in the unit cell.

All quantities introduced in Section 3.4 ($q$-factors, effective fields $\eta$, local density matrices $\mathbf{C}$) are defined analogously for each correlated site with impurity $I_n$, e.g., we can find an expression for the renormalization factor $q_{\sigma_1}^{\sigma_2}$, which depends on the variational parameter matrix $\lambda_{\Gamma}^{n}$, where $\sigma_1, \sigma_2 \in \mathbf{I}_n$.

Due to the symmetry present at the impurity sites, the off-site elements of the renormalization matrices $q$ vanish. We define $q_\sigma$, which includes the renormalization factors of all impurities,

$$q_\sigma := q_\sigma^\sigma, \quad \text{where} \quad q_\sigma = 1 \ \forall \sigma \notin \mathbf{C} \ . \tag{4.23}$$

With this notation we formulate the energy functional (per unit cell)

$$E_{\mathrm{G}}(\lambda_{\Gamma}^{1}, \lambda_{\Gamma}^{2}, \ldots, \lambda_{\Gamma}^{z}; \Phi_0) = \sum_{\sigma_1,\sigma_2} q_{\sigma_1} q_{\sigma_2} E_{\sigma_1,\sigma_2} + \sum_{n} \sum_{\Gamma} E_{\Gamma}^{n} \lambda_{\Gamma}^{n*} \lambda_{\Gamma}^{n} m_{\Gamma,\Gamma}^{n,0} \tag{4.24}$$

with

$$E_{\sigma_1,\sigma_2} \equiv \begin{cases} (1/L) \sum_{i \neq j} t_{i,j}^{\sigma_1,\sigma_2} \left\langle \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{j,\sigma_2} \right\rangle_{\Phi_0} & \text{if } \sigma_1 = \sigma_2 \ \vee \ \sigma_1 \in \mathbf{C} \\ (1/L) \sum_{i,j} t_{i,j}^{\sigma_1,\sigma_2} \left\langle \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{j,\sigma_2} \right\rangle_{\Phi_0} & \text{else,} \end{cases} \tag{4.25}$$

where $L$ is the number of unit cells in the crystal. Our minimization procedure for this functional is described as follows.

(i) We set the starting values. The iteration counter is set to $c = 0$; $q_\sigma = 1$ for all orbitals; $\eta_\sigma$ is set to the on-site energies obtained from the DFT tight-binding model.

(ii) The ground-state $|\Phi_0\rangle$ of

$$\hat{H}_0^{\mathrm{eff}} = \sum_{(i,\sigma)\neq(j,\sigma')} q_\sigma q_{\sigma'} t_{i,j}^{\sigma,\sigma'} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{j,\sigma'} + \sum_{i,\sigma} \eta_\sigma \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{i,\sigma} \tag{4.26}$$

is determined. The impurity $I_n$ is picked, where $n = (j \bmod z) + 1$. From the ground-state we calculate the energy tensor $E_{\sigma_1,\sigma_2}$ (4.25), and the local density matrix for impurity $I_n$

$$C_{\sigma,\sigma'}^{n} = \langle \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{i,\sigma'} \rangle, \qquad \text{where } \sigma, \sigma' \in \mathbf{I}_n \ . \tag{4.27}$$

(iii) An inner minimization is performed. We find the optimal set of variational parameters $\lambda_{\Gamma}^{n}$, fulfilling the constraints and minimizing (4.24) with all other sets of variational parameters kept fixed. We obtain new values $q_\sigma$, where $\sigma \in \mathbf{I}_n$.

(iv) Using $C_{\sigma,\sigma'}^{n}$ we obtain the effective fields $\eta_\sigma$ where $\sigma \in \mathbf{I}_n$.

(v) The new values for the effective fields $\eta_\sigma$ and the hopping reduction factors $q_\sigma$ for the correlated orbitals belonging to impurity $I_n$ are stored for the next iteration. The value $c$ is incremented. Go to step (ii).

As usual, we stop the minimization procedure when self-consistency is achieved, i.e., the newly calculated renormalization factors $q$ and the effective fields $\eta$ are close to the values of the previous iteration. We also apply a simple linear mixing scheme to achieve better convergence.

We implemented our program is such a way that we minimize two impurities simultaneously if we know from symmetry that they are equivalent. We also implemented the possibility to break the $s_z$-symmetry for the impurities by adding small spin-dependent fields for each correlated impurity-site in step (i). This allows us to investigate symmetry-broken ground-states. We often use this to calculate the ground-state energy differences between a ferromagnetically and an antiferromagnetically (Néel) aligned configuration of two impurities. Special care was taken to calculate ground-state expectation values of the effective Hamiltonian $\hat{H}_0^{\mathrm{eff}}$ accurately. We implemented a BZ integration routine based on a subdivision of the BZ into tetrahedra, see Appendix A.4. These more advanced routines are necessary to obtain convergence in our minimization scheme.

## 4.7 Overview of the algorithm

Fig. 4.8 summarizes the algorithm in a flow chart.

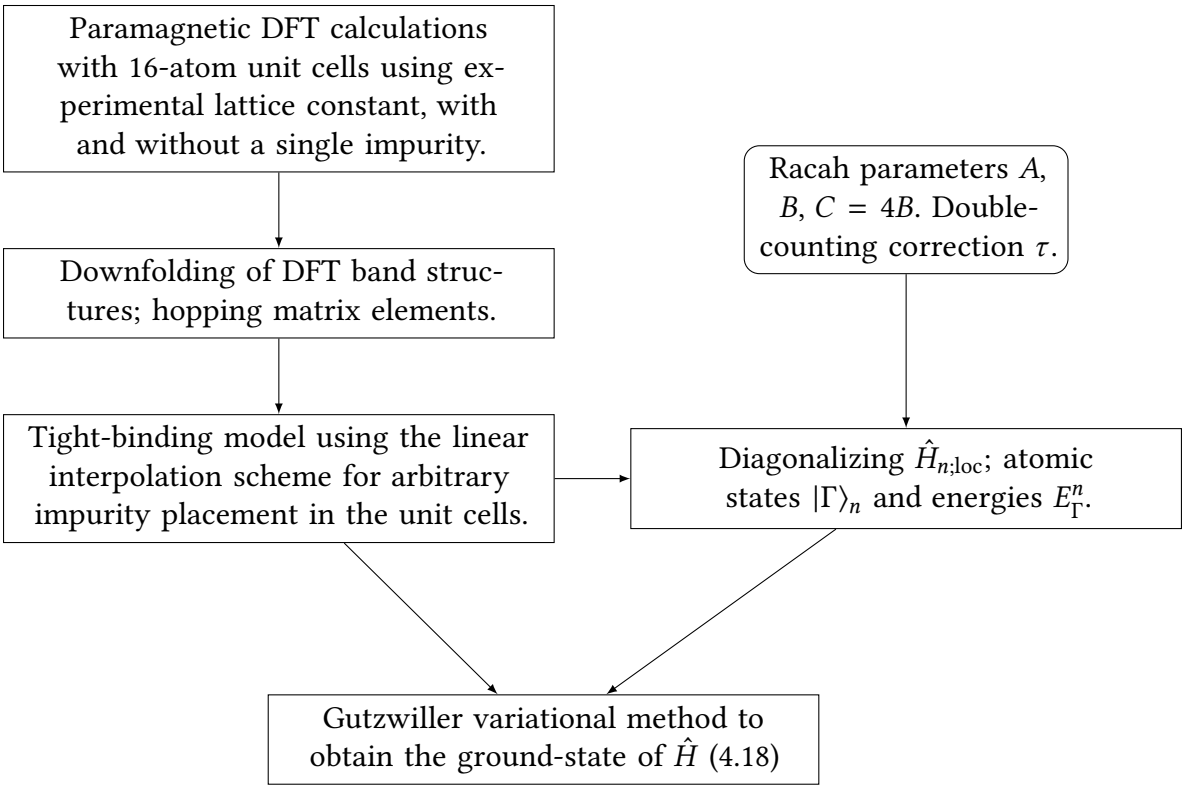


**Figure 4.8:** Flow diagram of the Gutzwiller DFT method for diluted magnetic semiconductors introduced in this chapter. Required input parameters are $A$, $B$, and $\tau$.

# 5 Manganese ions in II-VI semiconductors

*The results presented in this chapter summarize our studies of II-VI semiconductors doped with manganese. This chapter closely follows our publication in [74].*

Mn-doped II-VI semiconductors were among the first diluted magnetic semiconductors to be studied intensively [2, 75]. For small doping, the isovalent Mn ions replace the semiconductor cations. Early on, it was pointed out that the Mn ions possess magnetic moments whose couplings are mediated by the semiconductor host material: The $d$-shell of the manganese weakly hybridizes with neighboring orbitals so that the atomic picture of a $s = 5/2$ Hund's Rule ground-state for the half-filled $3d$-shell stays valid; the weak hybridzation causes an antiferromagnetic coupling between pairs of these localized "spins" due to superexchange. It is an experimental observation that pairs of Mn ions in II-VI semiconductors placed in an external magnetic field exhibit equidistant magnetization steps as one would expect for antiferromagnetically coupled $s = 5/2$ spins. This underpins the atomic picture and validates the description of the spins and their mutual interaction in terms of a Heisenberg model with antiferromagnetic pair couplings $J_n > 0$ at $n$th-neighbor distance.

Not only the exchange couplings between nearest neighbors but also those between Mn ions at 2nd, 3rd, and 4th neighbor distances were experimentally determined; all other couplings are negligibly small, $J_{n\geq 5} \ll J_4$. A modeling of the magnetization curves at very low temperatures leads to the surprising result that $J_4 > J_2, J_3$ [76, 77], i.e., the exchange couplings do not decay monotonously as a function of the geometrical distance. The unexpected non-monotonous decay of $J_n$ as a function of the Mn-Mn separation, and also the overall size of the exchange couplings, are unexplained. Only the nearest-neighbor exchange couplings $J_1$ for Mn-doped II-VI semiconductors were calculated using the superexchange approach [78, 79, 80, 81], or by using DFT [82, 83].

In this chapter we calculate the exchange couplings $J_{n\leq 4}$ using three itinerant-electron approaches, (i), the generalized-gradient approximation (GGA) to DFT, (ii) GGA+U as implemented in the FLEUR package [49], and, (iii), GGA+Gutzwiller for a suitable two-ion Hubbard model. We confirm that $J_4 > J_2, J_3$ and find a reasonable agreement with measured values for Cd(Mn)Te, Zn(Mn)Te, Zn(Mn)Se, and Zn(Mn)S. Furthermore, our analysis shows that the filling of the Mn $3d$-shell is not integer which challenges the notion of Mn ions carrying a spin $s = 5/2$. We study a few-ion toy model to show that the non-integer filling remains consistent with equidistant magnetization plateaus. The picture of a spatially distributed spin $s = 5/2$ emerges which includes the neighboring Wannier orbitals that hybridize with the Mn $3d$-states. We therefore conclude that the concept of interacting Heisenberg spins remains applicable for Mn ions diluted in II-VI semiconductors.

## 5.1 Ion pairs in a semiconductor host

We are interested in the properties of manganese atoms diluted in a II-VI host semiconductor at low temperatures and in sizable magnetic fields. To be definite, we shall focus on CdTe. For very small Mn concentrations $x$ in $Cd_{1-x}Mn_xTe$, we may safely assume that the $Mn^{2+}$ ions substitute the isovalent $Cd^{2+}$ ions. We tested that it is a reasonable approximation to neglect lattice distortions in the theoretical analysis because structural relaxations turned out to be small within the DFT(GGA) calculations, see Section 4.1.

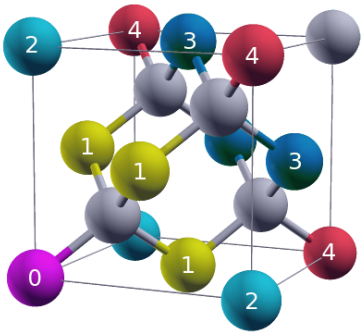


**Figure 5.1:** Eight-atomic unit cell with 1st, 2nd, 3rd and 4th nearest neighbors on the fcc cation sublattice of a (semiconductor) zincblende structure from the lower left corner.

The spin of an isolated Mn ion aligns with any finite magnetic field. The non-trivial magnetization curves seen in experiment are due to the exchange interaction between different Mn ions. Test calculations confirmed that the interaction of three or more Mn ions is given by the sum of pair interactions so that we can concentrate on the interaction between pairs of Mn ions as a function of their distance. We found in numerically expensive Gutzwiller calculations with $L = 512$ atoms in the unit cell that the interaction between two Mn ions beyond 4th-neighbor distance is negligibly small. In Fig. 5.1 we show the first, second, third, and fourth neighbors on the fcc sublattice.

### 5.1.1 GGA and GGA+U calculations

As discussed in Section 4.2, we would ideally study a single pair of Mn ions with Cd ions on all other sites of the fcc lattice. However practical band-structure calculations require translational symmetry. For our pure GGA and GGA+U we therefore start with large but finite supercells with $L$ atoms that contain two Mn ions, and link them together so that periodic boundary conditions apply in all three spatial directions. In this chapter we use supercells with $L = 128$ atoms, that are obtained by transforming the eight-atomic unit cell shown in Figure 4.1 by the matrix (022/202/220). This supercell is sufficient to study Mn pairs that are maximally fourth-nearest neighbors.

We run the FLEUR code using the following settings. We use the DFT with the PBE functional for the exchange-correlation energy. Since we are investigating a band insulator with a sizable gap, it is sufficient to use only 10 inequivalent **k**-points in the irreducible part of the Brillouin zone; depending on the impurity positions, this corresponds to 20 or 40 **k**-points in the full Brillouin zone. The basis functions inside the muffin tins are expanded in spherical harmonic functions with a cut-off of $L_{max} = 10$. The muffin tin radii are $R_{Cd} = R_{Mn} = 2.64$ a.u. and $R_{Te} = 2.58$ a.u. (1 atomic unit $= a_B = 0.529$ Å). We

use $R_{\mathrm{Te}}K_{\max} = 8.26$, where $K_{\max}$ is the plane wave cut-off. For the GGA+U calculations we use the standard FLL double-counting correction [38]. Due to the computational cost of the DFT calculations a rigourous analysis of the error in the coupling constants (see Section 5.1.2) is not feasable. However, for testing purposes, we also calculated $J_1$ and $J_4$ for CdTe with (i) 20 inequivalent $\mathbf{k}$-points and (ii) a larger plane wave cut-off of $R_{\mathrm{Te}}K_{\max} = 9.29$. The results for $J_1$ and $J_4$ in these calculations change by less than one percent.

In the GGA+Gutzwiller approach we use 512 atomic supercells constructed by the linear superposition scheme described in Section 4.2. For the GGA calculation of the 16-atom cells we use 120 $\mathbf{k}$-points in the irreducible part of the Brillouin zone (1/24 of the full Brillouin zone), $L_{\max} = 10$, and $R_{\mathrm{Te}}K_{\max} = 9.80$. The bandstructures obtained are shown in Fig. 4.4.

### 5.1.2 Definition of exchange couplings

The notion of an "exchange coupling" between the two Mn atoms hinges on the concept of a Heisenberg exchange between the two Mn impurity spins at $\mathbf{f}_1$ and $\mathbf{f}_2$,

$$\hat{H}^{\mathbf{f}_1,\mathbf{f}_2}_{\mathrm{Heis}} = 2J_{\mathbf{f}_1-\mathbf{f}_2}\mathrm{S}_{\mathbf{f}_1}\cdot\mathrm{S}_{\mathbf{f}_2}\ . \tag{5.1}$$

Here, we tacitly assume that the average filling of the $3d$-shell in the Mn atoms is close to integer filling, i.e., $n_d \approx 5$, and the Hund's-rule coupling fixes the ground-state spin to $s = 5/2$ on each ion. The exchange coupling is positive, $J_{\mathbf{f}_1-\mathbf{f}_2} > 0$, for an antiferromagnetic coupling.

Under the assumption that a Heisenberg model provides an adequate description of the ground state (and low-energy excitations) of our two Mn impurities, we can estimate their exchange coupling using the bandstructure and GGA+Gutzwiller approach. We orient the Mn spins into the $z$-direction, either parallel ('ferromagnetic alignment') or antiparallel ('Neél-antiferromagnetic alignment'). The algorithm converges to the corresponding (local) minima and provides $(\Delta E)_{\mathbf{f}_1-\mathbf{f}_2} = E_{\parallel} - E_{\perp}$ for the energy differences. This energy difference can also be calculated from the Heisenberg model (5.1),

$$\begin{aligned}(\Delta E)_{\mathbf{f}_1-\mathbf{f}_2} &= 2J_{\mathbf{f}_1-\mathbf{f}_2}\langle \mathrm{FM}|\mathrm{S}_{\mathbf{f}_1}\cdot\mathrm{S}_{\mathbf{f}_2}|\mathrm{FM}\rangle - 2J_{\mathbf{f}_1-\mathbf{f}_2}\langle \mathrm{AFM}|\mathrm{S}_{\mathbf{f}_1}\cdot\mathrm{S}_{\mathbf{f}_2}|\mathrm{AFM}\rangle \\ &= 4J_{\mathbf{f}_1-\mathbf{f}_2}(5/2)^2 = 25J_{\mathbf{f}_1-\mathbf{f}_2}\ ,\end{aligned} \tag{5.2}$$

with the spin states $|\mathrm{FM}\rangle = |5/2,5/2\rangle_{\mathbf{f}_1}|5/2,5/2\rangle_{\mathbf{f}_2}$ and $|\mathrm{AFM}\rangle = |5/2,5/2\rangle_{\mathbf{f}_1}|5/2,-5/2\rangle_{\mathbf{f}_2}$. Here we used that only the $z$-components contribute to the expectation values. In this way, the values $J_{\mathbf{f}_1-\mathbf{f}_2} = (\Delta E_{\mathbf{f}_1-\mathbf{f}_2})/25$ are accessible from approaches that employ itinerant electrons.

## 5.2 Exchange couplings

First, we show that the experimentally observed exchange couplings for Mn ion pairs up to 4th-neighbor distance can be reproduced from scaled DFT(GGA), GGA+U, and GGA+Gutzwiller. Second, we analyze the local occupancies as obtained in GGA+Gutzwiller.

| $J^{\mathrm{CdTe}}$ | Exp. | GGA | s·GGA | GGA+U | GGA+G |
|---|---|---|---|---|---|
| $J_1$ | 6.1 | 17.1 | 6.1 | 6.1 | 6.1 |
| $J_2$ | 0.06 | 0.30 | 0.11 | 0.10 | 0.10 |
| $J_3$ | 0.18 | 0.96 | 0.34 | 0.30 | 0.27 |
| $J_4$ | 0.39 | 1.44 | 0.51 | 0.49 | 0.61 |

**Table 5.1:** Heisenberg exchange couplings $J_n$ in K between Mn ions at $n$th neighbor distance on the Cd fcc lattice in **CdTe** from experiment, [77] and from DFT(GGA), DFT(GGA) scaled by a factor $s = 0.357$, GGA+U for $\overline{U} = U - J = 2.65\,\mathrm{eV}$, and GGA+Gutzwiller for $A = 4.4\,\mathrm{eV}$, $B = 0.1\,\mathrm{eV}$, $C = 0.4\,\mathrm{eV}$ and $n_d = 5.19$.

The values for the exchange couplings $J_n$ in Cd(Mn)Te are known from experiment for up to 4th neighbors on the cation fcc lattice. The values for the couplings have been determined from the steps in the magnetization as a function of the externally applied field for very low temperatures, $T \lesssim 0.1\,\mathrm{K}$. Their sequence, e.g., the fact that $J_4 > J_2, J_3$, has been extracted from a fit of the data to cluster spin models. Bindilatti et al. [77] find $J_1 = 6.1 \pm 0.3\,\mathrm{K}$, $J_2 = 0.06 \pm 0.01\,\mathrm{K}$, $J_3 = 0.18 \pm 0.01\,\mathrm{K}$, and $J_4 = 0.39 \pm 0.02\,\mathrm{K}$. In this section we derive and compare the exchange couplings from DFT(GGA), GGA+U and GGA+Gutzwiller calculations, and compare the resulting magnetization curves with experiment.

### 5.2.1 Coupling strengths

The DFT(GGA) calculation does not contain any specific parameters to adjust the exchange couplings. For large supercells, $L = 128$, the influence of Mn pairs between neighboring supercells in negligibly small.

As seen from Table 5.1, the value for the nearest-neighbor coupling from DFT(GGA) is too large by more than a factor of two, $J_1^{\mathrm{DFT}} = 17.1\,\mathrm{K} \approx J_1/0.36$. DFT(GGA) overestimates the size of the exchange coupling because it finds a too small charge-transfer gap $\Delta_{\mathrm{CT}}$ between occupied Te levels and unoccupied Mn-levels in Cd(Mn)Te. In super-exchange models [81], the exchange integral $J_1$ is inversely proportional to $\Delta_{\mathrm{CT}}$ so that the exchange integral $J_1$ becomes too large in DFT(LDA) and DFT(GGA), by almost a factor of three. GGA+U is frequently used to tackle gap problems in correlated insulators. When we apply a Hubbard-$U$ on the Mn sites, we find a larger charge-transfer gap which leads to smaller exchange couplings, see below. As mentioned in Section 2.4.2, the gap in pure CdTe is too small in DFT(GGA) calculations. This can also be corrected using GGA+U [50]. However, the exchange couplings between Mn ions are mediated by electron transfer processes between Mn and Te so that the precise value of the CdTe band gap is irrelevant for our considerations.

In Fig. 5.2 we show the dependence of $J_1^{\mathrm{GGA+U}}$ as a function of $U$ for various values of $J$. The exchange coupling only depends on the combination $\overline{U} = U - J$ [84]. For $\overline{U} = 2.65\,\mathrm{eV}$ we obtain $J_1^{\mathrm{GGA+U}} = 6.1\,\mathrm{K}$. The values for other exchange interactions for farther distances are collected in Table 5.1. The values for $J_{2,3,4}^{\mathrm{GGA+U}}$ are very similar, and even slightly closer to experiment, than those from the scaled DFT(GGA). This demonstrates that an adjustment

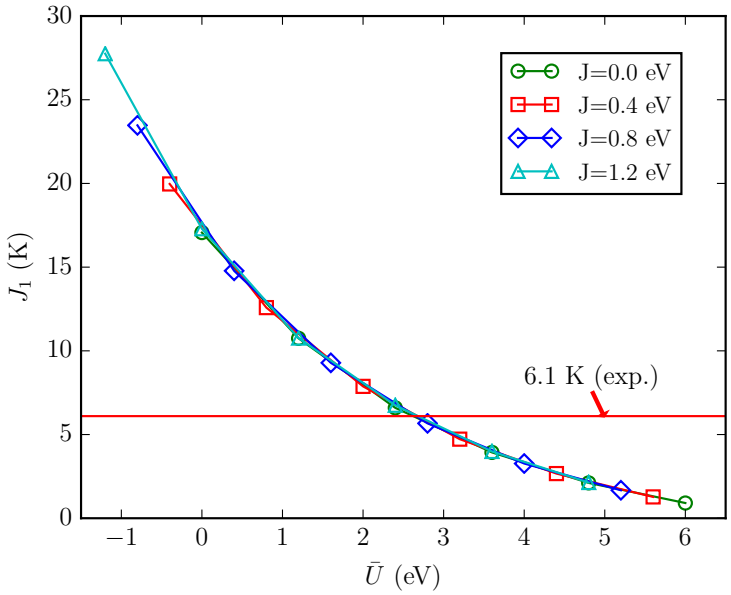


**Figure 5.2:** Heisenberg exchange coupling $J_1$ between two Mn ions in CdTe at nearest-neighbor distance from GGA+U as a function of $\overline{U} = U - J$, calculated for a supercell with $L = 128$ atoms using the FLEUR program package. The red horizontal line shows the experimental value $J_1^{\text{exp}} = 6.1\,\text{K}$.

of the charge-transfer gap cures in effect the overestimation of the exchange interactions in DFT(GGA).

Lastly, we discuss the results for $J_n$ as obtained from our GGA+Gutzwiller calculations. We set $C = 0.4\,\text{eV}$, in agreement with crystal-field theory for data from infrared spectroscopy for isolated $Mn^{2+}$ ions in CdTe [75]. Moreover, we use $C = 4B$, i.e., $B = 0.1\,\text{eV}$, as is a reasonable assumption for transition metals [70]. A similar set of values was used in a recent study of exchange integrals in Mn-doped II-VI semiconductors [81]. The Hubbard-parameter $U$ in transition metals is of the order of several eV [85]. In this work we set $A = 4.4\,\text{eV}$. Note that we have $U = A+4B+3C$ and $J = (5/2)B+C$ for the intra-orbital Hubbard interaction and Hund's-rule coupling, or, for the Slater-Condon parameters, we have $F^{(0)} = A + (7/5)C$, $F^{(2)} = 49B + 7C$, and $F^{(4)} = (63/5)C$ [86]. Therefore, the Hund's-rule exchange on the Mn sites is $J = 0.65\,\text{eV}$ and we employ $F^{(0)} = 4.96\,\text{eV}$ or $U = 6\,\text{eV}$.

In Fig. 5.3, we show $J_1$ as a function of the electron number $n_d$ in the Mn $3d$-shell. As seen from the figure, the curves for $4.0\,\text{eV} \leq A \leq 4.8\,\text{eV}$ and $0.3\,\text{eV} \leq C \leq 0.5\,\text{eV}$ essentially collapse onto each other in the region of interest, $J_1 = 6.1\,\text{K}$. Therefore, the specific choice of the Racah parameters is not crucial. As also seen from Fig. 5.3, the filling is not integer. Instead, we find that $n_d = 5.19$ reproduces the nearest-neighbor Heisenberg exchange coupling best for $A = 4.4\,\text{eV}$, $B = 0.1\,\text{eV}$, $C = 0.4\,\text{eV}$. The resulting values for the exchange couplings for Mn ions in CdTe are compiled in Table 5.1.

Our Gutzwiller calculations here are very close to a Hartree-Fock calculation. Correlation effects are small for the two fully polarized Mn atoms with their (anti-)parallel spins. We discuss this point further in Section 5.3.3. This agreement is specific for Mn

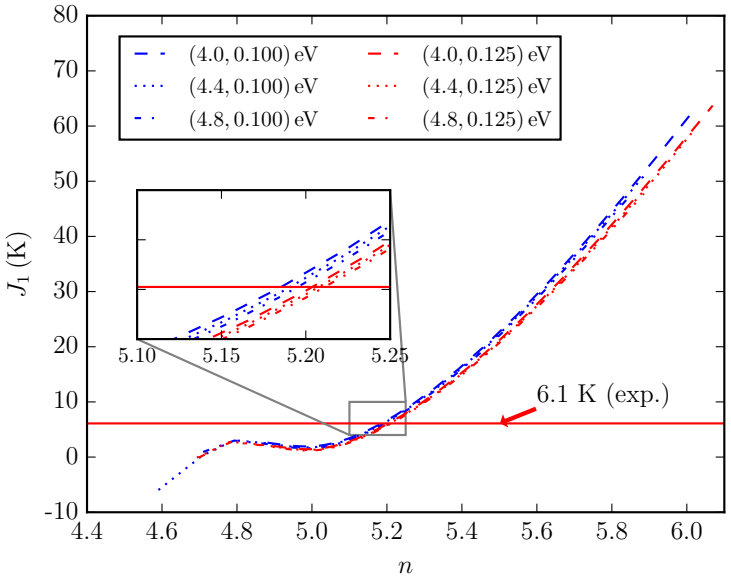


**Figure 5.3:** Heisenberg exchange coupling $J_1$ between two Mn ions in CdTe at nearest-neighbor distance from GGA+Gutzwiller as a function of the number of electrons in the Mn $3d$-shell for various values of the Racah parameters $A$, $B$, and $C = 4B$. The red horizontal line shows the experimental value, $J_1^{\text{exp}} = 6.1\,\text{K}$.

in II-VI semiconductors because we encounter a fully polarized, half-filled $3d$ shell in a wide-gap insulator. In other systems, correlation effects are more pronounced, as seen in calculations for Mn-doped GaAs, see Chapter 6.

We compile the exchange couplings for Zn(Mn)Se, Zn(Mn)Te, and Zn(Mn)S in Table 5.2. Note that the exchange couplings for $n > 4$ are at least an order of magnitude smaller than $J_2, J_3$, of the order of $J_{n\geq 5} = 0.01\,\text{K}$, or less. This justifies our restriction to $J_{n\leq 4}$.

As seen from the tables, the GGA+Gutzwiller method overestimates by some 20%-30% the nearest-neighbor exchange couplings $J_1$ for Zn-VI semiconductors (VI=Te, Se, S) when we use $A = 4.4\,\text{eV}$, $B = 0.1\,\text{eV}$, $C = 0.4\,\text{eV}$ and $n_d = 5.19$ for the Mn ions. With this parameter set, the method can be used to provide a reasonable estimate for the nearest-neighbor couplings for Mn ion pairs in II-VI semiconductors. GGA+Gutzwiller provides a much better estimate for the couplings $J_{n\geq 2}$ than DFT(GGA) but they are still systematically too large by a factor two to three.

### 5.2.2 Magnetization for small doping and low temperatures

As an application, we calculate the magnetization $M(B)$ as a function of the applied external field $B$ for $Cd_{1-x}Mn_xTe$ at small but finite doping $x = 0.005$. A Mn ion is placed in the center of a large but finite fcc lattice with $50^3$ sites. Then, Cd atoms in the surrounding of the 'seed site' are replaced by Mn atoms with probability $x$. As a first possibility, the central Mn ion remains isolated, i.e., with only Cd atoms on its 1st, 2nd, 3rd, and 4th neighbor shell ('maximal surrounding'). In the absence of spin-orbit coupling, the spin of

| $J^{\text{ZnTe}}$ | Exp. | GGA | s·GGA | GGA+G |
|---|---|---|---|---|
| $J_1$ | 9.0 | 41.2 | 9.0 | 11.45 |
| $J_2$ | 0.20 | 0.96 | 0.21 | 0.49 |
| $J_3$ | 0.16 | 2.61 | 0.57 | 0.54 |
| $J_4$ | 0.51 | 3.97 | 0.87 | 1.13 |

| $J^{\text{ZnSe}}$ | Exp. | GGA | s·GGA | GGA+G |
|---|---|---|---|---|
| $J_1$ | 12.2 | 48.1 | 12.2 | 14.97 |
| $J_2$ | 0.16 | 0.81 | 0.21 | 0.28 |
| $J_3$ | 0.07 | 1.61 | 0.41 | 0.42 |
| $J_4$ | 0.43 | 3.26 | 0.82 | 1.16 |

| $J^{\text{ZnS}}$ | Exp. | GGA | s·GGA | GGA+G |
|---|---|---|---|---|
| $J_1$ | 16.9 | 60.3 | 16.9 | 19.73 |
| $J_2$ | 0.27 | 0.99 | 0.28 | 0.47 |
| $J_3$ | 0.04 | 1.14 | 0.32 | 0.40 |
| $J_4$ | 0.41 | 2.85 | 0.80 | 0.97 |

**Table 5.2:** Heisenberg exchange couplings $J_n$ in K between Mn ions at $n$th neighbor distance on the cation fcc lattice in **ZnTe**, **ZnSe**, and **ZnS** from experiment, [77] from (scaled) DFT(GGA), and from GGA+Gutzwiller for $A = 4.4\,\text{eV}$, $B = 0.1\,\text{eV}$, $C = 0.4\,\text{eV}$ and $n_d = 5.19$.

such an isolated Mn ion aligns with any finite magnetic field so that its magnetic response is given by the Brillouin function. Note that neglecting the spin-orbit coupling is justified because of the full magnetic polarization of the Mn ions [63].

A second possibility are two-spin clusters with exactly one Mn ion in the maximal surrounding of the seed site. Two such clusters are equivalent when they can be mapped onto each other by applying some space-group transformations of the fcc lattice. Since equivalent clusters lead to the same magnetic response we only need to store one representative $C$ and determine its multiplicity $A_C$. Moreover, we need to calculate the probability $p_C$ that a lattice point is part of cluster $C$ [87, 76]. For example, for a nearest-neighbor cluster we have $A_C = 12$ and $p_C = x^2(1-x)^{72}$ (because in this case 72 sites must be unoccupied). This construction principle is readily generalized for clusters with three or more spins.

In this work we include clusters with one to four Mn atoms and thus find in total 1130 inequivalent clusters $C$. At doping $x = 0.005$, clusters with up to three Mn atoms cover 98.5% of all possible configurations, clusters with up to four Mn atoms cover 99.6% of all possible configurations. Therefore, clusters with five and more Mn atoms are irrelevant at $x = 0.005$.

For each cluster $C$, the interaction between the Mn spins is described by a Heisenberg model,

$$\hat{H}^C_{\text{Heis}}(B) = \sum_{\substack{\mathbf{f}_1,\mathbf{f}_2 \in C \\ (\mathbf{f}_1 \neq \mathbf{f}_2)}} J_{\mathbf{f}_1-\mathbf{f}_2} \hat{\mathbf{S}}_{\mathbf{f}_1} \cdot \hat{\mathbf{S}}_{\mathbf{f}_2} - g\mu_{\text{B}} B \sum_{\mathbf{f}\in C} \hat{S}^z_{\mathbf{f}} \,, \tag{5.3}$$

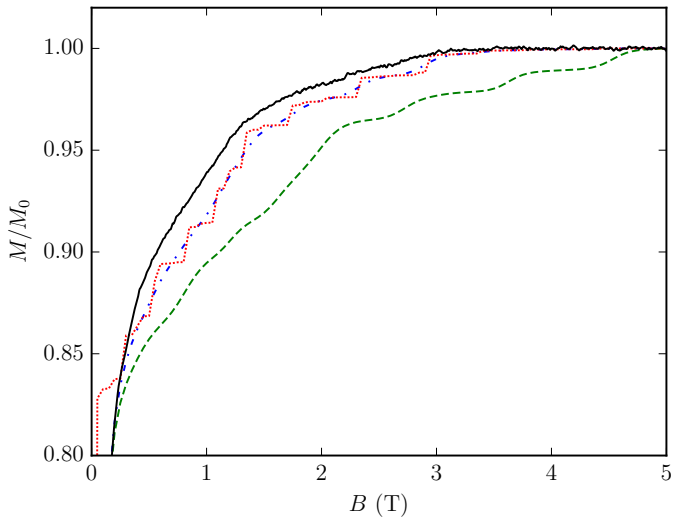


**Figure 5.4:** Magnetization $M(B)$ as a function of the external field $B$ normalized to its value at $B = 5\,\mathrm{T}$ for $Cd_{1-x}Mn_xTe$ at Mn-doping $x = 0.005$. Black line (solid): experimental curve [77]; red line (dotted): Heisenberg model with experimental coupling parameters at $T = 0$; blue line (dash-dotted): Heisenberg model with experimental coupling parameters at $T_{\mathrm{spin}} = 0.1\,\mathrm{K}$; green line (dashed): Heisenberg model with GGA+Gutzwiller parameters at $T_{\mathrm{spin}} = 0.1\,\mathrm{K}$. Clusters with up to four Mn ions are included.

where the sums run over all lattice sites $\mathbf{f}$ in cluster $C$, containing $n_C = 1, \ldots, 4$ spins. We include the interaction with the external field $B$ where $g = 2$ is the gyromagnetic ratio and $\mu_{\mathrm{B}}$ is the Bohr magneton. For our comparisons with experiment, we use the experimental values for $J_n$ from Table 5.1 and theoretical values from the GGA+Gutzwiller approach. However, the differences between scaled GGA, GGA+U, and GGA+Gutzwiller are fairly small.

For each cluster $C$, we determine its contribution to the magnetization per lattice site,

$$M^C(B) = \frac{1}{n_C}\mathrm{Tr}\left(\hat{\rho}_C \sum_{\mathbf{f}\in C} \hat{S}^z_{\mathbf{f}}\right), \quad \hat{\rho}_C = \frac{e^{-\beta \hat{H}^C_{\mathrm{Heis}}(B)}}{\mathrm{Tr}\left(e^{-\beta \hat{H}^C_{\mathrm{Heis}}(B)}\right)} \tag{5.4}$$

with $\beta = 1/(k_{\mathrm{B}}T_{\mathrm{spin}})$. The trace is readily calculated using the exact spectrum that we obtain from a complete diagonalization of the cluster Hamiltonian $\hat{H}^C_{\mathrm{Heis}}(B)$. The magnetization per lattice site is then given by the sum over all clusters weighted by their multiplicity $A_C$ and probability $p_C$,

$$M(B) = \sum_C A_C p_C M^C(B) \,. \tag{5.5}$$

We show the resulting magnetization in Fig. 5.4.

The curve for zero temperature shows the expected magnetization steps that occur when more and more Mn pairs (or clusters) align with the external field [88]. When we use the experimentally determined values for the exchange couplings from Table 5.1 and a

spin temperature $T_{\text{spin}} = 100\,\text{mK}$ that is somewhat higher than the environment temperature $T = 20\,\text{mK}$ [77], we find that the agreement between theory and experiment for $M(B)$ is very good. The agreement becomes slightly worse when we use the coupling parameters calculated by GGA+Gutzwiller. Note that the experimentally accessible magnetic fields probe mostly $J_2$, $J_3$ and $J_4$ because we have $k_\text{B}T, g\mu_\text{B}B \ll J_1$ and $k_\text{B}T, g\mu_\text{B}B \gg J_{n\geq 5}$.

## 5.3 Band structure, density of states and local moments

As a second result, we present the band structure and Density Of States (DOS) of the prototype material $Cd_3MnTe_4$ which we obtain using the local parameters for the on-site interaction and the double-counting correction that reproduce the experimental observed exchange couling $J_1$. For the sake of clarity we focus on this rather small unit cell with one impurity. The local physics at the impurity site is only marginally affected by the presence of other magnetic impurities in its surrounding, so that the results are comparable to results obtained from the full calculation with 128 atoms. Again, we only present data for Mn doped CdTe since the results for Mn doped Zn-based II-VI semiconductor are qualitative and even quantitatively very similar.

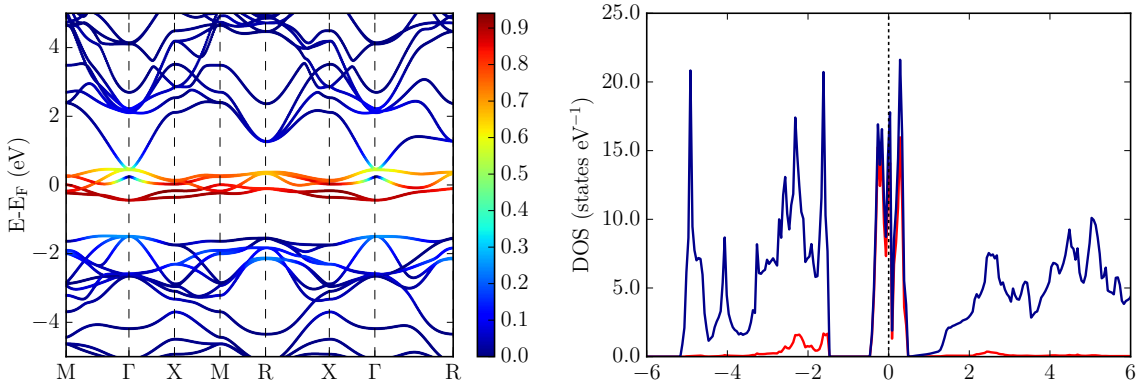


**Figure 5.5:** Band structure and DOS of $Cd_3MnTe_4$, obtained from downfolding a paramagnetic GGA calculation. Left: color coding of the band structure indicates the Mn $3d$-character of the bands. Right: Blue indicates the total DOS, red indicates the local DOS obtained by projecting on Mn $3d$ orbitals.

### 5.3.1 Paramagnetic band structure

Fig. 5.5 depicts the downfolded band structure of $Cd_3MnTe_4$ and the corresponding DOS, obtained from a paramagnetic GGA calculation. As described in Section 4.2 and Section 4.3 this results are used as input for our Gutzwiller variational method. The color coding in Fig. 5.5 indicates the amount of Mn $3d$-character. The Mn $3d$-bands are positioned at the Fermi energy, because the DFT calculation is not spin-polarized and the $d$-shell has to be approximatively half-filled. The $3d$-bands are flat with a bandwidth of about 1 eV, indicating only a small overlap with neighboring orbitals. The DOS shows a peak at the

| atom | type | $n_{\mathrm{GGA}}$ | $n^{\downarrow}_{\mathrm{GGA+G}}$ | $n^{\uparrow}_{\mathrm{GGA+G}}$ | $m$ |
|---|---|---|---|---|---|
| Mn | $sp^3$ | 1.640 | 0.904 | 0.831 | 0.073 |
| Mn | $t_{2g}$ | 2.055 | 2.969 | 0.199 | 2.769 |
| Mn | $e_g$ | 3.718 | 1.986 | 0.038 | 1.947 |
| Cd | $sp^3$ | 7.001 | 3.495 | 3.385 | 0.110 |
| Te | $sp^3$ | 22.587 | 11.647 | 11.547 | 0.101 |
| $\sum$ | | 37.000 | 21.000 | 16.000 | 5.000 |

**Table 5.3:** Occupation numbers for $Cd_3MnTe_4$ obtained from the paramagnetic GGA, $n_{\mathrm{GGA}}$, and the spin-polarized GGA+Gutzwiller method, $n^{\uparrow,\downarrow}_{\mathrm{GGA+G}}$. Note that the Cd atoms and Te atoms can be mapped by symmetry transformation of the crystal onto each other. The table shows the total charge for equivalent atoms.

Fermi energy that is mainly of Mn $3d$-character. Due to crystal field splitting, the peak at the Fermi level has the typical $d$-substructure with two subpeaks originating from the $e_g$ and $t_{2g}$ states. The $d$-orbitals are hybridized with the Te $p$-orbitals of the valence band. In Table 5.3 we show the corresponding occupation numbers $n_{\mathrm{GGA}}$. As seen from the table, the $d$-type Wannier orbitals are occupied by $n_{\mathrm{GGA,CdTe}} = 5.77$ electrons in total.

### 5.3.2 Spin-split bands and Mn moments

Fig. 5.6 shows the results of our Gutzwiller calculation with the on-site Racah parameters $A = 4.4\,\mathrm{eV}$, $B = 0.1\,\mathrm{eV}$, $C = 0.4\,\mathrm{eV}$ and a double-counting correction of $\tau = 23.06\,\mathrm{eV}$ which adjusts the filling of the $d$-shell to $n_d = 5.19$ electrons. The $d$-orbitals are split by strong exchange fields $\eta_\sigma$, c.f (4.26), shifting the majority bands by about 7 eV against the minority bands. The formation of a local magnetic moment is evident. The corresponding occupation numbers $n^{\uparrow,\downarrow}_{\mathrm{GGA+G}}$ are collected in Table 5.3. The induced magnetic moment at the Mn site is not excactly $5\mu_B$, as one could expect from a naive atomic Hund's rule picture, but $4.71\,\mu_B$.

At first glance it is surprising that the induced magnetic moment over the whole unit cell is exactly $5\,\mu_B$. This observation is readily understood by noting that the total induced magnetic moment has to be integer because the majority bands and the minority bands are both gapped at the Fermi level, resulting in integer numbers in the total occupation of both spin-channels. Therefore, a magnetization of $\approx 0.29\mu_B$ has to come from spin-polarized neighboring atoms. Surprisingly, not only the Te atoms show a small polarization, but also the Cd sites, that are next-nearest neighbors. Since the magnetic moment is not fully localized at the impurity site, reported local magnetic moments from DFT calculations [89] are always below $5\,\mu_{\mathrm{B}}$,

As we will see in the next subsection, the Gutzwiller calculations are close to Hartree-Fock calculations, because the Mn $d$-shell is close to half-filling and strong exchange fields correctly "select" the true many-body ground state which is close to an atomic Hund's rule $s = 5/2$ state. Because of the absence of many-body effects we expect the GGA+U calculation to produce similiar results.

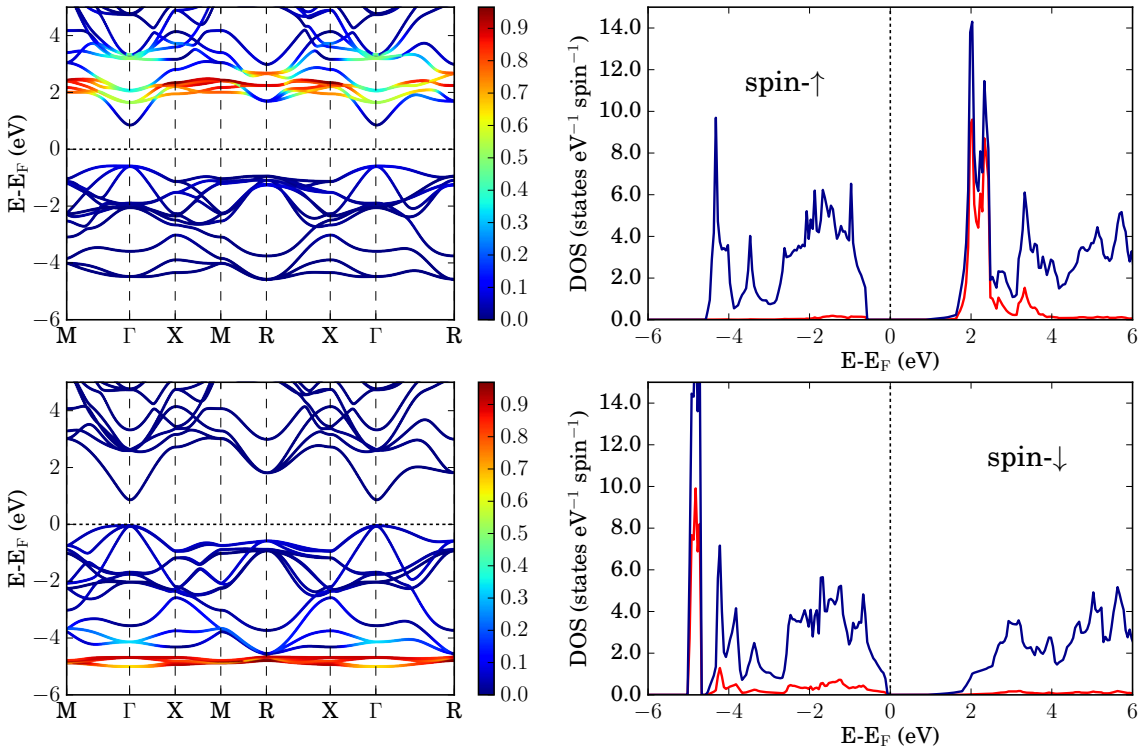


**Figure 5.6:** Band structure and DOS of $Cd_3MnTe_4$ as obtained from a spin-polarized GGA+Gutzwiller calculation using the on-site Coulomb parameters $A = 4.4$ eV, $B = 0.1$ eV, $C = 0.4$ eV and a double-counting correction of $\tau = 23.06$ eV. The top figures show the minority spin channel, the bottom figures the majority spin channel. The band structures (left) are color coded, showing the amount of Mn $3d$-character in the bands. The DOS plots (right) the blue lines indicate the total DOS, the red lines indicate the DOS projected onto the Mn $3d$-states.

In Fig. 5.7 we show the DOS for a spin-polarized GGA calculation and a spin-polarized GGA+U calculation with $\bar{U} = 2.65$ eV which was used to obtain the experimentally observed exchange coupling for $J_1$. The spin-polarized GGA+U calculation also shows an exchange splitting of about 7 eV, with band positions comparable to the ones obtained in the GGA+Gutzwiller calculation. The pure GGA calculation, however, shows a spin-splitting of about 4.5 eV, which underestimates the charge-transfer gap and leads to too small exchange couplings.

### 5.3.3 Local density and spin distributions

To gain further insight into the nature of the ground-state of the Mn ion, we present results for the local occupancies.

We start our discussion with the probability distribution $p_c(n)$ to find $n$ $3d$ electrons on the Mn ion on site **f** ($0 \leq n \leq 10$). As seen from Fig. 5.8(a) the distribution peaks at $n = 5$ which reflects the fact that the average particle number is $n_d = 5.19$, see Section 5.2.1. Correspondingly, there also is a sizable probability to find $3d^6$ configurations on the Mn

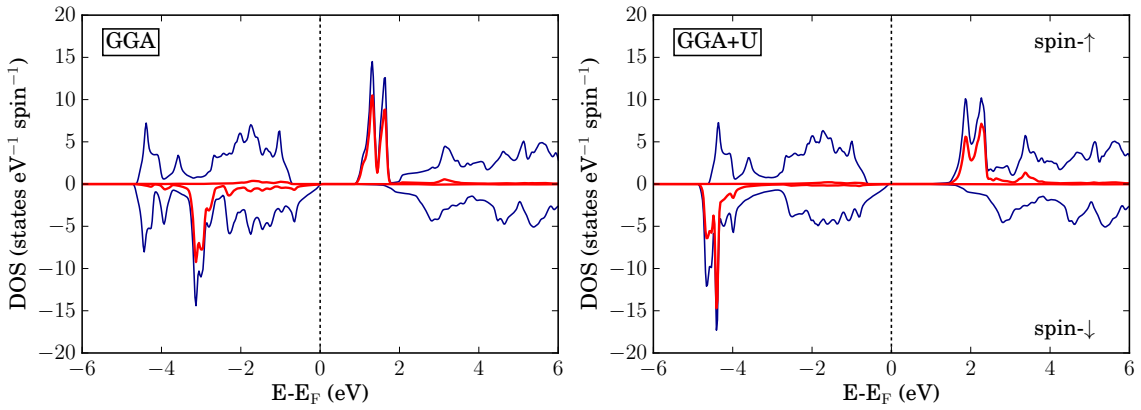


**Figure 5.7:** DOS and projected DOS for a spin-polarized GGA (left) and a spin-polarized GGA+U calculation (right) of $Cd_3MnTe_4$. For the GGA+U calculation $U = 3.25\,\text{eV}$ and $J = 0.60\,\text{eV}$ is used. The blue lines indicate the total DOS, the red-lines the DOS projected onto the Mn $3d$-states.

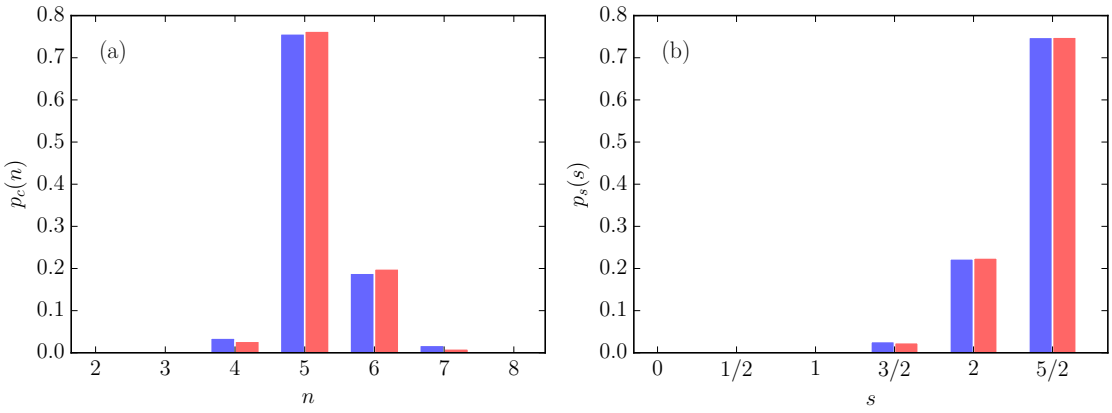


**Figure 5.8:** Mn charge distribution $p_c(n)$ as a function of the number $n$ of $3d$-electrons and Mn spin distribution $p_s(s)$ as a function of the spin $s$ of $3d$-electrons for $A = 4.4\,\text{eV}$, $B = 0.1\,\text{eV}$, $C = 0.4\,\text{eV}$ for $n_d = 5.19$ in GGA+Gutzwiller (red columns), in comparison with the Hartree–Fock result (blue columns).

ion whereas the probability for all other occupation numbers is negligible. Note that this distribution is not the result of electronic correlations because the corresponding Hartree–Fock state displays almost the same distribution function.

The probability distribution function $p_s(s)$ for finding local spins with size $0 \leq s \leq 5/2$ is very similar to the distribution in the single-particle product state $|\Phi_0\rangle$, i.e., the correlation enhancement of the local spin moment is also small for the spin distribution function, see Fig. 5.8(b). The average local spin is $\langle \hat{S}^z_{\mathbf{f}} \rangle_{\text{loc}} = 2.33$ because the admixture of spin $s = 2$ to the dominant configurations with $s = 5/2$ is not negligibly small.

The Mn ions do not show integer filling $n_d = 5$, nor does the spin moment correspond to the atomic spin $s = 5/2$. This observation puts into question the concept of a Heisenberg-

model description that we employed in Section 5.1.2 to derive the exchange couplings. Even if we accept a non-integer filling of the Mn ions' 3$d$-shell, we are actually far from a local-moment regime that is implicit in the Heisenberg-model description. Our analysis of the local occupations shows that the magnetic moment of 5 $\mu_B$ is distributed over the unit cell. The analysis of a toy model in the next section confirms that an extended $s = 5/2$ spin is formed, which is not exactly located at the Mn site.

## 5.4 Magnetic response of ion pairs at non-integer filling

In order to reconcile the finding of a non-integer Mn 3$d$ filling and the notion of a spin $s = 5/2$ effective Heisenberg model, we study the magnetic response in a simplified toy model of two Mn atoms close to their Hund's rule ground states that are coupled to three uncorrelated sites. The uncorrelated sites serve two purposes, namely, (i), they act as a reservoir to adjust the average particle number on the Mn sites away from integer filling and, (ii), they serve as an intermediate charge-transfer (exchange) site to mimic the super-exchange mechanism.

### 5.4.1 Model Hamiltonian

The Hamiltonian for our few-site toy-model, illustrated in Fig. 5.9, is readily formulated. We use the local Hamiltonian $\hat{H}_f^{AT}$ defined in (4.8) for the Mn atoms at $\mathbf{f}_l$ and $\mathbf{f}_r$, and

$$\hat{H}_j^{loc} = \epsilon_j \sum_\sigma \hat{n}_{j,\sigma} \tag{5.6}$$

for the local Hamiltonians of the three uncorrelated orbitals. Here, $\epsilon_j$ (j = l, e, r) are the local chemical potentials that permit the adjustment of the average electron number in the left (l) and right (r) bath orbitals and the exchange (e) orbital, and $\hat{n}_{j,\sigma} = \hat{c}^\dagger_{j,\sigma}\hat{c}_{j,\sigma}$ counts the number of electrons in the uncorrelated orbitals. The sites are coupled via the kinetic terms

$$\begin{aligned} \hat{T}_{l/e,\mathbf{f}_l} &= \sum_{c,\sigma} T^{l/e,\sigma}_{\mathbf{f}_l,c,\sigma}\hat{c}^\dagger_{\mathbf{f}_l,c\sigma}\hat{c}_{l/e,\sigma} + \text{h.c.}\,, \\ \hat{T}_{e/r,\mathbf{f}_r} &= \sum_{c,\sigma} T^{e/r,\sigma}_{\mathbf{f}_r,c,\sigma}\hat{c}^\dagger_{\mathbf{f}_r,c\sigma}\hat{c}_{e/r,\sigma} + \text{h.c.}\,. \end{aligned} \tag{5.7}$$

The full model Hamiltonian reads

$$\hat{H} = \hat{H}^{AT}_{\mathbf{f}_l} + \hat{H}^{AT}_{\mathbf{f}_r} + \hat{H}^{loc}_l + \hat{H}^{loc}_e + \hat{H}^{loc}_r + \hat{T}_{l,\mathbf{f}_l} + \hat{T}_{e,\mathbf{f}_l} + \hat{T}_{e,\mathbf{f}_r} + \hat{T}_{r,\mathbf{f}_r}\,. \tag{5.8}$$

The maximal dimension of the corresponding Fock space is $\dim H = 1024^2 \cdot 4^3$. It is too large to be handled exactly.

From our analysis in Section 5.3.3 we know that, for large $U, J$, those Mn configurations are dominantly occupied that, in the sectors with $n_d = 4, 5, 6$ electrons, have maximal spin $s_{max} = 2, 5/2, 2$ and maximal orbital momentum $l_{max} = 2, 0, 2$, which is a good quantum number in spherical approximation. Therefore, we restrict the Hilbert space of our two

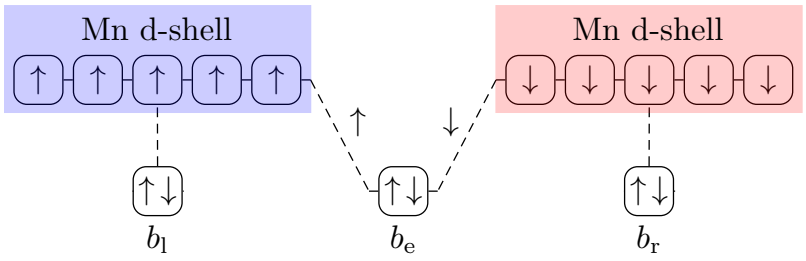


**Figure 5.9:** Toy model for the study of two Mn atoms with two reservoir sites and an indirect Mn-Mn coupling via a charge-transfer site.

Mn atoms to these atomic subspaces. To this end, we introduce the projection operators $\mathcal{P}^{\mathrm{H}}_{\mathrm{g},n_d}$ onto the lowest-lying $(2s_{\max}+1)(2l_{\max}+1)$ Hund's-rule states for fixed electron number $n_d$,

$$\begin{aligned}
\hat{H}^{\mathrm{loc}}_{\mathrm{g}}|\Gamma_n\rangle_{\mathrm{g}} &= E^{\mathrm{loc}}_n|\Gamma_n\rangle_{\mathrm{g}}\,, \\
\hat{n}_{\mathrm{g}}|\Gamma_n\rangle_{\mathrm{g}} &= n_d|\Gamma_n\rangle_{\mathrm{g}}\,, \\
\hat{\mathbf{S}}^2_{\mathrm{g}}|\Gamma_n\rangle_{\mathrm{g}} &= s_{\max}(s_{\max}+1)|\Gamma_n\rangle_{\mathrm{g}}\,, \\
\hat{\mathbf{L}}^2_{\mathrm{g}}|\Gamma_n\rangle_{\mathrm{g}} &= l_{\max}(l_{\max}+1)|\Gamma_n\rangle_{\mathrm{g}}\,, \\
\mathcal{P}^{\mathrm{H}}_{\mathrm{g},n_d} &= \sum_{\Gamma_n} |\Gamma_n\rangle_{\mathrm{g}}\,{}_{\mathrm{g}}\langle\Gamma_n|\,.
\end{aligned} \tag{5.9}$$

Then, we define the total projection operator

$$\mathcal{P}^{\mathrm{H}}_{4,5,6} = \left(\sum_{n_d=4}^{6} \mathcal{P}^{\mathrm{H}}_{\mathrm{f_l},n_d}\right)\left(\sum_{n_d=4}^{6} \mathcal{P}^{\mathrm{H}}_{\mathrm{f_r},n_d}\right)\,, \tag{5.10}$$

and we limit ourselves to the investigation of our model Hamiltonians in the projected form

$$\mathcal{H} = \mathcal{P}^{\mathrm{H}}_{4,5,6}\hat{H}\mathcal{P}^{\mathrm{H}}_{4,5,6}\,. \tag{5.11}$$

The dimension of the partial Fock space on the Mn atoms is $(2s_{\max}+1)(2l_{\max}+1)$ so that the maximal Fock-space dimension is $\dim\mathcal{H} = (25+6+25)^2\cdot 4^3 = 200704$. This partial Fock space is accessible using the Lanczos technique.

### 5.4.2 Magnetization plateaus

The magnetic field couples to the spin-component of the Mn atoms in $z$-direction,

$$\hat{H}_B = -g\mu_{\mathrm{B}}B(\hat{S}^z_{\mathrm{f_l}} + \hat{S}^z_{\mathrm{f_r}})\,. \tag{5.12}$$

The magnetization is obtained from

$$M(B) = \langle\Psi_0|\hat{S}^z_{\mathrm{f_l}} + \hat{S}^z_{\mathrm{f_r}}|\Psi_0\rangle\,, \tag{5.13}$$

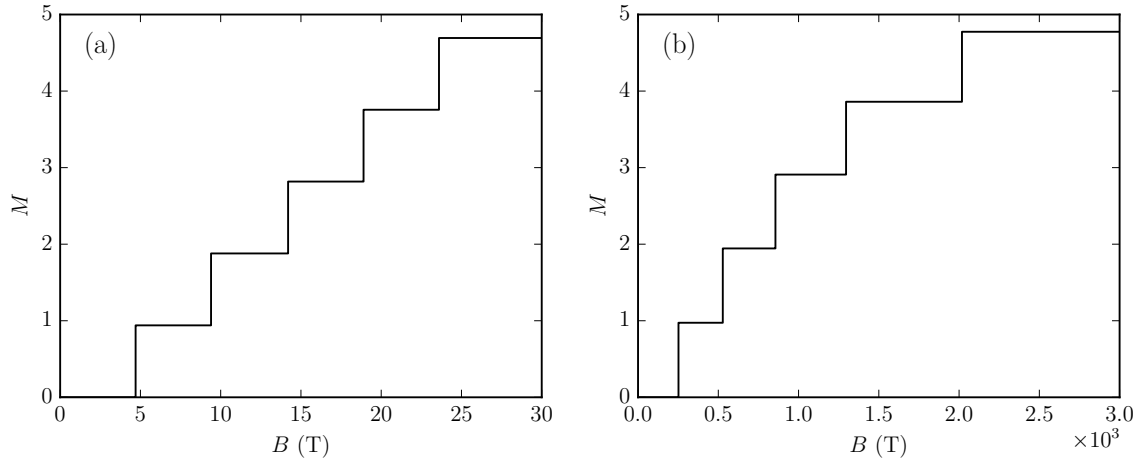


**Figure 5.10:** Magnetization $M(B)$ as a function of the external field for the toy model with an almost filled exchange site, $n_e = 1.98$.

where $|\Psi_0\rangle$ is the ground-state of our model Hamiltonian in the presence of a magnetic field,

$$\mathcal{H}(B) = \mathcal{H} + \hat{H}_B \ . \tag{5.14}$$

We employ the Lanczos algorithm to find $|\Psi_0(B)\rangle$.

We fix the total number of electrons in the system to $n_{\text{tot}} = 16$, and choose the local chemical potentials $\epsilon_l = \epsilon_r$ to adjust the average electron number on the Mn sites so that we have an average number of $n_d = 5.30$ electrons. Note that this number marginally changes as a function of the magnetic field. We set all electron transfer matrix equal in (5.7), $T_{\cdot\cdot}^{\cdot\cdot} = 1\,\text{eV}$.

In the following case (i), we set $\epsilon_e = 8.0\,\text{eV}$ and $\epsilon_{l,r} = 23.1\,\text{eV}$ so that we have $n_e = 1.98$ electrons in the exchange orbital and $n_{l,r} = 1.71$ electrons in each bath orbital in the ground-state. The resulting magnetization steps are equidistant, as shown in Fig. 5.10(a), despite the fact that the Mn filling is far from integer.

The width of the magnetization steps become non-uniform in case (ii) when the exchange site is not almost filled. To illustrate this case, we choose $\epsilon_e = 22.6\,\text{eV}$ and $\epsilon_{l,r} = 19.5\,\text{eV}$ so that we have $n_{l,r} = 1.93$ electrons in each bath orbital and $n_e = 1.51$ electrons in the exchange orbital. Now, the lengths of the corresponding magnetization plateaus are inequivalent, as shown in Fig. 5.10(b). As seen from the figure, the required magnetic fields for the observation of inequidistant magnetization steps become very large and generally unreachable in experiment. Therefore, experimentally realizable magnetic fields which are strong enough to "unlock" the anti-parallel alignment of two antiferromagnetically coupled ions must result in equidistant plateaus.

The toy model shows that equidistant plateaus are possible even though the occupation of the Mn sites is not integer. Our numerical observations can be readily understood using perturbative arguments. For negligible couplings to the exchange orbital, the ground-state of each Mn ion and its attached bath site has spin $s = 5/2$. Note that this spin is not solely located on the Mn site but also partly on the corresponding bath site. In case (i), the

exchange orbital introduces only a small coupling between the left and the right spin-5/2 systems, and perturbation theory leads to a dominant term of the usual antiferromagnetic Heisenberg form (5.1). Consequently, the magnetization steps are equidistant [90, 88]. In case (ii), charge fluctuation contributions invalidate the simple spin-only picture. This results in non-equidistant magnetization steps as seen in Fig. 5.10(b).

When we apply the Gutzwiller variational method to case (i) of our toy model, we find an exchange coupling $J_1$ that is very close to the exact value derived from the width of the magnetization plateaus. This corroborates our finding in Section 5.1.2 and further justifies the applicability of our toy model.

Due to the large gap for charge excitations, the situation of Mn ions in CdTe resembles scenario (i) in our toy model and explains the experimental observation of equidistant magnetization plateaus. The filling of the Mn 3$d$-shell is not integer but the total spin of the Mn ion and its surrounding atoms still is essentially $s = 5/2$.

## 5.5 Summary

In this chapter we used three band structure methods, DFT(GGA), GGA+U, and GGA+Gutzwiller, to derive the exchange couplings between Mn ions diluted in II-VI semiconductor host materials such as CdTe. First, we calculate the energy of the configurations with parallel and antiparallel alignments of the Mn spins. Next, we interpret the energy difference in terms of a two-spin Heisenberg model and thereby deduce the exchange couplings as a function of the Mn-Mn separation for up to fourth neighbor distances.

For the GGA calculations we employ the FLEUR code with the PBE functional for large supercells with $L = 128$ atoms where two of the Cd ions are replaced by isovalent Mn ions. The ab-initio results for the exchange couplings are too large by a factor of two to three which is related to the fact that DFT(GGA) underestimates gaps in II-VI semiconductors systematically. The nearest-neighbor couplings $J_1$ for Mn ions in II-VI semiconductors can be reconciled with experiment by using the GGA+U and GGA+Gutzwiller methods. These methods employ adjustable parameters that are used to match the experimental value for $J_1$ in Cd(Mn)Te. The exchange couplings $J_{2,3,4}$ for 2nd, 3rd, and 4th neighbor distances are then predictions from theory.

In general, the values for $J_{n\geq 2}$ agree qualitatively with experiment, i.e., band theory recovers $J_4 > J_2, J_3$ and $J_{n\geq 5} \lesssim 0.01\,\mathrm{K}$. However, the values for the couplings do not agree perfectly, i.e., we observe quantitative deviations up to a factor of two. About the same level of accuracy can be obtained by a simple rescaling of the DFT(GGA) data that fits the nearest-neighbor coupling $J_1$, see Table 5.2. The bare energy scale in our itinerant-electron description are of the order of several eV, i.e., of the order of $10^5$ K, whereas the exchange couplings $J_{n\geq 2}$ are one Kelvin and below. Therefore, it does not come as a surprise that the band structure methods reach their accuracy limits.

The notion of exchange couplings and the applicability of the super-exchange approach hinges on the mapping of the low-energy degrees of freedom of the itinerant-electron problem to those of a spin-5/2 Heisenberg model. This mapping successfully explains the equidistant magnetization plateaus as a function of applied magnetic field, as seen in experiment. However, the analysis of the Gutzwiller ground state for the two-ion Hubbard

model shows that the filling of the Mn 3$d$-shell is *not* integer which seemingly invalidates the whole concept of localized spins. The analysis of an exactly solvable few-site toy-model reassures that an integer filling is not a prerequisite for equidistant magnetization plateaus. Due to the hybridization of the Mn 3$d$ orbitals with its insulating environment, a slightly delocalized spin-5/2 magnetic moment is formed combing Mn $3d^5$ and $3d^6$ with neighboring valence band states. Our picture of an extended spin-5/2 magnetic moment interacting with each other reconciles the usage of an effective spin-5/2 Heisenberg model to explain the experimentally observed magnetization steps and simultaneously a non-integer valence of the Mn 3$d$ shell.

In the case of Mn-doped II-VI semiconductors, the Gutzwiller method and the Hartree-Fock approach to the two-ion Hubbard model lead to essentially the same results for an (anti-)ferromagnetic alignment of the Mn spins. Our results in the next chapter show that this is not the case for Mn in GaAs where the dopant electrons are more itinerant than in the case of Mn doping.

# 6 Manganese ions in III-V semiconductors: Mn doped GaAs

In this chapter we present Gutzwiller-DFT results for Mn-doped GaAs. Without any doubt, this material is the most prominent and most studied representative of a class called "ferromagnetic DMS". Materials in this class are characterized by the presence of free charge carriers that ultimately cause ferromagnetism at rather small doping concentrations. High scientific effort is devoted to finding and explaining mechanisms that enhance the Curie temperature $T_C$. As mentioned in the introduction, high Curie temperatures are a prerequisite for spintronic devices that operate using charge and spin degrees of freedom. Before we present the Gutzwiller-DFT results for $Ga_{1-x}Mn_xAs$ we give a brief overview of the current theoretical understanding and the experimental results.

## 6.1 (Ga,Mn)As: prospects and problems

The indirect exchange coupling between spins of localized $d$-shells mediated by free carriers can be described by the Ruderman-Kittel-Kasuya-Yosida (RKKY) exchange mechanism [91, 92, 93]. This mechanism takes the itinerant character of the electrons and the Friedel-oscillations around the localized electrons into account. However, in a seminal paper by Dietl et al. [94], it was argued that a Zener model[95] provides a sufficient description of the material if the mean distance between the carriers is larger than that between the spins. In such a case, the Friedel-oscillations average to zero and the RKKY approach becomes equivalent to the Zener model approach. Of course, the model is semiphenomenological: The Mn $d$-orbitals are assumed to form localized spins and the hole states in the valence bands are described using a Kohn-Luttinger parametrization for GaAs. The approach uses the experimentally derived value for $J_{pd}$, which is the exchange between the localized $s = 5/2$ spin and the $s = 1/2$ spin of the carriers. The publication raised enormous attention, because it predicts ferromagnetism in DMS with Curie temperatures above room temperature. For $Ga_{1-x}Mn_xAs$, room temperature ferromagnetism is predicted for a doping concentration of $x = 10\%$ [96].

However, the experimental situation is ambiguous because the fabrication of "good" samples of Mn-doped GaAs poses a difficult task. Unlike in II-VI semiconductors, Mn ions tend to introduce unwanted defects, such as Mn interstitial sites, arsenic anti-sites or clustering of Mn ions. These defects dramatically increase with increasing Mn concentration, limiting the concentrations of available Mn doped GaAs samples well below 10%. Consequently, the sample quality is the main cause of indecisive experiments results. Nevertheless, since the first measurement, the highest experimentally reported Curie temperature of $Ga_{1-x}Mn_xAs$ rose from 60 K [3] to around 170 K [4].

The theoretical description of the defects is complicated by their variety. For example, the role of interstitial Mn ions acting as double donors is discussed in [97, 98], the role of As anti-sites as hole-compensating defects is addressed in [99].

Although Mn-doped GaAs is intensively studied, the fundamental microscopic mechanisms of the induced ferromagnetic correlation are still vigorously debated, see for example [100] and [101]. Therefore, ab-initio calculations constitute a valuable theoretical tool to investigate the magnetic properties of DMS from a more fundamental point of view. In the past, Mn-doped GaAs was subject to various DFT(+U) supercell studies [102, 103, 104, 105]. In the following, we apply the Gutzwiller-DFT method to study $Ga_{1-x}Mn_xAs$ with a focus on the local impurity physics due to intra-atomic Coulomb interactions.

## 6.2 Ab-initio tight-binding results for Mn-doped GaAs

### 6.2.1 DFT band structure

We use the following settings for the 16-atomic DFT calculations that provide the starting point for the Gutzwiller-DFT method, see Section 4.2. We use the PBE functional for the exchange-correlation energy. The muffin tin radii are $R_{\mathrm{As}} = R_{\mathrm{Mn}} = 2.25$ a.u. and $R_{\mathrm{Ga}} = 2.31$ a.u.. We set $L_{\max} = 10$ and $R_{\mathrm{As}}K_{\max} = 8.8$.

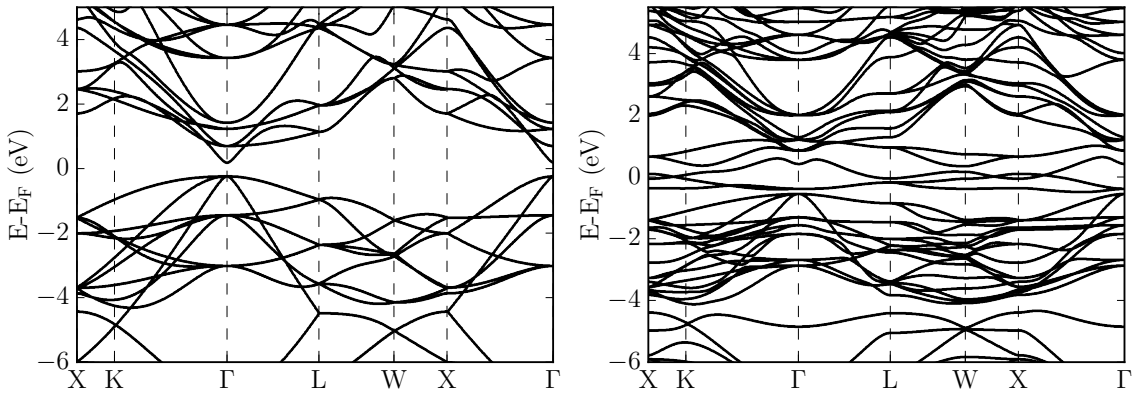


**Figure 6.1:** Band structures of $Ga_8As_8$ (left) and $Ga_7MnAs_8$ (right) for the 16-atom unit cell shown in Fig. 4.3 obtained from paramagnetic DFT calculations utilizing the non-relativistic PBE functional.

Fig. 6.1 shows the obtained band structures for both calculations. We observe a direct band-gap in GaAs at the Γ-point of 0.41 eV, which is in accordance with standard DFT calculations, but only a fraction of the experimental value. The band structure for the calculation containing one Mn impurity shows Mn-3*d* bands lying at the Fermi energy. They have a much more dispersive character compared to the Mn-3*d* bands in the wide-gap semiconductor CdTe. Their bandwidth is about 1.8 eV. For the downfolding of the band-structures into 64 and 69 bands we use the disentanglement routine in Wannier90 with an upper energy window of $E - E_{\mathrm{F}} = 11$ eV. In contrast to the method for Mn-doped CdTe, we do not have to project out *d*-type cation semi-core states directly in-between

the the $s$-type and $p$-type valence bands, see Section 2.6.6. This simplifies the downfolding procedure.

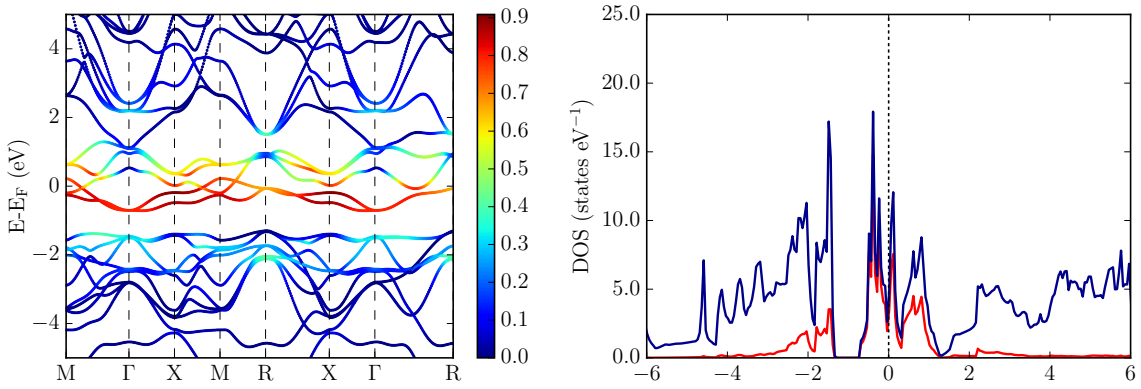


**Figure 6.2:** Band structure and DOS of $Ga_3MnAs_4$, obtained from downfolding a paramagnetic GGA calculation. Left: color coding of the band structure indicates the Mn $3d$-character of the bands. Right: Blue indicates the total DOS, red indicates the local DOS obtained by projecting on Mn $3d$ orbitals.

Figure 6.2 depicts the band structure and DOS obtained from the linear interpolation scheme with a color-coding indicating the admixture of $3d$-states. The Mn $t_{2g}$ bands are strongly hybridized with the $p$-like valence bands below the Fermi energy. We show the corresponding occupations numbers $n_{\mathrm{GGA}}$ in Table 6.1 for a filling of 36 atoms/unit cell. As seen from the table, the Mn $d$-type Wannier orbitals are occupied by a total of $n_{\mathrm{GGA,GaAs}} = 5.59$ electrons in the paramagnetic state.

## 6.3 Band structure and density of states from Gutzwiller-DFT

We investigate the hypothetical eight atomic-unit cell $Ga_3MnAs_4$. Despite the small size of the unit cell we find that the local occupation values of the impurity are very similar to the values obtained in a calculation involving huge cells and low doping concentrations. For the local Coulomb interaction we employ the same values we used in the II-VI calculations. The Racah parameters are set to $A = 4.4\,\mathrm{eV}, B = 0.1\,\mathrm{eV}$, and $C = 0.4\,\mathrm{eV}$. The double-counting correction is set to $\tau = 22.18$ eV to obtain a $d$-shell occupation of $n_{\mathrm{d}} = 5.19$. The double-counting correction is slightly lower than the corresponding value for Mn-ions in CdTe, which is in accordance with the lower value of $n_{\mathrm{GGA,GaAs}} = 5.59$, as compared to $n_{\mathrm{GGA,CdTe}} = 5.77$, see Section 4.12.

As seen from the results shown in Fig. 6.3 the material becomes half-metallic, i.e. only one spin-channel has states within the Fermi energy. The $d$-bands sink into the valence states and push the hybridized valence $p$-orbitals into the Fermi energy. The gap in the majority bands closes whereas the minority bands show a gap of about 1.5 eV. The $d$-derived structures in the DOS are positioned at about $E - E_{\mathrm{F}} = -3.8\,\mathrm{eV}$ in the valence-bands, and at about $E - E_{\mathrm{F}} = 2\,\mathrm{eV}$ in the conduction bands. The valence-band position is

| atom | type | $n_{\mathrm{GGA}}$ | $n^{\downarrow}_{\mathrm{GGA+G}}$ | $n^{\uparrow}_{\mathrm{GGA+G}}$ | $m$ |
|---|---|---|---|---|---|
| Mn | $sp^3$ | 1.737 | 0.999 | 0.912 | 0.087 |
| Mn | $t_{2g}$ | 2.255 | 2.762 | 0.419 | 2.343 |
| Mn | $e_g$ | 3.333 | 1.948 | 0.063 | 1.885 |
| Ga | $sp^3$ | 9.448 | 4.765 | 4.623 | 0.142 |
| As | $sp^3$ | 19.226 | 9.525 | 9.983 | -0.458 |
| $\sum$ | | 36.000 | 20.000 | 16.000 | 4.000 |

**Table 6.1:** Occupation numbers for $Ga_3MnAs_4$ obtained from the paramagnetic GGA, $n_{\mathrm{GGA}}$, and the spin-polarized GGA+Gutzwiller method, $n^{\uparrow,\downarrow}_{\mathrm{GGA+G}}$. Note that the Ga atoms and As atoms can be mapped by symmetry transformation of the crystal onto each other. The table shows the total charge for equivalent atoms.

in good agreement with results obtained from photo-emission experiments [106, 107, 96, 108].

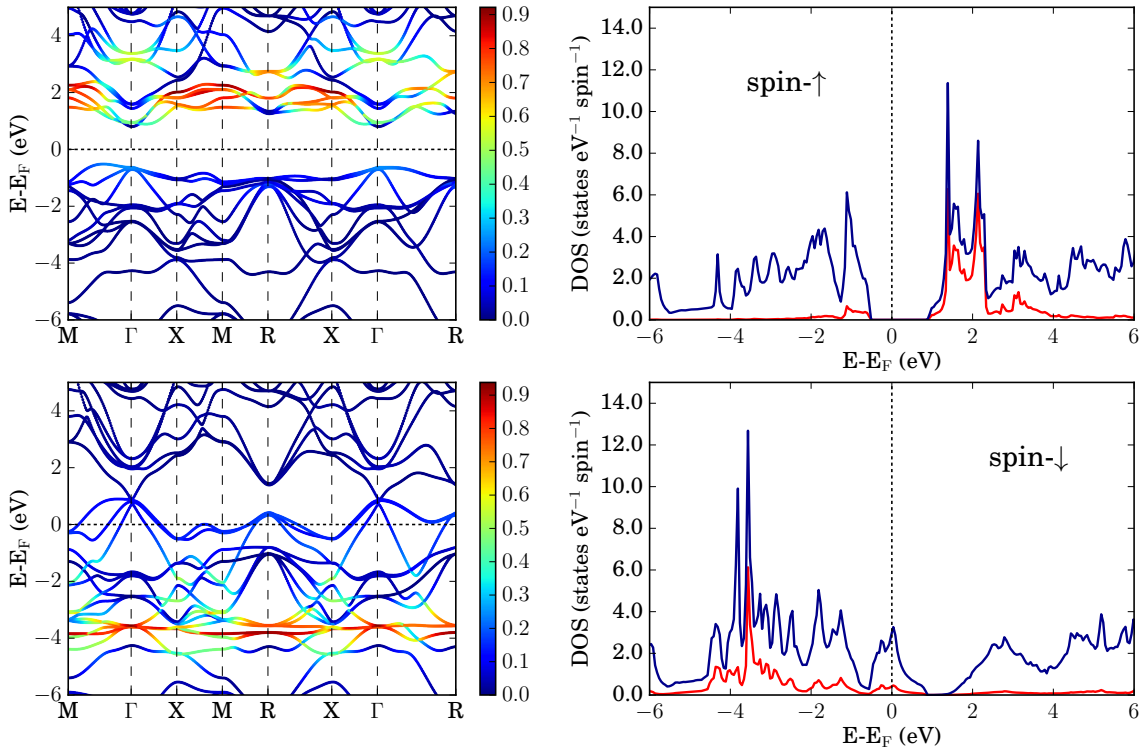


**Figure 6.3:** Band structure and DOS of $Ga_3MnAs_4$ as obtained from a spin-polarized GGA+Gutzwiller calculation using the on-site Coulomb parameters $A = 4.4\,\mathrm{eV}$, $B = 0.1\,\mathrm{eV}$, $C = 0.4\,\mathrm{eV}$ and a double-counting correction of $\tau = 22.18\,\mathrm{eV}$. The top figures show the minority spin-channel, the bottom figures the majority spin-channel. The band structures (left) are color coded, showing the amount of Mn $3d$-character in the bands. In the DOS plots (right), the blue lines indicate the total DOS, the red lines indicate the DOS projected onto the Mn $3d$-states.

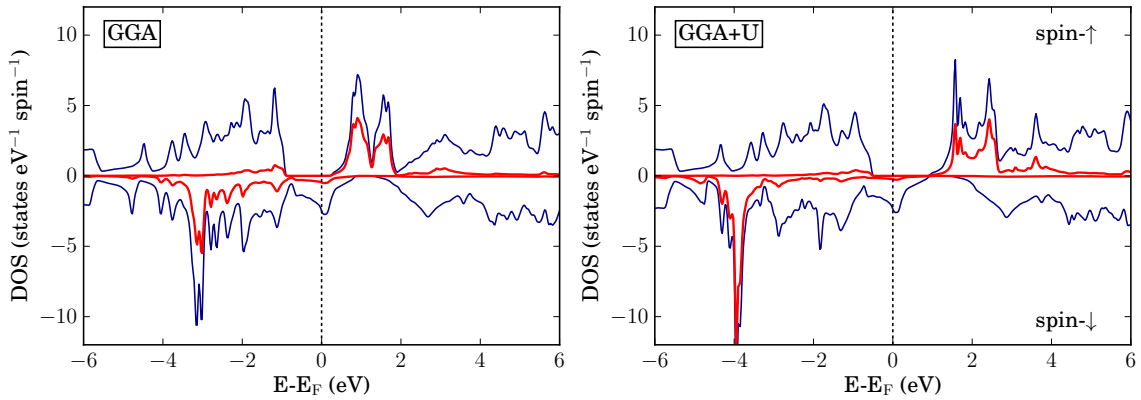


**Figure 6.4:** DOS and projected DOS for a spin-polarized GGA (left) and a spin-polarized GGA+U calculation (right) of $Ga_3MnAs_4$. For the GGA+U calculation $U = 3.25\,\text{eV}$ and $J = 0.60\,\text{eV}$ is used. The blue lines indicate the total DOS, the red-lines the DOS projected onto the Mn $3d$-states.

We show the occupations $n^{\uparrow,\downarrow}_{\text{GGA+G}}$ of the Wannier orbitals in Table 6.1. Since the total number of electrons is fixed, and one spin-channel is gapped, the magnetic moment over the whole unit cell is integer. The impurity introduces a magnetic moment of $4\,\mu_B$ in the unit cell. The local moment of the $d$-shell is $4.2\,\mu_B$, inducing an opposing moment on the As sites. We compare our results with a spin-polarized GGA and a spin-polarized GGA+U calculation where we have chosen $U = 3.25\,\text{eV}$ and $J = 0.60\,\text{eV}$, see Figure 6.4. Our results for the electronic structure are in good agreement with previously reported DFT results, see for example [102, 103, 104, 105]. The GGA+U DOS match the Gutzwiller DOS remarkably well.

## 6.4 Density and spin distributions

We now focus on Gutzwiller-specific results that address the question of the many-body state of the $d$-shell on the impurity. We show the probability distribution to find a many-body state in spin-state $s$, as well as the probability distribution to find a many-body state with $n$ electrons in Fig. 6.5. For the calculations we use a supercell with 216 atoms, in which one impurity is replaced by a Mn atom. Compared to probabilities for Mn in the wide-gap semiconductor CdTe, Fig. 5.8, the distribution is broader and includes a significant amount of states with $s = 3/2$ and $s = 1$ states. The states with $s = 5/2$ and $s = 2$ states are high-spin (hs) states of $n = 5$ electrons and $n = 4, 6$ electrons, respectively. The $n = 5$ Hund's rule state has a probability of $p_{\text{GW}}(|n = 5, s = 5/2\rangle_{\text{hs}}) = 49.5\%$, whereas we find $p_{\text{GW}}(|n = 6, s = 2\rangle_{\text{hs}}) = 23.2\%$ and $p_{\text{GW}}(|n = 4, s = 2\rangle_{\text{hs}}) = 8.4\%$ for the high-spin states with $s = 2$. These probabilities add up to $p_{\text{hs}} = 81\,\%$. In comparison we find $p_{\text{hs}} = 98\,\%$ for Mn doped CdTe. These results clearly show that there is a different physical situation present in III-V semiconductors, attributed to the itinerant character of the electrons. For the expectation value of the spin we get $s_{\text{GW}} = 2.13$. The value derived from experiment is $s_{\text{exp}} = 2.2$ [109]. As seen from Fig. 6.5, the Gutzwiller probabilities differ noticeable

from the Hartree-Fock results. The Mn sites display stronger correlations in GaAs than in CdTe. Nevertheless, our calculation yields a minor renormalization of the hopping matrix elements, with renormalization factors of $q \approx 0.95$. For this reason, the correlation effects on the band structure are only minor.

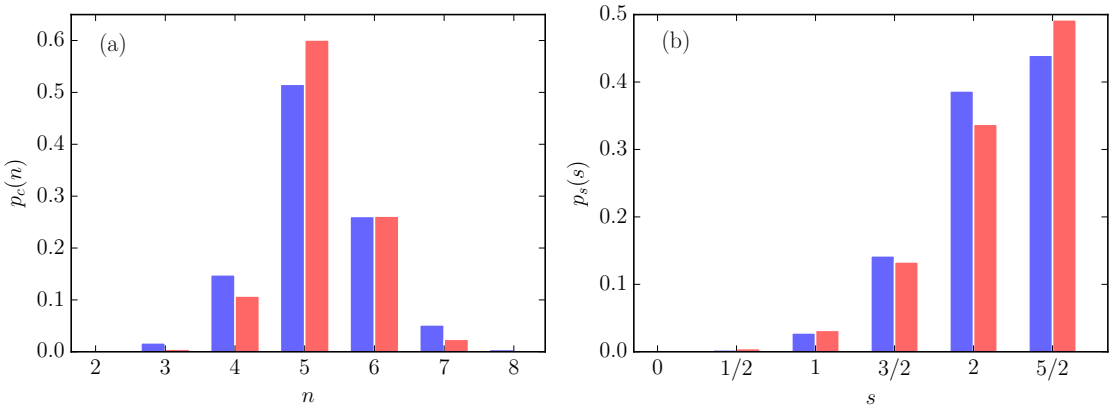


**Figure 6.5:** Mn charge distribution $p_c(n)$ as a function of the number $n$ of $3d$-electrons and Mn spin distribution $p_s(s)$ as a function of the spin $s$ of $3d$-electrons for $A = 4.4\,\text{eV}$, $B = 0.1\,\text{eV}$, $C = 0.4\,\text{eV}$ for $n_d = 5.19$ in GGA+Gutzwiller (red columns), in comparison with the Hartree–Fock result (blue columns).

## 6.5 Ferromagnetic correlations

The results from the previous subsections encourage us to study larger unit cells in order to investigate the correlation between ion pairs, as done in Chapter 5. For convenience we continue to use the notion of an "exchange coupling" $J^{\text{GaAs}}$ that we define as in Section 5.1.2. We note, however, that the strong hybridization and the itinerant nature of the magnetism puts the underlying Heisenberg-spin picture into question.

The following results were obtained using a cubic 512-atom supercell (the (400/040/004) replica of the eight atomic cell show in Fig. 5.1) with two Mn impurities placed in a 1st,2nd,...,8th and 28th neighbor configuration. If we assume that the corners of the cubic unit cells are occupied by one of the two impurities (the corners are occupied by equivalent atoms due to translational invariance) then we have to take into account that in the 28th neighbor configuration the second impurity occupies the center of the unit cell. In the 8th neighbor configuration the second impurity is located at the center of an edge. We therefore "correct" the obtained couplings $J_{28}$ and $J_8$ by a factor of $1/8$ and $1/2$, respectively. All other couplings are not scaled. We denote the scaled exchange couplings by $\tilde{J}_n$. The 28th neighbor configuration serves as test to check for long-ranged correlations. Since our results show non-vanishing couplings even for 28th-neighbor pairs, the results for $\tilde{J}_n$ do only approximatively match the true values $J_n$.

Our tight-binding Hamiltonian has a basis of 4116 orbitals and a total of 2056 electrons per unit cell. Since the material is not insulating we depend on an accurate sampling of the

Fermi-surface in order to resolve small energy differences correctly. We use a mesh with 2560 **k**-points in the full Brillouin zone. We note in passing that the employed supercell sizes and dense **k**-meshes rule out a pure DFT (FLAPW) calculation. It is an accepted theoretical result that the effective exchange couplings are concentration dependent. The impurity concentration in our calculation is given by $x = 2/254 \approx 0.8\,\%$. To not further complicate our studies, we restrict our calculations to the simple two-ion case described above, even though, at least in principle, our method does allow the study of additional impurities.

To investigate the influence of the on-site Coulomb interaction, we repeat the calculation for different Racah parameters $A$, $B$ and $C = 4B$. In each calculation we adjust the double-counting correction so that the number of electrons in the $d$-shell is fixed. As seen from Fig. 6.6, we obtain a ferromagnetic exchange. Our values show a fluctuation in the strength of the coupling constants but we do not observe antiferromagnetic exchange couplings, as expected from a pure RKKY mechanism. Even for the 28th neighbor we observe a sizeable value for the exchange coupling. Changing the Racah parameter $A$ has only a minor influence on the coupling strength. The adjusted double-counting potential that fixes the number of total electrons in the Mn $d$-shell seems to partly compensate the effect of the increased values $U(c, c')$ in the atomic Hamiltonian (4.8) on a mean-field level. Our calculations show that $B$ and $C$ alter the magnetic moment at the impurity sites. Because our double-counting correction scheme is not spin-dependent there is no counter-term that compensates for the induced magnetization in the Gutzwiller treatment. For $B = 0.10\,\text{eV}$, the local magnetic moment $m_\text{d}$ is about $0.1\,\mu_\text{B}$ larger than for $B = 0.08\,\text{eV}$, independent of $A$. The close proximity of the curves which differ only in the parameter $A$ is a consequence of the fixed particle constraint.

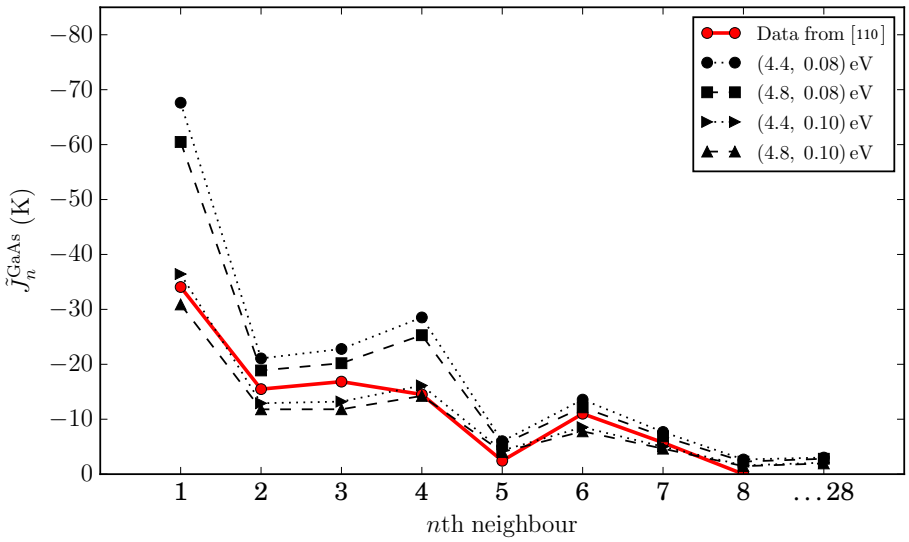


**Figure 6.6:** Scaled exchange couplings $\tilde{J}_n$ (see text) in Mn doped GaAs for different values of the Racah parameters $A$ and $C = 4B$. The double-counting is adjusted to result in an occupation of 5.19 electrons in the $d$-shell. The results were obtained for a 512-atom Gutzwiller-DFT calculation, corresponding to a doping of $x \approx 0.8\,\%$. Red lines: Results from [110] for a Mn concentration of $x = 1\%$.

In Fig. 6.6 we also depict results from a Korringa-Kohn-Rostoker coherent-potential approximation (KKR-CPA) calculation [110] utilizing a formula by Lichtenstein et al. [111] to extract $J_{ij}$ for a Mn concentration of $x = 1\%$. The results are in good qualitative and even quantitative agreement with our results for $A = 4.4\,\mathrm{eV}$, $B = 0.1\,\mathrm{eV}$.

In the present calculations each Mn atom introduces one (weakly bound) electron hole. The exchange strength is strongly dependent on the hole doping concentration [94], which is of course directly related to the Mn concentration. An additional effect of higher Mn-doping concentrations originates from disorder; in some semi-phenomenological models this is attributed for by potential scattering terms [112, 113]. Using our method for a systematic study of multiple Mn impurities, e.g. by means of configuration averaging of various impurity distributions, is numerically feasible but not within the scope of this thesis. The Gutzwiller-specific results indicate a slight enhancement of the local expectation value of the spin compared to Hartree-Fock results but the overall ab-initio picture of $Ga_{1-x}Mn_xAs$ remains unchanged.

# 7 Conclusion and outlook

In this work we combined conventional band-theory with the Gutzwiller variational method to investigate the magnetic properties of manganese based DMS in their ground-state. The basis for our calculations is a tight-binding model obtained from downfolding a paramagnetic DFT band structure. We added a local interaction Hamiltonian to the tight-binding model that accounts for the correlated $d$-orbitals of the impurities. The ground state of the final Hubbard Hamiltonian was studied using the Gutzwiller variational method.

The description of the local correlations using the full interaction Hamiltonian (4.8) is clearly beyond the mean-field approach employed in the DFT+U formalism (2.71). Unlike in the DFT+U method, we did not use a fully self-consistent scheme in which a possible charge flow from the Gutzwiller many-body treatment is fed back into the DFT calculation until self-consistency is achieved. Fully self-consistent Gutzwiller-DFT calculations are a recent development [114, 115, 116, 117]; their use for supercell calculations with several hundred atoms, including multiple magnetic impurities, is out of our current abilities. Without charge self-consistency, our method belongs to the so-called "one-shot" methods. These methods are based on the assumption that the DFT calculation yields a suitable (paramagnetic) single-particle description as a starting-point for a subsequent many-particle (Gutzwiller) treatment. A rigorous test of this assumption is not possible. However, our calculations for Mn doped II-VI and III-V semiconductors show consistent results, as seen in Chapter 5 and Chapter 6.

Although our method is based on ab-initio results we still have to provide the Racah parameters $A$, $B$, $C$ and a double-counting parameter $\tau$. A certain loss of the "ab-inito" character seems inevitable when combining DFT with model-based calculations. The ratio between $B$ and $C$ is assumed to be fixed, leaving us with three adjustable parameters. In our calculations we avoided to connect the double-counting correction potential $\tau$ to the occupation number of the paramagnetic DFT result and the Racah parameters, as it is done in standard double-counting correction schemes. In this way the influence of the Coulomb-interaction strength was studied in dependence of the filling of the $d$-shell.

In our studies of Mn-doped CdTe we found that calculations with a certain filling of the $d$-shell reproduce the nearest-neighbor exchange couplings derived from experiment. The precise value of the Racah parameters plays a minor role. Using these values we calculated coupling constants for other Mn-doped II-VI DMS that turn out to be in reasonable agreement with experiment. We have shown that a DFT+U treatment of Mn-doped II-VI semiconductors is sufficient to account for the Coulomb interaction in the $d$-shell, since the "correct" $s = 5/2$ ground-state is selected by huge exchange-fields on a single-particle level.

We do not observe a hopping renormalization in the Gutzwiller calculations of Mn-doped II-VI semiconductors. The expectation values for the local multiplet states are

equal to Hartree-Fock results. The DFT (GGA) systematically overestimates exchange couplings in the wide-gap II-VI semiconductors because of the underestimation of the charge-transfer gap. Our calculations suggest that this effect is cured by the +U treatment. Consequently, a comparison of GGA+U with GGA+Gutzwiller band structures shows a good agreement.

For $Ga_{1-x}Mn_xAs$ our calculations yield minor renormalization factors of $q \approx 0.95$. The local occupation and spin-probabilities show slight deviations from the Hartree-Fock limit. This is a hint for correlation effects that can only be described on a many-body level. However, we do not expect the small effects to alter significantly results of existing DFT+U calculations for $Ga_{1-x}Mn_xAs$. We expect that the correlation effects in ferromagnetic semiconductors are more pronounced when they are doped with $3d$-transition metal ions other than Mn, as indicated by exploratory studies we performed for Cr doped CdTe. Correlated electron approaches are advisable for the study of these materials.

The applicability of $s = 5/2$ spin-models for large gap Mn doped II-VI semiconductors was shown using a toy model. The $s = 5/2$ spin is not located on the Mn $3d$ impurity site but resides in its vicinity due to hybridization effects. The character of the magnetism in the ferromagnetic DMS $Ga_{1-x}Mn_xAs$ shows a different picture. The Gutzwiller calculation yields an expectation value of $s = 2.1$ for the local spin in the itinerant system, in close agreement with experiment.

As a final remark we comment on a possible treatment of spin-orbit coupling that gives rise to magnetic anisotropy effects. The effect can be accounted for on the band-structure level by employing scalar-relativistic DFT calculations. In recent publications [63, 118] the Gutzwiller method was extended to account for non-diagonal entries in the variational matrix $\lambda_{\Gamma,\Gamma'}$ which is a prerequisite for the inclusion of spin-orbit coupling within the Gutzwiller calculation. Strain and hole-dopant concentration dependent magnetic anisotropy is experimentally observed in Ga(Mn)As [119].

In 1931 Pauli made the famous statement: “One shouldn’t work on semiconductors, that is a filthy mess; who knows whether any semiconductors exist.” [120]. The pessimistic statement has to be seen in the historical context. The upcoming advances in theory and experiment that allowed to create the first tailored semiconductors were hard to anticipate. Today, the situation for diluted magnetic semiconductors seems to be comparable. Spin properties already found application in data storage (via the effect of giant magnetoresistance) but until today a full inclusion of ferromagnetic DMS in existing semiconductor heterostructure keeps to be a future vision. The search for suitable ferromagnetic room-temperature DMS is an interesting challenge that will continue to be an important field of research in which theoretical investigations based on correlated-electron approaches remain to be of great value.

# A Appendix

## A.1 Coulomb matrix elements for d-orbitals in a cubic environment

The Coulomb matrix elements are derived in [70].

$$\begin{aligned}
U(u,u) = U(v,v) &= U(u,v) + 2J(u,v)\ , \\
U(\xi,u) = U(\eta,u) &= (U(\zeta,u) + 3U(\zeta,v))/4\ , \\
U(\xi,v) = U(\eta,v) &= (3U(\zeta,u) + U(\zeta,v))/4\ , \\
J(\xi,u) = J(\eta,u) &= (J(\zeta,u) + 3J(\zeta,v))/4\ , \\
J(\xi,v) = J(\eta,v) &= (3J(\zeta,u) + J(\zeta,v))/4\ , \\
T(\eta;u,v) = -T(\xi;u,v) &= \sqrt{3}(U(\zeta,u) - U(\zeta,v))/4\ , \\
A(\eta;u,v) = -A(\xi;u,v) &= \sqrt{3}(J(\zeta,u) - J(\zeta,v))/4\ , \\
S(\xi,\eta;\zeta,u) &= S(\eta,\xi;\zeta,u)\ , \\
S(\zeta,\xi;\eta,u) &= -2S(\eta,\xi;\zeta,u)\ , \\
S(\xi,\eta;\zeta,v) &= -\sqrt{3}S(\eta,\xi;\zeta,u)\ , \\
S(\zeta,\xi;\eta,u) &= \sqrt{3}S(\eta,\xi;\zeta,u)\ .
\end{aligned} \tag{A.1}$$

If we employ the "spherical approximation" (the same radial part for the $t_{2g}$-orbitals and for the $e_g$-orbitals) we may express all matrix elements using the three Racah parameters $A$, $B$, and $C$ that are related to the Slater-Condon parameters by $A = F^{(0)} - F^{(4)}/9$, $B = (F^{(2)} - 5F^{(4)}/9)/49$, and $C = 5F^{(4)}/63$.

$$\begin{aligned}
U(u,v) &= A - 4B + C\ , \\
J(u,v) &= 4B + C\ , \\
U(\zeta,\zeta) &= A + 4B + 3C\ , \\
U(\xi,\eta) &= A - 2B + C\ , \\
J(\xi,\eta) &= 3B + C\ , \\
U(\zeta,u) &= A - 4B + C\ , \\
U(\zeta,v) &= A + 4B + C\ , \\
J(\zeta,v) &= C\ , \\
J(\zeta,u) &= 4B + C\ , \\
S(\eta,\xi;\zeta,u) &= -\sqrt{3}B\ .
\end{aligned} \tag{A.2}$$

## A.2 Symmetry conservation for projected Wannier functions

We show that the Wannier functions $|W_{\mathbf{R},n}\rangle$ obtained by the projection method, see Section 2.6.2, transform under the same symmetry operations as the trial orbitals $|g_n\rangle$ from which they are constructed. We have

$$|\Phi'_{\mathbf{k},n}\rangle = \sum_m |\Phi_{\mathbf{k},m}\rangle\langle\Phi_{\mathbf{k},m}|g_n\rangle = \sum_m A^{(\mathbf{k})}_{m,n}|\Phi_{\mathbf{k},m}\rangle \,, \tag{A.3a}$$

$$|\tilde{\Phi}_{\mathbf{k},n}\rangle = \sum_m \left((S'^{(\mathbf{k})})^{-1/2}\right)_{mn} |\Phi'_{k,m}\rangle \,, \tag{A.3b}$$

$$|W_{\mathbf{R},n}\rangle = \sqrt{\frac{1}{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{R}}|\tilde{\Phi}_{\mathbf{k},n}\rangle \,, \tag{A.3c}$$

$$\text{with } S'^{(\mathbf{k})}_{m,n} = \langle\Phi'_{\mathbf{k},m}|\Phi'_{\mathbf{k},n}\rangle \,. \tag{A.3d}$$

The proof is very similar to the proof given by Slater and Koster in the context of the method of Linear Combination Of Atomic Orbitals (LCAO) [121].

Suppose the wave functions $|g_n\rangle$ transform under some symmetry operation *of the crystal* $\hat{O}$ given by

$$\hat{O}|g_n\rangle = \sum_l \Gamma_{l,n}(\hat{O})|g_l\rangle \,, \tag{A.4}$$

where $\mathbf{\Gamma}(\hat{O})$ is a unitary matrix. It directly follows that $|\Phi'_{\mathbf{k},n}\rangle$ (A.3a) transforms under the same operation

$$\hat{O}|\Phi'_{\mathbf{k},n}\rangle = \sum_l \Gamma_{l,n}(\hat{O})|\Phi'_{\mathbf{k},l}\rangle \,. \tag{A.5}$$

We apply $\hat{O}$ to $|\tilde{\Phi}_{\mathbf{k},n}\rangle$ (A.3b) and obtain

$$\begin{aligned} \hat{O}|\tilde{\Phi}_{\mathbf{k},n}\rangle &= \sum_m ((S'^{(\mathbf{k})})^{-1/2})_{m,n}\hat{O}|\Phi'_{\mathbf{k},m}\rangle \\ &= \sum_{l,m} ((S'^{(\mathbf{k})})^{-1/2})_{m,n}\Gamma_{l,m}(\hat{O})|\Phi'_{\mathbf{k},l}\rangle \,. \end{aligned} \tag{A.6}$$

Our goal is to proof that $|\tilde{\Phi}_{\mathbf{k},n}\rangle$ (and therefore $|W_{\mathbf{R},n}\rangle$) transforms according to

$$\hat{O}|\tilde{\Phi}_{\mathbf{k},n}\rangle = \sum_l \Gamma_{l,n}(\hat{O})|\tilde{\Phi}_{\mathbf{k},l}\rangle \,. \tag{A.7}$$

Using (A.6) we rewrite (A.7) as

$$\sum_l \Gamma_{l,n}(\hat{O})|\tilde{\Phi}_{\mathbf{k},l}\rangle = \sum_{l,m} ((S'^{(\mathbf{k})})^{-1/2})_{m,n}\Gamma_{l,m}(\hat{O})|\Phi'_{\mathbf{k},l}\rangle \tag{A.8}$$

$$\Leftrightarrow \sum_{l,m} \Gamma_{l,n}(\hat{O})((S'^{(\mathbf{k})})^{-1/2})_{m,l}|\Phi'_{\mathbf{k}m}\rangle = \sum_{l,m} ((S'^{(\mathbf{k})})^{-1/2})_{m,n}\Gamma_{l,m}(\hat{O})|\Phi'_{\mathbf{k},l}\rangle \,. \tag{A.9}$$

To proof (A.9) it is sufficient to show that

$$\sum_m ((S'^{(\mathbf{k})})^{-1/2})_{l,m}\Gamma(\hat{O})_{m,n} = \sum_m ((S'^{(\mathbf{k})})^{-1/2})_{m,n}\Gamma(\hat{O})_{l,m} \ , \tag{A.10}$$

which can be written more conveniently using matrix notation as

$$\begin{aligned} &(\mathbf{S}'^{(\mathbf{k})})^{-1/2}\mathbf{\Gamma}(\hat{O}) = \mathbf{\Gamma}(\hat{O})(\mathbf{S}'^{(\mathbf{k})})^{-1/2} \\ \Leftrightarrow &\mathbf{\Gamma}^*(\hat{O})(\mathbf{S}'^{(\mathbf{k})})^{-1/2}\mathbf{\Gamma}(\hat{O}) = (\mathbf{S}'^{(\mathbf{k})})^{-1/2} \ . \end{aligned} \tag{A.11}$$

The only thing left to show is that the matrix $(\mathbf{S}^{(\mathbf{k})})^{-1/2}$ is invariant under a unitary transformation given by $\mathbf{\Gamma}(\hat{O})$. We require that the matrix $(\mathbf{S}^{(\mathbf{k})})^{-1/2}$ can be expanded according to

$$(\mathbf{S}'^{(\mathbf{k})})^{-1/2} = (\mathbb{1} + \mathbf{M}^{(\mathbf{k})})^{-1/2} = \mathbb{1} - \frac{1}{2}\mathbf{M}^{(\mathbf{k})} + \frac{3}{8}(\mathbf{M}^{(\mathbf{k})})^2 - \dots \ . \tag{A.12}$$

Therefore $(\mathbf{S}'^{(\mathbf{k})})^{-1/2}$ is invariant under the unitary transformation if $\mathbf{S}'^{(\mathbf{k})}$ is invariant, which is readily shown because

$$\begin{aligned} \sum_{m,n} \Gamma_{m,s}(\hat{O})^* S'^{(\mathbf{k})}_{m,n}\Gamma_{n,l}(\hat{O}) &= \sum_{m,n} \langle \Phi'_{\mathbf{k},m}|\Gamma^*_{m,s}(\hat{O})\Gamma_{n,l}(\hat{O})|\Phi'_{\mathbf{k},n}\rangle \\ &= \sum_m \langle \Phi'_{\mathbf{k},m}|\Gamma^*_{m,s}(\hat{O}) \sum_n \Gamma_{n,l}(\hat{O})|\Phi'_{\mathbf{k},n}\rangle \\ &= \langle \Phi'_{\mathbf{k},s}|\hat{O}^\dagger\hat{O}|\Phi'_{\mathbf{k},l}\rangle = S'^{(\mathbf{k})}_{s,l} \ . \end{aligned} \tag{A.13}$$

## A.3 Gutzwiller minimization: two-site model

As an application of the general formulation of the Gutzwiller method, we investigate the ground-state energy of the two-site model,

$$\hat{H} = t \sum_{s\in\{\uparrow,\downarrow\}} \left(\hat{c}^\dagger_{1,s}\hat{c}_{2,s} + \hat{c}^\dagger_{2,s}\hat{c}_{1,s}\right) + U\sum_i \hat{n}_{i,\uparrow}\hat{n}_{i,\downarrow} \tag{A.14}$$

for the case of half filling. We follow [59] and add some additional remarks on the asymptotic behavior of the ground-state energy in the following.

For the minimization of the Gutzwiller energy functional with respect to $|\Phi_0\rangle$ we choose the ground-state of the effective Hamiltonian

$$\hat{H}^0_{\text{eff}} = t\sum_s \begin{pmatrix} \hat{c}^\dagger_{1,s} \\ \hat{c}^\dagger_{2,s} \end{pmatrix}^T \begin{pmatrix} +\eta_s & t \\ t & -\eta_s \end{pmatrix} \begin{pmatrix} \hat{c}_{1,s} \\ \hat{c}_{2,s} \end{pmatrix} \ . \tag{A.15}$$

The field $\eta_s$ breaks the symmetry between the spin states. We assume that the antiferromagnetic ground-state is lower in energy and set

$$\eta_\uparrow = -\eta_\downarrow =: \eta \ . \tag{A.16}$$

The ground-state of (3.52) is given by

$$|\Phi_0\rangle = \hat{h}_1^\dagger \hat{h}_2^\dagger |0\rangle \quad \text{with} \quad \hat{h}_{1(2)}^\dagger = \cos(\theta)\hat{c}_{1,\uparrow(2,\downarrow)}^\dagger + \sin(\theta)\hat{c}_{2,\uparrow(1,\downarrow)}^\dagger \,, \tag{A.17}$$

where

$$\tan(\theta) = \frac{\eta - \sqrt{\eta^2 + t^2}}{t} \,. \tag{A.18}$$

Due to the symmetries of the Hamiltonian the occupations $n_{b,s}^0 = \langle \hat{c}_{b,s}^\dagger \hat{c}_{b,s} \rangle_{\Phi_0}$ fulfil the relation

$$n_{1,\uparrow}^0 = 1 - n_{1,\downarrow}^0 = n_{2,\downarrow}^0 = 1 - n_{2,\uparrow}^0 \,. \tag{A.19}$$

We define the magnetization per site as

$$\Delta = n_{1,\uparrow}^0 - n_{1,\downarrow}^0 = \cos^2(\Theta) - \sin^2(\Theta) \,, \tag{A.20}$$

and proceed by expressing the constraints (3.32) and (3.33) using the variational parameter $\lambda_{\{d,s,\varnothing\}}$

$$1 = \lambda_\varnothing^2 m_\varnothing^0 + \sum_s \lambda_s^2 m_s^0 + \lambda_d^2 m_d^0 \,, \tag{A.21}$$

$$n_s^0 = \lambda_d^2 m_d^0 + \lambda_s^2 m_s^0 \,, \tag{A.22}$$

where we used the same notation for the expectation values as in (3.18). Equation (3.35) allows us to express the correlated expectation values using the uncorrelated values,

$$m_d = \langle \hat{m}_d \rangle_{\Psi_G} = \lambda_d^2 \hat{n}_\uparrow^0 \hat{n}_\downarrow^0 \,, \tag{A.23}$$

$$m_s = \langle \hat{m}_s \rangle_{\Psi_G} = \lambda_s^2 \hat{n}_s^0 (1 - \hat{n}_s^0) \,, \tag{A.24}$$

$$m_\varnothing = \langle \hat{m}_\varnothing \rangle_{\Psi_G} = \lambda_\varnothing^2 (1 - \hat{n}_\uparrow^0)(1 - \hat{n}_\uparrow^0) \,. \tag{A.25}$$

With this relations the constraints read

$$1 = m_d + \sum_s m_s + m_\varnothing \,, \tag{A.26}$$

$$n_s^0 = m_d + m_s = n_s \,. \tag{A.27}$$

Using equation (3.39) the $q$-matrix for the simple model is diagonal and given by

$$\begin{aligned} q_s := q_s^s &= \frac{1}{n_s^0}(\lambda_\varnothing \lambda_s m_s^0 + \lambda_d \lambda_{\bar{s}} m_d^0) \\ &= \frac{1}{\sqrt{n_s^0(1 - n_s^0)}} \left( \sqrt{m_\varnothing m_s} + \sqrt{m_d m_{\bar{s}}} \right) \,. \end{aligned} \tag{A.28}$$

Using the constraints (A.27) and the relation (A.19) we find

$$q = q_{(1,\uparrow)} q_{(2,\downarrow)} = \frac{2m_d \left( \sqrt{1 - 2m_d + \Delta(\eta)} + \sqrt{1 - 2m_d - \Delta(\eta)} \right)^2}{1 - \Delta^2(\eta)} \,. \tag{A.29}$$

The energy functional (3.40) which has to be minimized with respect to the expectation value of the double occupation and the field $\eta$ reads

$$E_G^{\mathrm{GW}}(\eta, d) = 4tq(\eta, d)\sin(\Theta)\cos(\Theta) + 2Um_d \ . \tag{A.30}$$

From (A.30) we can easily derive the non-magnetic Gutzwiller functional, as well as the magnetic and non-magnetic Hartree-Fock solutions. By setting $\eta = 0$ we get the paramagnetic solution

$$E_G^{\mathrm{GW,\,para}}(\eta, d) = -2tq(\eta = 0, d) + 2Um_d \ . \tag{A.31}$$

Setting $q = 1$ (uncorrelated limit) and using $m_d = m_d^0 = n_{1,\uparrow}^0 n_{1,\downarrow}^0$ we arrive at the Hartree-Fock solutions

$$E_G^{\mathrm{HF}}(\eta, d) = 4t\sin(\Theta)\cos(\Theta) + 2U\sin^2(\Theta)\cos^2(\Theta) \tag{A.32}$$

and

$$E_G^{\mathrm{HF,\,para}}(\eta, d) = -2t + \frac{U}{2} \ . \tag{A.33}$$

It is important to note that the Gutzwiller wave function was developed to study the effect of local correlations in infinite, translational invariant systems such as the one-band Hubbard model. The complexity of the formalism presented in Section 3.3.3 stems from the complexity of combining the physics of delocalized electrons with the effect of local correlations. In the case presented here, we apply the Gutzwiller method to a rather simple model which can be easily solved exactly. However, the comparison between the Gutzwiller, the Hartree-Fock, and the exact results lead to interesting insights, as discussed in the following.

The two-site model and its exact ground state in (A.14) possesses $SU(2)$ symmetry which is broken for $\eta \neq 0$ in the effective model (3.52). The exact ground state with $S = 0$ is given by

$$|\Psi_G^{\mathrm{exact}}\rangle = 1/\sqrt{2 + 2\chi^2}\,(|\uparrow;\downarrow\rangle - |\uparrow;\downarrow\rangle + \chi|\uparrow\downarrow;\circ\rangle + \chi|\circ;\uparrow\downarrow\rangle) \ , \tag{A.34}$$

$$E_G^{\mathrm{exact}} = -2t\chi = \frac{1}{2}\left(U - \sqrt{U^2 + 16t^2}\right) \overset{U \gg t}{\sim} -\frac{4t^2}{U} \ . \tag{A.35}$$

We minimize the energy functionals (A.30) and (A.32) numerically and compare them to the exact ground-state energy, see Fig. A.1. In the intermediate regime of the Coulomb interaction strength ($U/t$ = 2-3), the improvement of the ground-state energy of the Gutzwiller method over the Hartree-Fock result is most significant. However, both energies differ from the exact result. At $U/t = 2$ a symmetry-broken wave function becomes energetically favourable in the Hartree-Fock method. It is well known that the Hartree-Fock theory tends to break symmetries to avoid energetically unfavourable double occupations. The Gutzwiller method predicts this transition at $U/t \approx 3.3$, which is completely absent in the exact solution. If we enforce $\eta = 0$ to avoid this transition, the Hartree-Fock ground-state energy (A.33) turns out to be not bounded from above, whereas the

Gutzwiller result (A.31) approaches the correct value of zero for $U \to \infty$. In Fig. A.1 we compare the double occupation $m_d$ and the hopping renormalization $q$ for the magnetic and non-magnetic Gutzwiller solutions. We observe that the renormalization factor $q$ in the magnetic Gutzwiller solution approaches the uncorrelated limit of $q = 1$ in the high-$U$ limit, reproducing the Hartree-Fock solution.

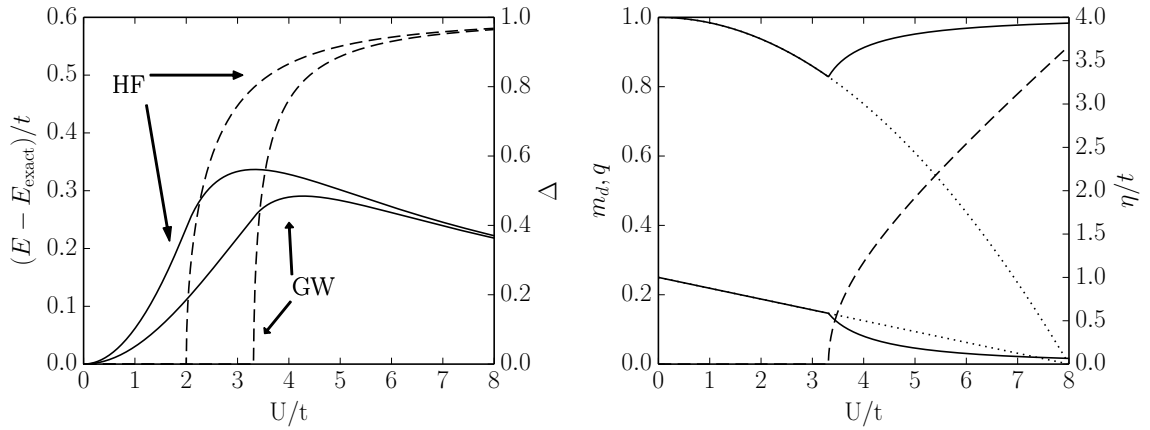


**Figure A.1:** Left: Difference in energy of the exact solution and the Gutzwiller (GW) and Hartree-Fock (HF) solution of the two-site Hubbard model in units of the hopping $t$ (solid lines). The magnetization $\Delta$ for the Gutzwiller/Hartree-Fock ground-state (dashed lines). After [59]. Right: Hopping renormalization factor $q$ (solid line, starting at unity for $U = 0$), double occupation $m_d$ (solid line, starting at $1/4$ for $U = 0$) and field $\eta$ (dashed line) in units of the hopping $t$ for the two-site Hubbard model in the magnetic Gutzwiller calculation. Dotted lines show results of the non-magnetic Gutzwiller solution.

Both the Hartree-Fock and the Gutzwiller solution do not capture the correct ground-state physics since the wave functions do not possess $SU(2)$ symmetry. The consequences can also been seen in the asymptotic behavior of the ground-state energies

$$|\Psi_G^{\text{GW,HF}}\rangle \overset{U \gg t}{=} -\frac{2t^2}{U} \neq -\frac{4t^2}{U} , \tag{A.36}$$

which differ by a factor of two from the exact result. This mismatch can be readily understood if we map the low-energy states of the two-site Hubbard model in the high-$U$ limit to an effective spin model with Heisenberg-type interaction

$$\hat{H}_{\text{spin}} = J \left( \hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 - \frac{\hat{n}_1 \hat{n}_2}{4} \right) , \tag{A.37}$$

where $\hat{\mathbf{S}}_{1,2}$ are $s = 1/2$ spin operators acting on both sites. This model obviously possesses $SU(2)$ symmetry. The exchange parameter is found to be $J = 4t^2/U$, which is the energy gap between the ground-state and the triplet states

$$\begin{aligned} |\Psi_{t-1}^{\text{exact}}\rangle &= |\uparrow;\uparrow\rangle , \\ |\Psi_{t0}^{\text{exact}}\rangle &= \frac{1}{\sqrt{2}} \left( |\uparrow;\downarrow\rangle + |\downarrow;\uparrow\rangle \right) , \\ |\Psi_{t+1}^{\text{exact}}\rangle &= |\downarrow;\downarrow\rangle ; \end{aligned} \tag{A.38}$$

with $E_t^{\text{exact}} = 0$. The field $\eta$ applied in $z$-direction breaks the symmetry and we erroneously find the ground-state energy to be

$$\begin{aligned} E_G^z &= J \langle\uparrow;\downarrow | \hat{H}_{\text{spin}} | \uparrow;\downarrow\rangle \\ &= J \langle\uparrow;\downarrow | \hat{S}_1^z \hat{S}_2^z | \uparrow;\downarrow\rangle - J/4 \\ &= -J(s^2 + 1/4) = -J/2 \,, \end{aligned} \tag{A.39}$$

instead of $E_G = -J$ for the correct $S_{\text{total}} = 0$ state. Note, that we can trivially improve the Gutzwiller description by treating both sites as "local" and minimize with respect to the $2^2 \times 2^2$ eigenstates, which results in the exact solution of the problem.

## A.4 Brillouin zone integration using the tetrahedron method

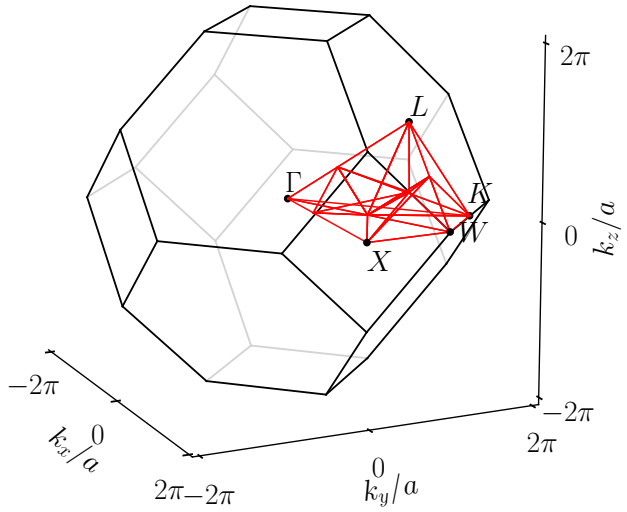


**Figure A.2:** Example of the BZ of a fcc-lattice (black lines), including the IBZ and high-symmetry points. The IBZ is constructed and divided into non-overlapping tetrahedra (red lines). The full BZ is recovered by applying all 48 symmetry operations of the $O_h$ group to the irreducible wedge of the BZ.

The evaluation of expectation values by integrating over the BZ zone is common to all band structure methods. The expectation value of a not necessarily hermitian operator $\hat{A}$ is obtained by evaluating

$$\langle \hat{A} \rangle = \frac{1}{V_G} \sum_n \int_{V_G} A_n(\mathbf{k}) f(\epsilon_n(\mathbf{k})) \, \mathrm{d}\mathbf{k} \tag{A.40}$$

with the matrix elements

$$A_n(\mathbf{k}) = \langle \Phi_n(\mathbf{k}) | \hat{A} | \Phi_n(\mathbf{k}) \rangle \ . \tag{A.41}$$

As usual, $\epsilon_n(\mathbf{k})$ denotes the $n$-th eigenenergy ("band-index") of an $\mathbf{k}$-dependent Hamiltonian, $V_G$ is the volume of the BZ, $f(\epsilon)$ the occupation number. For the occupation number we only consider the Fermi distribution at $T = 0$, which is equal to unity for $\epsilon < E_F$ and zero for $\epsilon > E_F$. In the typical integration methods, $A_n(\mathbf{k})$ is evaluated at discrete points within the BZ. The integral in (A.40) is evaluated using a weighted sum,

$$\langle \hat{A} \rangle = \sum_{j,n} A_n(\mathbf{k}_j) w_{j,n} \ , \tag{A.42}$$

where $j$ sums over the set of discrete $\mathbf{k}$-points ("$\mathbf{k}$-mesh") with the weights $w_{j,n}$. In order to obtain these weights, the irreducible part of the BZ is divided into non-overlapping tetrahedra [122], compare Fig. A.2. In our program, the irreducible wedge of the BZ is automatically created from the $N_{\text{sym}}$ symmetries present in the crystal. The next step is to approximate the number of states in each tetrahedron. The energies $\epsilon_{t,n,1}$, $\epsilon_{t,n,2}$, $\epsilon_{t,n,3}$, $\epsilon_{t,n,4}$ are the eigenenergies $\epsilon_n(\mathbf{k})$ of band $n$ evaluated at each of the four vertices of a tetrahedron $t$ with volume $V_t$. The energies are sorted, so that we have $\epsilon_{t,n,1} < \epsilon_{t,n,2} < \epsilon_{t,n,3} < \epsilon_{t,n,4}$. We introduce the notation $\epsilon_{t,n,ij} = \epsilon_{t,n,i} - \epsilon_{t,n,j}$.

For the sake of simplicity we drop all band and tetrahedra indices in the following expressions. Given a Fermi-energy $\epsilon := E_{\text{F}}$, the number of states in a volume of $\mathbf{k}$-space enclosed by a tetrahedron is approximated using the following expressions [123]:

$\underline{\epsilon < \epsilon_1}$.

$$n(\epsilon) = 0 \ . \tag{A.43a}$$

$\underline{\epsilon_1 < \epsilon < \epsilon_2}$

$$n(\epsilon) = \frac{V}{V_G} \frac{(\epsilon - \epsilon_1)^3}{\epsilon_{21}\epsilon_{31}\epsilon_{41}} \ . \tag{A.43b}$$

$\underline{\epsilon_2 < \epsilon < \epsilon_3}$

$$n(\epsilon) = \frac{V}{V_G} \frac{1}{\epsilon_{31}\epsilon_{41}} \left[ \epsilon_{21}^2 + 3\epsilon_{21}(\epsilon - \epsilon_2) + 3(\epsilon - \epsilon_2)^2 - \frac{\epsilon_{31} + \epsilon_{42}}{\epsilon_{32}\epsilon_{42}} (\epsilon - \epsilon_2)^3 \right] \ . \tag{A.43c}$$

$\underline{\epsilon_3 < \epsilon < \epsilon_4}$

$$n(\epsilon) = \frac{V}{V_G} \left( 1 - \frac{(\epsilon_4 - \epsilon)^3}{\epsilon_{41}\epsilon_{42}\epsilon_{43}} \right) \ . \tag{A.43d}$$

$\underline{\epsilon > \epsilon_4}$

$$n(\epsilon) = \frac{V}{V_G} \ . \tag{A.43e}$$

The first step in our band structure method is to determine the Fermi-energy. The "filling", i.e., the number of electrons $N_e$ per unit cell, is known. Using a bisection-based method we determine $E_\mathrm{F}$, which satisfies

$$N_e = N_\mathrm{sym} \sum_{n,t} n_{t,n}(E_F) \ , \tag{A.44}$$

where we reintroduced the band index $n$ and the tetrahedron index $t$. We store the calculated number of states $n_{t,n}(E_F)$. Next we determine the geometrical center of each tetrahedron. The centers have the special property that they are not lying on any high-symmetry lines, which would otherwise cause problems in the numerical evaluation of expectation values, due to the ambiguity of degenerated eigenvectors. All symmetry operations of the crystal are applied to each of the centers in the IBZ, spreading points over the whole BZ. Note that all of the points are unique by construction. We store these points $\mathbf{k}_j$ together with their weights $w_{j,n}$, which are given by the number of states of the tetrahedron from which they are constructed. With this at hand, the energy tensor (4.25) and the local density matrix (4.27) can be calculated using (A.42).

# Bibliography


[1] T. Dietl and H. Ohno. "Dilute ferromagnetic semiconductors: Physics and spintronic structures". In: *Rev. Mod. Phys.* 86 (1 Mar. 2014), pp. 187–251. DOI: 10.1103/RevModPhys.86.187.

[2] J. K. Furdyna and N. Samarth. "Magnetic properties of diluted magnetic semiconductors: A review (invited)". In: *Journal of Applied Physics* 61.8 (1987), pp. 3526–3531. DOI: 10.1063/1.338714.

[3] H. Ohno. "Making Nonmagnetic Semiconductors Ferromagnetic". In: *Science* 281 (Aug. 1998), pp. 951–956. DOI: 10.1126/science.281.5379.951.

[4] K. Y. Wang, R. P. Campion, K. W. Edmonds, M. Sawicki, T. Dietl, C. T. Foxon, and B. L. Gallagher. "Magnetism in (Ga,Mn)As Thin Films With $T_C$ Up To 173K". In: *AIP Conference Proceedings* 772.1 (2005), pp. 333–334. DOI: 10.1063/1.1994124.

[5] S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnár, M. L. Roukes, A. Y. Chtchelkanova, and D. M. Treger. "Spintronics: A Spin-Based Electronics Vision for the Future". In: *Science* 294.5546 (2001), pp. 1488–1495. ISSN: 0036-8075. DOI: 10.1126/science.1065389.

[6] P. A. M. Dirac. "Quantum Mechanics of Many-Electron Systems". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 123.792 (1929), pp. 714–733. ISSN: 09501207. DOI: 10.1098/rspa.1929.0094.

[7] W. Kohn. "Nobel Lecture: Electronic structure of matterwave functions and density functionals". In: *Rev. Mod. Phys.* 71 (5 Oct. 1999), pp. 1253–1266. DOI: 10.1103/RevModPhys.71.1253.

[8] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389.20 (1927), pp. 457–484. ISSN: 1521-3889. DOI: 10.1002/andp.19273892002.

[9] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136 (3B Nov. 1964), B864–B871. DOI: 10.1103/PhysRev.136.B864.

[10] M. Levy. "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the $v$-representability problem". In: *Proceedings of the National Academy of Sciences* 76.12 (1979), pp. 6062–6065.

[11] K. Capelle. "A bird's-eye view of density-functional theory". In: *eprint arXiv:cond-mat/0211443* (Nov. 2002).

[12] P. A. M. Dirac. "Note on Exchange Phenomena in the Thomas Atom". In: *Proceedings of the Cambridge Philosophical Society* 26 (1930), p. 376. DOI: 10.1017/S0305004100016108.

[13] D. M. Ceperley and B. J. Alder. "Ground State of the Electron Gas by a Stochastic Method". In: *Phys. Rev. Lett.* 45 (Aug. 1980), pp. 566–569. DOI: 10.1103/PhysRevLett.45.566.

[14] P. Bagno, O. Jepsen, and O. Gunnarsson. "Ground-state properties of third-row elements with nonlocal density functionals". In: *Phys. Rev. B* 40 (3 July 1989), pp. 1997–2000. DOI: 10.1103/PhysRevB.40.1997.

[15] S.-K. Ma and K. A. Brueckner. "Correlation Energy of an Electron Gas with a Slowly Varying High Density". In: *Phys. Rev.* 165 (1 Jan. 1968), pp. 18–31. DOI: 10.1103/PhysRev.165.18.

[16] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais. "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation". In: *Phys. Rev. B* 46 (11 Sept. 1992), pp. 6671–6687. DOI: 10.1103/PhysRevB.46.6671.

[17] U. von Barth and L. Hedin. "A local exchange-correlation potential for the spin polarized case. i". In: *Journal of Physics C: Solid State Physics* 5.13 (1972), p. 1629. DOI: 10.1088/0022-3719/5/13/012.

[18] N. Ashcroft and N. Mermin. *Solid State Physics*. Philadelphia: Saunders College, 1976. DOI: 10.1002/piuz.19780090109.

[19] H. Hellmann. "A New Approximation Method in the Problem of Many Electrons". In: *The Journal of Chemical Physics* 3.1 (1935), pp. 61–61. DOI: 10.1063/1.1749559.

[20] B. Meyer. "The pseudopotential plane wave approach". In: *Computational Nanoscience: do it yourself*. NIC series 31 (2006), pp. 71–83.

[21] D. J. Singh and L. Nordstrom. *Planewaves, Pseudopotentials, and the LAPW method*. Springer Science & Business Media, 2006. ISBN: 978-1-4419-3954-8. DOI: 10.1007/978-0-387-29684-5.

[22] J. C. Slater. "Wave Functions in a Periodic Potential". In: *Phys. Rev.* 51 (10 May 1937), pp. 846–851. DOI: 10.1103/PhysRev.51.846.

[23] S. Blügel and G. Bihlmayer. "Full-Potential Linearized Augmented Planewave Method". In: *Computational Nanoscience: do it yourself*. NIC series 31 (2006), pp. 85–129.

[24] J. C. Slater. "An Augmented Plane Wave Method for the Periodic Potential Problem". In: *Phys. Rev.* 92 (3 Nov. 1953), pp. 603–608. DOI: 10.1103/PhysRev.92.603.

[25] M. M. Saffren and J. C. Slater. "An Augmented Plane-Wave Method for the Periodic Potential Problem. II". In: *Phys. Rev.* 92 (5 Dec. 1953), pp. 1126–1128. DOI: 10.1103/PhysRev.92.1126.

[26] Shick, A. B. and Lichtenstein, A. I. and Pickett, W. E. "Implementation of the LDA+U method using the full-potential linearized augmented plane-wave basis". In: *Phys. Rev. B* 60 (15 Oct. 1999), pp. 10763–10769. DOI: 10.1103/PhysRevB.60.10763.

[27] R. Parr and Y. Weitao. *Density-Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry. Oxford University Press, 1994. ISBN: 9780195357738. DOI: 10.1002/qua.560470107.

[28] J. P. Perdew, A. Ruzsinszky, L. A. Constantin, J. Sun, and G. I. Csonka. "Some Fundamental Issues in Ground-State Density Functional Theory: A Guide for the Perplexed". In: *Journal of Chemical Theory and Computation* 5.4 (2009). PMID: 26609599, pp. 902–908. DOI: 10.1021/ct800531s.

[29] J. P. Perdew. "Density functional theory and the band gap problem". In: *International Journal of Quantum Chemistry* 28.S19 (1985), pp. 497–523. ISSN: 1097-461X. DOI: 10.1002/qua.560280846.

[30] J. F. Janak. "Proof that $\partial E/\partial n_i = \epsilon$ in density-functional theory". In: *Phys. Rev. B* 18 (12 Dec. 1978), pp. 7165–7168. DOI: 10.1103/PhysRevB.18.7165.

[31] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz. "Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy". In: *Phys. Rev. Lett.* 49 (23 Dec. 1982), pp. 1691–1694. DOI: 10.1103/PhysRevLett.49.1691.

[32] T. Koopmans. "Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms". In: *Physica* 1.1 (1934), pp. 104–113. ISSN: 70031-8914. DOI: 10.1016/S0031-8914(34)90011-2.

[33] J. P. Perdew and M. Levy. "Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities". In: *Phys. Rev. Lett.* 51 (20 Nov. 1983), pp. 1884–1887. DOI: 10.1103/PhysRevLett.51.1884.

[34] B. Himmetoglu, A. Floris, S. de Gironcoli, and M. Cococcioni. "Hubbard-corrected DFT energy functionals: The LDA+U description of correlated systems". In: *International Journal of Quantum Chemistry* 114.1 (2014), pp. 14–49. ISSN: 1097-461X. DOI: 10.1002/qua.24521.

[35] P. Kuiper, G. Kruizinga, J. Ghijsen, G. A. Sawatzky, and H. Verweij. "Character of Holes in $Li_xNi_{1-x}O$ and Their Magnetic Behavior". In: *Phys. Rev. Lett.* 62 (2 Jan. 1989), pp. 221–224. DOI: 10.1103/PhysRevLett.62.221.

[36] V. I. Anisimov, I. V. Solovyev, M. A. Korotin, M. T. Czyżyk, and G. A. Sawatzky. "Density-functional theory and NiO photoemission spectra". In: *Phys. Rev. B* 48 (23 Dec. 1993), pp. 16929–16934. DOI: 10.1103/PhysRevB.48.16929.

[37] I. V. Solovyev, P. H. Dederichs, and V. I. Anisimov. "Corrected atomic limit in the local-density approximation and the electronic structure of *d* impurities in Rb". In: *Phys. Rev. B* 50 (23 Dec. 1994), pp. 16861–16871. DOI: 10.1103/PhysRevB.50.16861.

[38] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein. "First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA + U method". In: *Journal of Physics: Condensed Matter* 9.4 (1997), p. 767. DOI: 10.1007/978-3-642-57834-2_9.

[39] A. I. Lichtenstein, V. I. Anisimov, and J. Zaanen. "Density-functional theory and strong interactions: Orbital ordering in Mott-Hubbard insulators". In: *Phys. Rev. B* 52 (8 Aug. 1995), R5467–R5470. DOI: 10.1103/PhysRevB.52.R5467.

[40] G. H. Wannier. "The Structure of Electronic Excitation Levels in Insulating Crystals". In: *Phys. Rev.* 52 (3 Aug. 1937), pp. 191–197. DOI: 10.1103/PhysRev.52.191.

[41] E. Blount. "Formalisms of Band Theory". In: ed. by F. Seitz and D. Turnbull. Vol. 13. Solid State Physics. Academic Press, 1962, pp. 305–373. DOI: 10.1016/S0081-1947(08)60459-2.

[42] R. D. King-Smith and D. Vanderbilt. "Theory of polarization of crystalline solids". In: *Phys. Rev. B* 47 (3 Jan. 1993), pp. 1651–1654. DOI: 10.1103/PhysRevB.47.1651.

[43] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt. "Maximally localized Wannier functions: Theory and applications". In: *Rev. Mod. Phys.* 84 (4 Oct. 2012), pp. 1419–1475. DOI: 10.1103/RevModPhys.84.1419.

[44] N. Marzari and D. Vanderbilt. "Maximally localized generalized Wannier functions for composite energy bands". In: *Phys. Rev. B* 56 (20 Nov. 1997), pp. 12847–12865. DOI: 10.1103/PhysRevB.56.12847.

[45] N. Marzari, I. Souza, and D. Vanderbilt. "An introduction to maximally-localized Wannier functions". In: *Psi-K Newsletter* 57 (2003), p. 129.

[46] F. Freimuth, Y. Mokrousov, D. Wortmann, S. Heinze, and S. Blügel. "Maximally localized Wannier functions within the FLAPW formalism". In: *Phys. Rev. B* 78 (3 July 2008), p. 035120. DOI: 10.1103/PhysRevB.78.035120.

[47] P.-O. Löwdin. "On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals". In: *The Journal of Chemical Physics* 18.3 (1950), pp. 365–375. DOI: 10.1063/1.1747632.

[48] I. Souza, N. Marzari, and D. Vanderbilt. "Maximally localized Wannier functions for entangled energy bands". In: *Phys. Rev. B* 65 (3 Dec. 2001), p. 035109. DOI: 10.1103/PhysRevB.65.035109.

[49] *FLEUR, The Jülich FLAPW code family*. http://www.flapw.de/. Accessed: Feb. 2015.

[50] Y. Wu, G. Chen, Y. Zhu, W.-J. Yin, Y. Yan, M. Al-Jassim, and S. J. Pennycook. "LDA+U/GGA+U calculations of structural and electronic properties of CdTe: Dependence on the effective U parameter". In: *Computational Materials Science* 98 (2015), pp. 18–23. ISSN: 0927-0256. DOI: 10.1016/j.commatsci.2014.10.051.

[51] D. G. Thomas. "Excitons and Band Splitting Produced by Uniaxial Stress in CdTe". In: *Journal of Applied Physics* 32 (1961), pp. 2298–2304. DOI: 10.1063/1.1777063.

[52] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari. "Wannier90: A tool for obtaining maximally-localised Wannier functions". In: *Comp. Phys. Comm.* 178.9 (2008), pp. 685–699. DOI: 10.1016/j.cpc.2007.11.016.

[53] F. Lechermann, A. Georges, A. Poteryaev, S. Biermann, M. Posternak, A. Yamasaki, and O. K. Andersen. "Dynamical mean-field theory using Wannier functions: A flexible route to electronic structure calculations of strongly correlated materials". In: *Phys. Rev. B* 74 (12 Sept. 2006), p. 125120. DOI: 10.1103/PhysRevB.74.125120.

[54] M. C. Gutzwiller. "Correlation of Electrons in a Narrow *s* Band". In: *Phys. Rev.* 137 (6A Mar. 1965), A1726–A1735. DOI: 10.1103/PhysRev.137.A1726.

[55] M. C. Gutzwiller. "Effect of Correlation on the Ferromagnetism of Transition Metals". In: *Phys. Rev.* 134 (4A May 1964), A923–A941. DOI: 10.1103/PhysRev.134.A923.

[56] M. C. Gutzwiller. "Effect of Correlation on the Ferromagnetism of Transition Metals". In: *Phys. Rev. Lett.* 10 (5 Mar. 1963), pp. 159–162. DOI: 10.1103/PhysRevLett.10.159.

[57] F. Gebhard. *The Mott Metal-Insulator Transition: Models and Methods*. Springer Tracts in Modern Physics. Springer Berlin Heidelberg, 2003. ISBN: 9783540148586. DOI: 10.1007/3-540-14858-2.

[58] J. Bünemann, W. Weber, and F. Gebhard. "Multiband Gutzwiller wave functions for general on-site interactions". In: *Phys. Rev. B* 57 (12 1998), pp. 6896–6916. DOI: 10.1103/PhysRevB.57.6896.

[59] J. Bünemann. "The Gutzwiller Variational Theory and Related Methods for Correlated Electron Systems". Habilitationsschrift. Philipps-Universität Marburg, 2009.

[60] F. Gebhard and D. Vollhardt. "Correlation functions for interacting fermions in the Gutzwiller ansatz". In: *Phys. Rev. B* 38 (10 Oct. 1988), pp. 6911–6927. DOI: 10.1103/PhysRevB.38.6911.

[61] W. Metzner and D. Vollhardt. "Analytic calculation of ground-state properties of correlated fermions with the Gutzwiller wave function". In: *Phys. Rev. B* 37 (13 May 1988), pp. 7382–7399. DOI: 10.1103/PhysRevB.37.7382.

[62] J. Bünemann, F. Gebhard, T. Schickling, and W. Weber. "Numerical minimisation of Gutzwiller energy functionals". In: *physica status solidi (b)* 249.6 (2012), pp. 1282–1291. ISSN: 1521-3951. DOI: 10.1002/pssb.201147585.

[63] J. Bünemann, T. Linneweber, U. Löw, F. B. Anders, and F. Gebhard. "Interplay of Coulomb interaction and spin-orbit coupling". In: *Phys. Rev. B* 94 (3 July 2016), p. 035116. DOI: 10.1103/PhysRevB.94.035116.

[64] J. Bünemann, F. Gebhard, and R. Thul. "Landau-Gutzwiller quasiparticles". In: *Phys. Rev. B* 67 (7 Feb. 2003), p. 075103. DOI: 10.1103/PhysRevB.67.075103.

[65] W. F. Brinkman and T. M. Rice. "Application of Gutzwiller's Variational Method to the Metal-Insulator Transition". In: *Phys. Rev. B* 2 (10 Nov. 1970), pp. 4302–4304. DOI: 10.1103/PhysRevB.2.4302.

[66] P. G. J. van Dongen, F. Gebhard, and D. Vollhardt. "Variational evaluation of correlation functions for lattice electrons in high dimensions". In: *Zeitschrift für Physik B Condensed Matter* 76.2 (1989), pp. 199–210. ISSN: 1431-584X. DOI: 10.1007/BF01312685.

[67] R. Wyckoff. *Crystal Structures*. Crystal Structures Vol. 1. Wiley, 1971.

[68] W. Limmer, A. Koeder, S. Frank, M. Glunk, W. Schoch, V. Avrutin, K. Zuern, R. Sauer, and A. Waag. "Electronic and magnetic properties of GaMnAs: annealing effects". In: *Physica E: Low-dimensional Systems and Nanostructures* 21.2–4 (2004). Proceedings of the Eleventh International Conference on Modulated Semiconductor Structures, pp. 970–974. ISSN: 1386-9477. DOI: 10.1016/j.physe.2003.11.173.

[69] D. Krause and P. Thörnig. "JURECA: General-purpose supercomputer at Jülich Supercomputing Centre." In: *Journal of large-scale research facilities* 2 (A62 2016). DOI: 10.17815/jlsrf-2-121.

[70] S. Sugano, Y. Tanabe, and H. Kamimura. *Multiplets of transition-metal ions in crystals*. Academic Press New York, 1970. ISBN: 0126760500. DOI: 10.1016/b978-0-12-676050-7.x5001-x.

[71] P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai. "Ground States of Constrained Systems: Application to Cerium Impurities". In: *Phys. Rev. Lett.* 53 (26 Dec. 1984), pp. 2512–2515. DOI: 10.1103/PhysRevLett.53.2512.

[72] I. Schnell, G. Czycholl, and R. C. Albers. "Hubbard-$U$ calculations for Cu from first-principle Wannier functions". In: *Phys. Rev. B* 65 (7 Jan. 2002), p. 075103. DOI: 10.1103/PhysRevB.65.075103.

[73] M. Karolak, G. Ulm, T. Wehling, V. Mazurenko, A. Poteryaev, and A. Lichtenstein. "Double counting in LDA+DMFT – The example of NiO". In: *Journal of Electron Spectroscopy and Related Phenomena* 181.1 (2010). Proceedings of International Workshop on Strong Correlations and Angle-Resolved Photoemission Spectroscopy 2009, pp. 11–15. ISSN: 0368-2048. DOI: 10.1016/j.elspec.2010.05.021.

[74] T. Linneweber, J. Bünemann, U. Löw, F. Gebhard, and F. Anders. "Exchange couplings for Mn ions in CdTe: Validity of spin models for dilute magnetic II-VI semiconductors". In: *Phys. Rev. B* 95 (4 Jan. 2017), p. 045134. DOI: 10.1103/PhysRevB.95.045134.

[75] M. Jain. *Diluted magnetic semiconductors*. World Scientific, Singapore, 1991. ISBN: 978-981-02-0176-0. DOI: 10.1142/9789814368216.

[76] V. Bindilatti, E. ter Haar, N. F. Oliveira, Y. Shapira, and M. T. Liu. "Distant-Neighbor Exchange Constants in Dilute Magnetic Semiconductors". In: *Phys. Rev. Lett.* 80 (24 June 1998), p. 5425. DOI: 10.1103/PhysRevLett.80.5425.

[77] H. Malarenko Jr., V. Bindilatti, N. F. Oliveira Jr., M. T. Liu, Y. Shapira, and L. Puech. "Distant neighbors exchange interactions in $Cd_{1-x}Mn_xTe$". In: *Physica B: Condensed Matter* 284–288 (2000), p. 1523. DOI: 10.1016/S0921-4526(99)02728-3.

[78] J. Spałek, A. Lewicki, Z. Tarnawski, J. K. Furdyna, R. R. Galazka, and Z. Obuszko. "Magnetic susceptibility of semimagnetic semiconductors: The high-temperature regime and the role of superexchange". In: *Phys. Rev. B* 33 (5 Mar. 1986), pp. 3407–3418. DOI: 10.1103/PhysRevB.33.3407.

[79] B. E. Larson, K. C. Hass, H. Ehrenreich, and A. E. Carlsson. "Theory of exchange interactions and chemical trends in diluted magnetic semiconductors". In: *Phys. Rev. B* 37 (8 Mar. 1988), pp. 4137–4154. DOI: 10.1103/PhysRevB.37.4137.

[80] S. Barthel, G. Czycholl, and G. Bouzerar. "Effective Heisenberg exchange integrals of diluted magnetic semiconductors determined within realistic multi-band tight-binding models". In: *Eur. Phys. J. B* 86.1 (2013), p. 11. DOI: 10.1140/epjb/e2012-30795-4.

[81] A. Savoyant, S. D'Ambrosio, R. O. Kuzian, A. M. Daré, and A. Stepanov. "Exchange integrals in Mn- and Co-doped II-VI semiconductors". In: *Phys. Rev. B* 90 (7 Aug. 2014), p. 075205. DOI: 10.1103/PhysRevB.90.075205.

[82] T. Chanier, F. Virot, and R. Hayn. "Chemical trend of exchange coupling in diluted magnetic II-VI semiconductors: *Ab initio* calculations". In: *Phys. Rev. B* 79 (20 May 2009), p. 205204. DOI: 10.1103/PhysRevB.79.205204.

[83] T. Chanier, F. Virot, and R. Hayn. "Erratum: Chemical trend of exchange coupling in diluted magnetic II-VI semiconductors: *Ab initio* calculations [Phys. Rev. B **79** , 205204 (2009)]". In: *Phys. Rev. B (Erratum)* 83 (23 June 2011), p. 239903. DOI: 10.1103/PhysRevB.83.239903.

[84] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton. "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study". In: *Phys. Rev. B* 57 (3 1998), pp. 1505–1509. DOI: 10.1103/PhysRevB.57.1505.

[85] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger. "Calculations of Hubbard $U$ from first-principles". In: *Phys. Rev. B* 74 (12 Sept. 2006), p. 125106. DOI: 10.1103/PhysRevB.74.125106.

[86] T. Schickling, J. Bünemann, F. Gebhard, and L. Boeri. "Quasiparticle bands and structural phase transition of iron from Gutzwiller density-functional theory". In: *Phys. Rev. B* 93 (20 May 2016), p. 205151. DOI: 10.1103/PhysRevB.93.205151.

[87] M. Liu, Y. Shapira, E. ter Haar, V. Bindilatti, and E. J. McNiff. "Magnetization steps of spin quartets". In: *Phys. Rev. B* 54 (9 Sept. 1996), p. 6457. DOI: 10.1103/PhysRevB.54.6457.

[88] S. Foner, Y. Shapira, D. Heiman, P. Becla, R. Kershaw, K. Dwight, and A. Wold. "Magnetization steps in dilute magnetic semiconductors to 55T: $Mn^{2+}$ pair saturation in $Cd_{1-x}Mn_xTe$ and steps in $Zn_{1-x}Mn_xSe$, $Zn_{1-x}Mn_xTe$, $Cd_{1-x}Mn_xSe$". In: *Phys. Rev. B* 39 (1989), p. 11793. DOI: 10.1103/PhysRevB.39.11793.

[89] A. Merad, M. Kanoun, and S. Goumri-Said. "Ab initio study of electronic structures and magnetism in ZnMnTe and CdMnTe diluted magnetic semiconductors". In: *Journal of Magnetism and Magnetic Materials* 302.2 (2006), pp. 536–542. ISSN: 0304-8853. DOI: 10.1016/j.jmmm.2005.10.018.

[90] B. E. Larson, K. C. Hass, and R. L. Aggarwal. "Effects of internal exchange fields on magnetization steps in diluted magnetic semiconductors". In: *Phys. Rev. B* 33 (3 Feb. 1986), pp. 1789–1796. DOI: 10.1103/PhysRevB.33.1789.

[91] M. A. Ruderman and C. Kittel. "Indirect Exchange Coupling of Nuclear Magnetic Moments by Conduction Electrons". In: *Phys. Rev.* 96 (1 Oct. 1954), pp. 99–102. DOI: 10.1103/PhysRev.96.99.

[92] T. Kasuya. "A Theory of Metallic Ferro- and Antiferromagnetism on Zener's Model". In: *Progress of Theoretical Physics* 16.1 (1956), pp. 45–57. DOI: 10.1143/PTP.16.45.

[93] K. Yosida. "Magnetic Properties of Cu-Mn Alloys". In: *Phys. Rev.* 106 (5 June 1957), pp. 893–898. DOI: 10.1103/PhysRev.106.893.

[94] T. Dietl, H. Ohno, F. Matsukura, J. Cibert, and D. Ferrand. "Zener Model Description of Ferromagnetism in Zinc-Blende Magnetic Semiconductors". In: *Science* 287.5455 (2000), pp. 1019–1022. ISSN: 0036-8075. DOI: 10.1126/science.287.5455.1019.

[95] C. Zener. "Interaction Between the *d* Shells in the Transition Metals". In: *Phys. Rev.* 81 (3 Feb. 1951), pp. 440–444. DOI: 10.1103/PhysRev.81.440.

[96] T. Jungwirth, K. Y. Wang, J. Mašek, K. W. Edmonds, J. König, J. Sinova, M. Polini, N. A. Goncharuk, A. H. MacDonald, M. Sawicki, A. W. Rushforth, R. P. Campion, L. X. Zhao, C. T. Foxon, and B. L. Gallagher. "Prospects for high temperature ferromagnetism in (Ga,Mn)As semiconductors". In: *Phys. Rev. B* 72 (16 Oct. 2005), p. 165204. DOI: 10.1103/PhysRevB.72.165204.

[97] K. M. Yu, W. Walukiewicz, T. Wojtowicz, I. Kuryliszyn, X. Liu, Y. Sasaki, and J. K. Furdyna. "Effect of the location of Mn sites in ferromagnetic $Ga_{1-x}Mn_xAs$ on its Curie temperature". In: *Phys. Rev. B* 65 (20 Apr. 2002), p. 201303. DOI: 10.1103/PhysRevB.65.201303.

[98] S. C. Erwin and A. G. Petukhov. "Self-Compensation in Manganese-Doped Ferromagnetic Semiconductors". In: *Phys. Rev. Lett.* 89 (22 Nov. 2002), p. 227201. DOI: 10.1103/PhysRevLett.89.227201.

[99] R. C. Myers, B. L. Sheu, A. W. Jackson, A. C. Gossard, P. Schiffer, N. Samarth, and D. D. Awschalom. "Antisite effect on hole-mediated ferromagnetism in (Ga,Mn)As". In: *Phys. Rev. B* 74 (15 Oct. 2006), p. 155203. DOI: 10.1103/PhysRevB.74.155203.

[100] M. Dobrowolska, K. Tivakornsasithorn, X. Liu, J. K. Furdyna, M. Berciu, K. M. Yu, and W. Walukiewicz. "Controlling the Curie temperature in (Ga,Mn)As through location of the Fermi level within the impurity band". In: *Nat. Mat.* 11 (5 2012), pp. 444–449. DOI: 10.1038/nmat3250.

[101] K. W. Edmonds, B. L. Gallagher, M. Wang, A. W. Rushforth, O. Makarovsky, A. Patane, R. P. Campion, C. T. Foxon, V. Novak, and T. Jungwirth. "Correspondence on "Controlling the Curie temperature in (Ga,Mn)As through location of the Fermi level within the impurity band" by M. Dobrowolska et al". In: *ArXiv e-prints* (Nov. 2012).

[102] M. Jain, L. Kronik, J. R. Chelikowsky, and V. V. Godlevsky. "Electronic structure and spin polarization of Mn-containing dilute magnetic III-V semiconductors". In: *Phys. Rev. B* 64 (24 Dec. 2001), p. 245205. DOI: 10.1103/PhysRevB.64.245205.

[103] E. Kulatov, H. Nakayama, H. Mariette, H. Ohta, and Y. A. Uspenskii. "Electronic structure, magnetic ordering, and optical properties of GaN and GaAs doped with Mn". In: *Phys. Rev. B* 66 (4 July 2002), p. 045203. DOI: 10.1103/PhysRevB.66.045203.

[104] Y.-J. Zhao, W. T. Geng, K. T. Park, and A. J. Freeman. "Electronic and magnetic properties of $Ga_{1-x}Mn_xAs$ : Role of Mn defect bands". In: *Phys. Rev. B* 64 (3 June 2001), p. 035207. DOI: 10.1103/PhysRevB.64.035207.

[105] S.-H. Wei, X. G. Gong, G. M. Dalpian, and S.-H. Wei. "First-principles study of Mn-induced local magnetic moments in host semiconductors". In: *Phys. Rev. B* 71 (14 Apr. 2005), p. 144409. DOI: 10.1103/PhysRevB.71.144409.

[106] J. Okabayashi, A. Kimura, O. Rader, T. Mizokawa, A. Fujimori, T. Hayashi, and M. Tanaka. "Angle-resolved photoemission study of $Ga_{1-x}Mn_xAs$". In: *Phys. Rev. B* 64 (12 Sept. 2001), p. 125304. DOI: 10.1103/PhysRevB.64.125304.

[107] H. Åsklund, L. Ilver, J. Kanski, J. Sadowski, and R. Mathieu. "Photoemission studies of $Ga_{1-x}Mn_xAs$ : Mn concentration dependent properties". In: *Phys. Rev. B* 66 (11 Sept. 2002), p. 115319. DOI: 10.1103/PhysRevB.66.115319.

[108] J. Okabayashi, A. Kimura, T. Mizokawa, A. Fujimori, T. Hayashi, and M. Tanaka. "Mn 3*d* partial density of states in $Ga_{1-x}Mn_xAs$ studied by resonant photoemission spectroscopy". In: *Phys. Rev. B* 59 (4 Jan. 1999), R2486–R2489. DOI: 10.1103/PhysRevB.59.R2486.

[109] H. Ohno, A. Shen, F. Matsukura, A. Oiwa, A. Endo, S. Katsumoto, and Y. Iye. "(Ga,Mn)As: A new diluted magnetic semiconductor based on GaAs". In: *Applied Physics Letters* 69.3 (1996), pp. 363–365. DOI: 10.1063/1.118061.

[110] K. Sato, W. Schweika, P. H. Dederichs, and H. Katayama-Yoshida. "Low-temperature ferromagnetism in (Ga, Mn)N: *Ab initio* calculations". In: *Phys. Rev. B* 70 (20 Nov. 2004), p. 201202. DOI: 10.1103/PhysRevB.70.201202.

[111] A. Lichtenstein, M. Katsnelson, V. Antropov, and V. Gubanov. "Local spin density functional approach to the theory of exchange interactions in ferromagnetic metals and alloys". In: *Journal of Magnetism and Magnetic Materials* 67.1 (1987), pp. 65–74. ISSN: 0304-8853. DOI: 10.1016/0304-8853(87)90721-9.

[112] S. Barthel, G. Czycholl, and G. Bouzerar. "Effective Heisenberg exchange integrals of diluted magnetic semiconductors determined within realistic multi-band tight-binding models". In: *The European Physical Journal B* 86.1 (2013), p. 11. ISSN: 1434-6036. DOI: 10.1140/epjb/e2012-30795-4.

[113] R. Bouzerar, G. Bouzerar, and T. Ziman. "Non-perturbative V-J pd model and ferromagnetism in dilute magnets". In: *EPL (Europhysics Letters)* 78.6 (2007), p. 67003. DOI: 10.1209/0295-5075/78/67003.

[114] X. Y. Deng, L. Wang, X. Dai, and Z. Fang. "Local density approximation combined with Gutzwiller method for correlated electron systems: Formalism and applications". In: *Phys. Rev. B* 79 (Feb. 2009), p. 075114. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.79.075114.

[115] K. Ho, J. Schmalian, and C. Wang. "Gutzwiller density functional theory for correlated electron systems". In: *Phys. Rev. B* 77 (2008), p. 073101. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.77.073101.

[116] R. Dong, X. Wan, X. Dai, and S. Y. Savrasov. "Orbital-dependent electronic masses in Ce heavy-fermion materials studied via Gutzwiller density-functional theory". In: *Phys. Rev. B* 89 (16 Apr. 2014), p. 165122. DOI: 10.1103/PhysRevB.89.165122.

[117] T. Schickling, J. Bünemann, F. Gebhard, and W. Weber. "Gutzwiller density functional theory: a formal derivation and application to ferromagnetic nickel". In: *New Journal of Physics* 16.9 (2014), p. 093034. DOI: 10.1088/1367-2630/16/9/093034.

[118] J. Bünemann, T. Linneweber, and F. Gebhard. "Approximation schemes for the study of multi-band Gutzwiller wave functions". In: *physica status solidi (b)* 254.1 (2017). 1600166, 1600166–n/a. ISSN: 1521-3951. DOI: 10.1002/pssb.201600166.

[119] M. Glunk, J. Daeubler, L. Dreher, S. Schwaiger, W. Schoch, R. Sauer, W. Limmer, A. Brandlmaier, S. T. B. Goennenwein, C. Bihler, and M. S. Brandt. "Magnetic anisotropy in (Ga,Mn)As: Influence of epitaxial strain and hole concentration". In: *Phys. Rev. B* 79 (19 May 2009), p. 195206. DOI: 10.1103/PhysRevB.79.195206.

[120] W. Pauli. *Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a. Band II: 1930–1939*. Springer-Verlag Berlin Heidelberg, 1985. ISBN: 978-3-662-31100-4. DOI: 10.1007/978-3-540-78801-0.

[121] J. C. Slater and G. F. Koster. "Simplified LCAO Method for the Periodic Potential Problem". In: *Phys. Rev.* 94 (6 June 1954), pp. 1498–1524. DOI: 10.1103/PhysRev.94.1498.

[122] J. Rath and A. J. Freeman. "Generalized magnetic susceptibilities in metals: Application of the analytic tetrahedron linear energy method to Sc". In: *Phys. Rev. B* 11 (6 Mar. 1975), pp. 2109–2117. DOI: 10.1103/PhysRevB.11.2109.

[123] P. E. Blöchl, O. Jepsen, and O. K. Andersen. "Improved tetrahedron method for Brillouin-zone integrations". In: *Phys. Rev. B* 49 (23 June 1994), pp. 16223–16233. DOI: 10.1103/PhysRevB.49.16223.

# Acronyms

| | |
|---|---|
| AMF | Around Mean Field. |
| APW | Augmented Plane Wave. |
| ARPES | Angle-Resolved Photoemission Spectroscopy. |
| BZ | Brillouin Zone. |
| DFT | Density Functional Theory. |
| DMFT | Dynamical Mean Field Theory. |
| DMS | Diluted Magnetic Semiconductors. |
| DOS | Density Of States. |
| FLAPW | Full Potential Linear Augmented Plane Wave. |
| FLL | Fully Localized Limit. |
| GEA | Gradient-Expansion Approximations. |
| GGA | Generalized Gradient Approximation. |
| HEG | Homogenous Electron Gas. |
| IBZ | Irreducible Brillouin Zone. |
| KKR-CPA | Korringa-Kohn-Rostoker coherent-potential approximation. |
| LAPW | Linear Augmented Plane Wave. |
| LDA | Local Density Approximation. |
| LSDA | Local Spin Density Approximation. |
| MBE | Molecular Beam Epitaxy. |
| MLWF | Maximally Localized Wannier Function. |
| PBE | Perdew–Burke–Enzernhof. |
| PW | Plane Wave. |
| PWF | Projected Wannier Function. |
| RKKY | Ruderman-Kittel-Kasuya-Yosida. |
| WF | Wannier Function. |

# Publications

- T. Linneweber, J. Bünemann, U. Löw, F. Gebhard and F. Anders.
  "Exchange couplings for Mn ions in CdTe: validity of spin models for dilute magnetic II-VI semiconductors", Phys. Rev. B **95**, 045134 **(January 2017)**. [74]
- J. Bünemann, T. Linneweber, U. Löw, F. Anders and F. Gebhard.
  "Interplay of Coulomb interaction and spin-orbit coupling", Phys. Rev. B **94**, 035116 **(July 2016)**. [63]
- J. Bünemann, T. Linneweber and F. Gebhard.
  "Approximation schemes for the study of multi-band Gutzwiller wave functions", physica status solidi(b) **254**, 1521-3951 (**January 2017**). [118]

# Danksagung

Zunächst bedanke ich mich bei Prof. Frithjof Anders, der mir die Promotion an seinem Lehrstuhl ermöglicht hat. Seine Anmerkungen und Fragen bei internen Vorträgen waren für mich von großem Wert.

Bei Priv.-Doz. Dr. Ute Löw möchte ich mich herzlich für die gute Betreuung bedanken. Ihr entgegengebrachtes Vertrauen hat es ermöglicht, selbstständig Fragestellungen anzugehen. Auch ein Antrag auf Rechenleistung am Forschungszentrum Jülich wäre ohne ihre Hilfe nicht möglich gewesen.

Besonderer Dank gilt dem verstorbenen Prof. Werner Weber, der diese Themenstellung gewählt hat und mich im ersten Jahr meiner Promotionszeit an Bandstrukturrechnungen und die Gutzwillermethode herangeführt hat. Sein Enthusiasmus und seine Liebe zur Festkörperphysik, und eigentlich zur gesamten Wissenschaft bleibt unvergessen.

Prof. Florian Gebhard bin ich zu besonderem Dank verpflichtet. Sein stetiges Interesse und seine Unterstützung haben entscheidend zum Abschluss dieser Arbeit beigetragen.

Dr. Jörg Bünemann danke ich für die Zusammenarbeit und die großartige Unterstützung bei allen Problemen, die im Zusammenhang mit der Gutzwillermethode stehen.

Den Mitarbeitern der Lehrstühle T1 und T2 sei gedankt für die wunderbare Arbeitsatmosphäre. Dr. Carsten Nase danke ich für die unterhaltsame Einführung in die Computersysteme des Lehrstuhls. Einige Tassen Kaffee/Tee mit Lars Gravert, Dr. Jörn Krones, Dr. Wouter Beugeling und Daniel May sorgten für tiefere Einsichten, die nicht zwangsläufig physikalischer Natur waren.

Ebenso danke ich Marcel Strzys, der diese Arbeit Korrektur gelesen hat.

Großer Dank geht an meine Eltern und meine Brüder, für ihre Unterstützung und ihren Rückhalt.

Meiner Freundin Maja danke ich für ihre Liebe und Geduld.

The author gratefully acknowledges the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA [69] at Jülich Supercomputing Centre (JSC) under project no. HDO08.