Theoretical Studies of Two-Dimensional Magnetism and Chemical Bonding

OLEKSIY GRECHNYEV
Dissertation presented at Uppsala University to be publicly examined in Siegbahnsalen, Ångström Laboratory, Uppsala, Friday, April 1, 2005 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

This thesis is divided into two parts. In the first part we study thermodynamics of the two-dimensional Heisenberg ferromagnet with dipolar interaction. This interaction breaks the conditions of the Mermin-Wagner theorem, resulting in a finite transition temperature. Our calculations are done within the framework of the self-consistent spin-wave theory (SSWT), which is modified in order to include the dipolar interaction. Both quantum and classical versions of the Heisenberg model are considered.

The second part of the thesis investigates the chemical bonding in solids from the first principles calculations. A new chemical bonding indicator called balanced crystal orbital overlap population (BCOOP) is developed. BCOOP is less basis set dependent than the earlier indicators and it can be used with full-potential density-functional theory (DFT) codes. We apply BCOOP formalism to the chemical bonding in the high-Tc superconductor MgB2 and the theoretically predicted MAX phase Nb3SiC2. We also study how the chemical bonding results in a repulsive hydrogen–hydrogen interaction in metal hydrides. The role of this interaction in the structural phase transition in Ti3SnHx is investigated.

Keywords: spin Hamiltonians, quantized spin models, Heisenberg model, spin waves, self-consistent spin-wave theory, dipolar interaction, density functional theory, chemical bonding, overlap population, MAX phases, metal hydrides

Oleksiy Grechnyev, Department of Physics, Box 530, Uppsala University, SE-75121 Uppsala, Sweden

© Oleksiy Grechnyev 2005

ISSN 1651-6214
ISBN 91-554-6164-6
urn:nbn:se:uu:diva-4815 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4815)
Dedicated to Gary Gygax, Dave Arneson, George Lucas and Bioware Corp.
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I  Thermodynamics of a two-dimensional Heisenberg ferromagnet with dipolar interaction
A. Grechnev, V. Yu. Irkhin, M. I. Katsnelson and O. Eriksson

II Geometry of the valence transition induced surface reconstruction of Sm(0001)

III Balanced crystal orbital overlap population-a tool for analysing chemical bonds in solids
A. Grechnev, R. Ahuja and O. Eriksson

IV A new nanolayered material, Nb\textsubscript{3}SiC\textsubscript{2}, predicted from First-Principle s Theory
A. Grechnev, S. Li, R. Ahuja, O. Eriksson, U. Jansson and O. Wilhelmsson

V Elastic properties of Mg\textsubscript{1−x}Al\textsubscript{x}B\textsubscript{2} from first principles theory
P. Souvatzis, J. M. Osorio-Guillen, R. Ahuja, A. Grechnev and O. Eriksson

VI H-H interaction and structural phase transition in Ti\textsubscript{3}SnH\textsubscript{x}
A. Grechnev, P. H. Andersson, R. Ahuja, O. Eriksson, M. Vennström and Y. Andersson

VII Phase relations in the Ti\textsubscript{3}SnD system
M. Vennström, A. Grechnev, O. Eriksson and Y. Andersson
Reprints were made with permission from the publishers.

The following papers are not included in the thesis

I  Ab initio calculation of depth-resolved optical anisotropy of the Cu(110) surface
P. Monachesi, M. Palummo, R. Del Sole, A. Grechnev and O. Eriksson

II  Unusual magnetism and magnetocrystalline anisotropy of CrPt₃
P. M. Oppeneer, I. Galanakis, A. Grechnev and O. Eriksson

III  Many-body projector orbitals for electronic structure theory of strongly correlated electrons
O. Eriksson, J.M. Wills, M. Colarieti-Tosti, S. Lebégue and A. Grechnev
Submitted to Int. J. Qu. Chem.

Comments on my contribution

In the papers where I am the first author I am responsible for the main part of the work: theory, calculations and writing the paper. In the other papers I have contributed in different ways, such as ideas, calculations, code development or analysis.
Contents

1 Introduction .......................................................... 1
2 Thermodynamics of a 2D ferromagnet with dipolar interaction .... 5
  2.1 Heisenberg model ............................................... 7
  2.2 Magnons and Mermin-Wagner theorem ....................... 10
  2.3 Free magnons with dipolar interaction ....................... 13
  2.4 Self-consistent spin-wave theory ............................. 19
  2.5 Calculation of the lattice sums: Ewald method ............... 25
  2.6 SSWT results .................................................. 26
3 Interlude: Density functional theory ............................... 33
  3.1 Hohenberg-Kohn theorem ..................................... 34
  3.2 Kohn-Sham equation .......................................... 35
  3.3 The FP-LMTO method ......................................... 39
  3.4 Pseudopotentials and the OPW method ....................... 40
  3.5 Projector Augmented-Wave (PAW) method ................... 44
  3.6 Quasiparticle excitations vs Kohn-Sham eigenvalues ........ 45
4 Balanced crystal orbital overlap population (BCOOP) ............. 49
  4.1 Definition of COOP and COHP ............................... 49
  4.2 Orbital population of an H$_2$-like molecule ................ 51
  4.3 Definition of BCOOP ........................................ 52
  4.4 BCOOP for Si, TiC, Ru, Na and NaCl ....................... 53
  4.5 Nb$_3$SiC$_2$ – a theoretically predicted new MAX phase .... 59
  4.6 Chemical bonding in MgB$_2$ .................................. 62
5 Hydrogen–hydrogen interaction and structural stability of Ti$_3$SnH$_x$ 65
  5.1 Hydrogen-metal and hydrogen-hydrogen interactions ......... 66
  5.2 Ti$_3$SnH$_x$ as a model system ............................... 67
  5.3 The H–H pair potential ...................................... 70
  5.4 Physics behind the H–H repulsion ........................... 73
6 Sammanfattning på svenska ....................................... 77
1 Introduction

What is physics? The best definition I can come up with is "Physics is the part of natural science which is not chemistry, biology, geology or any other discipline". And, I must add, physics uses a lot of math, in contrast to philosophy. So, physics studies everything about the world, except for the things belonging to other sciences. Physics also includes many branches: particle physics, solid state (or condensed matter) physics, nonlinear dynamics etc. This thesis only deals with the solid state physics.

Physics is divided into experimental, theoretical and computational physics. Experimental physics is the oldest kind of physics. It started with our everyday experience (an apple falls onto your head), and later evolved into true experiment (shake the apple tree to achieve certain effect). Theoretical physics uses mathematical language to systematize experimental data. Modern physics could not exist without precisely defined experimentally measurable quantities (e.g. velocity, mass, resistivity), precisely defined but not directly measurable ones (e.g. entropy, wave function), and also more intuitive, qualitative concepts (e.g. chemical bonding). We need this language to understand the results of any modern experiment (or even to perform one). With our apple example, we must first define mass, force and distance in order to come up with Newton’s law of gravity. A physical law is a generalization of the experimental data, usually written as a mathematical equation, which we believe to be valid under certain circumstances.

Once the laws are established, the true theoretical physics begins. The point is to predict new results from known laws without doing the experiment. In the process, new concepts, formalisms and even new laws are created. When we cannot get result with pen and paper, the computational physics, historically the latest of the three, comes into play. Computational physics is based on quantum mechanics. The Schrödinger equation for the nonrelativistic case and Dirac equation in the relativistic case are the basic equations of quantum mechanics. In principle, they describe all matter, from nuclei to galaxies. However, quantum-mechanical description employs wavefunction \( \Psi \) which depends on coordinates of all particles of the system. For solids, the number of particles \( N \) is of the order of \( 10^{23} \). Even the best supercomputers cannot treat functions depending on \( 10^{23} \) variables. The revolution in computational physics happened in 1960’s when the theory known as the Density Functional
Theory (DFT) was developed. DFT uses the particles density \( n(r) \) which depends only on 3 variables instead of \( 3N \) variables. DFT is the foundation of modern computational physics. Like in experiment, the end result of a computation is a bunch of numbers. The most interesting part of the computational work is analysis of the results, and it always involves a bit of theory.

Physical science works best when theoretical, experimental and computational physics all work together, compare their results all the time (and find them to disagree, which is the normal situation in practical work). It is really a pity when theorist, experimentalist and computationalist speak only their own language and don’t understand each other.

I do not perform any experiments. My work lies in between theory and computation: while performing calculations I create simple theoretical models whenever I need them. These “theoretical miniatures” are not as big and important as quantum mechanics or general relativity, but they suit well the problems I am trying to solve.

The first part of my thesis is devoted to the magnetism of two-dimensional (2D) systems. A 2D magnet is the heart of any magnetic data recording. Our work is based not on DFT but on the model Hamiltonian called Heisenberg Hamiltonian. We have studied thermodynamics of a 2D ferromagnet with the long-range dipolar interaction, namely we wanted to know the magnetization curve \( M(T) \) and the Curie temperature \( T_c \). The existing theoretical approaches were found to be insufficient for the task. Our task was to construct a formalism to treat both classical and quantum Heisenberg models with dipolar interaction. Note that for 2D systems one should use quantum and not classical Heisenberg model even for large spins such as \( S = 7/2 \).

The second part of my thesis deals with the problem of describing chemical bonding from the first principles. Chemical bonding is the force that holds molecules and solids together. Basically, it always comes from the Coulomb attraction between electrons and nuclei, however, in solids it can take many forms: covalent bonding, metallic bonding, ionic bonding etc. When studying a solid, it is very important to know the precise nature of the chemical bonding. We have found that existing chemical bonding indicators (COOP, COHP) do not work well with full-potential DFT methods. Therefore, we had to develop a less basis set dependent indicator which we now call Balanced Crystal Orbital Overlap Population (BCOOP). The test calculations for different systems (Si, Ru, TiC, Na and NaCl) with BCOOP has been performed.

We have also applied BCOOP to a possible new MAX phase \( \text{Nb}_3\text{SiC}_2 \). MAX phases are newly discovered layered materials with a great ability to withstand thermal shock and other unique properties. Our calculation show that \( \text{Nb}_3\text{SiC}_2 \) is unstable, but the formation energy is only about 0.2 eV/atom, therefore its existence in metastable form is likely. Another BCOOP example is \( \text{MgB}_2 \), a recently discovered superconductor with \( T_c = 39 \text{ K} \). This com-

2
pound is unique since it combines covalent, ionic and metallic bonding in a single solid.

Another project involves hydrogen–hydrogen interaction in metals. It is known that H atoms in a metal repel each other at short distances, with the minimal H–H distance being about 2Å. For the hydrogen storage applications, however, the materials with small H–H distances are desired. We have analyzed the H–H interaction using the pair-potential model. By the example of Ti$_2$SnH$_x$, we have studied how the H–H interaction stems from the metal–hydrogen chemical bonding and how it can be responsible for structural phase transitions in metal hydrides.
2 Thermodynamics of a 2D ferromagnet with dipolar interaction

In the early 80’s the ultrathin magnetic films and magnetic multilayers became a very active field of research[1]. Surprisingly, they are still a hot topic in 2005. That means that these two-dimensional (2D) magnetic systems have important technological applications. Indeed, such systems not only form the media of every hard disc, but also can be found in the read/write heads of the same disc. They are also crucial for the relatively novel field of spintronics[2, 3]. All these technological achievements stem from the unique physical properties of the 2D magnetic systems. For example, magnetic multilayers demonstrate oscillating exchange coupling [4] and giant magnetoresistance (GMR)[5]. Many people believe that GMR is the most important discovery in physics in the last 50 years.

But what is a 2D lattice? What is a quasi-2D lattice? Different types of 2D and quasi-2D crystal lattices are shown in Fig. 2.1 in the “cross section”. A 2D lattice is simply a crystal lattice periodical in two dimensions and confined in the third dimension. It is known that 2D and 1D crystals cannot be stable at nonzero temperature. Nature tries to find the equilibrium by folding the low-
D crystal. Indeed, 2D carbon sheets fold into fullerenes, while 1D polymer molecules (including proteins and DNA) fold into spirals and even more complicated structures. How can thin films (which are 2D systems) exist in real life then? The answer is that this theorem applies only to infinite 2D crystals, while stable films have finite size and they are also sufficiently thick (as compared to their other dimensions). Really thin films, on the other hand, must be deposited on a thicker substrate. One particular type of 2D lattice is the multilayer, which consists of several layers of different materials. A quasi-2D lattice is, strictly speaking, 3D and not 2D lattice. However, its layered crystal structure is strongly anisotropic. As a result, certain interactions (e.g. the exchange interaction responsible for the magnetism) are much stronger within each atomic layer than between the layers. In other words, the systems behaves like an array of 2D crystals only weakly interacting with each other. A special case of the quasi-2D lattice is the superlattice, the periodic multilayer.

The Heisenberg model is the most common theoretical model for magnetic materials. It can describe dynamics and thermodynamics of a magnetic system, in particular, we can calculate the magnetization curve $M(T)$ and the Curie temperature $T_c$. The spin variable on each lattice site can be either classical or quantum, therefore we speak of the classical Heisenberg model and the quantum Heisenberg model. The Heisenberg model, in general, has no exact solution. Therefore we need either a direct numerical solution (Monte Carlo method), or some sort of approximation.

It is known (Mermin-Wagner theorem[6]) that 2D and 1D magnetic systems cannot have any long-range magnetic order at $T > 0$, provided that only an isotropic, short-range exchange interaction is included (not a big surprise, really, when even the crystal lattice is unstable). In real life even low-D magnets still have a finite $T_c$ (although it is generally lower than the $T_c$ of the corresponding bulk materials). One reason for this is the finite size effect described above. A more common reason is that the magnetocrystalline anisotropy and the dipolar interaction, despite being “small”, have a crucial effect on the 1D or 2D system by breaking the conditions of the Mermin-Wagner theorem. Moreover, the latter interaction has a strong effect on spin waves in thin films[7, 8]. The competition between perpendicular anisotropy and the dipolar interaction often results in the reorientation transition[9]. While it is relatively easy to include anisotropy in most theoretical approaches to the Heisenberg model, the dipolar interaction presents much more challenge. While the classical 2D Heisenberg model with the dipolar interaction has been studied rather extensively[10–19], the quantum results are scarce. To the best of our knowledge, the only approaches applied in the quantum case are spin-wave theory (free magnons) [20–22] and Tyablikov approximation [23, 24]. The motivation of the present work (Paper I) is to study the latter case (thermodynamics of a quantum 2D Heisenberg model with dipolar interaction) in more detail.
The chosen formalism is the self-consistent spin-wave theory (SSWT), which did not include dipolar interaction until the present work.

Paper II presents a completely different approach to the 2D magnetism. Besides experimental results, it includes a density functional theory calculation of the samarium surface. It is included in the thesis to illustrate the variety of methods applied to the 2D magnetic systems.

2.1 Heisenberg model

There is no single model which would describe all magnetic systems and cover all physical effects. The Heisenberg model (see e.g. [25, 26]) is a simple theoretical model for a magnetic solid. This model assumes that each lattice site $i$, located at position $R_i$, bears a localized magnetic moment, described by a quantum-mechanical spin operator $S_i$ (or total angular momentum operator $J_i$).

The Heisenberg Hamiltonian consists of up to four terms
\[ H = H_{ex} + H_{an} + H_{dd} - \hbar \sum_i S_i. \] (2.1)

Let us look at each one in turn.

The exchange term has the form
\[ H_{ex} = -\frac{1}{2} \sum_{i \neq j} J_{ij} S_i S_j, \] (2.2)

where $J_{ij}$ is called exchange integral. There are various mechanisms for this interaction[25], and it is normally a short-range one, i.e. $J_{ij}$ should decrease sufficiently fast with distance $R_{ij} \equiv R_i - R_j$. Positive, or ferromagnetic, $J_{ij}$ favours parallel spin alignment, while the negative, or antiferromagnetic, $J_{ij}$ favours antiparallel spin alignment (although the AFM ground state is a difficult problem by itself). One needs to include exchange interaction with many neighbor shells (sometimes 100 or more) to describe magnetic solids quantitatively. However, theorists often study simpler models with only one or two shells included.

The exchange term is isotropic in the sense that it is invariant under any rotation applied to each and every spin in the system. Indeed, if the same rotation is applied to $S_i$ and $S_j$, then the scalar product $S_i S_j$ does not change. Therefore, directions in the spin space are completely uncoupled to the direction of the crystal lattice. This symmetry is broken by the anisotropy term, which can have various forms, for example
\[ H_{an} = -K \sum_i (S_i^z)^2 \] (2.3)
is called single-ion quadratic uniaxial anisotropy. Anisotropy comes from the relativistic effect called spin-orbit interaction.

The dipolar, or magnetostatic term

\[ H_{dd} = -\frac{1}{2} \sum_{i \neq j} Q_{ij}^{ab} \mathbf{S}_i^a \mathbf{S}_j^b, \]  

(2.4)

where

\[ Q_{ij}^{ab} = J_d \left( 3\mathbf{R}_{ij}^a \mathbf{R}_{ij}^b - \delta_{ab} \mathbf{R}_{ij}^2 \right) R_{ij}^{-5}, \]  

(2.5)

describes the magnetostatic interactions between two magnetic moments. This is essentially the same effect as the interaction between two permanent magnets. The interaction constant \( J_d \) is equal to \( J_d \equiv (g\mu_b)^2 \), where \( \mu_b \) is the Bohr magneton, and \( g \) is the gyromagnetic ratio \( (g = 2 \text{ for spin magnetic moments}) \). If we chose to measure length in the units of \( a \), with \( a \) of the order of the lattice constant, then vectors \( \mathbf{R}_{ij} \) become unitless, and \( J_d = (g\mu_b)^2/a^3 \) will have the same dimension as \( J_{ij} \) (energy). However, dipolar interaction is a long range one (namely it decreases as \( R^{-3} \)). It leads to new physical effects even though normally \( J_d \) is much smaller than the nearest-neighbor exchange integral.

The last term is simply the Zeeman coupling to the external magnetic field \( H \), and \( h \equiv -g\mu_b H \).

To summarize, the Heisenberg model assumes that

- The magnetic moment at each site is localized
- It comes from a single quantum-mechanical spin operator \( \mathbf{S} \) (or total angular momentum operator \( \mathbf{J} \)), and the spin value \( S \) is the fixed parameter of the model.

This model is applicable to most f-electron magnets (although one might need to use total angular momentum \( \mathbf{J} \) instead of the spin \( \mathbf{S} \)), most insulating d-electron magnets, and (to less extent) many metallic d-electron magnets (such as iron). However, other metallic d-electron magnets (such as nickel) have poorly localized magnetic moments. The picture opposite to the Heisenberg model is the itinerant picture, which considers the magnetism of nearly-free band electrons.

The classical Heisenberg model uses classical spin (a fixed-length vector) instead of the quantum-mechanical spin operator. It is often said that the classical model is valid if \( S \gg 1 \), however, this is not always true. We are going to come back to this question soon.

As we have mentioned above, the exact solution of the Heisenberg model is not known, and various approximations are used. Some approximations have limited applicability: for example they can be used only for classical model, or only for 2D lattices, or only for the model without dipolar interaction. There are several classes of approximate theories, which we are now going to mention briefly, focusing on the theories most suitable for the 2D systems.
Monte Carlo methods are formally exact and give a numerical solution to the Heisenberg model with any desired accuracy. Quantum Monte Carlo (QMC) calculations have been carried out for small values of spin (such as $S = 1/2$ and $S = 1$) for 2D systems [27–30], as well as for 3D systems. QMC calculations are quite resource demanding and they are not yet available for larger values of spin or for the Heisenberg model with a dipolar term. Classical Monte Carlo calculations are less time consuming and they are available even for 2D systems with dipolar interaction [11, 12, 15–17, 19]. A very interesting theory called pure-quantum self-consistent harmonic approximation (PQSCHA) [31, 32] gives a quantitative solution to the quantum Heisenberg model with a computational effort comparable to that of a classical Monte Carlo calculation (but no dipolar interaction yet).

One group of approximate theories is based on magnon operators (magnons are interacting bosons, the quanta of spin waves), introduced via Dyson-Maleev, Holstein-Primakoff, or some other bosonic transformation. In the next two sections I am going to speak about magnons and discuss the problem of extra states. Free-magnon (spin-wave, SW) theory is only a very rough starting point that normally overestimates $T_c$ by a factor of 2–4. In order to get better results, we must take magnon-magnon interaction into account. Interacting magnon theories use various small parameters, such as temperature, $1/S$ or $1/N$ of the SU(N) model. The magnon-magnon interaction is often treated as a perturbation and sometimes the Feynman diagram technique is used for magnons. In two dimensions, the most useful magnon-based theory is the so-called self-consistent spin-wave theory (SSWT). It was first formulated for the Mermin-Wagner situation [33–36], but it was later generalized to systems with long-range order [37, 38]. SSWT can be formulated as the best possible one-magnon theory [35, 38], the zeroth order term in the $1/N$ expansion of the SU(N) theory [36, 39] or as the mean-field magnon theory [38]. Note that here and in the following the words “mean field” are applied to magnon occupation number operators and have nothing to do with the Weiss mean field for spin operators. Loosely speaking, SSWT can be called “Hartree-Fock theory for magnons”. The SSWT result can be further improved by renormalizing the magnon-magnon vertex [38], often providing quantitative agreement with experiment everywhere except the narrow critical region. The known weak point of SSWT is the erroneous critical behavior in the vicinity of the phase transition: SSWT gives either a spin-wave second-order phase transition with

---

1 Rotations in the spin space (for any $S$) are given by the SU(2) group. Different values of spin correspond, therefore, to different representations of SU(2). SU(N) model generalizes the concept of spin and the Heisenberg model by using the SU(N) group instead. $1/N$ theories use $1/N$ as a formal small parameter and then put $N = 2$ in the final equations.

2 This approximation is often called the random phase approximation (RPA). Again, do not confuse with “RPA” for spin operators, which is the Tyablikov approximation.
\( \beta = 1^3 \), or a first-order transition (\( \beta = 0 \)). However, SSWT describes perfectly the wide region of the short-range order (SRO) above \( T_c \) [40], which is the unique feature of low-D systems. Until our work, the SSWT formalism did not include dipolar interaction.

Another group of theories is based directly on spin operators, without introducing any magnon operators. The simplest one is the Weiss mean-field theory. This theory is not good enough for low-D systems, since it is not consistent with the Mermin-Wagner theorem. Instead, it gives \( T_c \) of the order of \( JS^2 \) (from now on we measure temperature in the energy units). Other well-known approximations based on spin operators are Tyablikov[41] and Anderson-Callen[42] decouplings. There is even a perturbation theory for spin operators. But this time the unperturbed system is just the lattice of uncoupled spins in the external magnetic field, while all spin-spin interactions (including the exchange) are treated as perturbations. There is no Wick’s theorem for spin operators, therefore it is much more difficult to build Feynman diagram technique for spin operators than for magnon operators.

2.2 Magnons and Mermin-Wagner theorem

The most important elementary excitations of the Heisenberg model are magnons, the quanta of spin waves. Consider, for example, the ferromagnet with all spins pointing along the \( z \) axis in the ground state. In the Heisenberg representation of the quantum mechanics, spin dynamics is described by the equation of motion

\[
\frac{dS_j}{dt} = \frac{i}{\hbar} [H, S_j].
\] (2.6)

With the notation

\[
S_j^\pm = S_j^x \pm iS_j^y,
\] (2.7)

we seek the single-magnon solution of the form[25]

\[
S_j^- = \frac{A}{\sqrt{N}} e^{i(\omega_k t + kr_j)},
\] (2.8)

With the exchange-only Heisenberg Hamiltonian, the magnon energy as function of \( k \) (dispersion law) is

\[
E_k \equiv \hbar \omega_k = S(J_0 - J_k),
\] (2.9)

\( \beta \) is the critical exponent in the magnetization vs temperature dependence: \( M(T) \sim (T_c - T)^\beta \) when \( T \to T_c - 0 \).
where

\[ J_k = \sum_i e^{-ikR_i}J_{qi}. \quad (2.10) \]

Alternatively, this result can be obtained by the method of the next section (Holstein-Primakoff transformation).

There is one big difference between magnetism in 3D and 2D systems. The 3D Heisenberg model always has long-range magnetic order at sufficiently low temperatures, with a transition temperature of the order of \( J_S^2 \), where \( J \) is a typical value of the exchange integral. However, the Mermin-Wagner theorem[6] states that no 2D or 1D Heisenberg system can have long-range order at \( T > 0 \), provided that only isotropic short-range exchange interaction is included. Although the theorem can be proven rigorously, here we only present the simple explanation of this phenomenon given by Bloch[43] long before the paper by Mermin and Wagner.

If the temperature is sufficiently low, the magnons can be approximately treated as independent bosons (free magnon theory). In that case, the magnetization is given by

\[ \langle S_z \rangle = S - \frac{1}{V_{BZ}} \int_{BZ} \frac{d\mathbf{k}}{\exp(E_\mathbf{k}/T) - 1}. \quad (2.11) \]

The magnon energy (2.9) is proportional to \( k^2 \) for small \( k \), for example \( E_\mathbf{k} \approx JSk^2 \) for nearest-neighbor only exchange. For the 3D Heisenberg model

\[ \frac{d\mathbf{k}}{\exp(E_\mathbf{k}/T) - 1} \approx \frac{4\pi k^2 d\mathbf{k}}{JSk^2}, \quad (2.12) \]

and the integral in (2.11) converges at small \( k \), giving a finite number of magnons at low \( T \). In two dimensions however, this is not the case. Namely, for small \( k \)

\[ \frac{d\mathbf{k}}{\exp(E_\mathbf{k}/T) - 1} \approx \frac{2\pi kd\mathbf{k}}{JSk^2}, \quad (2.13) \]

thus the integral in (2.11) diverges logarithmically at the lower limit. This means that the magnetic order breaks down at temperature above zero.

This result directly contradicts the experiment. In real life even 1D and 2D magnetic systems have finite \( T_c \). The reason for this is the presence of additional interactions (anisotropy, dipolar interaction or interlayer exchange in quasi-2D systems), and also the finite size of the sample. Each of these factors breaks the conditions of the Mermin-Wagner theorem (see e.g. Ref. [38]), resulting in a finite \( T_c \ll J_S^2 \). It is interesting that the short-range order, does not disappear at or close to \( T_c \), as for the 3D systems, but stays up to \( T \sim J_S^2 \) (Ref. [40]) for the low-D systems.

In terms of Eq. (2.11) any of the additional interactions mentioned above
makes the integral (2.11) convergent in 2D. The easy-axis anisotropy simply introduces a gap $J S \Delta$ in the magnon spectrum, while for the dipolar interaction the situation is a bit more complicated (see below). In all cases we can introduce the "effective gap", or low energy cutoff $\Delta \ll 1$ and the magnetization is given by (for a 2D square lattice)

$$S - \langle S_z \rangle \approx \frac{T}{4\pi JS} \int_{J S \Delta}^{T} \frac{dE}{E} = \frac{T}{4\pi JS} \ln \left( \frac{T}{JS \Delta} \right),$$

(2.14)
giving a spin-wave expression for $T_c$

$$T_c \approx \frac{4\pi JS^2}{\ln (T_c/JS \Delta)} \ll 4\pi JS^2.$$  

(2.15)

Thus $T_c$ in the 2D case is indeed much smaller than in the 3D case, so we can still speak of the "Mermin-Wagner scenario". If we seek an approximate solution to the Heisenberg model, we must be sure that the approximation used is consistent with the Mermin-Wagner theorem. Some approximations, for example Weiss mean-field model, give finite $T_c$ even for 1D and 2D systems. Therefore these approximations cannot be applied to the low-dimensional case. In order to give the Mermin-Wagner behavior, the approximation in question must have a correct form of magnon dispersion law at low $T$ and $k \to 0$.

We are now back to the question of applicability of the classical Heisenberg model. First note that the upper integration limit in Eq. (2.14) comes from the Bose function, and it is equal to $T$ only for the quantum model with $T \ll JS$. For the classical model, or for the case $T \gg JS$, the Bose function can be replaced by $T/E$ and the upper limit is equal to $32JS$, the value that stems from the finite size of the Brillouin zone. The classical description is thus appropriate if the Curie temperature is much larger than any spin-wave frequency, namely if $T_c \gg JS$. For the 3D Heisenberg model $T_c \sim JS^2$, and this criterion takes the well-known form $S \gg 1$. That is why the classical Heisenberg model is immensely useful for the 3D systems, especially for the Monte-Carlo calculations. However, for the 2D Heisenberg model $T_c \ll JS^2$, and the classical description is only valid if $S \gg \ln (1/\Delta) \gg 1$ (see Ref. [38]). This does not hold even for the largest spins such as $S = 7/2$, therefore the quantum effects are never negligible for 2D magnetic systems, at least for a single monolayer. However, if we increase film thickness, there is a crossover from 2D to 3D behavior, and the classical description gets better.
2.3 Free magnons with dipolar interaction

Now that the long introductory part is over, it is time to formulate the model of Paper I, which this chapter is based upon. We consider the quantum Heisenberg model on the simple 2D square lattice in the $xz$ plane with lattice constant $a$. We include nearest-neighbor ferromagnetic exchange $J$ and the dipolar interaction $J_d = (g\mu_b)^2/a^3$ (and no anisotropy or Zeeman terms for simplicity). As in the previous sections, we take the lattice constant $a$ to be the unit of length. If $J_d$ is sufficiently small, the ground state is ferromagnetic with an $xz$ easy plane, and we take the $z$-axis direction for the ground state magnetization. The parameters of the model are therefore $J, J_d, S$ (the value of spin) and $T$ (temperature). The goal of this section is to study this model within the free-magnon (spin-wave) approximation. This has been first done by Maleev[20]. The spin-wave theory is only valid at $T \ll T_c$ and normally gives a very bad value for $T_c$, however this is a reasonable place to start.

The Hamiltonian (2.1) becomes

$$H = H_{ex} + H_{dd} = -\frac{1}{2}\sum_{i \neq j} \left( J_{ij} \delta_{\alpha\beta} + Q_{ij}^{\alpha\beta} \right) S_i^\alpha S_j^\beta,$$  

(2.16)

where

$$Q_{ij}^{\alpha\beta} = J_d \left( 3R_{ij}^\alpha R_{ij}^\beta - \delta_{\alpha\beta} R_{ij}^2 \right) R_{ij}^{-5}.$$  

(2.17)

With the notation (2.7) the scalar product $S_i S_j$ is equal to

$$S_i S_j = \frac{1}{2} \left( S_i^+ S_j^- + S_i^- S_j^+ \right) + S_i^z S_j^z,$$  

(2.18)

and the exchange term becomes

$$H_{ex} = -\frac{1}{2} \sum_{i \neq j} J_{ij} S_i S_j = -\frac{1}{2} \sum_{i \neq j} J_{ij} \left[ \frac{1}{2} \left( S_i^+ S_j^- + S_i^- S_j^+ \right) + S_i^z S_j^z \right] =$$

$$-\frac{1}{2} \sum_{i \neq j} J_{ij} \left[ S_i^- S_j^+ + S_i^+ S_j^- \right],$$  

(2.19)

since we can always assume $J_{ji} = J_{ij}$. 

13
The dipole-dipole part is equal to
\[ H_{dd} = \frac{1}{2} J_d \sum_{i \neq j} R^{-3}_{ij} \left[ S_i^- S_j^+ + S_i^+ S_j^- \right] - \]
\[ \frac{3}{2} J_d \sum_{i \neq j} R^{-5}_{ij} \left( S_i^- Z_{ij} + \frac{1}{2} S_i^+ R_{ij} + \frac{1}{2} S_i^- R_{ij} \right) \left( S_j^- Z_{ij} + \frac{1}{2} S_j^+ R_{ij} + \frac{1}{2} S_j^- R_{ij} \right) = \]
\[ \frac{1}{2} J_d \sum_{i \neq j} R^{-3}_{ij} \left[ S_i^- S_j^+ + S_i^+ S_j^- \right] - \frac{3}{2} J_d \sum_{i \neq j} R^{-5}_{ij} \left\{ S_i^- S_j^+ Z_{ij} + S_i^+ S_j^- Z_{ij} + S_i^- S_j^+ R_{ij} Z_{ij} + S_i^+ S_j^- R_{ij} Z_{ij} \right\}, \tag{2.20} \]

where
\[ R_{ij}^\pm \equiv X_{ij} \pm iY_{ij}. \tag{2.21} \]

Combining Eqs. (2.19) and (2.20) gives
\[ H = \sum_{i \neq j} \left\{ \Phi_{ij}^0 S_i^- S_j^+ + \Phi_{ij}^z S_i^- S_j^+ + \Phi_{ij}^+ S_i^+ S_j^- + \Phi_{ij}^- S_i^+ S_j^- \right\}, \tag{2.22} \]

where
\[ \Phi_{ij}^0 = -\frac{J_{ij}}{2} - \frac{1}{4} J_d R^{-3}_{ij} \left[ 1 - 3 \frac{Z_{ij}^2}{R_{ij}^6} \right], \tag{2.23} \]
\[ \Phi_{ij}^z = -\frac{J_{ij}}{2} + \frac{1}{2} J_d R^{-3}_{ij} \left[ 1 - 3 \frac{Z_{ij}^2}{R_{ij}^6} \right], \tag{2.24} \]
\[ \Phi_{ij}^\pm = -\frac{3}{8} J_d \frac{R_{ij}^2}{R_{ij}^5}, \tag{2.25} \]
\[ \Phi_{ij}^{\pm} = -\frac{3}{8} J_d \frac{R_{ij}^2 Z_{ij}}{R_{ij}^3}. \tag{2.26} \]

The expression (2.22) is, in fact, still valid for any 2D or 3D crystal lattice, but from now on we restrict ourselves to the 2D square net.

After the Fourier transform
\[ S_k^\alpha = \frac{1}{\sqrt{N}} \sum_R e^{-i k R} S_i^\alpha, \quad S_i^\alpha = \frac{1}{\sqrt{N}} \sum_k e^{i k R} S_k^\alpha, \tag{2.27} \]
with

\[ \sum_k \equiv V \int \frac{d^2k}{(2\pi)^2}, \quad \sum_k 1 = N, \]  

(2.28)

the Hamiltonian becomes

\[ H = \sum_k \left\{ \Phi_k^0 S^+_k S^-_{-k} + \Phi_k^+ S^+_k S^-_{-k} + \Phi_k^- S^-_k S^+_{-k} \right. \]
\[ + \left. \Phi_k^+ S^-_k S^-_{-k} + \Phi_k^- S^+_k S^-_{-k} + \Phi_k^+ S^-_k S^-_{-k} \right\}, \]  

(2.29)

where various \( \Phi_k \) are Fourier transformed versions of respective \( \Phi_{ij} \)

\[ \Phi_k = \sum_{R_i \neq 0} e^{-ikR_i} \Phi_{0i}, \quad \Phi_{ij} = \frac{1}{N} \sum_k e^{ik(R_j - R_i)} \Phi_k. \]  

(2.30)

They are equal to

\[ \Phi_k^0 = -\frac{J_k}{2} - \frac{1}{4} J_d \left[ S_1(k) - \frac{1}{2} S_3 \right], \]  

(2.31)

\[ \Phi_k^+ = \Phi_k^- = -\frac{3}{8} J_d \left[ S_2(k) + \frac{1}{2} S_3 \right], \]  

(2.33)

\[ \Phi_k^{z+} = \Phi_k^{z-} = -\frac{3}{2} J_d \sum_{R_i \neq 0} \frac{e^{ikR_i} X_i Z_i}{R_i^3}, \]  

(2.34)

where three lattice sums have been introduced

\[ S_3 \equiv \sum_{R_i \neq 0} R_i^{-3} \approx 9.034, \]  

(2.35)

\[ S_1(k) \equiv \sum_{R_i \neq 0} \left( e^{ikR_i} - 1 \right) R_i^{-3} \left[ 1 - \frac{3 Z_i^2}{R_i^2} \right], \]  

(2.36)

\[ S_2(k) \equiv \sum_{R_i \neq 0} \left( e^{ikR_i} - 1 \right) \frac{X_i^2}{R_i^3}, \]  

(2.37)

and

\[ J_k \equiv \sum_i e^{-ikR_i} J_{0i}. \]  

(2.38)
For the nearest neighbor exchange $J_k$ is equal to $2J \cos k_x + 2J \cos k_z$. From now on and to the rest of this chapter our general equations are valid for any form of $J_{ij}$ (and therefore $J_k$), but all asymptotics and numerical results are presented for the nearest-neighbor exchange. For small $k$ the lattice sums (2.36),(2.37) have the asymptotical form

$$S_1(k) \approx 2\pi \frac{k^2}{k}, \quad S_2(k) \approx -\frac{2\pi}{3k} (2k_x^2 + k_z^2).$$ (2.39)

Remember that the particular expressions for the different $\Phi_k$-s and lattice sums above are only valid for the simple square lattice.

Now we formally introduce magnons by the Holstein-Primakoff transformation

$$S_+^i = \sqrt{2S} \left(1 - a_i^\dagger a_i / 2S\right)^{1/2} a_i$$
$$S_-^i = \sqrt{2Sa_i^\dagger} \left(1 - a_i^\dagger a_i / 2S\right)^{1/2} a_i^\dagger$$
$$S_z^i = S - a_i^\dagger a_i,$$ (2.40) (2.41) (2.42)

and expand the Hamiltonian (2.16) into the series of $S^{-1/2}$

$$H = S^2 N_0 + S^4 N_2(a_i^\dagger, a_k) + S^{1/2} N_3(a_i^\dagger, a_k) + S^4 N_4(a_i^\dagger, a_k) + S^{-1/2} N_5(a_i^\dagger, a_k) + \ldots,$$ (2.43)

where $N_n(a_i^\dagger, a_k)$ means a certain $n$-th order polynomial of the Bose operators $a_i^\dagger$, $a_k$ in the normal form (creation operators to the left). This expansion gives rise to the problem of extra states (with $\langle S_z \rangle < -S$) and it does not work at all for e.g. systems with easy-plane anisotropy. The problem of extra states has been studied in Ref. [38] using pseudo-fermions and was found to be not crucial for most 2D systems. The question of convergence of the series (2.43) is irrelevant since we are only using terms up to $N_4$ in SSWT (see below).

The free-magnon Hamiltonian is

$$H_0 \equiv S^4 N_2(a_k^\dagger, a_k) - \mu \sum_k a_k^\dagger a_k$$
$$= \sum_k \left\{ A^0_k a_k^\dagger a_k + \frac{1}{2} B^0_k a_k^\dagger a_{-k} + \frac{1}{2} B^0_{-k} a_k a_{-k} \right\},$$ (2.44)

where

$$A^0_k = S(J_0 - J_k) - \frac{1}{2} J_d S \left[ S_1(k) - \frac{3}{2} S_3 \right] - \mu$$ (2.45)
and

\[ B^0_k = -\frac{3}{2} J_d S \left[ S_2(k) + \frac{1}{2} S_3 \right]. \] (2.46)

The “magnon chemical potential” \( \mu \) is the Lagrange multiplier used in spin-wave theory and SSWT to enforce the condition \( \langle S_z \rangle = 0 \) in the paramagnetic (PM) phase (in the ferromagnetic phase one has \( \mu = 0 \)). The next step is to get rid of the “anomalous” terms \( a^\dagger_k a^\dagger_{-k} \) and \( a_k a_{-k} \) by the Bogoliubov transformation

\[
\begin{align*}
\{ & a_k = \cosh(\xi_k) b_k - \sinh(\xi_k) b^\dagger_{-k} \\
& a^\dagger_k = \cosh(\xi_k) b^\dagger_k - \sinh(\xi_k) b_{-k}
\} \tag{2.47}
\end{align*}
\]

with

\[ \tanh(2 \xi_k) = \frac{B^0_k}{A^0_k}. \] (2.48)

The Hamiltonian in the new magnon operators \( b^\dagger_k, b_k \) has a trivial free-bosons form

\[ H_0 = \text{const} + \sum_k \varepsilon^0_k b^\dagger_k b_k, \quad \varepsilon^0_k = \sqrt{\left( A^0_k \right)^2 - \left| B^0_k \right|^2}, \] (2.49)

therefore the thermodynamic average number of the “new magnons” is given simply by the Bose function

\[ \langle b^\dagger_k b_k \rangle = N_k \equiv \left[ \exp(\varepsilon^0_k/T) - 1 \right]^{-1}. \] (2.50)

The thermodynamic averages involving the “old magnon” operators are found using the Bogoliubov transformation (2.47)

\[
\begin{align*}
\langle a^\dagger_k a_k \rangle &= \frac{A^0_k}{\varepsilon^0_k} \left( \langle b^\dagger_k b_k \rangle + \frac{1}{2} \right) - \frac{1}{2}, \quad \tag{2.51} \\
\langle a^\dagger_k a_{-k} \rangle &= \langle a_k a_{-k} \rangle = -\frac{B^0_k}{\varepsilon^0_k} \left( \langle b^\dagger_k b_k \rangle + \frac{1}{2} \right). \tag{2.52}
\end{align*}
\]

Alternatively, the expectation value (2.51) can be obtained from the free-magnon Matsubara Green’s function (as it has been done in Ref. [20])

\[ G^0_k(i\omega_n) = \frac{i\omega_n + A^0_k}{(i\omega_n)^2 - (\varepsilon^0_k)^2}, \quad \omega_n \equiv 2\pi n T \] (2.53)
through the frequency summation
\[
\left\langle a^+_k a_k \right\rangle = \lim_{\tau \to +0} T \sum_{i, \omega_n} e^{i \omega_n \tau} G^0_k (i \omega_n).
\] (2.54)

If \( J_d = 0 \), then \( B^0_k = 0 \) and \( A^0_k = \epsilon^0_k \) is given by Eq. (2.9).

The magnetization is given by
\[
<S_z> \equiv \bar{S} = S - \frac{1}{N} \sum_k \left\langle a^+_k a_k \right\rangle = S - \frac{1}{N} \sum_k \left[ \frac{A^0_k}{\epsilon^0_k} \left( N_k + \frac{1}{2} \right) - \frac{1}{2} \right].
\] (2.55)

Let us define
\[
j_d \equiv J_d / J.
\] (2.56)

For the case \( j_d \ll 1 \) and in the quantum regime \( (JS_j)^{3/2} \ll T \ll JS \) the free-magnon (SW) magnetization is approximately equal to (after a challenging textbook exercise in integration)
\[
\bar{S} = S - \frac{T}{4\pi JS} \ln \left[ \frac{2T}{\pi JS \sqrt{\pi j_d}} \right] \quad \text{(SW)},
\] (2.57)

where \( f \equiv (3/8\pi)S_3 \approx 1.078 \), and our notation corresponds to that of Ref. [20] as
\[
D = JS, \quad \Omega_0 = 2\pi S j_d, \quad \alpha = 2f = (3/4\pi)S_3.
\] (2.58)

It gives the equation for the free-magnon \( T_c \) as
\[
\frac{4\pi JS^2}{T_c} = \ln \left[ \frac{4S}{\sqrt{\pi} j_d} \right] - \ln \left[ \frac{T_c}{4\pi JS^2} \right] \quad \text{(SW)}.
\] (2.59)

Note that the convergence of the integral in (2.55) at \( k \to 0 \) limit is ensured not only by tricky \( |k| \) dependence of \( \epsilon_k \) at \( k \to 0 \) (as for the case of anisotropy), but rather by both radial and angular dependence of \( \epsilon_k \) together with the fact that the expectation value (2.51) is no longer just Bose function.

For the classical Heisenberg model the Bose function is replaced by
\[
N_k \to \frac{T}{\epsilon_k}.
\] (2.60)

The analytical expressions for \( \bar{S} \) and \( T_c \) are obtained by replacing \( T / JS \to 32 \) under the logarithm in Eqs. (2.57),(2.59), yielding the classical spin-wave expressions
\[
\bar{S} = S - \frac{T}{4\pi JS} \ln \left[ \frac{32}{\pi \sqrt{\pi} j_d} \right] \quad \text{(Classic SW)},
\] (2.61)
The free-magnon (SW) transition temperature $T_c$ versus dipolar interaction $J_d$. The symbols are numerical results, while the curves are the asymptotical formulas (2.59), (2.62).

$$\frac{4\pi JS^2}{T_c} = \ln \left[ \frac{32}{\sqrt{\pi} J_d^{3/2}} \right] \quad \text{(Classic SW)}.$$ (2.62)

In Figure 2.2 the free-magnon transition temperature is presented as a function of $j_d$ for three values of spin: $S = 1/2$, $S = 7/2$, and the classical spin. One can immediately see that the quantum asymptotic expression (2.59) works very well for small $j_d$ and small $S$ (but not for $S = 7/2$). The classical asymptotic form (2.62) is also very good at small $j_d$. It can also be observed that even for such a large spin as $S = 7/2$, the transition temperature still differs by about 10% from its classical value, in agreement with our discussion in the previous section.

### 2.4 Self-consistent spin-wave theory

Self-consistent spin-wave theory (SSWT) can be most easily formulated using the Feynman-Peierls-Bogoliubov variational principle [41]. For any Hamiltonian $H$ and any trial Hamiltonian $H_t$, the free energy $F = -\ln \text{Tr}(e^{-\beta H})$ satisfies the inequality

$$F \leq F' = F_t + \langle H - H_t \rangle,$$ (2.63)

19
where $F_t$ and the expectation values are calculated using $H_t$. This inequality becomes equality if $H_t = H$. SSWT is defined as the best possible one-magnon theory (according to this variational principle). Namely, we take $H_t$ to have the free-magnon form with arbitrary dispersion law. For a ferromagnet in the absence of dipolar interaction this means

$$H_t = \sum_k E_k a_k^\dagger a_k,$$

(2.64)

but in our case we must also include the “anomalous” terms, and

$$H_t = \sum_k \left\{ A_k a_k^\dagger a_k + \frac{1}{2} B_k a_k^\dagger a_{-k}^\dagger + \frac{1}{2} B_k^* a_k a_{-k} \right\},$$

(2.65)

where $A_k$ and $B_k$ are variational functions. They are found from the variational equations

$$\frac{\delta F'}{\delta A_k} = 0, \quad \frac{\delta F'}{\delta B_k^*} = 0.$$

(2.66)

This procedure is analogous to the Hartree-Fock method for fermions, and it can be shown to be equivalent to the mean-field (MF) treatment of the magnon-magnon interaction. This simplest, “mean-field” form of SSWT does not always give meaningful results. For example, it works fine for exchange-only 2D and quasi-2D Heisenberg model, but fails in the presence of anisotropy [38]. We cannot know in advance whether the MF theory is applicable to the case of dipolar interaction.

In the following, we include only the 3- and 4-operator terms in the magnon-magnon interaction: $H \equiv H + V$, where $V \equiv \frac{S^{1/2}N_3}{2}(a_k^\dagger a_k) + S^0N_4(a_k^\dagger a_k)$ (and the $N_3$ term does not give any contribution). Such truncation of the Hamiltonian can be justified by comparison with the case of SSWT without dipolar interaction, where the truncated Holstein-Primakoff Hamiltonian is equivalent to the Dyson-Maleev Hamiltonian. For the dipolar case, the Dyson-Maleev representation is not suitable due to the essentially non-Hermitian form of the Hamiltonian derived, and we use the truncated Holstein-Primakoff Hamiltonian instead.
Thus the magnon-magnon interaction is
\[ V \equiv S^{1/2}N_3(a_k^+a_k) + S^0N_4(a_k^+a_k) \]
\[ = \sum_q \left\{ \Phi_q^0 \left[ -\frac{1}{2N} \sum_{k_1,k_2} a_{k_1}^+a_{k_2}^+a_{k_2}a_{k_1-k_2} - \frac{1}{2N} \sum_{k_1,k_2} a_{k_1}^+a_{k_2}^+a_{k_2}a_{k_1+k_2-a-q} \right] \right. \\
+ \Phi_q^+ \left[ \frac{1}{N} \sum_{k_1,k_2} a_{k_1}^+a_{k_2}^+a_{k_1+q}a_{k_2-q} \right] \\
+ \Phi_q^- \left[ -\frac{1}{2N} \sum_{k_1,k_2} a_{k_1}^+a_{k_2}a_{k_2-k_1-a-q} - \frac{1}{2N} \sum_{k_1,k_2} a_{k_1}^+a_{k_2}a_{k_2}a_{k_1-k_2} \right] \\
+ \Phi_q^{*-} \left[ -\sqrt{\frac{2S}{N}} \sum_k a_{k}^+a_{k} + \Phi_q^+ \left[ -\sqrt{\frac{2S}{N}} \sum_k a_{k}^+a_{k+q} \right] \right\}. \tag{2.67} \]

From now on we normalize the Hamiltonian per one atom \((H \rightarrow H/N)\), and also redefine the BZ integration as
\[ \sum_k \equiv \int \frac{d^3k}{(2\pi)^3}, \quad \sum_k 1 = 1. \tag{2.68} \]

The mean field Hamiltonian takes the form (2.65) with
\[ A_k^{MF} = S \left[ J_0 - J_k - \frac{1}{2} J_d S_1(k) + \frac{3}{4} J_d S_3 \right] - \mu + \frac{3}{4} J_d \sum_q \langle a_{q}a_{-q} \rangle \left[ 2S_2(q) + S_2(k) + \frac{3}{2} S_3 \right] + \sum_q \langle a_{q}^+a_{q} \rangle \left[ J_q + J_k - J_q - J_{q-k} + \frac{1}{2} J_d S_1(q) + \frac{1}{2} J_d S_1(k) + J_d S_1(q-k) - \frac{3}{2} J_d S_3 \right], \tag{2.69} \]
\[ B_k^{MF} = -\frac{3}{2} J_d \left[ S_2(k) + \frac{1}{2} S_3 \right] + \frac{3}{4} J_d \sum_q \langle a_{q}^+a_{q} \rangle \left[ 2S_2(k) + S_2(q) + \frac{3}{2} S_3 \right] + \sum_q \langle a_{q}a_{-q} \rangle \left[ J_k - \frac{1}{2} J_q - J_{q-k} + \frac{1}{4} J_d S_1(k) + \frac{1}{4} J_d S_1(q) + J_d S_1(q-k) - \frac{3}{4} J_d S_3 \right], \tag{2.70} \]

and \((B_k^{MF})^* = B_k^{MF}\). This Hamiltonian is meaningful provided that \(|B_k^{MF}| \leq |A_k^{MF}|\) in the entire Brillouin zone. It can be diagonalized by the Bogoliubov transformation (2.47)–(2.52), with \(A_k^{MF}, B_k^{MF}\) instead of \(A_k^0, B_k^0\). Note that, while \(A_k^{MF}\) and \(B_k^{MF}\) are functionals of the expectation values \(\langle a_{k}^+a_{k} \rangle\) and \(\langle a_{k}a_{-k} \rangle\),
the expectation values in turn depend on $A_k^{MF}$ and $B_k^{MF}$ via Eqs. (2.51)–(2.52). This means that we have a system of equations which must be solved in a self-consistent cycle.

Unfortunately, our numerical tests reveal that this system of MF equations has no physically reasonable solutions (except for very high $T$ in the paramagnetic phase) in the presence of dipolar interaction. This is not very surprising, knowing that the MF theory does not work for anisotropic FM either. We need to construct a more elaborate theory, for example to restrict the variational freedom of the functions $A_k$ and $B_k$ to a certain functional form. We have considered two different ways to construct such a theory.

In the first one (we call it $\gamma\delta$-model) we restrict the function $A_k, B_k$ in Eq. (2.65) to the free-magnon form (2.45),(2.46) with exchange and dipolar interactions renormalized by parameters $\gamma$ and $\delta$ respectively

$$J \rightarrow \gamma J, \quad J_d \rightarrow \delta J_d.$$ (2.71)

The variational freedom is therefore reduced from two trial functions to just two real numbers. Note that the free-magnon form of $A_k, B_k$ gives a dispersion law that has no gap at $k \rightarrow 0$. The trial Hamiltonian of the $\gamma\delta$-model is

$$H_t = \sum_k \left\{ A_k^\dagger a_k^\dagger a_k + \frac{1}{2} B_k^\dagger a_k^\dagger a_{-k}^\dagger + \frac{1}{2} B_k a_k^\dagger a_{-k} \right\},$$ (2.72)

$$A_k^\dagger = \gamma S (J_0 - J_k) - \frac{1}{2} \delta J_d S \left[ S_1(k) - \frac{3}{2} S_3 \right] - \mu,$$ (2.73)

$$B_k^\dagger = -\frac{3}{2} \delta J_d S \left[ S_2(k) + \frac{1}{2} S_3 \right].$$ (2.74)

Since $\gamma$ renormalizes the short-range exchange interaction, it has the physical meaning of a short-range order (SRO) parameter. In the absence of the dipolar interaction, the nearest-neighbor spin correlation function is equal to [38]

$$\langle S_i S_{i+\delta} \rangle = \gamma^2.$$ (2.75)

For $0 < j_d \ll 1$ the equality (2.75) is no longer exact, but it still holds to a high degree of accuracy. The parameter $\delta$ renormalizes the long-range dipolar interaction and has the meaning of some long-range order parameter, different from $\tilde{S}/S$.

The variational procedure now constitutes of minimizing the trial free energy $F'$ defined by (2.63) with respect to two parameters $\gamma$ and $\delta$. The varia-
tional equations are

\[ 0 = \frac{\partial F'}{\partial \gamma} = \sum_k \left\{ A_k \frac{\partial \langle a_k^\dagger a_k \rangle}{\partial \gamma} + B_k \frac{\partial \langle a_k a_{-k} \rangle}{\partial \gamma} - \epsilon_k \frac{\partial N_k}{\partial \gamma} \right\}, \quad (2.76) \]

\[ 0 = \frac{\partial F'}{\partial \delta} = \sum_k \left\{ A_k \frac{\partial \langle a_k^\dagger a_k \rangle}{\partial \delta} + B_k \frac{\partial \langle a_k a_{-k} \rangle}{\partial \delta} - \epsilon_k \frac{\partial N_k}{\partial \delta} \right\}, \quad (2.77) \]

where the functionals \( A_k^{MF} \) and \( B_k^{MF} \) have been defined above. The equations (2.76), (2.77) should be solved self-consistently together with Eqs. (2.50)–(2.52). The Bogoliubov transformation should now employ \( A'_k \) and \( B'_k \), of course.

However, for reasons stated below, we are going to concentrate on the second approach to SSWT, which we call the \( \gamma\bar{s}^2 \)-model. In this approach we give up attempts to obtain \( \delta \) from the SSWT equations. Instead, we renormalize the dipolar interaction with a phenomenological multiplier \( \bar{s}^2 \equiv (\bar{S}/S)^2 \)

\[ J_d \rightarrow J_d \bar{s}^2 \quad (2.78) \]

in the original Hamiltonian\(^4\) and ignore the dipolar contribution to the magnon-magnon interaction \( N_3(a_k^\dagger, a_k) + S^2 N_4(a_k^\dagger, a_k) \).

This renormalization “by hand” requires a physical explanation. The effective dipolar interaction can, generally speaking, have different temperature-dependent renormalization for different distances \( R_{ij} \). Since the systematic attempt to build a \( R_{ij} \)-dependent renormalization (magnon mean-field theory) does not seem to work, some other approximation is required. In particular, for \( j_d \ll 1 \) one can neglect the specific character of the short-range dipolar interactions, since they are negligible compared to the short-range exchange interaction, and construct a renormalization which is valid in the \( R_{ij} \rightarrow \infty \) limit. In the latter limit, the macroscopic theory can be applied, and therefore the effective dipolar interaction is proportional to the square of magnetization, i.e. Eq. (2.78). According to this approximation the effective dipolar interaction vanishes in the paramagnetic (PM) phase. In real life, it does not vanish, but it becomes a short-range one (due to the finite correlation length), and can be neglected compared to the exchange interaction if \( j_d \ll 1 \). The approximation (2.78) is very similar to the way the anisotropy is treated in Ref. [38].

\(^4\)If we wanted to truly follow the scheme of the constrained variational approach, we would put \( \delta \equiv \bar{s}^2 \) and still solve Eq. (2.76) of the \( \gamma\delta \)-theory for \( \gamma \). The \( \gamma\bar{s}^2 \)-model described in the text gives almost identical results, but it is numerically more efficient.
The initial Hamiltonian of the $\gamma \bar{s}^2$ model is therefore
\[
H = \sum_k \left\{ \tilde{A}_k^0 a_k^\dagger a_k + \frac{1}{2} \tilde{B}_k^0 a_k^\dagger a_{-k} + \frac{1}{2} \tilde{B}_k^0 a_k a_{-k} \right\} + \tilde{V}, \tag{2.79}
\]
where
\[
\tilde{A}_k^0 = S(J_0 - J_k) - \frac{1}{2} \bar{s}^2 J_d S \left[ S_1(k) - \frac{3}{2} S_3 \right] - \mu, \tag{2.80}
\]
and
\[
\tilde{B}_k^0 = -\frac{3}{2} \bar{s}^2 J_d S \left[ S_2(k) + \frac{1}{2} S_3 \right], \tag{2.81}
\]
is the exchange-only version of the magnon-magnon interaction vertex (2.67).

The variational equation is
\[
0 = \frac{\partial F'}{\partial \gamma} = \sum_k \left\{ \tilde{A}_k^{MF} \frac{\partial \langle a_k^\dagger a_k \rangle}{\partial \gamma} + \tilde{B}_k^{MF} \frac{\partial \langle a_k a_{-k} \rangle}{\partial \gamma} - \tilde{V}_k \frac{\partial N_k}{\partial \gamma} \right\}, \tag{2.85}
\]
where
\[
\tilde{A}_k^{MF} = S(J_0 - J_k) - \frac{1}{2} \bar{s}^2 J_d S \left[ S_1(k) - \frac{3}{2} S_3 \right] - \mu + \sum_q \langle a_q^\dagger a_q \rangle [J_q + J_k - J_0 - J_{q-k}] \tag{2.86}
\]
and
\[
\tilde{B}_k^{MF} = -\frac{3}{2} \bar{s}^2 J_d S \left[ S_2(k) + \frac{1}{2} S_3 \right] + \sum_q \langle a_q a_{-q} \rangle \left[ \frac{J_k}{2} + \frac{J_q}{2} - J_{q-k} \right], \tag{2.87}
\]
Equation (2.85) should be solved self-consistently in order to obtain $\gamma$ and $\bar{s}$.
2.5 Calculation of the lattice sums: Ewald method

In order to solve the SSWT equations numerically, we have developed an ad hoc computer code. Many technical problems appeared in the process of the code development. One of them was the problem of calculating lattice sums $S_1(k)$, $S_2(k)$ and $S_3$ efficiently. It is a bad idea to calculate such sums directly since the convergence is very slow. Instead we used a common trick known as the Ewald method. We demonstrate the technique by the example of the sum

$$
\sum_{R_i \neq 0} \frac{e^{ikR_i}}{R_i^k},
$$

where $n > 2$. First note the identity

$$
R^{-n} = \frac{2}{\Gamma(n/2)} \int_0^\infty d\rho e^{-\rho^2 R^2} \rho^{n-1},
$$

which follows directly from the definition of the $\Gamma$-function. We can split the integration range with an arbitrary parameter $\eta > 0$

$$
R^{-n} = \frac{2}{\Gamma(n/2)} \times \left\{ \int_0^\eta d\rho \rho^{n-1} \sum_{R_i \neq 0} e^{ikR_i - \rho^2 R_i^2} + \int_\eta^\infty d\rho \rho^{n-1} \sum_{R_i \neq 0} e^{ikR_i - \rho^2 R_i^2} \right\}.
$$

The sum (2.88) is equal to

$$
\sum_{R_i \neq 0} \frac{e^{ikR_i}}{R_i^k} = \frac{2}{\Gamma(n/2)} \left\{ \int_0^\eta d\rho \rho^{n-1} \sum_{R_i \neq 0} e^{ikR_i - \rho^2 R_i^2} + \int_\eta^\infty d\rho \rho^{n-1} \sum_{R_i \neq 0} e^{ikR_i - \rho^2 R_i^2} \right\},
$$

If $\eta$ is of the order of unity, then the second term in this expression includes a rapidly convergent sum, but we still have to do something with the first term. It can be made rapidly convergent using the Fourier transform with respect to the variable $R_i$. The final expression for the sum (2.88) is

$$
\sum_{R_i \neq 0} \frac{e^{ikR_i}}{R_i^k} = \frac{2}{\Gamma(n/2)} \left\{ \pi \int_0^\eta d\rho \rho^{n-1} \sum_G \exp \left[ -\frac{(k - G)^2}{4\rho^2} \right] - \frac{\eta^n}{n} + \int_\eta^\infty d\rho \rho^{n-1} \sum_{R_i \neq 0} \exp \left( i k R_i - \rho^2 R_i^2 \right) \right\},
$$

where $G$ are the reciprocal lattice vectors.

Instead of one slowly convergent sum we got two rapidly convergent sums under the integral sign. Even though the numerical integration requires eval-
uating these sums many times for different values of $\rho$, the efficiency is increased greatly. The sums $S_1, S_2$ and $S_3$ are directly related to the sum (2.88), for example

$$\tilde{S}_2(k) \equiv \sum_{\mathbf{R} \neq 0} e^{i \mathbf{k} \mathbf{R}} \frac{X_2}{R^5_i} = - \frac{\partial^2}{\partial k^2} \sum_{\mathbf{R} \neq 0} \exp(i \mathbf{k} \mathbf{R}) R^{-5}_i, \quad (2.93)$$

and

$$S_2(k) = \tilde{S}_2(k) - \tilde{S}_2(0). \quad (2.94)$$

This technique gives the value of $S_3 = 9.03362178$ (cf. Ref. [19]). However, in order to make the SSWT code yet more efficient, we have parametrized the lattice sums $S_1(k)$ and $S_2(k)$ over the entire Brillouin zone (BZ). Our analytical expressions can be viewed as an improvement of the earlier expressions (2.39). They have the correct asymptotical form for $k \to 0$ (up to the $k^2$ terms) and an 1% accuracy over the entire BZ.

![Figure 2.3: SSWT relative magnetization $\bar{s}$ and short-range order parameter $\gamma$ versus temperature for $S = 1/2$ and $j_d = 10^{-3}$. For comparison, SW magnetization for $j_d = 10^{-3}$, and $\gamma$ from SSWT for $j_d = 0$ (Mermin-Wagner situation) are also shown.](image)

### 2.6 SSWT results

Finally we are ready to present results of our SSWT calculations. We start with a particular example ($S = 1/2$ and $j_d = 10^{-3}$). Next, we construct the analytical form of SSWT valid in the $j_d \to 0$ limit and finally we investigate
the spin and \( j_d \) dependence of \( T_c \).

In Figure 2.3 we present the relative magnetization curves \( \bar{s}(T) \) given by \( \gamma \bar{s}^2 \) SSWT and SW for \( S = 1/2 \) and \( j_d = 10^{-3} \). The two magnetization curves are rather different. SW theory gives an almost linear \( \bar{s}(T) \) dependence and a spin-wave phase transition (second-order phase transition with \( \beta = 1 \)). On the contrary, SSWT gives a first-order phase transition (formally \( \beta = 0 \)). This means that the magnetization reaches a finite minimal value \( \bar{s}_{\text{min}} \approx 0.199 \) at \( T_c/J \approx 0.1976 \). After that point the ferromagnetic solution to the SSWT equations ceases to exist abruptly and the system goes to the paramagnetic state. Both kinds of critical behavior are completely nonphysical. However, outside the narrow critical region, SSWT is definitely superior to SW theory, and the SSWT \( T_c \) is much smaller than the obviously overestimated SW \( T_c \). In particular, all realistic (experimental and Monte Carlo) magnetization curves have a sharp fall at \( T \to T_c \) and resemble much more the SSWT curve, with a step, compared to the linear SW curve. In other words, a 2D critical behavior (e.g., with \( \beta = 1/8 \)) is much closer to \( \beta = 0 \) than to \( \beta = 1 \).

The SRO parameter \( \gamma \) as the function of \( T \) is also presented in Fig. 2.3 for \( j_d = 10^{-3} \) and for \( j_d = 0 \) (Mermin-Wagner situation). It is close to unity in a wide range of temperatures, until it finally falls to zero at \( T_{\text{SRO}}/J \approx 0.75 \). Thus SSWT describes correctly the experimentally confirmed[40] wide region with considerable short-range order above \( T_c \). The two \( \gamma(T) \) curves practically coincide, hence SRO is rather insensitive to the strength of the dipolar interaction and to the presence or absence of the long-range order. For \( S = 1/2 \), \( j_d = 10^{-3} \) we have \( \gamma(T_c) \approx 0.989 \), therefore we can say that practically \( \gamma = 1 \) up to \( T_c \). However, for larger (and classical) spins \( \gamma(T_c) \) takes values of the order of 0.7–0.9, depending on \( j_d \). In the latter case, SSWT renormalization of the exchange interaction (i.e. \( \gamma \)) and not only of the dipolar interaction (i.e. \( \bar{s}^2 \)) is important. The same trend has been observed earlier for quasi-2D magnets, see Fig. 3 of Ref. [38], which shows stronger \( \gamma(T) \) dependence for larger values of \( S \).

In the \( j_d \to 0 \) limit the \( \gamma \bar{s}^2 \)-SSWT takes a particularly simple analytical form. We can put \( \gamma = 1 \) and the SSWT magnetization is given by the Maleev’s formula (2.57) with \( J_d \to J_d \bar{s}^3 \)

\[
\bar{S} = S - \frac{T}{4\pi JS} \ln \left[ \frac{1}{\pi J\sqrt{4\pi f}} \bar{s}^{3/2} \right]. \tag{2.95}
\]

The solution of this equation \( \bar{s}(T) \) gives the magnetization curve. Although \( \bar{s} \) decreases with increasing \( T \), it does not reach zero, but only reaches the minimal value \( \bar{s}_{\text{min}} > 0 \) at \( T = T_c \) (this is our definition of the transition temperature). There is no solution of Eq. (2.95) for \( T > T_c \). \( \bar{s}_{\text{min}} \) is found as the
The transition temperature $T_c$ versus dipolar interaction $J_d$ from different approaches for $S = 1/2$. The symbols are numerical results and the curves are the asymptotical formulas (2.59), (2.98).

Minimum of the function

$$\bar{s} - \frac{3T}{4\pi JS^2} \ln(\bar{s}),$$

namely

$$\bar{s}_{\text{min}} = \frac{3T_c}{4\pi JS^2}.$$  (2.97)

Combining Eqs. (2.95) and (2.97) gives the SSWT equation for the $T_c$

$$\frac{4\pi JS^2}{T_c} = \ln \left( \frac{4S}{\sqrt{\pi} J_d} \frac{3}{2} \right) - 2 \ln \left( \frac{T_c}{4\pi JS^2} \right) + 3 \left( 1 - \ln 3 \right).$$  (2.98)

Note that the first-order character of the SSWT phase transition is already contained in a simple equation (2.95).

For the classical case, one uses Eq. (2.61) with $J_d \to J_d s^2$ and obtains the equation for classical SSWT $T_c$

$$\frac{4\pi JS^2}{T_c} = \ln \left( \frac{32}{\pi \sqrt{\pi} J_d} \frac{3}{2} \right) - 3 \ln \left( \frac{T_c}{4\pi JS^2} \right) + 3 \left( 1 - \ln 3 \right).$$  (2.99)

The equations (2.98) and (2.99) are the SSWT equivalents of the free-magnon asymptotical formulas (2.59) and (2.62). Note that the coefficient before the $\ln(T_c/4\pi JS^2)$ term has changed its value from +1 to -2 as compared to the free-magnon theory in the quantum case, and from 0 to -3 in the classical case.
Figure 2.5: The transition temperature $T_c$ versus dipolar interaction $J_d$ from different approaches for classical spin. The symbols are numerical results and the curve is the free-magnon asymptotical formula (2.62). The Monte Carlo result for $J_d/J = 0.1$ is taken from Ref. [12].

The transition temperature as a function of $j_d$ is shown in Fig. 2.4 for $S = 1/2$. Several different approximations are presented. The $\gamma s^2$-SSWT and SW curves are qualitatively similar, with SSWT-$T_c$ being 1.5–2.5 times lower than the SW one. The asymptotical formulas (2.59),(2.98) work very well for $S = 1/2$. The $\gamma \delta$-SSWT, however, gives a crossover from SW to $\gamma s^2$-SSWT behavior upon changing $j_d$. We find this result completely nonphysical and therefore abandon the $\gamma \delta$-SSWT in favor of the $\gamma s^2$-SSWT (which we from now on call simply SSWT). We have also presented the data from the SSWT with renormalized magnon vertex (RPA theory), see Paper I for details. This theory is supposed to improve the SSWT result much in the same way as the normal RPA improves the Hartree-Fock result for fermions. However, in our case the RPA correction to SSWT is negligibly small. This is surprising, since such correction was previously found to be important for quasi-2D and anisotropic 2D systems [38].

Although the Tyablikov approximation is derived from the spin-operator formalism, it can be formulated as a magnon theory with the renormalization $J \rightarrow sJ, J_d \rightarrow sJ_d$. This approximation is very useful for 3D systems and, to some extent, even for 2D systems[30]. However, in the latter case, this approximation does not account for the short-range order above $T_c$. Tyablikov’s result for $T_c$ is also presented in Fig. 2.4 alongside the SW and SSWT results.
This $T_c$ is much smaller than the SSWT one, especially for small values of $j_d$. Tyablikov approximation predicts a first-order phase transition with an enormous step of $\tilde{s}_{\text{min}} \approx 1/2$ at $T_c$ (which immediately follows from Eq. (2.57) upon the substitution $J \rightarrow \tilde{s}J, J_d \rightarrow \tilde{s}J_d$).

In Fig. 2.5 the $T_c$ values from various approximations are presented again, this time for classical spins. Now we can compare our theoretical findings to the classical Monte Carlo (MC) calculation[12] for $j_d = 0.1$ ($T_c/JS^2 \approx 0.85$). One can see that the SSWT value for $T_c$ lies within 9% of the MC result, which is a good agreement for such relatively simple and parameter-free approximation as SSWT. The RPA correction lowers the $T_c$ by 1–5%, which is still a surprisingly small difference compared to the anisotropic FM[38]. Since the SSWT-$T_c$ is already lower that the Monte Carlo $T_c$ for $j_d = 0.1$, RPA apparently does not improve the SSWT result. In contrast to the SSWT, the free-magnon and Tyablikov approximations are much less accurate. Because of this, and the factors mentioned above, the usefulness of the Tyablikov approximation for the 2D systems with dipolar interaction can be questioned. Note that the present discussion refers to the simplest form of Tyablikov decoupling, while more elaborate spin Green’s function approaches, such as e.g. Anderson-Callen decoupling[24, 42, 44], might give better results.

Figure 2.6 summarizes the spin dependence of the SSWT Curie temperature. As for the case of the SW theory (Fig. 2.2), the quantum effects cannot be ignored, even for such a large spin as $S = 7/2$. The formulas (2.98), (2.99),
which work fine for small spins, fail for the large and classical ones, mainly due to the $\gamma = 1$ approximation. In the latter case, the complete (numerical) form of $\gamma^2$-SSWT must be used.

In the conclusion to this chapter, I would say that SSWT is a relatively simple, computationally cheap and reliable theory for studying 2D spin systems with dipole-dipole interaction, with the ability to treat quantum spins being its strongest advantage.
3 Interlude: Density functional theory

Quantum mechanics states that a solid or any other system of particles is described by the Schrödinger equation

$$\hat{H}\Psi = E\Psi,$$  \hspace{1cm} (3.1)

or some analogous relativistic equation if relativistic effects are important.

Here the wavefunction $\Psi = \Psi(r_1, \ldots, r_N)$ depends on the positions of all particles (both electrons and quarks). A few approximations can immediately be made. First, for the solid state applications the internal structure of the nuclei (held together with strong nuclear forces) is of no interest and each nucleus can be treated as one particle. Thus we are left with positively charged nuclei and negatively charged electrons, which interact with each other through electromagnetic forces only. Second, the mass of any nucleus $M_N$ is much larger than the electron mass $m$, therefore we can fix the positions of the nuclei (Born-Oppenheimer approximation), and consider only electron degrees of freedom in Eq. 3.1.

The Hamiltonian can be written as

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V}_{ext},$$  \hspace{1cm} (3.2)

where

$$T = \sum_{i=1}^{N} -\frac{\hbar^2}{2m} \nabla_i^2$$  \hspace{1cm} (3.3)

is the kinetic energy,

$$V_{ee} = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|r_i - r_j|}$$  \hspace{1cm} (3.4)

is the electron-electron Coulomb interaction,

$$V_{ext} = \sum_{i,N} \frac{Ze^2}{|r_i - R_N|}$$  \hspace{1cm} (3.5)

is the external Coulomb potential generated by the nuclei, and $R_N$ is the position of the nucleus $N$.

The idealized (defect-free and infinite) crystalline solids possess transla-
tional symmetry, i.e. the Hamiltonian is invariant under transformation
\[ r_i \rightarrow r_i + T. \] (3.6)

The translation vector \( T \) has the form
\[ T = n_1a_1 + n_2a_2 + n_3a_3, \] (3.7)
where \( a_i \) are the lattice vectors, and \( n_i \) are arbitrary integer numbers.

In principle, there exist methods for solving Eq. 3.1 with any given accuracy, such as the Configuration Interaction (CI) method (see. e.g. [45]) and the Quantum Monte-Carlo (QMC) method. These methods cannot make any use of the translational symmetry and cannot, in fact, be applied to infinite systems at all. The computer resources needed for a CI or QMC calculation grow so rapidly with the number of electrons \( N \), that only molecules and clusters of rather moderate size can be treated in practice. Solids, on the other hand, have about \( N \sim 10^{23} \) electrons, and there is no guarantee that a small cluster will be a suitable model of the solid, hence the use of CI and QMC in the solid-state physics is still rather limited.

Is there a method which has better scaling with the increase of \( N \) and makes use of the translational symmetry? Such method indeed exists. It is called Density Functional Theory (DFT).

### 3.1 Hohenberg-Kohn theorem

In 1964 Hohenberg and Kohn [46] proposed to use the particle density \( n(r) \) instead of the wavefunction \( \Psi \) as the main variable of a many-particle system. For this work Walter Kohn was awarded the Nobel prize 1998 in chemistry. While the wavefunction \( \Psi \) depends on \( 10^{23} \) variables, and only a real mathematician can imagine such a function, the electron density \( n \) depends only on 3 variables. It has been used long before, in the Thomas-Fermi theory and Slater's \( X\alpha \) method, but Hohenberg and Kohn provided the formal background. Namely, they proved the theorem which was later named after them.

The main statement of the Hohenberg-Kohn theorem is that the external potential \( v_{\text{ext}} \) is uniquely determined by the ground state particle density \( n(r) \)
\[ v_{\text{ext}} = v_{\text{ext}}[n(r)]. \] (3.8)

Therefore a given \( n(r) \) uniquely determines the Hamiltonian and hence the total energy of the ground state
\[ E[n] = T[n] + V_{\text{ext}}[n] + \int d\mathbf{r} n(\mathbf{r}) v_{\text{ext}}(\mathbf{r}). \] (3.9)
Further, they proved that the functional (3.9) is minimized by the true ground state density \( n(r) \) for a fixed \( v_{\text{ext}} \). Therefore, the ground state density can be found from the variational equation

\[
\delta E[n] = 0. \tag{3.10}
\]

In principle, we must consider only such trial functions \( n(r) \) which can be represented as the particle density of a certain interacting system (\( N- \) and \( v\)-representability).

The H-K theorem is a very general statement. It is valid for basically all many particle systems (both fermionic and bosonic), both for non-relativistic and relativistic Hamiltonians. The problem is that the functional \( F[n] \equiv T[n] + V_{ee}[n] \) is not known. The Hohenberg-Kohn theorem states the existence of such functional, but says nothing about its explicit form. All simple approximations to \( F[n] \) (such as the one made in the Thomas-Fermi model) give rather bad results.

### 3.2 Kohn-Sham equation

One year after the paper by Hohenberg and Kohn, Kohn and Sham [47] proposed a way to partially overcome this difficulty for a system of fermions (e.g. electrons). They rewrote the total energy functional \( E[n] \) as

\[
E[n] \equiv T_0[n] + \frac{e^2}{2} \int d\mathbf{r} d\mathbf{r}' \frac{n(r)n(r')}{|r-r'|} + \int d\mathbf{r} n(r) v_{\text{ext}}(r) + E_{\text{xc}}[n], \tag{3.11}
\]

where the first term \( T_0[n] \) is the kinetic energy of the noninteracting electron gas, the second term is the classical form of \( V_{ee} \) (called Hartree term), the 3-rd term is the interaction with the external field (Coulomb field generated by the nuclei), and, finally, the last term \( E_{\text{xc}}[n] \), which is called exchange-correlation energy, is defined by Eq. (3.11). The ground state density is given by the variational equation

\[
0 = \frac{\delta E[n]}{\delta n} = \frac{\delta T_0[n]}{\delta n} + e^2 \int d\mathbf{r} n(r) \frac{n(r')}{|r-r'|} + v_{\text{ext}}(r) + \frac{\delta E_{\text{xc}}[n]}{\delta n}. \tag{3.12}
\]

The explicit form of \( T_0[n] \) and \( E_{\text{xc}}[n] \) is not known. We still need some approximations for \( E_{\text{xc}}[n] \), but if we manage to avoid knowing \( T_0[n] \) at all, we certainly surpass Thomas-Fermi-like theories, where both \( T_0[n] \) and \( E_{\text{xc}}[n] \) are approximated.

Consider the noninteracting electron system with the Hamiltonian

\[
H = \sum_i h_i. \tag{3.13}
\]
where the one-electron Hamiltonian $h_i$ is equal to

$$h_i = -\frac{\hbar^2}{2m} \nabla_i^2 + v(r_i), \quad (3.14)$$

$r_i$ are the coordinates of $N$ particles and $v(r)$ is the one-electron potential. Such system has the electron density

$$n(r) = \sum_{i \in occ} |\psi_i(r)|^2, \quad (3.15)$$

where $\{\epsilon_i, |\psi_i\rangle\}$ is the eigenspectrum of the one-particle Hamiltonian (3.14) and the sum is taken over all occupied states $i$. Similarly, the noninteracting kinetic energy $T_0$ is given by

$$T_0 = \sum_{i \in occ} \langle \psi_i | -\frac{\hbar^2}{2m} \nabla^2 | \psi_i \rangle. \quad (3.16)$$

In other words, although we do not know $T_0[n]$ as the functional of the density $n(r)$, both $T_0$ and $n(r)$ are readily obtained from the one-electron eigenspectrum $\{\epsilon_i, |\psi_i\rangle\}$.

The idea of Kohn and Sham was to introduce a reference noninteracting system with the the same electron density $n(r)$ as the interacting system, and to use the one-electron eigenspectrum $\{\epsilon_i, |\psi_i\rangle\}$ as the main variable instead of the density $n(r)$. We assume that such system can be constructed for a given $n(r)$ (noninteracting $N$- and $v$-representability), and we call its potential the effective potential $[v_{eff}(r)]$. The total energy $E_0[n]$ of the reference system can be written as

$$E_0[n] = T_0[n] + \int d\mathbf{r} n(\mathbf{r}) v_{eff}(\mathbf{r}). \quad (3.17)$$

The electron density $n(r)$ can be obtained from the variational equation for the noninteracting system

$$0 = \frac{\delta E_0[n]}{\delta n} = \frac{\delta T_0[n]}{\delta n} + v_{eff}(r). \quad (3.18)$$

However, Eqs. (3.12) and (3.18) must both give the same $n(r)$. This means that

$$v_{eff} = v_h + v_{xc} + v_{ext}, \quad (3.19)$$

where

$$v_h(r) = e^2 \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (3.20)$$
is the Hartree potential and
\[ v_{xc}(r) = \frac{\delta E_{xc}[n]}{\delta n} \]  
(3.21)

is called the exchange-correlation potential.

Now instead of solving the variational equation (3.12) we solve the effective one-electron Schrödinger equation of the noninteracting system, which is called Kohn-Sham equation
\[ \left( -\frac{\hbar^2}{2m} \nabla^2 + v_{\text{eff}} - \varepsilon \right) \psi(r) = 0. \]  
(3.22)

Once we have the Kohn-Sham (KS) eigenspectrum, we construct the electron density \( n(r) \) via Eq. (3.15) and the noninteracting kinetic energy \( T_0 \) via Eq. (3.16). However, we still do not know the functional \( T_0[n] \), since if \( v_{xc} \) is not known exactly, then the Kohn-Sham eigenvalues and the noninteracting kinetic energy \( T_0 \) for a given density \( n(r) \) will neither be exact.

We still need an approximation for \( v_{xc} \). The simplest one is called Local Density Approximation (LDA) [48, 49]
\[ E_{xc}[n] = \int d\mathbf{r} e_{xc}(n(\mathbf{r})) n(\mathbf{r}). \]  
(3.23)

In this case \( v_{xc} \) becomes simply a function of one variable \( n \), rather than a functional of \( n(r) \)
\[ v_{xc}(n) = \frac{\partial e_{xc}(n)}{\partial n} n + e_{xc}(n). \]  
(3.24)

The function \( v_{xc}(n) \) is the same for all many-electron systems, and it is the exact exchange-correlation potential for the uniform electron gas with density \( n \). It was calculated once and for all times with the Quantum Monte Carlo method [50, 51]. In practice, different parameterization of this function are used. Many attempts have been made to improve the LDA expression for \( v_{xc} \) by including terms dependent on the density gradient \( \nabla n(\mathbf{r}) \). Such methods are known under the common name of Generalized Gradient Approximation (GGA). Görling and co-workers have developed a formally exact perturbation expansion of \( E_{xc} \)[52], although only the truncated version known as Exact Exchange (EXX)[53] is widely used at present.

Within the Kohn-Sham formulation of DFT we have succeeded to reduce the many body problem to the one-electron Schrödinger-like equation with the effective potential \( v_{\text{eff}}(\mathbf{r}) \). For the crystals, the potential \( v_{\text{eff}}(\mathbf{r}) \) is periodic due to the translational symmetry (3.6). The Bloch theorem states that every solutions \( \psi_{k,n} \) of the KS equation must have the Bloch form
\[ \psi_{k,n}(\mathbf{r}) = U_{k,n}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \]  
(3.25)
where $k$ is the crystal momentum within the first Brillouin zone (BZ), $n$ is the band index (it numbers the solutions for given $k$) and the function $U_{n,k}(r)$ is periodic (invariant under lattice translation operations). If we have a spatially localized (e.g. atomic like) function $\phi(r)$, we can build a Bloch sum which has the Bloch form

$$
\psi_k(r) = \sum_T e^{ikT} \phi(r - T),
$$

(3.26)

with the summation over all possible translation vectors (3.7). Instead of fixing the number of particles, we introduce the chemical potential (or Fermi energy) $\mu$ which is chosen so that $n(r)$ gives the right number of electrons $N_e$ per unit cell

$$
\int_{UC} d r n(r) = N_e.
$$

(3.27)

Practical methods of solving Kohn-Sham equation introduce some kind of basis set in the Hilbert space of the one-electron Kohn-Sham states (wavefunctions). Although this Hilbert space is infinite-dimensional, only finite size (truncated) basis sets are used in practice. The state vector (wavefunction) is expanded in a basis set $\{|i\rangle\}$

$$
|\psi\rangle = \sum_i c_i |i\rangle,
$$

(3.28)

and the Kohn-Sham equation becomes

$$
0 = \sum_j (H_{kj} - \epsilon S_{kj}) c_j,
$$

(3.29)

where

$$
H_{ij} \equiv \langle i | H | j \rangle
$$

(3.30)

is the Hamiltonian matrix and

$$
S_{ij} \equiv \langle i | j \rangle
$$

(3.31)

is the overlap matrix. The only requirement for the basis vectors $\{|i\rangle\}$ is that they must be linear independent. They do not necessarily have to be orthogonal to each other or normalized.

A typical DFT calculation is the self-consistent cycle

1. Build a starting guess for the electron density $n(r)$. Usually it is the superposition of atomic densities. Then $\nu_{eff}(r)$ is calculated as the sum of $\nu_{ext}$ (nuclei), $\nu_h$ (Poisson equation is solved) and $\nu_{xc}$ (LDA or some other approximation).

2. Solve the KS equation for every $k$-point in the Brillouin zone (BZ). In practice a finite $k$-mesh is used. Point group symmetry allows to treat
the irreducible part of the BZ instead of the full BZ. The core states are usually treated separately and only valence states are included in the KS equation.

- Construct a basis set $|i(k)\rangle$ for a given k-point.
- Calculate matrices $H_{ij}$ and $S_{ij}$.
- Solve the equation (3.29) by numerical matrix diagonalization. Eigenvalues $\varepsilon_{k,n}$ and eigenvectors $|\psi_{k,n}\rangle$ are obtained.

3. Calculate the new electron density $n(r)$ as

$$n(r) = 2V_{uc} \sum_n \int_{BZ} \frac{d^3k}{(2\pi)^3} \Theta(\mu - \varepsilon_{k,n}) |\psi_{k,n}(r)|^2,$$

(3.32)

Where $V_{uc}$ stands for the unit cell volume. The coefficient 2 stands for spin (for non-spin polarized system). There are many methods for doing $k$-integration, among them the Methfessel-Paxton method [54], the tetrahedron method [55] and the modified tetrahedron method of Blöchl [56].

4. Calculate a new $v_{eff}(r)$ from $n(r)$. 
5. Mix old and new $v_{eff}$ to find a next guess for $v_{eff}$. If the new $v_{eff}$ was used alone, the self consistent cycle would not converge.
6. Check for convergence, e.g. by comparing total energies or $n(r)$ between the last two iterations. If convergence to requested accuracy is not yet achieved, go to step 2.
7. Output the results. Calculate additional quantities (total energy, density of states, dielectric tensor) if required.

The Kohn-Sham form of DFT is the main working horse of modern first-principle calculations. However, the problem of approximating exchange-correlation functional still remains. There is a wide class of strongly correlated systems (systems with sufficiently localized d- or f-orbitals), for which none of the regular functionals (LDA, GGA, EXX) is good enough.

### 3.3 The FP-LMTO method

The Full-Potential Linear Muffin-Tin Orbital (FP-LMTO) method is the modern development of the LMTO method [57, 58]. It solves the KS equation (3.29) directly using an elaborately constructed, but small, basis set. We use the FP-LMTO implementation written by John Wills [59] in our calculations. The FP-LMTO method uses the fact that $v_{eff}(r)$ approximately equals the atomic potential close to the nucleus, while far away from the nucleus it is relatively flat. Then the basis set is constructed in the following way.

The space is divided into non-overlapping muffin-tin spheres around the nuclei and the rest of space is called the interstitial. We want each basis function to
• Be localized on a certain atom \( \tau \) in the unit cell and be small elsewhere. Then the Bloch sum (3.26) over all unit cells is made.

• Have \( Y_{lm}(\theta, \phi) \) angular dependence within the selected sphere and be close to atomic \( l, m \) valence state of this atom.

\[
\psi^{MT}_{\tau,l,m}(k, r) = \sum_T e^{i k T} Y_{lm}(r - T - \tau) R_l(|r - T - \tau|). \tag{3.33}
\]

The radial function \( R_l(r) \) is calculated numerically (see below). Here we wrote the atom (or, more precisely, the nucleus) position \( R_\nu \) as \( R_\nu = T + \tau \), where \( \tau \) is the atom position in the unit cell and \( T \) is the translation vector.

• Have a simple mathematical form in the interstitial region. In practice, a Hankel or Neumann functions are used

\[
\psi^{int}_{\tau,l,m}(k, r) = \sum_T e^{i k T} (-i h^+ \left( \frac{\kappa}{|r - T - \tau|} \right), \kappa^2 \leq 0 \left( n_l(\kappa|r - T - \tau|), \kappa^2 > 0 \right).
\]

The parameter \( \kappa^2 \) is called \textit{tail energy}.

• Be continuous and differentiable at the sphere boundary. This puts a boundary condition on the radial function \( R_l(r) \) within the muffin tin sphere. If \( \phi(\epsilon, r) \) is the numerical solution of the radial KS equation for the spherically averaged potential, then the radial function is chosen as

\[
R_l(r) = \alpha \phi(\epsilon_\nu, r) + \beta \frac{\partial \phi(\epsilon, r)}{\partial \epsilon} \bigg|_{\epsilon = \epsilon_\nu}, \tag{3.35}
\]

where \( \epsilon_\nu \) is called \textit{linearization energy}. The numbers \( \alpha \) and \( \beta \) are chosen to make the basis function continuous and differentiable.

One can define several basis functions with the same \( \tau, l, m \) but with different \( \kappa^2 \) (and sometimes even with different \( \epsilon_\nu \)). This allows us to enlarge the basis set thus improving the accuracy of the FP-LMTO method systematically.

The procedure described above applies only to the valence states. The core states are calculated separately on each iteration. The hybridization between core and valence states is neglected. The valence basis states are enforced to be orthogonal to the core states when the radial function \( R_l(r) \) is calculated. We are not going to focus any further on technical details of the FP-LMTO method. Curious reader can see Refs. [59, 60].

3.4 Pseudopotentials and the OPW method

As we have noted already, the Kohn-Sham crystal potential \( V(r) \) is approximately equal to the atomic potential \( V_{\text{at}}(r) \) close to the nucleus. The solutions of the Kohn-Sham equation can be divided into the low-lying core states and the valence states, with all of them being orthogonal to each other. For the
atomic problem (spherically symmetric potential \( V_{at}(|r|) \)) the quantum numbers \( l, m, n \) label all bound KS eigenstates, and the KS eigenfunctions have the form.

\[
\psi_{lmn}(r) = R_{ln}(r)Y_{lm}(\hat{r}).
\] (3.36)

The radial wavefunction \( R_{ln}(r) \) has \( n - l - 1 \) nodes, and therefore for valence electrons it can oscillate rapidly. For example, for the 7s shell it has six (!) nodes.

To summarize, we have a hard (i.e. not flat) potential \( V(r) \) (which is roughly equal to \( -Z e^2 / r \) close to the nucleus), and the rapidly oscillating wavefunctions. Therefore, in order to solve the KS equation efficiently, one needs to use some elaborately constructed basis set, such as the FP-LMTO basis set described above. Any mathematically simple basis set, e.g. plane waves, would be extremely inefficient.

However, there are other ways to handle the problem. The potential \( V(r) \) is relatively smooth outside the core region\(^1\), and the nodes of the valence states wavefunctions are also located in the core region. The idea of the pseudopotential [61] approach is to exclude core electrons from the play and map the KS problem for valence electrons into an effective Schrödinger equation

\[
(H_{KS} - \varepsilon) |\psi\rangle = 0 \quad \longrightarrow \quad (H_{ps} - \varepsilon) |\psi_{ps}\rangle = 0,
\] (3.37)

where \( \psi_{ps}(r) \) is called pseudo wavefunction,

\[
H_{ps} = -\frac{\hbar^2}{2m} \nabla^2 + V_{ps}
\] (3.38)

is the pseudo Hamiltonian and the pseudopotential, \( V_{ps} \), is smooth in the whole space. No orthogonality conditions are imposed on \( \psi_{ps}(r) \), thus it has no or few nodes. There is certain freedom in defining \( V_{ps} \). It is only required that each energy eigenvalue \( \varepsilon \) must be equal to the corresponding KS eigenvalue, and that \( \psi_{ps}(r) \) must be equal to \( \psi(r) \) outside the core region with smooth matching at the core boundary. This freedom is used to make \( V_{ps} \) as smooth as possible.

In the normal KS formalism the valence electrons are repelled from the core region due to the orthogonality requirement, while in the pseudopotential formalism there is no such requirement and the valence electrons are repelled from the core because of the shape of \( V_{ps} \). Since the pseudopotential \( V_{ps} \) is now smooth, and it goes to a constant value for \( r \to 0 \) (while \( V(r) \) behaves

\(^1\)Core region is the spherical region (its radius is called core radius) which contains the core states of an atom. Since the wavefunctions formally extend to infinity, this definition is somewhat arbitrary.
like \(-Z^2/r\), we can use the plane waves basis set

\[
\psi_{G,k}(r) = \frac{1}{\sqrt{V}} e^{i(k+G)r},
\]

where \(G\) is a reciprocal lattice vector. The plane waves basis is orthonormal by definition, therefore the overlap matrix is equal to the identity matrix. Sometimes, other smooth basis functions, e.g. gaussians, are used.

A mathematically exact, though not very efficient, way to introduce pseudopotentials is the Orthogonalized Plane Waves (OPW) method [62]. An orthogonalized plane wave is a plane wave \(|k+G\rangle\) explicitly orthogonalized to all core states \(|\psi_c\rangle\)

\[
|k + G, ps\rangle = |k + G\rangle - \sum_c |\psi_c\rangle \langle \psi_c | k + G\rangle. \quad (3.40)
\]

The valence-state wavefunction is expanded in this basis set\(^{2}\)

\[
|\psi\rangle = \sum_G C_G |k + G, ps\rangle. \quad (3.41)
\]

The Kohn-Sham equation becomes

\[
0 = (H - \epsilon) |\psi\rangle = \sum_G C_G (H - \epsilon) |k + G, ps\rangle
\]

\[
= \sum_G C_G (H - \epsilon) \left[ |k + G\rangle - \sum_c |\psi_c\rangle \langle \psi_c | k + G\rangle \right]
\]

\[
= \left[ H - \epsilon - \sum_c (\epsilon_c - \epsilon) |\psi_c\rangle \langle \psi_c | \right] \sum_G C_G |k + G\rangle = \left( \frac{p^2}{2m} + V_{ps} - \epsilon \right) |\psi_{ps}\rangle. \quad (3.42)
\]

We got the effective Kohn-Sham equation for the pseudo wavefunction

\[
|\psi_{ps}\rangle = \sum_G C_G |k + G\rangle \quad (3.43)
\]

and the pseudopotential

\[
V_{ps} = V(r) - \sum_c (\epsilon_c - \epsilon) |\psi_c\rangle \langle \psi_c |. \quad (3.44)
\]

The use of the plane waves here is purely historical. Any complete basis set used in the derivation above will result in the pseudopotential (3.44). The OPW pseudopotential (3.44) is nonlocal, i.e. \(V\) is a quantum-mechanical op-

---

\(^{2}\) The OPW basis set is complete in the subspace of the valence states. It can be made complete in the entire Hilbert space by adding core eigenstates \(|\psi_c\rangle\) to the basis set.
erator rather than simply a function of coordinate as in the Schrödinger and KS formalisms. It also depends explicitly on the energy $\varepsilon$, and, finally, it involves projection to the core states which is not numerically efficient. Therefore, although the OPW method is essentially an all-electron full-potential method (since no additional approximation has been done by introducing the pseudopotential), it is inefficient and therefore rarely used nowadays.

The other way to introduce the pseudopotential is to postulate

$$V_{ps}(\mathbf{r}) = v_h(\mathbf{r}) + v_{xc}(\mathbf{r}) + \sum_{\nu} V_{at,\nu}(\mathbf{r} - \mathbf{R}_\nu),$$

where the Hartree and exchange-correlation terms are calculated for the valence electrons only and the (possibly nonlocal) atomic pseudopotential $V_{at,\nu}$ includes all effects of the nucleus at position $\mathbf{R}_\nu$ and its core electrons. The approximation made here is that $V_{at,\nu}$ does not depend on the surroundings which we insert our atom into (kind of a frozen-core approximation).

**Empirical pseudopotentials** have simple mathematical form, with a few parameters chosen for each atom in order to fit results to the experiment. For instance, the local Ashcroft empty core pseudopotential [63] is defined as

$$V_{at}(\mathbf{r}) = \begin{cases} 0, & \text{if } r < R_c - N_{v/2} \exp(-r/d) \frac{1}{r} \\ -N_v e^2 \exp(-r/d) \frac{1}{r}, & \text{if } r > R_c, \end{cases}$$

where $N_v = Z - N_c$ is the number of valence electrons, $d$ is the screening length (the screening exponent is not really necessary if the exchange-correlation term is properly included in Eq. 3.45), and $R_c$ is the core radius. The idea is that valence electrons feel the Coulomb potential, generated by the charge $N_v |e|$ (nucleus+core electrons), while the core region is excluded in order to account for the valence-core repulsion. The Ashcroft empty core pseudopotential and other empirical pseudopotentials are useful to build theoretical models of metallic bonds [64], but they hardly give very accurate results.

The more accurate pseudopotentials are the **first-principle pseudopotentials**. They have the form

$$V_{at} = \sum_l V_l(\mathbf{r}) |l\rangle \langle l|.$$  

They are nonlocal, since they have different functional form $V_l(\mathbf{r})$ for different values of the angular momentum $l$. The functions $V_l(\mathbf{r})$ are calculated from the atomic KS problem, and are chosen to reproduce the KS spectrum for the atomic valence shell. Among practical implementations I should perhaps mention norm conserving pseudopotentials[65] and ultrasoft pseudopotentials[66]. First-principle pseudopotentials give much better results than the empirical ones.
3.5 Projector Augmented-Wave (PAW) method

The important question is, whether it is possible to develop an all-electron method with a smooth effective potential, or, in other words, to perform transformation (3.37) without making any uncontrolled approximations. The OPW method described above is one example. There’s a newer and more efficient method called the Projector Augmented-Wave (PAW) method [67, 68]. We first explain this method by the example of a single atom.

The Hilbert space of the KS valence states \( |\psi\rangle \) of an atom is called all-electron (AE) space. It is orthogonal to all core states. The Hilbert space of pseudo-states \( |\tilde{\psi}\rangle \), called PS space. Suppose, there exist a one-to-one correspondence between these two spaces given by a transformation \( T \)

\[
|\psi\rangle = T |\tilde{\psi}\rangle.
\]  

(3.48)

Now, in contrast to the OPW method, we demand that \( |\tilde{\psi}\rangle \) must be exactly equal to \( |\psi\rangle \) outside a sphere with an arbitrary chosen radius \( R_a \), which is called the augmentation sphere.

The explicit form of the operator \( T \) can be written if we introduce a basis \( \{ |\phi_i\rangle \} \) in the AE space (AE partial waves) and a basis \( \{ |\tilde{\phi}_i\rangle \} \) in the PS space (PS partial waves). Now we postulate that

\[
|\phi_i\rangle = T |\tilde{\phi}_i\rangle, \quad \forall i.
\]  

(3.49)

We are also going to need the conjugate basis \( \{ \langle \tilde{p}_i | \} \) in the PS space

\[
\langle \tilde{p}_i | \tilde{\phi}_j \rangle = \delta_{ij}, \quad \sum_i |\tilde{\phi}_i\rangle \langle \tilde{p}_i | = 1
\]  

(3.50)

(\( \langle \tilde{p}_i | = \langle \tilde{\phi}_i | \) if the basis \( \{ |\tilde{\phi}_i\rangle \} \) is orthogonal and normalized). The states \( \langle \tilde{p}_i | \) are sometimes called projector functions. Now the operator \( T \) takes the form

\[
T = \sum_i |\phi_i\rangle \langle \tilde{p}_i | = 1 + \sum_i \left( |\phi_i\rangle - |\tilde{\phi}_i\rangle \right) \langle \tilde{p}_i |,
\]  

(3.51)

The simplest choice for the AE partial waves \( |\phi_i\rangle \) is to take the solutions of the Kohn-Sham equation for the valence states. They have the form

\[
\phi_i(r) = Y_{lm}(\hat{r})\phi_{ln}(r)
\]  

with atomic quantum numbers \( i = (l, m, n) \). This states are orthogonal to each other and to the core states. Usually only one or two values of the main quantum number \( n \) are used in practice.

We still have a degree of freedom in defining PS partial waves \( |\tilde{\phi}_i\rangle \). We can
define them as
\[ \tilde{\phi}_i(r) = Y_{lm}(\hat{r}) \tilde{\phi}_{ln}(r), \] (3.53)
but we still have to choose radial function \( \tilde{\phi}_{ln}(r) \). We exploit this freedom to build most computationally convenient PS basis. Cutoff radii \( r_{lc} \leq R_{\nu a} \) are introduced for each value of \( l \), so that \( \tilde{\phi}_{ln}(r) = \phi_{ln}(r) \) if \( r \geq r_{lc} \). For \( r < r_{lc} \) we choose some suitable smooth function, which can differ for different particular implementations of the PAW method. For instance, we can use a linear combination of two spherical Bessel functions
\[ \tilde{\phi}_{ln}(r) = \begin{cases} \sum_{u=1}^{2} \alpha_u j_l(q_u r) & , r < r_c \\ \phi_{ln}(r) & , r > r_c. \end{cases} \] (3.54)
The \( ln \)-dependent values of \( \alpha_u \) and \( q_u \) are chosen so that \( \tilde{\phi}_{\nu ln}(r) \) is two times continuously differentiable at \( r = r_{lc} \).

Once we know the functions \( |\phi_i\rangle \) and \( |\tilde{\phi}_i\rangle \), we can construct the transformation operator \( \mathcal{T} \). The KS equation is transformed into the PS Kohn-Sham equation
\[ \left( \tilde{H} - \epsilon \tilde{S} \right) |\tilde{\Psi}\rangle = 0, \] (3.55)
where
\[ \tilde{H} = \mathcal{T}^+ H \mathcal{T} \] (3.56)
is the PS Hamiltonian, and
\[ \tilde{S} = \mathcal{T}^+ \tilde{S} \mathcal{T} \] (3.57)
is the PS overlap operator.

This entire construction is meaningless for the single atom, since in order to construct the PS basis we need to know the KS eigenspectrum first. However, in the frozen core approximation, we can construct the PAW potentials for the entire periodic table, and then construct potential of a solid from Eq. 3.45. Since the effective potential is now smooth, the plane wave basis can be used to solve the PS Kohn-Sham equation for a solid.

### 3.6 Quasiparticle excitations vs Kohn-Sham eigenvalues

Density functional theory gives the ground state of a many-electron system. But what can we say about its excited states? Knowledge of the excited states is vital for optics, transport, interpretation of various X-Ray experiments etc. The one-electron picture introduces concepts such as one-electron spectrum
(band structure) \( \varepsilon_n(k) \) and density of states (in this section I speak about systems with translational symmetry). Can these concepts be applied to a system of interacting electrons?

The problem of excited states is trivial for the noninteracting electrons (one-electron picture). Solution of the Schrödinger equation in a periodic lattice potential gives the one-electron spectrum \( \varepsilon_n(k) \), where \( k \) is the quasi-momentum and \( n \) is the band index which numbers eigenvalues for given \( k \) in the Brillouin zone. In the ground state the electrons fill all states with \( \varepsilon_n(k) < \mu \) (\( \mu \) is the Fermi level) while the states with \( \varepsilon_n(k) > \mu \) are left empty. The simplest elementary excitations are excited electrons and holes. If an electron excited from the Fermi level to a state \( n,k \) above the Fermi level, the energy of the excitation is \( \varepsilon_n(k) - \mu \). It can be regarded as a quasiparticle with charge \( -|e| \), spin \( \frac{1}{2} \) and energy \( \varepsilon_n(k) - \mu \). If an electron is excited from the state \( n,k \) below the Fermi level to the Fermi level, we have a hole excitation with charge \( +|e| \), spin \( \frac{1}{2} \) and energy \( \mu - \varepsilon_n(k) \). Any excited state can be viewed as the superposition of these elementary excitations. Therefore, the one-electron spectrum \( \varepsilon_n(k) \) completely describes excited states of a non-interacting electron system.

However, in real solids electrons are strongly interacting. Landau’s Fermi liquid theory states that low-energy excitations of a many-electron system often can be described as a gas of weakly interacting quasiparticles. Such quasiparticles have spin \( \frac{1}{2} \) and charge \( \pm e \) and can be labeled excited electrons and holes on the analogy of the non-interacting case. Many-body theory (see e.g. [69]) introduces the concept of spectral function. The spectral function \( A(k,\varepsilon) \) is the probability of an electron with crystal momentum \( k \) to have energy \( \varepsilon \). While for noninteracting electron there is a discrete spectrum of allowed one-electron energies \( \varepsilon_n(k) \) for a given \( k \), for the interacting case the spectrum is continuous. In most cases, however, the spectral function forms clear peaks reminiscent of the band structure, although the “bands” can be pretty smeared out[70]. Density of states (DOS) is defined rigorously as the \( k \)-integrated spectral function

\[
DOS(\varepsilon) = \sum_k A(k,\varepsilon),
\]

where

\[
\sum_k \equiv v_0 \int_{BZ} \frac{dk}{(2\pi)^3},
\]

and “BZ” means integration over the Brillouin zone. Although spectral function is formally a property of the thermodynamical equilibrium, it is often viewed as an excitation spectrum, with the width of each excitation peak being inversely proportional to excitation’s lifetime.

The Kohn-Sham method introduces a noninteracting reference system. Does
the KS spectrum have anything to do with the quasiparticle spectrum? Formally speaking, this is not the case. The eigenspectrum of the KS Hamiltonian is just an auxiliary construction used to build electron density via Eq. (3.15). It has nothing to do with the true excitation spectrum (which is continuous for each \( \mathbf{k} \)). However, the empirical knowledge from millions of DFT calculations tells us that the KS eigenvalues often bear significant resemblance to the true excitation energies for metals. For insulators and semiconductors the situation is worse, mainly due to the LDA or GGA approximation.

For a given eigenspectrum \( \{ \epsilon_n(\mathbf{k}), |n\mathbf{k}\rangle \} \) the density of states takes the form

\[
\text{DOS}(\epsilon) = \sum_{n, k} \delta(\epsilon - \epsilon_n(\mathbf{k})).
\]

(3.60)

The Partial density of states (PDOS) is a density of states projected to a certain state \( i \). Usually \( i \) means a certain atom (or type of atoms) and a certain value of the angular momentum \( l \). PDOS is given by

\[
\text{PDOS}_i(\epsilon) = \sum_{n, k} \langle n\mathbf{k} | \hat{P}_i | n\mathbf{k} \rangle \delta(\epsilon - \epsilon_n(\mathbf{k})),
\]

(3.61)

where \( \hat{P}_i \) is the projector to states \( i \) under consideration. There is no strict way to define \( \hat{P} \) in crystal, however, all suitable definitions give similar results. Obviously,

\[
\sum_i \hat{P}_i = 1
\]

(3.62)

if we sum over all states \( i \). Therefore

\[
\sum_i \text{PDOS}_i(\epsilon) = \text{DOS}(\epsilon).
\]

(3.63)

The density of states gives vast information about the electronic structure, and PDOS tells in addition on what atoms the electrons reside and what angular momentum character they have.
4 Balanced crystal orbital overlap population (BCOOP)

Solids are held together by the chemical bonding. For every system we study we want to know the bonding character (metallic, covalent or ionic) and the relative strength of different covalent bonds. We also want to know which bands of the electronic structure are responsible for the bonding. Can this information be extracted from a DFT calculation? One such chemical bonding indicator is the crystal orbital overlap population (COOP) proposed by Hughbanks and Hoffmann[71] in the early 80’s. Other COOP-like indicators include crystal orbital Hamilton population (COHP) [72] and \( E_{\text{cov}} \) [73].

Unfortunately, COOP is a basis set dependent quantity and therefore is not an absolute bonding indicator. It is not an expectation value of any quantum-mechanical operator. Traditionally, COOP has been mostly used with the tight-binding versions of Linear Combination of Atomic Orbitals (LCAO) or Linearized Muffin Tin Orbital (LMTO) methods[57, 58]. These methods use minimal (one basis function per every \( nlm \) state of an atom) and well-localized basis sets. However, in order to describe solids with open crystal structure properly, full-potential DFT methods should be used. Our study shows that for a non-minimal basis set the COOP of Ref. [71] is less suitable. In this chapter we propose a less basis set dependent modification of COOP which we name “Balanced COOP” or BCOOP. We will show a few examples of BCOOP analysis. This chapter is based upon Paper III, and the Nb3SiC2 and MgB2 examples are based upon Papers IV and V respectively.

4.1 Definition of COOP and COHP

COOP is the solid-state generalization of the quantum-chemical quantity called overlap population (OP). Consider a molecule with Kohn-Sham eigenspectrum \( \{ \varepsilon_n, |n\rangle \} \). Each eigenvector can be expanded in a basis set \( \{|i\rangle\} \)

\[ |n\rangle = \sum_i c_i |i\rangle. \]  

(4.1)

Overlap population between basis functions \( |i\rangle \) and \( |j\rangle \) is defined for each eigenvector \( |n\rangle \) as

\[ \text{OP}_{ij} = c_i^* c_j S_{ij}, \]  

(4.2)

49
where

\[ S_{ij} = \langle i|j \rangle \]  

(4.3)

is the overlap matrix. OP\(_{ij}\) describes this eigenvector’s contribution to the covalent bond between states \(|i\rangle\) and \(|j\rangle\). It is positive for bonding contributions and negative for the antibonding ones. The absolute value of OP gives the strength of the (anti)bonding. In order to give a meaningful chemical information, the basis set \{|i\rangle\} should consist of atomic-like orbitals with quantum numbers \(l, m, \sigma\) (\(\sigma\) is the spin projection).

For the crystalline solid the basis set \{|i\rangle\}, the eigenspectrum \{\(\varepsilon_n, |n\rangle\}\}, the coefficients \(c_i\) and the overlap matrix \(S_{ij}\) all depend on the crystal momentum \(k\). COOP is defined as the orbital-population weighted density of states

\[ \text{COOP}_{ij}(\varepsilon) = \sum_{nk} \delta(\varepsilon - \varepsilon_n(k)) c_i^* c_j S_{ij}, \]  

(4.4)

where, as usual,

\[ \sum_k \equiv v_0 \int_{BZ} \frac{dk}{(2\pi)^3}. \]  

(4.5)

COOP\(_{ij}\) describes the covalent bond between states \(|i\rangle\) and \(|j\rangle\). It is an energy resolved quantity, which is positive for bonding combinations of orbitals \(|i\rangle\) and \(|j\rangle\) and negative for antibonding ones. COOP has a sum rule

\[ \sum_{ij} \text{COOP}_{ij}(\varepsilon) = \text{DOS}(\varepsilon), \]  

(4.6)

where the sum is taken over all possible contributions to COOP, including the diagonal ones COOP\(_{ii}(\varepsilon)\), which do not correspond to any chemical bonds.

COHP is defined in a way similar to Eq. (4.4), but with Hamiltonian matrix elements \(H_{ij}\) instead of overlap matrix elements \(S_{ij}\)

\[ \text{COHP}_{ij}(\varepsilon) = \sum_{nk} \delta(\varepsilon - \varepsilon_n(k)) c_i^* c_j H_{ij}. \]  

(4.7)

The well-known drawback of COHP \([72, 73]\) is the fact that it depends on the position of the origin of the energy scale (energy zero) and cannot be defined in a unequivocable way (unless basis set \{|i\rangle\} is strictly orthogonal). Indeed, upon transformation \(\varepsilon \rightarrow \varepsilon + \Delta \varepsilon\), the Hamiltonian transforms as \(H_{ij} \rightarrow H_{ij} + \Delta \varepsilon S_{ij}\), and, consequently

\[ \text{COHP}_{ij}(\varepsilon) \rightarrow \text{COHP}_{ij}(\varepsilon) + \Delta \varepsilon \text{COOP}_{ij}(\varepsilon). \]  

(4.8)

The idea behind COHP is that \(H_{ij}\) and \(S_{ij}\) have opposite sign, and that negative COHP corresponds to bonding states. However, this fact is also dependent on the choice of the energy zero. Note that while the units of COOP are Energy\(^{-1}\)
COOP and COHP give good results for the tight-binding LCAO and LMTO methods. Other basis sets, e.g. Full Potential LMTO (FP-LMTO)[59], are less localized, and in addition the FP-LMTO method normally uses two or more basis functions (with different tail energies $\kappa^2$) per each $lm$ state. The main difference between these types of basis sets is that minimal basis sets are almost orthogonal while more complete basis sets are close to linear dependence. In particular, the off-diagonal overlap matrix elements are small in the former case $S^2_{ij} \ll S_{ii}S_{jj}$, while in the latter case often $S^2_{ij} \approx S_{ii}S_{jj}$ holds true.

4.2 Orbital population of an H₂-like molecule

We test the basis set dependence of the orbital population by the simplest possible example: a hydrogen-like molecule with just two basis functions, one per each atom. The LCAO equation is

$$(\hat{H} - \epsilon \hat{S}) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0$$

$(4.9)$

with

$$\hat{H} = \begin{bmatrix} 0 & -|h| \\ -|h| & 0 \end{bmatrix}, \hat{S} = \begin{bmatrix} 1 & S \\ S & 1 \end{bmatrix},$$

$(4.10)$

where $h = -|h|$ is the hopping parameter, and $S$ is the overlap integral.

This equation has two solutions: the bonding state

$$\epsilon_b = \frac{-|h|}{1+S}, |\Psi_b\rangle = \frac{1}{\sqrt{2(1+S)}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$(4.11)$

and the antibonding state

$$\epsilon_a = \frac{|h|}{1-S}, |\Psi_a\rangle = \frac{1}{\sqrt{2(1-S)}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$  

$(4.12)$

It is important that the eigenvectors in (4.11) and (4.12) are normalized.

The orbital populations for the two states are

$$OP_b = \frac{S}{2(1+S)}, OP_a = \frac{-S}{2(1-S)} \neq -OP_b.$$  

$(4.13)$

In the limit of linear dependence $S \rightarrow 1$ and the bonding OP becomes $+1/2$, while the OP for the antibonding state diverges as $-1/2(1-S)$. The reason for this may be seen from Fig. 4.1, where we present the eigenvectors $\Psi_b$ and $\Psi_a$. 

51
and $\Psi_a$ graphically as orthogonal and normalized state vectors expanded in non-orthogonal basis vectors $e_1$ and $e_2$. As Fig. 4.1 shows, $c_1$ and $c_2$ for the antibonding state must become very large when the angle between $e_1$ and $e_2$ is small, which results in a divergent OP for the antibonding state.

In order to remedy this problem we propose the balanced overlap population

$$BOP_{12} = \frac{c_1^* c_2 O_{12}}{O_{11} |c_1|^2 + O_{22} |c_2|^2}. \quad (4.14)$$

This quantity differs only in sign for bonding and antibonding states

$$BOP_b = -BOP_a = \frac{S}{2}, \quad (4.15)$$

and this symmetry between the two states motivates the name.

4.3 Definition of BCOOP

We now define the balanced crystal orbital overlap population (BCOOP) as the density of states weighted with the balanced overlap population. There are different possible definitions. The simplest version is the diagonal-balanced COOP

$$\text{DBCOP}_{ij}(\varepsilon) = \sum_{nk} \delta(\varepsilon - \varepsilon_n(k)) \frac{c_i^* c_j S_{ij}}{\sum_{f} S_{ff} |c_f^*|^2}. \quad (4.16)$$
Here OP in the numerator is balanced by the sum of all diagonal OPs in the denominator.

A more general definition is the balanced COOP

\[
BCOOP_{\alpha_1, \alpha_2}(\epsilon) = \sum_{nk} \delta(\epsilon - \epsilon_n(k)) \frac{\sum_{\alpha_1, \alpha_2} c_i^* c_j S_{ij}}{\sum_{\alpha} \sum_{j' \in A(\alpha)} c_i^* c_{j'} S_{ij'}}. \tag{4.17}
\]

Here the balancing denominator has the block-diagonal form. The quantum number \(\alpha\) includes all quantum numbers important for the chemical analysis (e.g. the atom type \(t\) and the orbital quantum number \(l\)). For example, for TiC (an example below), \(\alpha\) can take one of six different values \(\alpha = \text{Ti}_{s}, \text{Ti}_{p}, \text{Ti}_{d}, \text{Cs}, \text{C}_{p}, \text{orCd}\). These to fall basis functions with given \(\alpha\) (and any values of other quantum number) is called \(A(\alpha)\). Definition (4.17) can be rewritten in a more compact way

\[
BCOOP_{\alpha_1, \alpha_2}(\epsilon) = \sum_{nk} \delta(\epsilon - \epsilon_n(k)) \frac{\langle \alpha_1 | \alpha_2 \rangle}{\sum_{\alpha} \langle \alpha | \alpha \rangle} \tag{4.18}
\]

provided that we decompose the eigenvector \(|n\rangle\) as

\[
|\alpha\rangle = \sum_{i \in A(\alpha)} c_i |i\rangle, \quad \sum_{\alpha} |\alpha\rangle = |n\rangle. \tag{4.19}
\]

It is very important that vectors \(|\alpha\rangle\) are not orthogonal to each other.

In the limit of a nearly-orthogonal basis set the denominator in BCOOP becomes unity and BCOOP becomes identical to COOP. For an exactly orthogonal basis set COOP and BCOOP would vanish and one would rather use COHP as the chemical bonding indicator. The usefulness in BCOOP lies in the opposite limit of strongly non-orthogonal (or even near-linear-dependent) basis set.

4.4 BCOOP for Si, TiC, Ru, Na and NaCl

To test COOP, COHP and BCOOP with the FP-LMTO basis set [59], we have performed calculations for a broad spectrum of materials with different types of chemical bonding (within the local-density approximation). Our test systems include silicon (Si), titanium carbide (TiC), ruthenium (Ru), sodium (Na) and rock salt (NaCl).

Figures 4.2 and 4.3 present COOP and BCOOP for silicon, respectively, calculated between the two sublattices of the diamond structure (summed over all values of \(l\)). One can observe huge negative values of COOP (Fig. 4.2) in the conduction band, a direct consequence of the orbital population divergence for antibonding states discussed above. The \(l\)-resolved COOP curves (not shown) demonstrate that the huge negative peaks are almost exclusively
caused by the s-s contribution, since the overlap between two s orbitals is much larger than the s-p or p-p overlap. Also, as may be noted from the figure, the basis set dependence is significant, even for two different double basis sets, with some choices of tails giving negative values of COOP in some parts of the valence band. The negative antibonding peaks become much larger in magnitude for a double basis set, since they represent a situation closer to linear dependence, and also they are more extended for smaller negative $\kappa^2$ and for positive $\kappa^2$ (less localized orbitals).

On the other hand, BCOOP (Fig. 4.3) gives the same order of magnitude for bonding states (conduction band) and antibonding states (valence band) and a much weaker basis set dependence. The BCOOP curve has clearly defined bonding and antibonding regions, which is the sign of the covalent type of bonding. The two double basis set curves are almost identical and they are in excellent agreement with the COHP result from a TB-LMTO calculation [72]. The one-tail curve in Fig. 4.3 has an erroneous band gap width, but this is just the measure of the crudeness of the FP-LMTO band structure calculated using one tail. The DBCOOP curves (not shown) have a shape similar to BCOOP but with a smaller overall magnitude.

COHP curves for silicon (Fig. 4.4) have the same problems as the COOP curves at Fig. 4.2, including the nonphysical antibonding region just below the Fermi level. We presented COHP curves with a natural choice of energy zero (average potential of the FP-LMTO method [59]), plus one curve with the
Figure 4.3: BCOOP(ε) for silicon in the diamond structure. The Fermi level is at $E = 0$ (vertical line). The legend box contains tail energies $\kappa^2$ in Ry.

Figure 4.4: COHP(ε) for silicon in the diamond structure. The Fermi level is at $E = 0$ (vertical line). The legend box contains tail energies $\kappa^2$ in Ry. Also a curve with a shifted energy origin ($\Delta \varepsilon = -0.4563$ Ry) is presented.
energy zero shifted by the $\Delta \varepsilon = -0.4563$ Ry (this corresponds to taking the Fermi energy as zero). The latter curve has deeper antibonding peaks (down to -220), because, according to Eq. (4.8), there is a COOP contribution added to it.

In Fig. 4.5 Ti d–C p COOP and BCOOP curves for TiC are presented. This is an example of a covalent bond between atoms of different type. Like in the case of Si, there are huge negative peaks in the COOP curve that are avoided by the BCOOP. The electronic structure of TiC is dominated by the C p states in the energy range between -0.5 and 0 Ry and the Ti d states between 0 and 0.5 Ry, separated by the pseudogap. The hybridization between these orbitals is known to result[74] in the strong covalent bonds in this material. The transition between bonding and antibonding states takes place close to the Fermi level. The BCOOP curve reproduces all these features, with a positive contribution in the C-p band region and a negative contribution in the Ti-d band region.

COOP and BCOOP for ruthenium (Ru) are presented in Fig. 4.6 for the Ru d–Ru d bond. The BCOOP curve shows that the states corresponding to less than half filled d-band have a positive BCOOP, representing bonding states, and the states corresponding to more than half filling have negative BCOOP, representing antibonding states. Most of the bonding in Ru is known to be due to the d-states and the calculated BCOOP curve (4.6) is consistent with the Friedel model [64]. Although ruthenium is a metal, the bonding between d-states has mostly covalent character.
Figure 4.6: COOP(ε), BCOOP(ε) and density of states (DOS) for hcp ruthenium, illustrating the Ru-Ru bond between two sublattices. The height of the COOP peak at +1 Ry is -660 1/Ry. The Fermi level is at $E = 0$ (vertical line).

Figure 4.7: BCOOP(ε) and the density of states (DOS) for bcc sodium (Na), illustrating metallic bonding. The Fermi level is at $E = 0$ (vertical line).
Our next two examples deal with other types of bonding, namely metallic and ionic. Only the BCOOP curves are presented. Bcc sodium (Na) is a system with nearly-free electrons and a pure metallic bonding. DOS and BCOOP curves for Na are presented in Fig. 4.7. The DOS curve has almost free-electron (parabolic) shape. The total BCOOP curve is not unlike the covalent-bonding curves presented above. The most profound difference is that the total BCOOP for Na stays positive well above the Fermi level, while for the covalent-bonding systems there are clearly defined bonding and antibonding regions.

The rock salt (NaCl) is an insulator and a textbook example of the ionic bonding. The electronic structure of NaCl includes Cl 3s band at approximately -0.9 Ry, Cl 3p band just below the Fermi level and the unfilled Na 3s band starting at +0.4 Ry. In the ideal case of purely ionic bonding the BCOOP would be zero, since there is no hybridization in this limit. BCOOP is an indicator of the covalent bonding and it cannot measure the strength of the ionic bonding (the electrostatic Madelung energy). BCOOP of NaCl (Fig. 4.8) is small in magnitude and negative up to the Fermi level. This is the small covalent contribution to the bond, which can be neglected compared to the strong ionic bonding.

With these examples we have shown how BCOOP can distinguish between different types of chemical bonding. It also says which states give contribution to each covalent bond.
4.5 \( \text{Nb}_3\text{SiC}_2 \) – a theoretically predicted new MAX phase

Now it is time to apply BCOOP analysis to a more complicated systems. I am going to present the results for the \( \text{Nb}_3\text{SiC}_2 \) in this section (based on Paper IV) and for the \( \text{MgB}_2 \) in the next section (based on Paper V).

Recently a new set of layered metallic ceramics have been discovered [75]. They are called MAX phases because of their chemical composition \( \text{M}_{n+1}\text{AX}_n \), \( (n=1-3 \text{ have been reported}) \) where \( \text{M} \) is an early transition metal, \( \text{A} \) is usually a p-element and \( \text{X} \) is carbon or nitrogen. The nano-layered structure of MAX phases results in a combination of very interesting and useful properties, e.g. refractory ceramic properties, damage tolerance with the ability to withstand thermal shock, and also good electrical and tribological properties. The most known MAX phase is \( \text{Ti}_3\text{SiC}_2 \). Figure 4.9 shows its crystal structure. It can be described as a layered structure consisting of TiC layers interleaved with Si monolayers. There are two nonequivalent positions of titanium: Ti(1) and Ti(2).

In Paper IV we have investigated chemical stability of the MAX phase \( \text{Nb}_3\text{SiC}_2 \) from first principles theory. This phase has not been yet discovered experimentally. From the general trends in chemical bonding of transition metal carbides and silicides one would predict \( \text{Nb}_3\text{SiC}_2 \) to be significantly harder than \( \text{Ti}_3\text{SiC}_2 \). Our total energy calculation have been done within the local density approximation. They demonstrate that \( \text{Nb}_3\text{SiC}_2 \) is unstable,
namely it would decompose into binary compounds as

\[ \text{Nb}_3\text{SiC}_2 \rightarrow 2\text{NbC} + \frac{2}{7}\text{NbSi}_2 + \frac{1}{7}\text{Nb}_5\text{Si}_3. \]  

(4.20)

However, the energy gain from this reaction (formation energy of Nb$_3$SiC$_2$) is rather small, namely +0.02 eV/atom. Therefore we speculate that it might exist as a metastable phase.

Our speculation is based on the comparison to the Ti-Si-C phases. The MAX phases Ti$_3$SiC$_2$, Ti$_4$SiC$_3$, Ti$_5$Si$_2$C$_3$ and Ti$_7$Si$_2$C$_5$ have been recently obtained as thin films[76]. However, Ti$_5$Si$_2$C$_3$ and Ti$_7$Si$_2$C$_5$ are found to be unstable from the DFT calculations, with the formation energies +0.04 eV/atom and +0.03 eV/atom respectively. The latter phases are not observed in the bulk form, only as thin films. Since the formation energy per atom of Nb$_3$SiC$_2$ is even smaller than that of Ti$_5$Si$_2$C$_3$ and Ti$_7$Si$_2$C$_5$, it is likely that this compound can also be synthesized in thin film form.

The calculated bulk modulus for Nb$_3$SiC$_2$ is pretty high, 269 GPa (compared to the calculated value of 225 GPa for Ti$_3$SiC$_2$). Hence, our result is consistent with the speculation that Nb-based MAX phases should be stiffer than the Ti-based ones.

In order to analyze the electronic structure and the chemical bonding we show in Fig. 4.10a the density of states (DOS) of Nb$_3$SiC$_2$ and in Fig. 4.10b we present the BCOOP for the Nb d – C p (Si p) bonds. The DOS of Nb$_3$SiC$_2$ is fairly similar to that of Ti$_3$SiC$_2$ [76], except for the small gap between the C 2s band and the valence band, and also the Fermi level being situated higher up in the band structure for Nb$_3$SiC$_2$. In both cases, the valence band is dominated by the hybridizing M d – C p and M d – Si p states, which form covalent bonds.

From Fig. 4.10b we observe that the BCOOP curves have well-defined bonding (positive) and antibonding (negative) regions with the crossover point near the Fermi level, which is a sign of predominantly covalent bonding character. The states precisely at the Fermi level have mostly antibonding character. From the BCOOP curves one can see that the Nb 4d – C 2p bonds are stronger than the Nb 4d – Si 2p ones. There is also a slight difference between the Nb(1) 4d – C 2p and the Nb(2) 4d – C 2p bonds. The latter ones are weaker since they have a significant fraction of antibonding states below the Fermi level. The Nb 4d – C 2p BCOOP curves are very similar to the Ti 3d – C 2p BCOOP curves in TiC (Paper III) and Ti$_3$SiC$_2$ [76], hence the bonding character is also similar.
Figure 4.10: The partial DOS (a) and BCOOP (b) of Nb$_3$SiC$_2$. The thin vertical line shows the Fermi level ($E = 0$). All BCOOP curves are normalized per one nearest neighbor bond.

Figure 4.11: The crystal structure of MgB$_2$. The magnesium layer is displayed with gray circles and the boron layer is displayed with white circles.
4.6 Chemical bonding in MgB$_2$

In this section, which is based upon results of Paper V, we use BCOOP to analyze the chemical bonding in the famous superconductor MgB$_2$ [77], which has a rather high critical temperature of 39 K. MgB$_2$ has a layered crystal structure which is shown in Fig. 4.11.

The chemical bonding in MgB$_2$ is somewhat complicated. Previous studies [78, 79] suggest that the B—B in-plane bonding is of strong covalent character, the Mg—B inter-plane bonding has mainly ionic character, and the in-plane Mg—Mg bonding is essentially metallic. This means that all three main types of chemical bonding coexist in MgB$_2$, making it a tough test system for BCOOP.

We have calculated charge densities, PDOS and BCOOP for MgB$_2$ using the FP-LMTO method within the generalized gradient approximation (GGA). The charge density plots (Figs. 4–6 of Paper V) clearly support the picture outlined above. One can see obvious covalent bonds between B atoms, but no covalent bonds involving Mg atoms. Our calculation estimates that each Mg atom in MgB$_2$ loses about 0.5 electron compared to the case of pure metallic manganese. This charge transfer is a clear indication of ionic Mg–B bonding.

The partial density of states (PDOS) of MgB$_2$ is presented in Fig. 4.12, while the BCOOP is shown in Fig. 4.13. The B–B BCOOP curves (Fig. 4.13, middle panel) show a positive region at lower energies and a negative region.
at higher energies. This is a clear picture of covalent bonding, with bonding states (positive BCOOP) being separated from antibonding ones (negative BCOOP). Our calculations reveal that the biggest contribution to the B–B bonding comes from the p—p orbital overlap, while the s–s bond has some antibonding states below the Fermi level.

The upper panel of Fig. 4.13 shows the BCOOP between Mg and B. This is a more or less clear BCOOP picture of the ionic bonding (compare to our NaCl example above). Remember that BCOOP is unable to describe ionic bonds adequately, and the ionic BCOOP curve is mostly negative both above and below the Fermi level.

Finally, the Mg–Mg BCOOP curves are presented in the lower panel of Fig. 4.13. They do not show as clear metallic picture as in our previous example of Na (due to the much more complicated electronic structure of MgB₂), but at least there are no strong covalent bonds between Mg atoms, suggesting that the Mg–Mg bonding has predominantly metallic character. To summarize, BCOOP analysis is consistent with the previous studies of the chemical bonding in MgB₂.
5 Hydrogen–hydrogen interaction and structural stability of Ti$_3$SnH$_x$.

Metal–hydrogen (M–H) systems are well-known for their numerous technological applications, such as hydrogen storage[80, 81], fuel cells and nickel–hydrogen batteries. Many solids are able to absorb hydrogen in a certain pressure and temperature range. Hydrogenation can change literally any physical property of the host material. It can cause a metal-insulator transition[82], change the material from ductile to brittle, or from paramagnetic to ferromagnetic etc. This can be used to tune the properties of certain materials, for example Fe/VH$_x$ magnetic multilayers[83]. Since hydrogen atoms are small, they cannot substitute atoms of the host, instead they place themselves in the interstitial, usually occupying tetrahedral or octahedral voids between larger atoms. Thus, an M–H system is a kind of an interstitial alloy. Due to the smallness of the H atoms the volume change is minimal. Sometimes, however, the host crystal structure changes upon hydrogenation.

Hydrogen storage and other applications require answers to the questions "How many H atoms can a metal take?" and "How close can the H atoms come together?" The answer for the first question is: the largest known $H/M$ ratio (number of H atoms per metal atom) is 3.6 for thorium. The answer to the second question is the well-known "2 Å rule": the minimal H–H distance is 2.1 Å. However, a few known exceptions such as Th$_3$AlH$_4$ [84], K$_2$ReH$_6$ [85] and $RT$InH$_{4/3}$ ($R$ = La, Ce, Pr, or Nd; $T$ = Ni, Pd, or Pt) [86], require to put this threshold down to about 1.8 Å. However, even this H–H distance is still large compared to the one in the H$_2$ molecule (0.74 Å). In other words, there are no H$_2$ molecules inside the metal host. In this chapter, which is based upon Papers VI and VII, we investigate how the M–H bonding leads to an effective H–H interaction in crystals. This interaction is responsible for the two limits mentioned above. After the basics of the H–H interaction are given, we focus on the Ti$_3$SnH$_x$ system. Upon hydrogenation, Ti$_3$Sn undergoes a structural phase transition, which, as shall be demonstrated, reflects the H–H interaction in a very clear way. We investigate this system using first-principles calculations in order to learn the nature of the H–H interaction.
5.1 Hydrogen-metal and hydrogen-hydrogen interactions

What is the metal-hydrogen interaction? What is the hydrogen-hydrogen interactions? Suppose we put a single hydrogen atom into the metal lattice (Fig. 5.1 a). The total energy of such a system depends on the position of the H atom: it has minima in the voids between M atoms, while it becomes extremely large close to the metal nuclei. We write the total energy as

$$E(R) = E_M + E_{HM}(R),$$

where $E_M$ is the total energy of the host without hydrogen, and $R$ is the H atom position. The remaining contribution $E_{HM}(R)$ can be called hydrogen-metal interaction potential. This is shown schematically in Fig. 5.1 a. Note that if we put a hydrogen atom into a void, metal atoms actually shift outwards, and such structure relaxation somewhat lowers the total energy. Therefore, the hydrogen-metal interaction can be decomposed into electronic and elastic parts. To define the electronic part we assume that the structure relaxation does not occur, i.e. the metal atoms stay exactly where they were, and look at the total energy. The remaining contribution to $E_{HM}$ comes from the structure relaxation, hence it can be called elastic.

What happens if we put two H atoms into a metal? If two hydrogen atoms interacted with the metal independently, the total energy would be

$$E_{naive} = E_M + E_{HM}(R_1) + E_{HM}(R_2).$$

In reality H atoms feel the presence of each other. The correct total energy is

$$E = E_M + E_{HM}(R_1) + E_{HM}(R_2) + V_{HH}(R_1, R_2),$$

where the H–H pair potential $V_{HH}(R_1, R_2)$ is defined by Eq. 5.3, i.e. it is "everything beyond the independent H-atoms approximation" [64, 87].

It can be rewritten as a function of $R_1 - R_2$ and $(R_1 + R_2)/2$. Why does $V_{HH}$ depend on $R_1 - R_2$ and $(R_1 + R_2)/2$, and not just on the H–H distance $|R_1 - R_2|$? Because our H atoms are not placed in vacuum. The H–H interaction in vacuum is well known: it is repulsive at short distances, attractive...
at larger distances, with the equilibrium point at \( d_{HH} \approx 0.74 \) Å, where the \( H_2 \) molecule is formed. Our \( V_{HH} \), however, is an effective interaction of the two H atoms embedded into the metal, and it does depend on both \( R_1 - R_2 \) and \( (R_1 + R_2)/2 \).

What is the role of the H–H interaction? Independent H atoms would occupy any suitable voids they found, and there are plenty of crystal structures with voids located closer to each other that 2Å. However, a strong hydrogen-hydrogen repulsion or blocking exists at a range about 2Å, leading to the 2Å rule. Can this effect be reproduced by any simple theoretical model? The jellium model (uniform electron gas with a positive background) might seem to be a suitable model for a metal. However, two H atoms in jellium[88] form an \( H_2 \) molecule with the H–H separation close to 0.74 Å for all electron densities, hence jellium is not an adequate model in this case.

H–H interaction can be decomposed into elastic and electronic contributions, in the same way as the M–H interaction. The elastic H–H interaction is an extremely long range one, it can be both attractive and repulsive, depending on the material and strain[89]. In this chapter we focus mainly on the electronic H–H interaction, since, as it will be shown below, it is the one that is mainly responsible for blocking and structural phase transitions.

### 5.2 Ti\textsubscript{3}SnH\textsubscript{x} as a model system

The H–H repulsion makes it energetically favorable to maximize H–H distances. Therefore many metallic systems undergo a structural phase transition upon hydrogenation. As a result, a new crystalline hydride phase is formed with a structure different from that of the initial metal host. An example of such system is Ti\textsubscript{3}Sn, and we have chosen it to study the H–H interaction for the reasons described below. Ti\textsubscript{3}Sn has a hexagonal D0\textsubscript{19} structure (Fig. 5.2a). When we start do put hydrogen into Ti\textsubscript{3}Sn [90, 91], at first the structure does not change. Ti\textsubscript{3}SnH\textsubscript{x} exists in the hexagonal D0\textsubscript{19} structure for \( x < 0.7 \) as an

---

**Figure 5.2**: Unit cell of Ti\textsubscript{3}SnH in (a) the D0\textsubscript{19} structure, (b) the E2\textsubscript{1} structure.
Figure 5.3: Heat of formation of Ti$_3$SnH$_x$ versus volume. $\Delta E \equiv E(Ti_3SnH_x) - \frac{1}{2}E(H_2) - E_{hex}(Ti_3Sn)|_{V=V_{eq}}$. Squares and circles denote cubic and hexagonal phase respectively. Empty, striped and filled symbols denote $x=0$, $0.5$ and $1$ respectively.

interstitial solution, with H atoms occupying the Ti$_6$ octahedral voids. Upon further hydrogenation the Ti$_3$SnH hydride phase is formed. Ti$_3$SnH crystallizes in the cubic E$_2$1 structure (Fig. 5.2b), with H atoms still occupying the Ti$_6$ octahedral voids.

These two crystal structures are very similar despite completely different symmetries (D0$_{19}$ is hexagonal while E$_2$1 is cubic). The situation is similar to the fcc and hcp structures of transition metals$^1$. They have different symmetry, but almost identical nearest neighbor geometry. The biggest difference between the D0$_{19}$ and E$_2$1 structures is the H–H distance, which is quite large ($\approx 4.18$ Å) in the cubic phase and small ($\approx 2.3$ Å) in the hexagonal one (close, in fact, to the 2 Å limit). The Ti$_6$ octahedral voids form a cubic sublattice in the cubic phase, but they form a set of parallel chains in the hexagonal phase (Fig. 5.2). One can look at the phase transition as a way for the system to decrease its energy by increasing H–H distances. Since the two structures are otherwise very similar, the structural phase transition can be attributed to the H–H repulsion alone. This fact makes Ti$_3$SnH$_x$ a good choice to study the H–H interaction.

We have investigated the structural stability of Ti$_3$SnH$_x$ using the VASP plane wave code[68] with PAW potentials [67, 68]; within the LDA approx-

$^1$Under certain experimental conditions a third, orthorhombic form of Ti$_3$SnH$_x$ is formed. The latter structure is discussed in details in Paper VII.
Figure 5.4: Energy difference between cubic and hexagonal structures of Ti$_3$SnH$_x$ as a function of $x$.

The total energy of both phases have been calculated for $x = 0, 0.5$ and 1. Fig. 5.3 shows the calculated heat of formation of Ti$_3$SnH$_x$ with respect to the $H_2$ molecules, defined as

$$\Delta E = E(Ti_3SnH_x) - \frac{x}{2} E(H_2) - E_{hex}(Ti_3Sn)|_{V=V_{eq}},$$

(5.4)

where the total energy of the $H_2$ molecule has also been calculated. Note that the energy of Ti$_3$SnH$_x$ is lower than the energy of Ti$_3$Sn and $\frac{1}{2}$ $H_2$ gas molecules, hence Ti$_3$SnH$_x$ is a stable hydride at low temperatures. The calculated volumes of cubic and hexagonal Ti$_3$SnH$_x$ for the same $x$ are essentially identical.

In Fig. 5.4 we show the energy difference between the cubic and hexagonal phases as a function of $x$. At $x \approx 0.5$ the theoretical calculations predict the cubic phase to become stable over the hexagonal one. These results are in a good agreement with the experiment [90, 91]. The elastic hydrogen-hydrogen interaction can manifest itself in stoichiometric hydride phases only as a volume expansion or a change of the lattice geometry. To check if elastic effects are important for the structural phase transition in Ti$_3$SnH$_x$, we calculated the total energy of cubic and hexagonal Ti$_3$SnH$_x$ for $x = 0.5$ and 1, with exactly the same volume and the unit cell geometry as for Ti$_3$Sn (without H). The results are shown with empty symbols on Fig. 5.4. The elastic contribution to the energy difference is about an order of magnitude smaller than the electronic one. From now on we disregard the elastic effects completely and analyze the phase transition only in terms of the electronic structure.
Figure 5.5: Geometrical configurations of two hydrogen atoms displaced along the [011] direction. The unit cell is shown for the following configurations: (a) reference (b) hydrogen atoms centered on a Ti$_6$ octahedron (1b site) and (c) hydrogen atoms centered on a Ti$_4$Sn$_2$ octahedron (3d site). The crystal structure is cubic and shown for a cut spanned by the vectors [001] and [010]. The small black dot shows the Ti$_6$ site.

5.3 The H–H pair potential

In this section we analyze the structural phase transition in Ti$_3$SnH$_x$ in terms of the H–H interaction potential introduced above. In the spirit of Eq. (5.3), the total energy of this system can be written as

\[ E = E_M + \sum_i E_{HM}(R_i) + \frac{1}{2} \sum_{ij} V_{HH}(R_i, R_j). \]  

Here we ignore the triple and higher order interactions of the hydrogen atoms. Note that the host energy $E_M$ depends on the electron density, and thus implicitly on the hydrogen concentration. Can the observed phase transition be merely a band-filling effect? Our virtual crystal approximation (VCA) calculation demonstrate that the difference $E_{cub} - E_{hex}$ for Ti$_3$Sn is almost insensitive to the band filling. Hence the structural phase transition is caused by the second and third terms in Eq. (5.5). Since the nearest neighbor distances coincide for the two phases of Ti$_3$SnH$_x$, the nearest neighbor terms in $\sum_i E_{HM}(R_i)$ do not contribute to $E_{cub} - E_{hex}$, and the third term in Eq. (5.5) is expected to give the main contribution to the $x$-dependence of $E_{cub} - E_{hex}$. 

70
Our next step is to calculate the hydrogen-hydrogen pair potential from the first principles. We use supercells with displaced H atoms in the rigid Ti₃Sn matrix. The supercells have to be chosen carefully, our choice is presented in Fig. 5.5. In all three cases the unit cell of cubic Ti₃SnH is doubled along the z-direction to include two formula units, hence the supercell is small. The H atoms are displaced along the [110] direction, along the line connecting different octahedral voids, and there are no metal atoms in between. The H atoms can move away from each other up to $\sqrt{2}a \approx 5.9$ Å.

In order to exclude the hydrogen–metal interaction we use the supercell which we call the reference configuration (Fig. 5.5a). It keeps the H–H distance constant and equal to the lattice constant $a = 4.2$ Å. The other two configurations have the H–H distances varied, with the pair of H atoms centered at the 1b site (Fig. 5.5b) or at the 3d site (Fig. 5.5c). From now on we neglect the H–H interactions at distances larger than or equal to $a$. This means that the energy of the reference configuration is given by the H–M interaction alone

$$E_{\text{ref}} = E_M + E_{\text{HM}}(R_1) + E_{\text{HM}}(R_2), \quad (5.6)$$

where $R_{1,2}$ are the positions of the hydrogen atoms and the energy is defined per supercell (two formula units). For the configurations of Fig. 5.5(b,c) the H–H interaction within the pair should be also taken into account

$$E = E_M + E_{\text{HM}}(R_1) + E_{\text{HM}}(R_2) + V_{\text{HH}}(d), \quad (5.7)$$

where $d$ is the H–H distance and it is equal to $2x$ and $\sqrt{2}a - 2x$ for the configurations of Fig. 5.5b and c, respectively, where $x$ is the displacement of the H atom from the 1b site. The H–H pair potential can now be calculated from Eqs. (5.7) and (5.6) as

$$V_{\text{HH}}(d) = E(x) - E_{\text{ref}}(x). \quad (5.8)$$

The total energy as a function of $x$ for our three configurations is presented in Fig. 5.6a. The reference curve is basically the plot of the H–M potential. It has the global minimum at $x = 0$, which means that H atoms sit at the 1b (Ti₆) sites, which gives the cubic Ti₃SnH phase. It also has the local minimum at 3d (Ti₄Sn₂) sites and a potential barrier in between. The other two curves have higher total energy than the reference one. Both curves have a local minimum corresponding to a molecular hydrogen-like state at the respective octahedral void, with $d = 0.9$ Å and 1 Å for 1b and 3d sites respectively. This is a bit larger than the $d = 0.7$ Å for the free H₂ molecule. Note that these distances minimize the total energy, which includes both H–H and H–M interactions.

The calculated $V_{\text{HH}}(d)$ curves are presented in Fig. 5.6b. The 1b and 3d curves are not exactly the same which means that the pair potential is indeed a function of not only $R_1 - R_2$, but also of $(R_1 + R_2)/2$. However, both pair
Figure 5.6: (a) Total energy for the three H configurations along the [011] direction as function of the displacement $x$ from the 1b site. (b) H–H pair potentials in the [011] direction.
potentials have the same general structure. At short range there is strong repulsion between the H atoms, similar to that of an H$_2$ molecule. For larger distances the attraction caused by the covalent H–H bonding dominates. The minimum of the H–H pair potential is reached at about 1.6 and 1.2 Å for the 1b and 3d configurations, respectively. The pair potential curves at Fig. 5.6b can be viewed as a superposition of the 'normal' contribution (e.g. as calculated in a jellium host [88]), and a long-range repulsive component. The 'normal' contribution alone would result in a bond distance of about 0.7 Å. The origin of the repulsive component will be discussed below.

Let us apply the pair potential model to the structural phase transition in Ti$_3$SnH$_x$. Although our $V_{HH}(d)$ curves were calculated for specific cubic supercells of Fig. 5.5, we assume for the moment that these potentials have the general character and can be also applied to the hexagonal Ti$_3$SnH$_x$. We only take into account the strongest interaction, namely the interaction of an H atom in hexagonal Ti$_3$Sn with its two H neighbors within the chain $(E_{cub} - E_{hex})_{Ti_3Sn} = (E_{cub} - E_{hex})_{Ti_3Sn} - V_{HH}(d)$. (5.9)

With $d = 2.3$ Å (the H–H distance in the hexagonal Ti$_3$SnH phase), the obtained estimates of $(E_{cub} - E_{hex})_{Ti_3Sn} - (E_{cub} - E_{hex})_{Ti_3Sn}$ are -21 mRy and -9 mRy for the 1b and 3d curves respectively. As we remember, the total energy calculation gives the value of $(E_{cub} - E_{hex})_{Ti_3Sn} - (E_{cub} - E_{hex})_{Ti_3Sn} = -24$ mRy. Thus, the 1b pair potential gives basically a quantitative agreement with the direct DFT calculations, while the 3d one works much worse. This is not a surprise, since the H atoms in both cubic and hexagonal Ti$_3$SnH have only Ti nearest neighbors, while the 3d site has both Ti and Sn nearest neighbors.

### 5.4 Physics behind the H–H repulsion

To find the physics behind the H–H repulsion, we should take a look at the electronic structure of Ti$_3$SnH$_x$. The calculated density of states (DOS) and partial density of states (PDOS) of the cubic and hexagonal phases of Ti$_3$SnH$_x$ ($x = 0$ and 1) are presented in Fig. 5.7. The main change in DOS upon hydrogenation is the formation of a localized H s-like band about 0.5 Ry below the Fermi level, between the valence band and the Sn s-like band ($\approx -0.7$ Ry). From the PDOS curve one can make conclusions about the character of the chemical bonding. In a purely metallic picture hydrogen would donate one electron into the rigid valence band of Ti$_3$Sn. In a purely ionic picture the localized band of the H$^-$ ion with a pure H s character would be formed. And, finally, covalent bonding manifests itself in a strong hybridization between metal and hydrogen. In our PDOS curves the localized H-induced band at 0.5 Ry below $E_F$ has less than 50% H s character, with the rest coming from the
Figure 5.7: Partial density of states of (a) cubic Ti$_3$Sn, (b) cubic Ti$_3$SnH, (c) hexagonal Ti$_3$Sn, (d) hexagonal Ti$_3$SnH.
Figure 5.8: Partial density of states of cubic Ti$_3$SnH and Ti$_3$Sn(H$_2$)$_{0.5}$. Solid curve – total DOS, dashed – H 1s, dotted – Ti 3d.

metal states, mainly Ti d states. This indicates a strong covalent bonding between H atoms and their Ti nearest neighbors. As a result, Ti 3d states, which in Ti$_3$Sn are occupied in the region between -0.4 Ry and the Fermi level, can in the hydrogenated phase lower their energy substantially since they occupy states down to -0.6 Ry. From the PDOS curves we conclude that Ti$_3$SnH is a stable hydride due to the strong Ti–H covalent bonding.

To understand the microscopic nature of the H–H interaction, we must know how the electronic structure of Ti$_3$SnH depends on the H–H distance. In order to do that, we build a superstructure with an H$_2$ molecule immersed into the Ti$_6$ octahedron, and call it Ti$_3$Sn(H$_2$)$_{0.5}$ structure. The structure is quite similar to that of Fig. 5.5 b, but this time we align the H atoms along the [001] direction, purely for the purpose of convenience. We compare in Fig. 5.8 the DOS and PDOS of the cubic Ti$_3$SnH and of the Ti$_3$Sn(H$_2$)$_{0.5}$ structure (with H–H distance $d = 0.66$ Å). As one can see from Fig. 5.8, the main difference between the two PDOS curves is the position and character of the localized H s-like band. For cubic Ti$_3$SnH, it is a strongly hybridized band about 0.5 Ry below the Fermi level. For Ti$_3$Sn(H$_2$)$_{0.5}$, however, the H
1s states are situated below the Sn s band and the hybridization with all other states in the crystal is weak. The feature at 0.9 Ry below $E_f$ can be viewed as the bonding state of a H$_2$ molecule immersed in Ti$_3$Sn. This molecule is only weakly influenced by the surrounding Ti and Sn atoms. The hybridization between metal and hydrogen states is weak for Ti$_3$Sn(H$_2$)$_{0.5}$.

The reason for the H–H repulsion in Ti$_3$SnH$_x$ seems to be the competition between the H–H and the M–H covalent bonding. When the H atoms are far apart, there is a strong M–H covalent bonding. When they move closer, the H–H covalent bonding appears, which, by itself would favor short H–H distance. However, this H–H bonding weakens the stronger M–H bonds, increasing the total energy. Finally, the H$_2$ molecule in the metallic host is energetically unfavorable (although it may correspond to a local minimum of the total energy). This mechanism leads to the H–H repulsion for the H–H distances larger than approximately 2Å, and hence to the "2Å rule".
6 Sammanfattning på svenska

Teoretiska studier i tvådimensionell magnetism och kemisk bindning


minska, tills dess att vid en temperatur som kallas för *Curietemperatur* ($T_c$) försvinner alla tecken på magnetisk ordning. Det är denna "termodynamik" som vi har studerat med en Heisenbergmodell. Vi har undersökt hur magnetiseringen beror av temperatur och vi vill även veta Curietemperaturen på alla magnetiska material vi studerar. För praktiska tillämpningar vill man ha så stort $T_c$ som möjligt, åtminstone mycket högre än rumtemperatur. I vissa fall, t.ex. för elmotorer, krävs en ännu högre temperatur för att man kan använda det magnetiska materialet i tillämpningar.


I denna avhandling har vi behandlat både klassiska spinn och kvantspinn. För att studera den två-dimensionella Heisenbergmodellen har vi modifierat en teori som heter *självkonsistent spinnvågsteori* (SSWT) för att inkludera den dipolära växelverkan. Denna teori beskriver magnetens termodynamik genom att introducera magnoner och betraktar magon–magon växelverkan på ett visst enkelt sätt. Temperatur-renormalisering av dipolväxelverkan med kvadraten av magnetiseringen behövs också för att bygga en fungerande form av självkonsistent spinnvågsteori. Våra resultat är att Curietemperaturen beror på spinnvärde (i kvantfall) samt på styrkan hos dipolväxelverkan. För klassiska spinn ligger våra resultat inom 9% från exakta beräkningar med nu-

I den andra delen av avhandlingen har vi studerat kemisk bindning i fasta kroppar med hjälp av elektronstrukttureräkningar. Kemisk bindning är den kraft som håller ihop alla molekyler samt fasta kroppar. Den kommer ur sprungligen från Coulombattraktionen mellan de elektroner och kärnor som bygger upp materialet, men den kan ta många olika former i fasta kroppar (kovalent bindning, metallisk bindning, jonisk bindning etc.). Elektronstrukttureräkningar baseras på tätethetsfunktionalteori (DFT). Vad är nu detta för någonting?


När man utför DFT-beräkningar är målet ofta att veta vilken typ av kemisk binding materialet har, och vilka atomer deltar i t.ex. kovalenta bindingar (kovalent binding är när två atomer delar två elektroner och en "elektronpinne" skapas mellan dem). De finns så kallade indikatorer för kemisk binding som låter oss analysera den kemiska bindningen från elektronstrukturberäkningar. I den här delen av avhandlingen har vi utvecklat en ny kemisk "bindningsindikator" som vi kallat den balancerade kristallorbital överlapp populationen (BKOOP på svenska och BCOOP på engelska). Vår metod är en utveckling av en tidigare indikator, och vi visar att vår metod är lämpligare för praktiska DFT beräkningar. Vi har tillämpat den nya indikatortill vissa, ur materialvetenskaplig synpunkt, intressanta system, som t.ex. den supraledana MgB$_2$ föreningen och ett nytt keramiskt material, Nb$_3$SiC$_2$. MgB$_2$ blir en supraleddare med en kritisk temperatur $T_c = 39K$, som är högre än $T_-n$ för nästan alla material (förutom hög-$T_c$ kuprater). Detta material är också ett unikt ämne eftersom det består av element med tre typer av kemisk binding (metallisk, jonisk och kovalent). Nb$_3$SiC$_2$ tillhör en familj av
s.k. MAX-faser, metalliska keramiker med nano-laminat struktur och kemisk formel \( M_{n+1}AX_n \). Vi förutspår att \( \text{Nb}_3\text{SiC}_2 \), om det skulle lyckas syntetiseras, kommer att vara marginellt metastabil, och man kan förvänta sig att denna fas kan existera i tunna filmen.

I ett annat projekt har vi studerat väte–väte (H–H) växelverkan i metaller, som t. ex. i \( \text{Ti}_3\text{Sn} \). Detta problem är viktigt för väteförvaring. Vi visar att H–H växelverkan orsakas av en tävling mellan väte–metall och väte–väte bindning. När tillräckligt mycket väte går in i \( \text{Ti}_3\text{Sn} \), byter materialet kristallstruktur, en effekt som vi kan förklara både med en enkel modell samt via DFT beräkningar.
Acknowledgments

First I would like to thank my supervisor Olle Eriksson. This work would not be possible without his inexhaustible energy, original ideas and a lot of practical help. I am grateful to my other supervisors, Rajeev Ahuja and Mikhail Katsnelson, for helping me learn how to be a physicist. Thank you, Börje Johansson, for creating a unique "FYSIK4 atmosphere", which is still alive despite us being several groups now. I thank Lars Nordström, Susanne Mirbt, Igor Abrikosov and Peter Oppeneer for making our groups function.

I thank John Wills for creating the FP-LMTO code and for helping me to hack it. My gratitude to all my collaborators and people I have discussed science with: Marie Vennström, Yvonne Andersson, Valentin Irkhin, Helmut Rathgen, Petros Souvatzis, Per Andersson, Sebastien Lebegue, Sam Shallcross, Liviu "Lulu" Chioncel, Kay Dewhurst, Sangeeta Sharma, Anna Delin, and many others. My big thanks to Anders E., Per, Weine, Lars., Urban, Erik, Mamo, Adi, Björn, Thomas, Peter and Carlos for taking care of our computer system.

Last, but not least, my deepest thanks to my family, and for all my friends, coffee- and beermates. That includes the people I met in Sweden: Weine, Lars, Velimir, Arkady, Sam, Sangeeta, Kay, Bengt, Thomas & Katrin & Alex, Helmut, Till, Ola, Petros, Björn, Jorge, Natasha "Apollonia" P., Seb, . . . ; and also people I met elsewhere, Igor, Stas, Nick, Vasilyi, Andrey B., Alexei Soroka "AA" & Alexei Soroka "AK", Vera, Oksana K., Jaroslaw, Denis "Syoma" Semagin, Andrey S., Maria C., among others. I would not survive in a foreign country without your company.
References


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 21

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology").