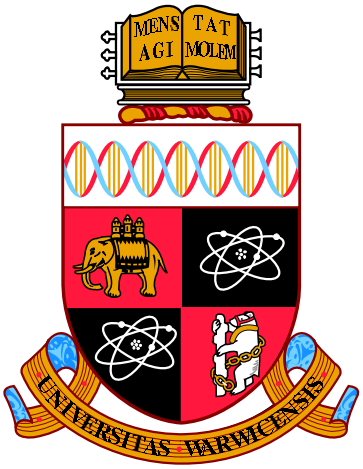

Ab–initio Theory of the Magnetic Properties of Heterogeneous Nanostructures

by

Manuel dos Santos Dias

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

September 2011

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Para a minha avó.

Acknowledgments

I would like to thank Prof. J. B. Staunton for her wise supervision of my PhD studies; only in hindsight could one say that this was the road leading to the final theoretical formalism, with so many fruitful applications. I am also very grateful to Prof. L. Szunyogh for giving me access to his SKKR code, and for many helpful discussions and practical remarks, which made the implementation of the theory and the subsequent applications possible. My thoughts also go to my friends and colleagues at the University of Warwick, too many to be named, but in particular to those that shared my office: Ian, Amitesh, Tilina, Alberto, Jonathan and Andrea. And there was life beyond Physics: for that, a very special thank you goes to Mariusz. To all my friends and family back in Portugal and throughout Europe, you were not forgotten, and are sorely missed.

The financial support of the Fundação para a Ciência e Tecnologia (Portugal), Grant No. SFRH/BD/35738/2007, using funding from the FSE/POPH, is gratefully acknowledged. The kind sponsorship of a collaboration visit to the group of Prof. Szunyogh by the Psi–k Network is also noted with appreciation. Finally, all this work would not have been possible without the resources and infrastructure provided by the Physics Department and the Centre for Scientific Computing of the University of Warwick.

Computers are fickle machines. They will not have my gratitude.

Declarations

This thesis is submitted to the University of Warwick as my application towards the degree of Doctor of Philosophy, and presents details of research carried out in the Theory Group of the Department of Physics between October 2007 and September 2011. The content of this thesis is my own work, unless stated otherwise, carried out under the supervision of Prof. J. B. Staunton, in collaboration with Prof. L. Szunyogh, of the Budapest University of Technology and Economics. No part of this thesis has previously been submitted for a research degree at any other institution.

The following material has been published or submitted for publication:

- M. dos Santos Dias, J. B. Staunton, A. Deak and L. Szunyogh, "Anisotropic spin-spin correlations in $Mn_1/X(111)$ (X = Pd, Pt, Ag, and Au)", Phys. Rev. B **83**, 054435 (2011)

- J. D. Aldous, C. J. Burrows, A. M. Sánchez, R. Beanland, I. Maskery, M. dos Santos Dias, M. K. Bradley, J. B. Staunton and G. R. Bell, "Spin polarisation in epitaxial MnSb polymorphs", submitted to Phys. Rev. Lett.

Abstract

This thesis presents a general formalism for a Density Functional Theory description of the magnetic state of a heterogeneous system, from its paramagnetic high–temperature state to the magnetically ordered ground state. The key ingredient is the identification of localised magnetic moments as emergent electronic degrees of freedom, and the evolution of their orientations under a classical spin Hamiltonian dictated by the underlying electronic structure, computed using Multiple Scattering Theory. The magnetic interactions are contained in the spin–spin correlation function, derived in a linear response approach, and are fully anisotropic, as relativistic effects are accounted for.

A range of applications is given. A Mn monolayer on the W(001) is shown to favour a cycloidal spin spiral as ground state, with a well–defined rotational sense due to unidirectional anisotropic interactions. The antiferromagnetism of Mn monolayers on X(111), with X = Pd, Pt, Ag and Au, is analysed, and the triangular Néel state is shown to possess a well–defined chirality pattern. Calculations for bulk zinc–blende MnSb characterise the half–metallic state as being robust against magnetic disorder; this particular polymorph was discovered in epitaxial form by experimentalists at the University of Warwick, and shows promise for spintronics applications. Bulk Co in the hcp and fcc structure reveals the strong feedback between the magnitude of the spin moment and the degree of magnetic order, through the corresponding electronic structure. To conclude, an FeRh film embedded in V is considered. The magnetisation profile corresponding to the imposed antiferromagnetic order is computed self–consistently, and shows the stabilisation of the Fe layers close to the interface with V by those in the interior of the film.

Abbreviations

AFM — Antiferromagnetic/antiferromagnetism

ASA — Atomic sphere approximation

bcc — body centred cubic (crystal structure)

BZ — Brillouin zone

CI — Configuration interaction (electronic structure method)

CPA — Coherent Potential Approximation

CW — Curie–Weiss (law)

DFT — Density Functional Theory

DLM — Disordered Local Moment (theory)

DM — Dzyaloshinksy–Moriya (unidirectional interaction)

DMFT — Dynamical Mean–Field Theory

DOF — Degree of freedom

DOS — Density of states

fcc — face centred cubic (crystal structure)

FM — Ferromagnetic/ferromagnetism

GF — Green function

GGA — Generalised Gradient Approximation

hcp — hexagonal close packed (crystal structure)

HM — Half–metal/half–metallic

KKR — Korringa–Kohn–Rostoker (electronic structure method)

LDA — Local Density Approximation

MBE — Molecular beam epitaxy

MBPT — Many–Body Perturbation Theory

ML — Monolayer

MOCVD — Metal–organic chemical vapour deposition

MST — Multiple Scattering Theory

MTA — Muffin–tin approximation

NLCPA — Non–local Coherent Potential Approximation

nn — Nearest neighbour

nnn — Next nearest neighbour

PLD — Pulsed laser deposition

QMC — Quantum Monte Carlo

RKKY — Ruderman–Kittel–Kasuya–Yosida (interaction)

RSA — Rigid Spin Approximation

$S^{(2)}$ — Spin–spin correlation function

SCF — Self–consistent field

SIC — Self–interaction correction

SKKR — Screened Korringa–Kohn–Rostoker (electronic structure method)

SOC — Spin–orbit coupling

SPO — Scattering path operator

STM — Scanning tunneling microscopy

zb — zinc–blende (crystal structure)

Chapter 1

Introduction

Magnetism is a quantum–mechanical effect [1]. The Bohr–van Leeuwen theorem shows that the magnetic susceptibility of a classical system is zero, by direct computation of the magnetic moment induced by an applied magnetic field; a non–zero answer requires quantisation of the orbital angular momentum. It was also recognised that electrons must possess intrinsic angular momentum, or spin, first experimentally and then mathematically, from the Dirac equation [2]. The classical theory of paramagnetism by Langevin had a quantisation hypothesis: the magnetic moments were assumed to be of fixed size, which has no classical justification. Quantum mechanics is thus at the heart of magnetism, and will constitute the cornerstone of the entire thesis.

Atomic magnetism is well–understood; the discrete nature of the energy levels determines the magnetic moment of the atom, following Hund's rules. The Langevin theory then becomes justified as a limit of the Brillouin theory of quantised atomic moments. When many atoms are brought together (to form a solid, a thin film, or a cluster) no such simple description is available: the electronic states from each atom are modified by hybridisation with their neighbours, and the simplicity of Hund's rules is lost, except in some special cases [3]. Simple models can be deployed to analyse the range of possible behaviour, with fitted parameters, but for a quantitative and independent analysis a first–principles theory is required.

There are several first–principles approaches to the quantum many–body problem, but the best compromise between accuracy and computational expedience is afforded by Density Functional Theory (DFT). Under the umbrella of DFT, many different computational methods are found [4]. Few, however, can deal with systems in which translational invariance does not apply in all three spatial directions[1]. One

[1] String theorists should read no further on this thesis.

is Multiple Scattering Theory (MST), the Korringa–Kohn–Rostoker (KKR) Green function (GF) method [5]. GFs were an exotic concept in band structure calculations, where the Bloch wavefunctions stood unchallenged, while familiar to practicioners of quantum many–body and quantum field theory. In recent years this separation is being eroded [6]: the two lines of research are merging, and modern extensions to DFT, or methods going beyond it, make extensive use of GFs, which are also a natural quantity in linear response and transport calculations. GFs will be seen to be more convenient for statistical mechanics than wavefunctions.

The magnetic properties of a condensed matter system are ultimately an aspect of the complex many–body ground state of a system of interacting electrons. Thus magnetism is an emergent property: the cooperative electronic motions result in the collective degrees of freedom thought of as net magnetic moments associated with each atom, and will be affected by them [7]. The different energy scales governing the electronic motions and the evolution of the magnetic moments can sometimes lead to an adiabatic approximation: the fast electronic degrees of freedom are integrated out, and the magnetic moments evolve under an effective Hamiltonian, valid for low (compared with the electronic) energies. The mutual feedback between the electrons and the magnetic moments they generate has to be accounted for, but if the electronic structure is fairly insensitive to the magnetic configuration it may be small, and perhaps neglected.

In this low energy regime, two opposite limits can be found: if the energy required to change the length of a magnetic moment is much larger than that of changing its orientation, the basic magnetic excitations will be transverse, precessional motions of the orientations of the moments about their equilibrium directions. These are the familiar magnons that emerge from the Heisenberg model for a collection of moments of fixed (quantum or classical) magnitude. If the longitudinal excitations are on the same energy scale as the transverse one, the Stoner picture is more appropriate: there are no local moments, the magnetism is delocalised throughout the system, and one thinks of spin fluctuations. Most systems will lie in–between these two idealised limits, and theoretical challenges remain in understanding them [7].

In metals *sp*–like electrons are considered to be delocalised, less so *d* electrons, and *f* electrons are usually completely localised. Delocalisation of the *d* electrons is not a synonym for the delocalisation of the associated magnetic moments: it took a few decades for this fact to be understood, with much controversy. It should then be possible to describe metallic systems which possess local magnetic moments in a way similar to the Heisenberg model picture [8]. Experimentalists were never shy of adopting this point of view (maybe not always in a justifiable way).

The interactions between magnetic moments are usually dominated by isotropic contributions, depending only on the relative angles between the moments which they couple. Anisotropies arise from relativistic effects, mainly spin–orbit coupling [2], which are usually weak compared with the isotropic exchange interaction. The anisotropic interactions, however, have a qualitative impact: they set the orientations of the magnetic moments in real space, and can lift degeneracies between magnetic states which are equivalent for the isotropic interactions. When inversion symmetry is broken, unidirectional interactions appear, usually of Dzyaloshinsky–Moriya (DM) form [9, 10]; these are the only ones which can lift chiral degeneracies.

The previous discussion should motivate the goals for the theory to be presented and its applications to be discussed in this thesis.

- It must be rooted in quantum mechanics, and have a sound first–principles basis. This leads to the use of DFT;
- Translational invariance should be a bonus, not a requirement. This argues in favour of the KKR–GF approach;
- The feedback between the electronic structure and the magnetic state must be included. This requires the incorporation of the statistical mechanics of the magnetic moments into DFT, which is made amenable by the use of GFs;
- Anisotropic effects should be treated on the same footing, thus the relativistic version of the KKR method is employed.

The outcome of combining all these ingredients is a theory which can be applied to an heterogeneous system, be it a solid, a thin film or a nanoparticle. The magnetic and electronic properties will be accessible from the fully saturated zero temperature regime up to the high–temperature paramagnetic state, and the effects of anisotropy can be tracked at the same time. Such an ambitious scheme will fall prey to the approximations necessary to make it computationally tractable: a thorough discussion will be given, and in the end it is hoped that the reader will be persuaded that there is much useful information to be extracted from the approximate version of the theory.

We begin by introducing the computational framework for electronic structure calculations, in Chapter 2. DFT is presented as a tractable scheme providing an approximate solution of the full quantum many–body problem. Then a method for the practical solution of the DFT equations is presented: MST—KKR. A self–contained presentation is given, motivating all the main results needed in the following.

The basic concepts of statistical mechanics are presented in Chapter 3. The variational property of the free energy is recalled, together with some approximation techniques. Conventional finite temperature DFT is briefly outlined, and the effective medium approach to the problem of computing the statistical averages is reviewed. This is immediately applied to electronic structure calculations, without further ado, as a consequence of working with GFs.

Chapter 4 collects the most useful concepts in magnetism for condensed matter systems. The main approximations and computational schemes are introduced both for the zero temperature and the finite temperature properties, with special emphasis on those deployed in the later theoretical developments and numerical applications.

The ingredients presented in the first three chapters are then combined into the disordered local moment theory of magnetism, in Chapter 5. The best single–site approximation to the statistical mechanical problem is explained, and given in its most general form. This is then cast into the KKR–DFT formalism, and practical considerations discussed. A linear response approach gives access to the magnetic interactions in the system, contained in the spin–spin correlation function, or $S^{(2)}$. The extreme limits of complete magnetic order and complete magnetic disorder are analysed, and the magnetic susceptibility emerges as a tool for detecting instabilities of the reference magnetic state.

The $S^{(2)}$'s computed from the paramagnetic state are used to understand the magnetism of Mn monolayers, in Chapter 6. The role of the DM interactions is highlighted. In Mn_1/W(001), considerations based on a simple model, combined with the information extracted from the $S^{(2)}$, explain the stabilisation of a cycloidal spin spiral. The chiral degeneracy is lifted by the DM anisotropy, and the spin spiral acquires a well–defined rotational sense [11]. Then the properties of Mn on X(111), with X = Pd, Pt, Ag and Au, are presented. This informs on the experimental finding of 120° antiferromagnetism in Mn_1/Ag(111), in contradiction with theoretical calculations. It is suggested that the different magnetic domains observed in experiment could be attributed to the two chirality patterns possible for the 120° state, which have their energy degeneracy broken by the DM anisotropy, as in the previous case [12].

Chapter 7 describes three other applications. The evolution of the half–metallic state in zincblende MnSb, which was grown epitaxially by experimentalists at the University of Warwick [13], is tracked as a function of the degree of magnetisation. The spin polarisation at the Fermi energy is shown to be more robust against magnetic disorder than that of NiMnSb [14], and shows promise for spintronics applications. Then the assumptions of the theory are put to the test in a system for

which the feedback between the magnitude of the local magnetic moments and the electronic structure is important, bulk Co. To conclude, an illustration of how the theory allows to compute the magnetic profile of a heterogeneous system is given, taking an FeRh film embedded in V as an example.

Final remarks and an outlook then conclude this thesis, with some helpful but not always essential information given in three appendices: on spherical harmonics, on the energetics of spin spirals, and on the specifics of working with an hexagonal lattice.

Chapter 2

Properties of real materials from *ab–initio* calculations

In a condensed matter system, one is faced with the daunting task of computing the physical properties of a large number of electrons and nuclei, all mutually interacting. There is no exact solution, and even if such a solution existed, its practical manipulation would be impossible. This realisation leads to approximation schemes, which may be more amenable to analytical or numerical treatment. Density Functional Theory (DFT) is one of such schemes, in which the fundamental quantities are the equilibrium densities (charge, magnetisation, etc.) of the physical system. It allows detailed comparison with experiment at an affordable computational cost, and provides a very flexible framework for theoretical and numerical developments. For this reason DFT is the foundation of all that follows.

Once the DFT formalism is established, practical methods for solving the Kohn–Sham equations for the systems of interest must be considered. Of the plethora of methods available, the one which will be discussed is Multiple Scattering Theory (MST), the Korringa–Kohn–Rostoker (KKR) Green function (GF) formalism. Working with the GF instead of the wave functions is very advantageous for several problems, including those to be treated in this thesis. At the same time, employing a DFT method which is naturally GF based also simplifies the numerical implementation of the theoretical developments. This justifies a self–contained presentation of the theory, as later chapters will make heavy use of these concepts.

2.1 Density Functional Theory

2.1.1 The quantum many–body problem

One could in principle attempt a solution of the system of coupled electrons and nuclei, as done in the Car–Parrinello molecular dynamics method [15]. However there is an intrinsic timescale separation in the problem: if the system is in thermal equilibrium, the electrons and nuclei have comparable kinetic energy. As the nuclear mass (M) is thousands of times larger than the electron mass (m), from the ratio of the respective kinetic energies the nuclear velocities (v_n) must be much smaller than the electronic velocities (v_e): $v_n/v_e \approx \sqrt{m/M}$. This is the essence of the Born–Oppenheimer approximation [4]: the electrons move in the field of the stationary nuclei, and follow them adiabatically.

The non–relativistic time–independent Hamiltonian operator describing this picture is, in the language of Quantum Field Theory [16],

$$\hat{H} = \hat{K} + \hat{V} + \hat{U} = -\frac{\hbar^2}{2m}\sum_\alpha \int \mathrm{d}\vec{r}\,\hat{\psi}^\dagger_\alpha(\vec{r})\,\nabla^2\,\hat{\psi}_\alpha(\vec{r}) + \sum_\alpha \int \mathrm{d}\vec{r}\,\hat{\psi}^\dagger_\alpha(\vec{r})\,V_{\mathrm{ext}}(\vec{r})\,\hat{\psi}_\alpha(\vec{r}) + \frac{1}{2}\sum_{\alpha\beta}\int \mathrm{d}\vec{r}\int \mathrm{d}\vec{r}'\,\hat{\psi}^\dagger_\alpha(\vec{r})\,\hat{\psi}^\dagger_\beta(\vec{r}')\,W(\vec{r},\vec{r}')\,\hat{\psi}_\beta(\vec{r}')\,\hat{\psi}_\alpha(\vec{r}) \quad , \tag{2.1}$$

$$\left\{\hat{\psi}_\alpha(\vec{r}),\hat{\psi}_\beta(\vec{r}')\right\} = \left\{\hat{\psi}^\dagger_\alpha(\vec{r}),\hat{\psi}^\dagger_\beta(\vec{r}')\right\} = 0 \;, \quad \left\{\hat{\psi}_\alpha(\vec{r}),\hat{\psi}^\dagger_\beta(\vec{r}')\right\} = \delta(\vec{r}-\vec{r}')\,\delta_{\alpha\beta} \quad . \tag{2.2}$$

Here $\hat{\psi}_\alpha(\vec{r})$ and $\hat{\psi}^\dagger_\alpha(\vec{r})$ are fermionic field operators, which obey the standard fermionic anticommutation relations, α, $\beta = \uparrow,\downarrow$ are spin labels, and $V_{\mathrm{ext}}(\vec{r})$ is the one–body external potential, in this case the attractive Coulomb potential generated by the nuclei. $W(\vec{r},\vec{r}') = W(\vec{r}',\vec{r})$ is the symmetric two–particle interaction between the electrons, usually the repulsive Coulomb interaction, which is the source of most difficulties for an approximate solution. Its second quantised form follows straighforwardly from the requirement that the Hamiltonian be a Hermitian operator. Both the one–body external potential and the two–body interaction can be spin–dependent, but this is an unnecessary complication for the moment. The Hamiltonian then naturally decomposes into the kinetic energy term $\hat{K}$, the external potential term $\hat{V}$ and the interaction term $\hat{U}$. Symbols with a hat denote operators, and the same symbols without the hat their expectation values.

2.1.2 Hohenberg–Kohn theorem and the Kohn–Sham equations

In DFT the focus is on the ground state density,

$$n(\vec{r}) = \langle\Psi|\,\hat{n}(\vec{r})\,|\Psi\rangle = \langle\Psi|\sum_{\alpha}\hat{\psi}^{\dagger}_{\alpha}(\vec{r})\,\hat{\psi}_{\alpha}(\vec{r})\,|\Psi\rangle\ , \qquad W(\vec{r},\vec{r}') = \frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r}'|} \tag{2.3}$$

where $|\Psi\rangle$ is the many–body ground state, and we consider explicitly Coulomb interactions. The Hohenberg–Kohn theorem [17] concerns the mappings from the allowable external potentials $V_{\rm ext}(\vec{r})$ to the respective ground states Ψ, and from these to the density $n(\vec{r})$. Suitable extensions to degenerate ground states and other less restrictive formulations exist, but for illustration purposes the original statements are sufficient:

- There is a one–to–one mapping between the external potential $V_{\rm ext}(\vec{r})$ and the one–particle density $n(\vec{r})$, through the many–body ground state;
- Variational principle: the total energy of the system is minimised by the ground state density, for a given external potential.

The total energy functional is given by

$$E[n] = \langle\Psi[n]|\,\hat{K}+\hat{U}+\hat{V}\,|\Psi[n]\rangle = F_{\rm HK}[n] + \int\!\mathrm{d}\vec{r}\,n(\vec{r})\,V_{\rm ext}(\vec{r}) \tag{2.4}$$

and $F_{\rm HK}[n]$ is the Hohenberg–Kohn functional, which is a universal functional of the density. The variational principle is given by the Euler–Lagrange equation

$$E[n] \longrightarrow E[n] + \mu\left(N - \int\!\mathrm{d}\vec{r}\,n(\vec{r})\right)\ , \quad \frac{\delta F_{\rm HK}[n]}{\delta n(\vec{r})} + V_{\rm ext}(\vec{r}) - \mu = 0 \tag{2.5}$$

by introducing the chemical potential μ, to ensure the correct number of electrons is obtained.

As the exact functional is unknown, the Euler–Lagrange equation cannot be solved as it stands. Kohn and Sham [18] suggested an approximation scheme, which connects the real interacting electron system to a fictitious non–interacting system, constructed so as to reproduce the one–particle density of the real system. Rewriting the universal functional,

$$F_{\rm HK}[n] = K_0[n] + E_{\rm H}[n] + E_{\rm xc}[n] \tag{2.6}$$

where $K_0[n]$ is the kinetic energy of the non–interacting system, the Hartree interaction energy term and the respective potential is introduced,

$$E_{\rm H}[n] = \frac{1}{2}\int \mathrm{d}\vec{r}\int \mathrm{d}\vec{r}'\, n(\vec{r})\,\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r}'|}\,n(\vec{r}') = \frac{1}{2}\int \mathrm{d}\vec{r}\, n(\vec{r})\, V_{\rm H}(\vec{r}) \tag{2.7}$$

and the remaining term is the unknown exchange–correlation functional, which is expected to be a small contribution, and which has to be approximated:

$$E_{\rm xc}[n] = K[n] - K_0[n] + U[n] - E_{\rm H}[n]\ ,\quad \frac{\delta E_{\rm xc}[n]}{\delta n(\vec{r})} = V_{\rm xc}(\vec{r}) \tag{2.8}$$

where the exchange–correlation potential is also defined. This functional must obey many sum–rules and satisfy properties which were theoretically derived for the exact functional [19]. The accuracy of a DFT calculation will depend crucially on the approximation used here.

This rewriting leads to an Euler–Lagrange equation for non–interacting electrons, which is equivalent to the Schrödinger Kohn–Sham equation,

$$\frac{\delta T_0[n]}{\delta n(\vec{r})} + V_{\rm eff}(\vec{r}) = \mu \implies \mathcal{H}_{\rm S}\,\phi_n(\vec{r}) = \left(-\frac{\hbar^2\nabla^2}{2m} + V_{\rm eff}(\vec{r})\right)\phi_n(\vec{r}) = \varepsilon_n\,\phi_n(\vec{r}) \tag{2.9}$$

with the effective potential and density,

$$V_{\rm eff}(\vec{r}) = V_{\rm ext}(\vec{r}) + V_{\rm H}(\vec{r}) + V_{\rm xc}(\vec{r})\ ,\quad n(\vec{r}) = \sum_n f_n\,\phi_n^\dagger(\vec{r})\,\phi_n(\vec{r})\ ,\quad \sum_n f_n = N \tag{2.10}$$

with occupation factors $f_n = 1$ if $\varepsilon_n < \mu$, $f_n = 0$ if $\varepsilon_n > \mu$ and $0 \le f_n \le 1$ if $\varepsilon_n = \mu$ (in case of degeneracy at the Fermi level). As the effective potential depends on the density, which in turn depends on the effective potential, we have a self–consistent field (SCF) problem. This leads to the familiar expression for the total energy:

$$E[n] = \sum_n f_n\,\varepsilon_n - E_{\rm H}[n] + E_{\rm xc}[n] - \int \mathrm{d}\vec{r}\, n(\vec{r})\, V_{\rm xc}(\vec{r}) \tag{2.11}$$

where the first contribution is known as the single–particle or band energy, and two double–counting terms are subtracted, to correct the contributions already included in the eigenvalues ε_n.

2.1.3 Relativistic theory and magnetism

So far the magnetic properties have been omitted from the discussion. Magnetism arises from the electronic spin, which is a natural concept in a relativistic formalism,

and somewhat *ad–hoc* in the previous non–relativistic Schrödinger picture. A full relativistic description should employ the four–component current density instead of the one–particle density, and include the electromagnetic Hamiltonian [19]. However, for the energies commonly encountered in condensed matter systems most of these effects are insignificant, and one can proceed with a more limited theory.

The main difficulties in establishing the relativistic version of DFT are the same one finds in developing relativistic quantum field theory. One must carefully renormalise the theory to ensure the removal of divergences, as explained in detail in Ref. [19], arriving at a formulation based on the conventional particle and vector current densities, and their couplings to the scalar and vector electromagnetic potentials. Next one invokes the Gordon decomposition of the current [20], which separates the total current into its orbital and spin parts (plus a gauge term). Keeping only the spin current (i.e., ignoring diamagnetic effects) leads to the Dirac version of the Kohn–Sham Hamiltonian [5]:

$$\mathcal{H}_{\rm D}\,\phi_n(\vec{r}) = \Big(-{\rm i}\hbar c\,\vec{\alpha}\cdot\vec{\nabla} + \beta mc^2 + V_{\rm eff}(\vec{r})\,\mathcal{I}_4 + \beta\vec{\Sigma}\cdot\vec{B}_{\rm eff}(\vec{r})\Big)\,\phi_n(\vec{r}) = \omega_n\,\phi_n(\vec{r}) \quad . \tag{2.12}$$

Now the Kohn–Sham orbitals are four–component spinors, the relativistic kinetic energy includes the rest mass term mc^2. From the familiar 2×2 Pauli matrices

$$\mathcal{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -{\rm i} \\ {\rm i} & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.13}$$

one introduces the vector of Pauli matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and constructs the necessary 4×4 matrices:

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} \mathcal{I}_2 & 0 \\ 0 & -\mathcal{I}_2 \end{pmatrix} \quad \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \quad \mathcal{I}_4 = \begin{pmatrix} \mathcal{I}_2 & 0 \\ 0 & \mathcal{I}_2 \end{pmatrix} \tag{2.14}$$

The scalar effective potential $V_{\rm eff}(\vec{r})$ is defined as before, and there is in addition an effective magnetic potential,

$$\vec{B}_{\rm eff}(\vec{r}) = \vec{B}_{\rm ext}(\vec{r}) + \vec{B}_{\rm xc}(\vec{r}) \;, \quad \vec{B}_{\rm xc}(\vec{r}) = \frac{e\hbar}{2mc}\frac{\delta E_{\rm xc}[n,\vec{m}]}{\delta\vec{m}(\vec{r})} = \mu_{\rm B}\frac{\delta E_{\rm xc}[n,\vec{m}]}{\delta\vec{m}(\vec{r})} \quad . \tag{2.15}$$

The exchange–correlation functional may now include relativistic corrections. The relativistic definition of the particle and magnetisation densities are as follows,

$$n(\vec{r}) = \sum_n f_n\,\phi_n^\dagger(\vec{r})\,\mathcal{I}_4\,\phi_n(\vec{r}) \;, \quad \vec{m}(\vec{r}) = \sum_n f_n\,\phi_n^\dagger(\vec{r})\,\beta\vec{\Sigma}\,\phi_n(\vec{r}) \;, \quad \sum_n f_n = N \tag{2.16}$$

where the four components of the spinors trace out. The summations are only over positive energy states. As done for the particle density, the magnetisation density is also defined without any prefactors (the Bohr magneton $\mu_{\rm B}$ is affected to the effective magnetic field).

The electrostatic interaction can be factored out into the Hartree term, but now there is also a magnetostatic contribution:

$$E_{\rm B}[\vec{m}] = \frac{1}{2}\int \mathrm{d}\vec{r}\int \mathrm{d}\vec{r}'\,\frac{\mu_{\rm B}^2}{|\vec{r}-\vec{r}'|^3}\,\vec{m}(\vec{r})\cdot\vec{m}(\vec{r}') - \frac{4\pi\mu_{\rm B}^2}{3}\int \mathrm{d}\vec{r}\,\vec{m}(\vec{r})\cdot\vec{m}(\vec{r}) \\ - \frac{1}{2}\int \mathrm{d}\vec{r}\int \mathrm{d}\vec{r}'\,\frac{3\,\mu_{\rm B}^2}{|\vec{r}-\vec{r}'|^5}\,\vec{m}(\vec{r})\cdot(\vec{r}-\vec{r}')\,\vec{m}(\vec{r}')\cdot(\vec{r}-\vec{r}') \tag{2.17}$$

which is one form of the Breit interaction [21]. This term is usually not considered explicitly, but it is an important contribution to the magnetic anisotropy, and to the stability of domain walls and domain structures, as is well–known in the micromagnetics community [22].

The total energy is the natural extension of the previous definition:

$$E[n] = \sum_n f_n\,\omega_n - E_{\rm H}[n] + E_{\rm xc}[n,\vec{m}] - \int \mathrm{d}\vec{r}\,n(\vec{r})\,V_{\rm xc}(\vec{r}) - \int \mathrm{d}\vec{r}\,\vec{m}(\vec{r})\cdot\vec{B}_{\rm xc}(\vec{r}) \tag{2.18}$$

and is the same expression found in conventional spin–polarised non–relativistic DFT [19]. The subtle effects of relativity in condensed matter physics will be considered in detail in later sections.

2.1.4 Properties of the exchange–correlation functional and the Local Density Approximation

A closer look at the ingredients of the quantum many–body problem gives useful information about the properties of the exchange–correlation functional. Consider the pair density,

$$n_2(\vec{r},\vec{r}') = \langle\Psi|\sum_{\alpha\beta}\hat{\psi}_\alpha^\dagger(\vec{r})\,\hat{\psi}_\beta^\dagger(\vec{r}')\,\hat{\psi}_\beta(\vec{r}')\,\hat{\psi}_\alpha(\vec{r})\,|\Psi\rangle \tag{2.19}$$

which is the probability of finding an electron at $\vec{r}'$, given that there is one at $\vec{r}$. Manipulating the field operators this can be rearranged as

$$n_2(\vec{r},\vec{r}') = n(\vec{r})\left[n(\vec{r}') + n_{\rm xc}(\vec{r},\vec{r}')\right] \tag{2.20}$$

where the quantity $n_{\rm xc}(\vec{r},\vec{r}')$ is known as the exchange–correlation hole. Integrating the pair density yields the number of pairs in the system, so

$$\int \mathrm{d}\vec{r} \int \mathrm{d}\vec{r}'\, n_2(\vec{r},\vec{r}') = N(N-1) \quad \Longrightarrow \quad \int \mathrm{d}\vec{r}'\, n_{\rm xc}(\vec{r},\vec{r}') = -1 \tag{2.21}$$

which shows where the exchange–correlation hole gets its name from. If the system is non–interacting, $W(\vec{r},\vec{r}') = 0$, then by definition there is no correlation among the electrons, and we find only the exchange hole, $n_{\rm x}(\vec{r},\vec{r}')$, which accounts for the quantum–mechanical nature of the electrons. Then we may split as follows,

$$n_{\rm xc}(\vec{r},\vec{r}') = n_{\rm x}(\vec{r},\vec{r}') + n_{\rm c}(\vec{r},\vec{r}') \;, \quad \int \mathrm{d}\vec{r}'\, n_{\rm x}(\vec{r},\vec{r}') = -1 \;, \quad \int \mathrm{d}\vec{r}'\, n_{\rm c}(\vec{r},\vec{r}') = 0 \tag{2.22}$$

a result which can be proved through the adiabatic connection between the interacting and the non–interacting system [19].

Introducing these definitions in the interaction functional one obtains

$$\begin{aligned} W[n] &= \frac{1}{2} \int \mathrm{d}\vec{r} \int \mathrm{d}\vec{r}'\, n_2(\vec{r},\vec{r}')\, W(\vec{r},\vec{r}') \\ &= \frac{1}{2} \int \mathrm{d}\vec{r} \int \mathrm{d}\vec{r}'\, n(\vec{r})\, W(\vec{r},\vec{r}') \left[n(\vec{r}') + n_{\rm xc}(\vec{r},\vec{r}') \right] = E_{\rm H}[n] + E'_{\rm xc}[n] \end{aligned} \tag{2.23}$$

which then decomposes into the Hartree energy and one contribution to the exchange–correlation functional (the other arising from the difference between the kinetic energy for the interacting and the non–interacting systems).

From the last expression one can see that the exchange–correlation functional is a non–local quantity. On the other hand, it should also be a universal functional. This reasoning leads to the Local Density Approximation (LDA): the exchange–correlation functional can be obtained for the homogeneous interacting electron gas [19]. Then one uses this result locally at each point for the inhomogeneous system:

$$E_{\rm xc}[n] = \int \mathrm{d}\vec{r}\, n(\vec{r})\, g_{\rm xc}[n] \approx \int \mathrm{d}\vec{r}\, n(\vec{r})\, g_{\rm xc}^{\rm LDA}(n(\vec{r})) \;\; . \tag{2.24}$$

The key for the success of the LDA is that is satisfies the sum rules indicated above, despite the local approximation. These ensure some cancellation of errors, but more refined approximations to the exchange–correlation functional exist and are necessary in many cases. The LDA was used in all the calculations to be described in the following.

2.1.5 Constrained calculations

Already in the development of the Kohn–Sham formalism a constraint was introduced in DFT, the chemical potential μ, ensuring the correct number of electrons is obtained. This is a very general way of incorporating constraints into DFT: the constraint and the respective Lagrange multiplier are added to the total energy functional, which is then minimised to obtain a modified Euler–Lagrange equation, and the respective Kohn–Sham equations [23]. The constraints relevant in the present context are those on magnetic properties.

Suppose one wishes to constraint the total magnetisation inside some volume V. The magnitude of the total magnetisation, M, and its orientation, $\hat{e}$, are defined as

$$M = \left| \int_V \mathrm{d}\vec{r}\, \vec{m}(\vec{r}) \right| , \quad \hat{e} = \frac{1}{M} \int_V \mathrm{d}\vec{r}\, \vec{m}(\vec{r}) , \quad \vec{M} = M\hat{e} \tag{2.25}$$

and let the desired values be M_0 and $\hat{e}_0$ (both constraints need not be simultaneously enforced). The functional to be minimised is given by

$$E[n, \vec{m}; M_0, \hat{e}_0] = E[n, \vec{m}] + h_L\big(M - M_0\big) + \vec{h}_T \cdot \big(\vec{M} - M\hat{e}_0\big) \ . \tag{2.26}$$

Varying with respect to the Kohn–Sham–Dirac orbitals, one arrives at

$$\big(\mathcal{H}_\mathrm{D} + \beta\vec{\Sigma}\cdot\vec{h}\big)\,\phi_n(\vec{r}) = \omega_n\,\phi_n(\vec{r}) , \quad \vec{h} = h_L\hat{e} + \vec{h}_T \ \text{ with } \ \vec{h}_T \cdot \hat{e} = 0 \ . \tag{2.27}$$

Here the Lagrange parameters are longitudinal (h_L) and transverse ($\vec{h}_T$) magnetic fields, with respect to the magnetisation direction $\hat{e}$.

The differential of the total energy functional with respect to the parameters M_0 and $\hat{e}_0$ is simply

$$\mathrm{d}E[n, \vec{m}; M_0, \hat{e}_0] = -h_L(M_0, \hat{e}_0)\,\mathrm{d}M_0 - M\,\vec{h}_T(M_0, \hat{e}_0)\cdot\mathrm{d}\hat{e}_0 \tag{2.28}$$

which is one instance of the ubiquitous Hellmann–Feynman theorem [19], and generalises the classical Zeeman energy. One can obtain the energy difference between constrained magnetic states by integrating this expression. Possible applications include energy differences between magnetic states, magnetic anisotropy energy calculations and magnetic exchange interaction parameters [23–26].

2.1.6 Extensions and going beyond Density Functional Theory

DFT using the LDA is a very successful scheme, giving results for lattice constants within a few percent of the experimental values for most systems. This is im-

proved further by the Generalised Gradient Approximation (GGA) [19, 27], which includes corrections due to the gradient of the density into the exchange–correlation functional. Research into improvements of the exchange–correlation functional is on–going, with new proposals being presented regularly. Both the LDA and the GGA, however, still fail for special classes of materials, most notably strongly correlated systems. This can be traced into two main points: the self–interaction error and the derivative discontinuity error. They are closely related to each other.

The derivative discontinuity error is the failure of the approximate exchange–correlation functional in reproducing a property of the exact one: the derivative of the total energy with respect to the number of electrons (generalised to be a real number) should be constant between integer values, with a jump at each integer [28, 29]. This derivative is just the chemical potential, and the jumps are related to the change in energy as an electron is added (affinity) or removed (ionisation) from the system.

The self–interaction error stems from approximating the true potential energy by the Hartree electrostatic energy, which includes the interaction of each electron with itself, and this is only partially corrected by the LDA functional. This was recognised a long time ago [30], and the self–interaction correction (SIC) scheme has been implemented successfully in various DFT codes, including KKR [31]. As an application, the magnetic properties of the rare–earth metals [32] and some transition metal oxides [33] were computed and found to be in very good agreement with experiment. It can be shown that the self–interaction error vanishes for a completely delocalised state. It should then be small for the delocalised *sp* states, but important for the fairly localised *d* and completely localised *f* states. An alternative correction scheme, the LDA+U [34, 35], adopts this point of view: electronic correlations in localised states are the important source of error, and so these are treated using a Hubbard–like model. The Hubbard parameter U can be determined from first–principles, but it is usually kept as an adjustable quantity. Dynamical Mean Field Theory (DMFT) [36] is the natural extension of the LDA+U method, moving from the static to a dynamic treatment of correlations. LDA+DMFT methods are becoming more common, also in KKR [37].

Ultimately a direct attempt at solving the full many–body problem could also be pursued, thanks to modern fast supercomputers. The GW method [38] is the first step towards this goal: starting from a DFT calculation, the screened Coulomb interaction W is computed, and the new ground state properties constructed. Many–Body Perturbation Theory (MBPT) [6] takes a similar approach. The most accurate and most computationally demanding methods are the quantum–chemical Configu-

ration Interaction (CI) [39] and the Quantum Monte Carlo (QMC) methods [40, 41]. This brief overview of methods beyond simple DFT+LDA is by no means exhaustive, and it is only meant to give a flavour of some challenging problems.

2.2 Multiple Scattering Theory

2.2.1 The Lippmann–Schwinger and Dyson equations

The core of any DFT calculation is the solution of the appropriate Kohn–Sham equations, to then reconstruct the density and proceed with the self–consistent cycle. In general terms, the problem is how to find the solution of the appropriate single–particle equations, given the Hamiltonian. This usually leads to an cumbersome eigenvalue problem, and much cunning is needed to accelerate this part of the computation. On the other hand, if one switches from wave functions to GFs, this can be avoided, and other benefits will follow, from the analytical properties of these operators. The application of these ideas to DFT calculations constitutes the KKR method of electronic structure [5, 42, 43].

The single–particle problem can be stated as such: given the Hamiltonian $\mathcal{H}$, find the eigenvalues ε_n and the corresponding eigenfunctions ψ_n. The Hamiltonian is then split into a reference part, $\mathcal{H}_0$, and a remainder, which is typically a potential term, V. $\mathcal{H}_0$ has its own set of eigenfunctions, ϕ_n, and presumably its spectrum overlaps with that of $\mathcal{H}$, so that in that energy range the eigenvalues are shared (which typically happens in the continuum part of the spectrum, for positive energies). If the eigenvalue ε_n is shared, then one can write

$$(\varepsilon_n - \mathcal{H})\,\psi_n = 0 \quad \text{and} \quad (\varepsilon_n - \mathcal{H}_0)\,\phi_n = 0 \quad \Longrightarrow \quad (\varepsilon_n - \mathcal{H})\,\delta\psi_n = V\phi_n \tag{2.29}$$

by putting $\psi_n = \phi_n + \delta\psi_n$. Defining the resolvents or GFs $G(z)$ and $G_0(z)$ at complex energy $z = \varepsilon + \mathrm{i}\delta$,

$$G(z) = \frac{1}{\varepsilon - \mathcal{H} + \mathrm{i}\delta}\,, \quad G_0(z) = \frac{1}{\varepsilon - \mathcal{H}_0 + \mathrm{i}\delta} \implies \begin{cases} \psi_n^{\pm} = \phi_n + G^{\pm}(\varepsilon_n)\,V\phi_n \\ \psi_n^{\pm} = G_0^{\pm}(\varepsilon_n)\,V\psi_n^{\pm} \end{cases} \tag{2.30}$$

where it is indicated that when the energy approaches an eigenvalue, one can choose two different side limits, $\delta \to \pm 0$. This result is known as the Lippmann–Schwinger equation [5], and has a simple physical interpretation. The total wavefunction ψ_n is the sum of the reference wavefunction ϕ_n, and its multiple scattering contribution from the potential V. Many experiments can be understood in this way.

As might be expected, the resolvents are also connected in a simple way, by the

so–called Dyson equation [5]:

$$\begin{aligned} G(z) &= G_0(z) + G_0(z)\, V G(z) = G_0(z) + G(z)\, V G_0(z) \\ &= G_0(z) + G_0(z)\, V G_0(z) + G_0(z)\, V G_0(z)\, V G_0(z) + \ldots \\ &= G_0(z) + G_0(z)\, \mathcal{T}(z)\, G_0(z) \end{aligned} \tag{2.31}$$

where the last equality makes the connection with the familiar Born series for multiple scattering from a given potential [5]. All the effects of scattering from the potential V were summed into the t–matrix operator, $\mathcal{T}$. The Lippmann–Schwinger equation can now also be written as

$$\psi_n^{\pm} = \phi_n + G_0^{\pm}(\varepsilon_n)\, \mathcal{T}(\varepsilon_n)\, \phi_n \quad \text{as} \quad G^{\pm}(\varepsilon_n)\, V = G_0^{\pm}(\varepsilon_n)\, \mathcal{T}(\varepsilon_n) \ . \tag{2.32}$$

The usefulness of the GF concept is already apparent in its spectral representation:

$$G^{\pm}(\varepsilon) = \frac{\mathcal{P}}{\varepsilon - \mathcal{H}} \mp \mathrm{i}\pi\, \delta(\varepsilon - \mathcal{H}) \implies \mp\frac{1}{\pi}\mathrm{Im\,Tr}\, G^{\pm}(\varepsilon) = \sum_n \delta(\varepsilon - \varepsilon_n) \equiv n(\varepsilon) \tag{2.33}$$

where $\mathcal{P}$ is the principal value, and we obtain the density of states (DOS) of the Hamiltonian, by definition. The number of electrons at energy ε is given by

$$N(\varepsilon) = \int^{\varepsilon} \mathrm{d}\varepsilon'\, n(\varepsilon') = \pm\frac{1}{\pi}\mathrm{Im\,Tr\ log}\, G^{\pm}(\varepsilon) \tag{2.34}$$

which follows directly from the previous equation and the definition of the GF. Using the Dyson equation and performing some manipulations yields

$$N(\varepsilon) = N_0(\varepsilon) + \Delta N(\varepsilon)\ , \quad \Delta N(\varepsilon) = \pm\frac{1}{\pi}\mathrm{Im\,Tr\ log}\, \mathcal{T}^{\pm}(\varepsilon) \tag{2.35}$$

with $N_0(\varepsilon)$ the number of electrons for the reference GF, and $\Delta N(\varepsilon)$ the excess number of electrons due to the scattering properties, the celebrated Lloyd formula [44].

In the position representation one writes

$$G^{\pm}(\vec{r}, \vec{r}'; \varepsilon) = \sum_n \frac{\psi_n(\vec{r})\, \psi_n^{\dagger}(\vec{r}')}{\varepsilon - \varepsilon_n \pm \mathrm{i}0} \implies n(\vec{r}) = \mp\frac{1}{\pi}\mathrm{Im\,Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\, G^{\pm}(\vec{r}, \vec{r}; \varepsilon) \tag{2.36}$$

as can be checked by comparing with the definition, Eq. 2.10. The occupation function is simply $f(\varepsilon) = 1$ if $\varepsilon \leq \mu$ and zero otherwise. Similar expressions can be derived for other physical properties, and will be presented later. In what follows

the approach to the real energy axis will always be taken from the upper half of the complex plane, so the sign in the GF is fixed to + and will henceforth be dropped.

As the GF $G(z)$ is analytic for $z = \varepsilon + \mathrm{i}|\delta|$, the energy integration contour can be deformed into the upper half of the complex plane. Computationally this is very advantageous, as the highly structured nature of the GF on the real energy axis is broadened when the energy acquires a finite imaginary part. The particle density can then be computed using only complex energies:

$$n(\vec{r}) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int^{\mu}\mathrm{d}\varepsilon\, G(\vec{r},\vec{r};\varepsilon) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int_{\curvearrowright}^{\mu}\mathrm{d}z\, G(\vec{r},\vec{r};z) \quad . \tag{2.37}$$

The integration along the real axis up to energy μ is deformed into a semicircular contour in the upper half of the complex plane, ending at the same energy.

More generally, consider some operator $\hat{A}$. Evaluating

$$-\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int_{\curvearrowright}^{\mu}\mathrm{d}z\,\hat{A}\,G(\vec{r},\vec{r};z) = \sum_{n}^{\varepsilon_n\leq\mu}\psi_n^{\dagger}(\vec{r}')\,\hat{A}\,\psi_n(\vec{r}) = A(\vec{r},\vec{r}') \tag{2.38}$$

where the trace was used to reorder the wavefunctions, and the contour integration becomes a summation over the simple poles corresponding to the eigenenergies. The expectation value of any operator can thus be computed by tracing it together with the GF, in the desired representation.

2.2.2 Scattering by a localised potential

The total potential is usually split into contributions surrounding each atom, $V = \sum_i V_i$. Then the calculation splits into two parts: solution of the scattering problem associated with each potential V_i, and combining all the scattering contributions into the full scattering properties of the system. The first part is described in this subsection, while the second part follows in the next subsection. Some comments on terminology [5]:

- Muffin–tin approximation (MTA): spherical potential in a spherical region, $V(\vec{r}) = V(r)$ for $r < r_{\mathrm{MT}}$ and $V(\vec{r}) = 0$ otherwise. When splitting the potential into cells, non–overlapping spheres are used. The region of constant potential outside the potential spheres is called interstitial region;

- Atomic sphere approximation (ASA): same conditions on the potential as the MTA. In the ASA the interstitial region is eliminated, by choosing overlapping spheres whose volume adds up to the total volume of the system;

- Full–potential: non–spherical potential in a non–spherical region. There is no interstitial region or overlap between potentials in adjacent cells.

Let $V_i(\vec{r})$ denote some local potential of finite range centred at position $\vec{R}_i$, and $\mathcal{S}_i$ the region where it is non–zero. The position vectors $\vec{r}$ and $\vec{r}'$ will be taken to refer to this origin. The following discussion applies to non–spherical potentials in regions of arbitrary shape [45], but the original theory made use of spherical potentials inside spherical regions [46]. The Lippmann–Schwinger equation now reads, in the position representation,

$$\psi_i(\vec{r};\varepsilon) = \phi(\vec{r};\varepsilon) + \int \mathrm{d}\vec{r}_1 \int \mathrm{d}\vec{r}'\, G_0(\vec{r},\vec{r}_1;\varepsilon)\, t_i(\vec{r}_1,\vec{r}';\varepsilon)\, \phi(\vec{r}';\varepsilon) \tag{2.39}$$

while the t–matrix for the potential $V_i(\vec{r})$ satisfies the Dyson equation:

$$t_i(\vec{r},\vec{r}';\varepsilon) = V_i(\vec{r})\,\delta(\vec{r}-\vec{r}') + V_i(\vec{r}) \int \mathrm{d}\vec{r}_1\, G_0(\vec{r},\vec{r}_1;\varepsilon)\, t_i(\vec{r}_1,\vec{r}';\varepsilon) \tag{2.40}$$

and $t_i(\vec{r},\vec{r}';\varepsilon) = 0$ if $\vec{r}$ or $\vec{r}'$ are outside the potential region.

Choosing the free electron Hamiltonian for $\mathcal{H}_0$, the wavefunctions $\phi(\vec{r};\varepsilon)$ are just plane waves, and can be written in spherical coordinates as a linear combination of spherical Bessel functions $j_\ell(x)$, using Bauer's identity [5]:

$$e^{\mathrm{i}\vec{k}\cdot\vec{r}} = 4\pi \sum_L \mathrm{i}^\ell\, j_\ell(kr)\, Y_L^*(\hat{r})\, Y_L(\hat{k}) \;, \quad k = \sqrt{\varepsilon} \;\; . \tag{2.41}$$

Here ℓ is the angular momentum quantum number, m is the magnetic quantum number, and $L = (\ell, m)$ is a combined index. The functions $Y_L(\hat{r})$ are spherical harmonics, the eigenfunctions of the orbital angular momentum operator, and we have $\hat{L}^2\, Y_L(\hat{r}) = \ell(\ell+1)\, Y_L(\hat{r})$ and $\hat{L}_z\, Y_L(\hat{r}) = m\, Y_L(\hat{r})$. This is the motivation for the *partial wave representation*, where all quantities are expressed in terms of linear combinations of spherical Bessel functions and spherical harmonics.

The GF corresponding to $\mathcal{H}_0$ is then the GF for the Helmholtz wave equation:

$$G_0(\vec{r},\vec{r}';\varepsilon) = -\frac{1}{4\pi}\frac{e^{\mathrm{i}k|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} = -\mathrm{i}k \sum_L j_\ell(kr_<)\, h_\ell^+(kr_>)\, Y_L(\hat{r})\, Y_L^*(\hat{r}') \tag{2.42}$$

with $r_> = \max\{r, r'\}$ and $r_< = \min\{r, r'\}$, and $h_\ell^+(x)$ a spherical Hankel function. For $\vec{r}$ outside the sphere bounding the region $\mathcal{S}_i$, we can write from Eq. 2.39

$$\psi_i(\vec{r};\varepsilon) \longrightarrow R_{i,L}(\vec{r};\varepsilon) = j_\ell(kr)\, Y_L(\hat{r}) - \mathrm{i}k \sum_{L'} h_{\ell'}^+(kr)\, Y_{L'}(\hat{r})\, t_{i,L'L}(\varepsilon) \tag{2.43}$$

which defines the angular momentum representation of the t–matrix:

$$t_{i,L'L}(\varepsilon) = \int \mathrm{d}\vec{r} \int \mathrm{d}\vec{r}' \, j_{\ell'}(kr') \, Y^*_{L'}(\hat{r}') \, t_i(\vec{r}', \vec{r}; \varepsilon) \, j_\ell(kr) \, Y_L(\hat{r}) \tag{2.44}$$

and this solution must match smoothly to the corresponding one for $\vec{r}$ inside the bounding sphere,

$$R_{i,L}(\vec{r}; \varepsilon) = j_\ell(kr) \, Y_L(\hat{r}) + \int \mathrm{d}\vec{r}_1 \int \mathrm{d}\vec{r}' \, G_0(\vec{r}, \vec{r}_1; \varepsilon) \, t_i(\vec{r}_1, \vec{r}'; \varepsilon) \, j_\ell(kr') \, Y_L(\hat{r}') \tag{2.45}$$

which is regular at the origin of the potential, $\vec{R}_i$. Likewise an irregular solution can be defined, by finding the solution that matches to

$$H_{i,L}(\vec{r}; \varepsilon) = -\mathrm{i}k \, h^+_\ell(kr) \, Y_L(\hat{r}) \tag{2.46}$$

outside the bounding sphere. There is an alternative set of solutions, which is also in common use, with different matching conditions:

$$Z_{i,L}(\vec{r}; \varepsilon) \longrightarrow \sum_{L'} j_{\ell'}(kr) \, Y_{L'}(\hat{r}) \, t^{-1}_{i,L'L}(\varepsilon) - \mathrm{i}k \, h^+_\ell(kr) \, Y_L(\hat{r}) \tag{2.47}$$

$$J_{i,L}(\vec{r}; \varepsilon) \longrightarrow j_\ell(kr) \, Y_L(\hat{r}) \ . \tag{2.48}$$

This approach leads naturally to matrix equations in angular momentum labels, where the angular momentum matrices are denoted with an underline. For instance, the Dyson equation for the t–matrix is then just

$$\underline{t}_i(\varepsilon) = \underline{V}_i(\varepsilon) + \underline{V}_i(\varepsilon) \, \underline{G}_{0,ii}(\varepsilon) \, \underline{t}_i(\varepsilon) \implies \underline{t}_i(\varepsilon) = \left[\underline{V}^{-1}_i(\varepsilon) - \underline{G}_{0,ii}(\varepsilon) \right]^{-1} \tag{2.49}$$

where the potential coefficients are

$$V_{i,L'L}(\varepsilon) = \int \mathrm{d}\vec{r} \int \mathrm{d}\vec{r}' \, j_{\ell'}(kr') \, Y^*_{L'}(\hat{r}') \, V_i(\vec{r}) \, \delta(\vec{r} - \vec{r}') \, j_\ell(kr) \, Y_L(\hat{r}) \ . \tag{2.50}$$

The GF is then, choosing the latter set of solutions,

$$G_i(\vec{r}, \vec{r}'; \varepsilon) = \sum_{LL'} Z_{i,L}(\vec{r}; \varepsilon) \, t_{i,LL'}(\varepsilon) \, Z^\times_{i,L'}(\vec{r}'; \varepsilon) - \sum_L Z_{i,L}(\vec{r}_<; \varepsilon) \, J^\times_{i,L}(\vec{r}_>; \varepsilon) \ . \tag{2.51}$$

The $^\times$ superscript means conjugate only the spherical harmonics, and $\vec{r}_>$ and $\vec{r}_<$ stand for the largest and the smallest of the vectors $\vec{r}$ and $\vec{r}'$, respectively.

The partial wave basis is a minimal basis for spherically symmetric potentials inside spherical regions. When dealing with non–spherical potentials and/or non–

spherical regions it can become quite cumbersome. The sum over angular momentum components is usually truncated at some value of ℓ. Typically $\ell_{\max} = 3$ or 4 is sufficient for most applications. The expansion of the potential also has to be truncated, but not necessarily at the same value of $\ell_{\max}$. For instance, the MTA and the ASA use $\ell_{\max} = 0$ for the potential, which is the spherical approximation.

When $\vec{r} \gg \vec{r}_1$ then $|\vec{r} - \vec{r}_1| \approx r - \hat{r} \cdot \vec{r}_1$, the Lippmann–Schwinger equation gives

$$\psi_i(\vec{r};\varepsilon) \approx e^{\mathrm{i}\vec{k}\cdot\vec{r}} - \frac{e^{\mathrm{i}kr}}{4\pi r}\int \mathrm{d}\vec{r}_1 \int \mathrm{d}\vec{r}'\, e^{\mathrm{i}(\vec{k}\cdot\vec{r}' - \vec{k}'\cdot\vec{r}_1)}\, t_i(\vec{r}_1,\vec{r}';\varepsilon) = e^{\mathrm{i}\vec{k}\cdot\vec{r}} + \frac{e^{\mathrm{i}kr}}{r}\, f(\vec{k}',\vec{k}) \quad (2.52)$$

by letting $\vec{k}' = k\hat{r}$. The initial planewave, with chosen $\vec{k}$ and $k = \sqrt{\varepsilon}$, is scattered into an outgoing spherical wave, with scattering amplitude from $\vec{k}$ to $\vec{k}'$ given by $f(\vec{k}',\vec{k})$, and its absolute value squared is the differential cross–section for the scattering process [45]. This asymptotic limit makes the connection with experiments in which a probing beam is fired at a sample, and its properties reconstructed from the signal measured at detectors placed around the sample.

2.2.3 Scattering by a collection of localised potentials

Suppose that the total potential is given as a sum of non–overlapping potentials V_i, centred at position vectors $\vec{R}_i$. If we work in the partial wave basis, operator equations become matrix equations in angular momentum labels. Eq. 2.49 then generalises to (omitting the energy dependence)

$$\underline{\mathcal{T}} = \sum_i \underline{V}_i + \sum_{ij} \underline{V}_i\, \underline{G}_{0,ij} \underline{V}_j + \sum_{ijk} \underline{V}_i\, \underline{G}_{0,ij} \underline{V}_j\, \underline{G}_{0,jk} \underline{V}_k + \dots \quad (2.53)$$

Making use of the single–site t–matrices introduced before, the Born series can be rewritten as [5, 45]

$$\underline{\mathcal{T}} = \sum_{ij} \underline{\tau}_{ij}\ , \quad \underline{\tau}_{ij} = \underline{t}_i\, \delta_{ij} + \underline{t}_i \sum_{k\neq i} \underline{G}_{0,ik}\, \underline{\tau}_{kj} = \underline{t}_i\, \delta_{ij} + (1-\delta_{ij})\, \underline{t}_i\, \underline{G}_{0,ij}\, \underline{t}_j + \cdots \quad (2.54)$$

where the scattering path operator (SPO) $\underline{\tau}_{ij}$ has been introduced. The scattering from each potential is contained in the $\underline{t}_i$'s, while the propagation between sites is given by $\underline{G}_{0,ij}$. The SPO then sums all scattering events which begin on site i and end on site j; each such sequence of scattering events is called a scattering path.

The other new quantity in this expression is the angular momentum representation of the free electron GF between different potential sites. This is given by

$$G_{0,ij,LL'}(\varepsilon) = -4\pi\,\mathrm{i}k \sum_{L_1} \mathrm{i}^{\ell-\ell'-\ell_1}\, C_{LL'}^{L_1}\, Y_{L_1}(\hat{R}_{ij})\, h_{\ell_1}^{+}(kR_{ij}) \tag{2.55}$$

where $\vec{R}_{ij} = \vec{R}_j - \vec{R}_i$ is the vector connecting the origins of the potentials V_i and V_j, R_{ij} its norm and $\hat{R}_{ij}$ the unit vector pointing from i to j. The coefficients $C_{LL'}^{L_1}$ are the Gaunt numbers,

$$C_{LL'}^{L_1} = \int \mathrm{d}\hat{r}\, Y_L(\hat{r})\, Y_{L'}^{*}(\hat{r})\, Y_{L_1}(\hat{r}) \tag{2.56}$$

which can also be expressed in terms of Clebsch–Gordan coefficients. The full GF is then written in terms of its blocks in site indices

$$G_{ij}(\vec{r},\vec{r}';\varepsilon) = Z_i(\vec{r};\varepsilon)\,\underline{\tau}_{ij}(\varepsilon)\,Z_j^{\times}(\vec{r}';\varepsilon) - \delta_{ij}\,Z_i(\vec{r}_<;\varepsilon)\,J_i^{\times}(\vec{r}_>;\varepsilon) \tag{2.57}$$

and omitting the summation over the angular momentum labels (cf. Eq. 2.51).

Now the main advantage of using MST and the KKR method becomes apparent: there is a clear separation between the potentials, whose local effects are summed up into the single–site t–matrices $\underline{t}_i$, and their arrangement in space, which is described by the reference GF connecting the potential sites, $\underline{G}_{0,ij}$.

2.2.4 Choice of reference system: free and screened propagators

The free electron GF has no special status in Multiple Scattering Theory. One could equally well choose a different reference Hamiltonian, and the theory would follow as before. This is the key idea to the screening transformation [47, 48]: the reference system is now taken as a set of repulsive potentials in the same geometry as the real system. For simplicity repulsive spherical muffin–tin potentials are usually chosen: $W_i(r) = W_0 > 0$ if $r < r_{\mathrm{MT}}$ and zero otherwise. The scattering properties of these potentials are encoded as usual in single–site t-matrices; $\underline{t}_r(\varepsilon)$ will then apply to all sites i. The GF for this new reference system has the following important property:

$$\underline{G}_{r,ij}(\varepsilon) \to \underline{0} \quad \text{exponentially when} \quad R_{ij} \to \infty \tag{2.58}$$

One can then consider only those sites which are less than some distance d apart, suitably chosen. The GF of the real system is connected to the one of the reference system by the usual Dyson equation:

$$G(z) = G_r(z) + G_r(z)\,\Delta\mathcal{T}(z)\,G_r(z) \quad . \tag{2.59}$$

Here $\Delta\mathcal{T}(z)$ denotes the difference in the scattering properties of the two systems. In terms of the single–site t–matrices this can be written as $\Delta\underline{t}_i(\varepsilon) = \underline{t}_i(\varepsilon) - \underline{t}_r(\varepsilon)$, where $\underline{t}_i$ are the previously described t–matrices, which in this context are referred to as bare or unscreened t–matrices. The corresponding SPOs are also connected in a simple way, as the transformation is diagonal in site indices [47]:

$$\underline{\tau}_{ij}(\varepsilon) = \underline{t}_i(\varepsilon)\left[\Delta\underline{t}_i(\varepsilon)^{-1}\Delta\underline{\tau}_{ij}(\varepsilon)\,\Delta\underline{t}_j(\varepsilon)^{-1} + \delta_{ij}\left(\underline{t}_i(\varepsilon)^{-1} - \Delta\underline{t}_i(\varepsilon)^{-1}\right)\right]\underline{t}_j(\varepsilon) \tag{2.60}$$

and the screened SPO is now obtained by inverting a sparse matrix, by construction:

$$\Delta\underline{\tau}_{ij}(\varepsilon) = \left[\Delta\underline{t}^{-1}(\varepsilon) - \underline{G}_r(\varepsilon)\right]^{-1}_{ij} \quad . \tag{2.61}$$

2.2.5 Bulk, surface and interface formulations, and the embedded cluster method

In most situations one is interested in a regular arrangement of single–site potentials. Examples are crystalline lattices, and their surfaces or interfaces. Here one can exploit two or three–dimensional periodicity to enormously reduce the computational burden, often in combination with the screening transformation. However, systems in which there is no periodicity, such as clusters, nanoparticles, and molecular junctions, are also within the remit of MST. This is in contrast to most other electronic structure methods, in which some supercell approximation must be made, in order to restore the necessary periodicity for calculations. Periodicity in a given spatial direction can be assumed if there is a macroscopic number of atoms in that direction: for instance, the number of atoms in a bulk sample is of the order of Avogadro's number, 10^{23}, and so the limit $N \to \infty$ at constant density is taken, the so–called thermodynamic limit.

Let us start from the most basic case: three–dimensional periodicity, which applies to bulk systems. The locations of the potentials are given in a regular fashion: $\vec{R}_{Ii} = \vec{R}_I + \vec{R}_i$. Here $\vec{R}_I$ sets the origin of sublattice I, and $\vec{R}_i = i_1\,\vec{a}_1 + i_2\,\vec{a}_2 + i_3\,\vec{a}_3$ spans the Bravais lattice associated with the lattice vectors $\vec{a}_1$, $\vec{a}_2$ and $\vec{a}_3$, common to all sublattices, with i_1, i_2 and i_3 running over all positive and negative integers. On every site of each sublattice the same potential can be found, $V_{Ii} = V_I$, and so the same single–site t–matrix $\underline{t}_I$ applies to all the sites in each sublattice. The reciprocal Bravais lattice is spanned by the vectors $\vec{b}_1$, $\vec{b}_2$ and $\vec{b}_3$, such that $\vec{a}_\alpha \cdot \vec{b}_\beta = 2\pi\delta_{\alpha\beta}$, for $\alpha, \beta = 1, 2, 3$. Due to periodicity, the crystal momentum can be restricted to $\vec{k} = k_1\,\vec{b}_1 + k_2\,\vec{b}_2 + k_3\,\vec{b}_3$, where k_1, k_2 and k_3 are reals, say between zero and one.

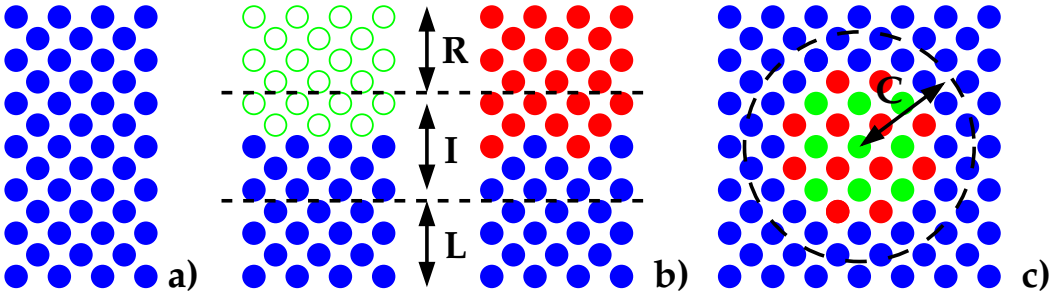

Figure 2.1: Commonly encountered geometries: a) — Simple bulk geometry, the same potential on all sites. b) — Two examples of semi–infinite structures. Semi–infinite left (L) and right (R) regions interact through an interfacial (I) region. The left diagram represents a surface, with vacuum depicted by empty circles, while the right diagram represents an interface between different materials; the dashed lines delimit the notional interface region, where the potentials deviate from their bulk counterparts. c) — A cluster embedded in a host material. The dashed circle represents the embedding region, including perturbed host atoms due to the presence of the neighbouring cluster atoms.

The SPO is given by Eq. 2.54. Summing over the Fourier factors yields

$$\underline{\tau}_{IJ}(\vec{k};\varepsilon) = \underline{t}_I(\varepsilon)\,\delta_{IJ} + \underline{t}_I(\varepsilon)\sum_K \underline{G}_{0,IK}(\vec{k};\varepsilon)\,\underline{\tau}_{KJ}(\vec{k};\varepsilon) \tag{2.62}$$

with the Fourier transforms defined as ($\vec{R}_{Ii,Jj} = \vec{R}_{Jj} - \vec{R}_{Ii} = \vec{R}_{IJ} + \vec{R}_{ij}$)

$$\underline{A}_{IJ}(\vec{k}) = \sum_j e^{\mathrm{i}\vec{k}\cdot\vec{R}_{Ii,Ij}}\underline{A}_{Ii,Jj}\ , \quad \underline{A}_{Ii,Jj} = \int \frac{\mathrm{d}\vec{k}}{V_{\mathrm{BZ}}}\, e^{-\mathrm{i}\vec{k}\cdot\vec{R}_{Ii,Jj}}\underline{A}_{IJ}(\vec{k}) \tag{2.63}$$

and noting the independence of the choice of origin for the lattice. Then the Fourier transformed SPO can be obtained by inverting a matrix in sublattice and angular momentum indices,

$$\underline{\tau}(\vec{k};\varepsilon) = \left[\underline{t}^{-1}(\varepsilon) - \underline{G}_0(\vec{k};\varepsilon)\right]^{-1} \Longrightarrow\ \underline{\tau}_{Ii,Jj}(\varepsilon) = \int \frac{\mathrm{d}\vec{k}}{V_{\mathrm{BZ}}}\, e^{-\mathrm{i}\vec{k}\cdot\vec{R}_{Ii,Jj}}\underline{\tau}_{IJ}(\vec{k};\varepsilon) \tag{2.64}$$

with V_{BZ} the volume of the Brillouin zone, the Wigner–Seitz cell of the reciprocal space Bravais lattice. These statements are equivalent to the traditional Bloch theorem of solid state physics. Site–diagonal properties, such as the density in each unit cell, are independent of the site index, but depend on the sublattice index.

In surface and interface calculations the system of interest is taken to be an infinite stack of layers, all of which must possess the same two–dimensional Bravais

lattice. Then only two–dimensional periodicity applies; in the previous definitions we now have the restrictions $\vec{R}_i = i_1\,\vec{a}_1 + i_2\,\vec{a}_2$ and $\vec{k} = k_1\,\vec{b}_1 + k_2\,\vec{b}_2$. The vectors $\vec{R}_I$ set the origin for each layer. The next step is to assume that the surface or interface region is of finite extent, being confined to a finite number N of layers (see Fig. 2.1):

- Left region (L): layers numbered $-\infty \leq P \leq 0$; a set of t–matrices $\underline{t}_L$ applies;
- Right region (R): layers numbered $N+1 \leq P \leq +\infty$; a set $\underline{t}_R$ applies;
- Surface/interface region (I): layers numbered $1 \leq P \leq N$; a set $\underline{t}_I$ applies.

This means that Eq. 2.64 is now infinite–dimensional. If however the left and right regions only couple through the interface region, the matrix to be inverted has the following structure:

$$\begin{pmatrix} \underline{t}_L^{-1}(\varepsilon) - \underline{G}_{0,LL}(\varepsilon;\vec{k}) & -\underline{G}_{0,LI}(\varepsilon;\vec{k}) & \underline{0} \\ -\underline{G}_{0,IL}(\varepsilon;\vec{k}) & \underline{t}_I^{-1}(\varepsilon) - \underline{G}_{0,II}(\varepsilon;\vec{k}) & -\underline{G}_{0,IR}(\varepsilon;\vec{k}) \\ \underline{0} & -\underline{G}_{0,RI}(\varepsilon;\vec{k}) & \underline{t}_R^{-1}(\varepsilon) - \underline{G}_{0,RR}(\varepsilon;\vec{k}) \end{pmatrix} \tag{2.65}$$

and the central block is of finite dimension in layer indices. This matrix can be formally inverted using the inversion formula for block matrices. Defining

$$\underline{\tau}_{LL}(\varepsilon;\vec{k}) = \left[\underline{t}_L^{-1}(\varepsilon) - \underline{G}_{0,LL}(\varepsilon;\vec{k})\right]^{-1}, \quad \underline{\tau}_{RR}(\varepsilon;\vec{k}) = \left[\underline{t}_R^{-1}(\varepsilon) - \underline{G}_{0,RR}(\varepsilon;\vec{k})\right]^{-1} \tag{2.66}$$

which are the SPOs for the left and right semi–infinite regions, and so can be obtained by the procedures applicable to bulk calculations, the SPO corresponding to the interface block is then given by the following finite–dimensional matrix inversion:

$$\begin{aligned} \underline{\tau}_{II}(\varepsilon;\vec{k}) = \Big[\underline{t}_I^{-1}(\varepsilon) - \underline{G}_{0,II}(\varepsilon;\vec{k}) - \underline{G}_{0,IL}(\varepsilon;\vec{k})\,\underline{\tau}_{LL}(\varepsilon;\vec{k})\,\underline{G}_{0,LI}(\varepsilon;\vec{k}) \\ - \underline{G}_{0,IR}(\varepsilon;\vec{k})\,\underline{\tau}_{RR}(\varepsilon;\vec{k})\,\underline{G}_{0,RI}(\varepsilon;\vec{k})\Big]^{-1} . \end{aligned} \tag{2.67}$$

The inverse of the first two terms would be the SPO for the isolated N interface layers; then the next two terms are the coupling to the left and right regions. Two–dimensional periodicity still applies.

All these computations are made very efficient by choosing the screened reference GF instead of the free electron one for $\underline{G}_0$. Then the matrix becomes sparse, and the method of Godfrin can be used [49].

If there is no overall periodicity to be exploited, but the system can be thought of as a perturbed region in some host material, one can make use of the embedded cluster method. The t–matrices for sites inside the cluster or perturbed region are

denoted $\underline{t}_c$ while for the unperturbed host atoms $\underline{t}_h$ applies. The SPO for the cluster region is then expressed in terms of the difference between the scattering properties of the atoms in the cluster and those of the host material [50, 51]:

$$\underline{\tau}_c(\varepsilon) = \underline{\tau}_h(\varepsilon) \left[\underline{1} + \left(\underline{t}_c^{-1}(\varepsilon) - \underline{t}_h^{-1}(\varepsilon)\right) \underline{\tau}_h(\varepsilon)\right]^{-1} \quad . \tag{2.68}$$

Here $\underline{\tau}_h$ is the SPO one obtains for the cluster region if all atoms are replaced with host atoms. As no use can be made of translational invariance, the matrix size is the product of the size of the angular momentum matrices and of the number of atoms in the cluster. Iterative methods can also be considered, again exploiting the short–range nature of the screened GFs. This geometry is depicted in the right panel of Fig. 2.1.

2.2.6 Relativistic formulation and magnetism

In the non–relativistic case, the angular momentum label $L = (\ell, m)$ combines the eigenvalues of $\hat{L}^2$ and $\hat{L}_z$, and can also include a spin label, for the spin–polarised case, according to the eigenvalues of $\hat{S}_z\,\Phi_s = s\,\Phi_s$, where Φ_s is a two–component spinor ($s = \pm 1/2$). These are the appropriate quantum numbers for a Schrödinger–type single–particle Hamiltonian. However, once relativistic effects are considered, in particular spin–orbit coupling (SOC), the spin and orbital degrees of freedom become coupled, and the more appropriate operators to consider are the total angular momentum $\hat{J} = \hat{L} + \hat{S}$ and the SOC operator $\hat{K} = 2\hat{L}\cdot\hat{S} + \hat{\mathcal{I}}$. Instead of spherical harmonics, augmented by spinor functions, one must now construct appropriate linear combinations, called spin–spherical harmonics, and denoted by $\chi_{\kappa\mu}(\hat{r})$. The new quantum numbers are obtained from $\hat{J}^2\,\chi_{\kappa\mu}(\hat{r}) = j(j+1)\,\chi_{\kappa\mu}(\hat{r})$, $\hat{J}_z\,\chi_{\kappa\mu}(\hat{r}) = \mu\,\chi_{\kappa\mu}(\hat{r})$ and $\hat{K}\,\chi_{\kappa\mu}(\hat{r}) = -\kappa\,\chi_{\kappa\mu}(\hat{r})$. The following relationship is to be noted: $\kappa = \ell$ if $j = \ell - 1/2$, and $\kappa = -\ell - 1$ if $j = \ell + 1/2$, so there is a unique connection between κ and j, through ℓ. The relativistic quantum numbers are thus usually chosen as $Q = (\kappa, \mu)$, and there is a simple correspondence between the (ℓ, m, s) and (κ, μ) representations, through the relevant Clebsch–Gordan coefficients [5].

Formally all the equations of MST are very similar in the non–relativistic and relativistic cases. However, there are two main caveats: working at complex energies and the presence of an effective magnetic field in the equations. These lead to several technical complications [5], but also to some symmetry considerations. While the non–relativistic Hamiltonian commutes with spin rotations, which leads to the generalised Bloch theorem used in magnetic spin spiral calculations [52], this is no longer true in the relativistic case. The Dirac Hamiltonian only remains invariant

to transformations which rotate simultaneous the spin and the coordinates [52].

In the relativistic case, the radial solutions become bi–spinors, and the particle and magnetisation densities in the cell corresponding to site i follow from Eq. 2.38, with $\varepsilon = \sqrt{\omega^2 - m^2c^4}$:

$$\begin{aligned} n_i(\vec{r}) &= -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\, G_{ii}(\vec{r},\vec{r};\varepsilon) \\ \vec{m}_i(\vec{r}) &= -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\, \beta\vec{\Sigma}\, G_{ii}(\vec{r},\vec{r};\varepsilon) \end{aligned} \tag{2.69}$$

which integrate to the total number of electrons and total magnetic moment in the cell. The DOS is given by

$$n_i(\varepsilon) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\vec{r} \sum_{QQ'} Z_{i,Q}(\vec{r};\varepsilon)\, \tau_{ii,QQ'}(\varepsilon)\, Z^{\times}_{i,Q'}(\vec{r};\varepsilon) + \ldots = \sum_{QQ'} n_{i,QQ'}(\varepsilon) \tag{2.70}$$

and the eigendecomposition of this angular momentum matrix provides information on the number of states (eigenvalue) of each symmetry (eigenvector). In particular, the amount of mixing betweeen spin–up and spin–down channels due to SOC can be assessed, after transforming back to the (ℓ, m, s) basis.

The bandstructure is given by the poles of the Bloch spectral function. If there is translational invariance, and for simplicity all sites are taken to be equivalent,

$$\begin{aligned} A_{\mathrm{B}}(\vec{k},\varepsilon) &= \sum_n \delta\big(\varepsilon - \varepsilon_n(\vec{k})\big) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\vec{r}\, G(\vec{r},\vec{r};\vec{k},\varepsilon) \\ &= -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\vec{r} \sum_{QQ'} Z_Q(\vec{r};\varepsilon)\, \underline{\tau}_{QQ'}(\vec{k},\varepsilon)\, Z^{\times}_{Q'}(\vec{r};\varepsilon) + \ldots \end{aligned} \tag{2.71}$$

which just requires the Fourier transform of the SPO.

Chapter 3

Ensemble averages and the effective medium approach

In DFT the original quantum–mechanical problem involving a large number of degrees of freedom (DOFs) was reduced to a much simpler problem, involving only the groundstate densities of the system of interest. The same approach is taken in statistical mechanics: the properties of the system are described by a small set of equilibrium variables. The parallel goes further: both formalisms can be cast into a variational form, and conventional DFT can be constructed as the zero–temperature limit of the electronic grand–potential, which will also be demonstrated here. The variational principle for statistical mechanics affords both a well–defined method of approximation, the Feynman–Peierls–Bogoliubov inequality, as well as a systematic perturbation theory, based on the generalised cumulant expansion.

Then the problem of computing the equilibrium properties of the interacting electron system is discussed. An introduction is given to several types of disorder, and the effective medium approach is presented as an alternative to the brute force calculation of the ensemble averages. It is shown that the GF is the appropriate quantity to work with, and the so–called single–site Coherent Potential Approximation (CPA) is derived, after some preliminaries on impurities in a host medium. The non–local extension of the method is outlined last.

3.1 Variational statistical mechanics

3.1.1 Basic results in statistical mechanics

The equilibrium properties of a system described by the Hamiltonian $\hat{\mathcal{H}}$ can be obtained from a variational principle. In the grand–canonical ensemble, where the

number of particles is allowed to vary, the partition function is given by

$$\mathcal{Z} = \operatorname{Tr} e^{-\beta\left(\hat{\mathcal{H}}-\mu\hat{N}\right)} = e^{-\beta\Omega} \tag{3.1}$$

which defines the thermodynamic grand–potential, tracing (i.e. summing or integrating) over all DOFs. Now define the grand–potential functional as

$$\widetilde{\Omega}[\hat{\rho}] = \langle\, \hat{\mathcal{H}} + \frac{1}{\beta}\log\hat{\rho} - \mu\hat{N}\,\rangle \quad \text{with} \quad \langle\, \hat{A}\,\rangle = \operatorname{Tr}\hat{\rho}\,\hat{A}\,, \quad \operatorname{Tr}\hat{\rho} = 1 \quad \text{and} \quad \hat{\rho} \geq 0 \tag{3.2}$$

for some trial probability density operator $\hat{\rho}$. Here $1/\beta = k_{\mathrm{B}}T$ is the temperature factor, μ is the chemical potential as before, and $\hat{N}$ is the particle number operator. Using the convexity properties of exponentials, it is straightforward to show that the minimum of $\widetilde{\Omega}[\hat{\rho}]$ is given by the thermodynamic grand–potential [53]:

$$e^{-\beta\Omega} \equiv \left\langle\, e^{-\beta\left(\hat{\mathcal{H}}-\mu\hat{N}\right)-\log\hat{\rho}}\,\right\rangle \geq e^{-\beta\langle\,\hat{\mathcal{H}}-\mu\hat{N}+1/\beta\log\hat{\rho}\,\rangle} = e^{-\beta\widetilde{\Omega}[\hat{\rho}]} \tag{3.3}$$

and it is attained for the correct probability density operator,

$$\hat{\rho} = \frac{e^{-\beta\left(\hat{\mathcal{H}}-\mu\hat{N}\right)}}{\mathcal{Z}} \quad\Longrightarrow\quad \Omega = \langle\,\hat{\mathcal{H}}\,\rangle + \frac{1}{\beta}\langle\,\log\hat{\rho}\,\rangle - \mu\,\langle\,\hat{N}\,\rangle = U - TS - \mu N \quad . \tag{3.4}$$

Here the internal energy U, the entropy S and the average number of particles N are also defined.

From the same inequality a variational bound on the grand–potential can be immediately derived. Suppose that we consider some Hamiltonian $\mathcal{H}_0$, parametrised in a convenient way. From the respective probability density operator we have

$$\Omega \leq \widetilde{\Omega}[\hat{\rho}_0] = \langle\,\hat{\mathcal{H}}\,\rangle + \frac{1}{\beta}\langle\,\log\hat{\rho}_0\,\rangle - \mu\,\langle\,\hat{N}\,\rangle \quad \text{with} \quad \langle\,\hat{A}\,\rangle = \operatorname{Tr}\hat{\rho}_0\,\hat{A} \quad . \tag{3.5}$$

Minimising $\widetilde{\Omega}[\hat{\rho}_0]$ with respect to the available variational parameters will then provide an upper bound to the true grand–potential. This result has been derived several times, and will be called the Feynman–Peierls–Bogoliubov inequality [54]. It is valid in both classic and quantum statistical mechanics.

3.1.2 Generalised cumulant expansion method

Another way of developing an approximate treatment of the statistical mechanical problem is by employing the generalised cumulant expansion method [55]. The classical statistical mechanics case is considered, so the operator notation is dropped. Similarly to the GF approach of referring the problem to a suitable reference system,

we formally split the Hamiltonian as follows:

$$\mathcal{H} = \mathcal{H}_0 + (\mathcal{H} - \mathcal{H}_0) = \mathcal{H}_0 + \Delta\mathcal{H} \ . \tag{3.6}$$

In the canonical ensemble the partition function is given by

$$\mathcal{Z} = \mathrm{Tr}\, e^{-\beta\mathcal{H}} = \mathrm{Tr}\, e^{-\beta\mathcal{H}_0}\, \frac{\mathrm{Tr}\, e^{-\beta\mathcal{H}_0}\, e^{-\beta\Delta\mathcal{H}}}{\mathrm{Tr}\, e^{-\beta\mathcal{H}_0}} = \mathcal{Z}_0 \, \langle\, e^{-\beta\Delta\mathcal{H}} \,\rangle \ , \quad \rho_0 = \frac{e^{-\beta\mathcal{H}_0}}{\mathrm{Tr}\, e^{-\beta\mathcal{H}_0}} \tag{3.7}$$

where the ensemble averages are taken with the probability density ρ_0. This is an exact statement. The exact free energy is then

$$\mathcal{F} = -\frac{1}{\beta} \log \mathcal{Z} = -\frac{1}{\beta} \log \mathcal{Z}_0 - \frac{1}{\beta} \log \langle\, e^{-\beta\Delta\mathcal{H}} \,\rangle = \mathcal{F}_0 + \Delta\mathcal{F} \tag{3.8}$$

and the difference can be expanded in a power series (if it converges):

$$\Delta\mathcal{F} = -\frac{1}{\beta} \log \Big(1 - \Big\langle 1 - e^{-\beta\Delta\mathcal{H}} \Big\rangle \Big) = -\frac{1}{\beta} \sum_{n=1}^{\infty} \frac{1}{n} \Big\langle \sum_{m=1}^{\infty} \frac{(-\beta)^m}{m!} (\Delta\mathcal{H})^m \Big\rangle^n \tag{3.9}$$

which corresponds to the Ursell–Mayer expansion [55]. The first two terms are:

$$\Delta\mathcal{F}_1 = \langle\, \Delta\mathcal{H} \,\rangle \ , \quad \Delta\mathcal{F}_2 = \Delta\mathcal{F}_1 - \frac{\beta}{2} \Big(\langle\, \Delta\mathcal{H}^2 \,\rangle - \langle\, \Delta\mathcal{H} \,\rangle^2 \Big) \tag{3.10}$$

which require the first and the second cumulant, $\langle\, x \,\rangle$ and $\langle\, x^2 \,\rangle - \langle\, x \,\rangle^2$. The general expression sums all cumulants, which gives the method its name. The quantum version is called the linked cluster expansion [55], but the commutation properties of the operators make the result more involved.

Taking $\mathcal{H}_0 = 0$ we regain the high–temperature series expansion of the free energy, which is a power series in the inverse temperature β. Approximating $\Delta\mathcal{F} \approx \Delta\mathcal{F}_1$ and minimising the resulting free energy with respect to parameters in $\mathcal{H}_0$ reproduces the previous variational argument, Eq. 3.5:

$$\mathcal{F} \leq \mathcal{F}_1 = \mathcal{F}_0 + \langle\, \mathcal{H} - \mathcal{H}_0 \,\rangle \ . \tag{3.11}$$

If one takes more terms it is no longer clear if the true free energy is bounded from above or below. However one can attempt to minimise $|\Delta\mathcal{F}|$, or a suitably truncated version, as indicated before.

In connection with lattice models, the variational method combined with the $\Delta\mathcal{F}_1$ approximation yields the conventional mean–field theory of statistical mechanics. More refined approximations then lead to the variational cumulant expansion, which

has been used with some success in treating Heisenberg and Ising models [56], and also to the variational cluster expansion [57, 58] and spin cluster expansion [59, 60] used in the theories of chemical/compositional ordering in alloys and magnetic ordering.

3.1.3 The Fluctuation–Dissipation theorem

A very important result in Statistical Mechanics is the fluctuation–dissipation theorem, which connects the energy dissipated in the system to its statistical fluctuations. Let $\mathcal{H}$ depend on a discrete set of DOFs $\{x\}$; this formulation is appropriate for lattice models, but it can easily be extended to continuum models. Consider also the conjugate fields to those DOFs, $\{h\}$ defined by the way they enter the Hamiltonian: $\mathcal{H}(\{x\}) \longrightarrow \mathcal{H}(\{x\}) - \sum_i h_i x_i$. The free energy is given by

$$\mathcal{F} = -\frac{1}{\beta} \log \mathrm{Tr}\, e^{-\beta \mathcal{H}(\{x\})} = -\frac{1}{\beta} \log \mathcal{Z} \tag{3.12}$$

and the mean value of x_i and the pair–correlation function follow,

$$\langle x_i \rangle = \frac{\mathrm{Tr}\, x_i\, e^{-\beta \mathcal{H}(\{x\})}}{\mathcal{Z}} = -\frac{\partial \mathcal{F}}{\partial h_i} \tag{3.13}$$

$$\beta(\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle) = -\frac{\partial^2 \mathcal{F}}{\partial h_i \partial h_j} = \frac{\partial \langle x_i \rangle}{\partial h_j} \equiv \chi_{ij} \quad . \tag{3.14}$$

The last equation can be interpreted as follows: the susceptibility quantifies the change in the order parameter $\langle x_i \rangle$ when there is a change in the conjugate field, while the pair–correlation $\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$ measures the fluctuations away from the order parameters. Thus the amount of energy dissipated into the system by the conjugate fields is connected to the fluctuations by the temperature factor β; this is one version of the general fluctuation–dissipation theorem [61].

3.1.4 Density Functional Theory at finite temperature

The finite temperature extension of DFT follows immediately from the properties of the grand–potential functional, by considering the Hamiltonian with the added external potential, as shown by Mermin [53]. The particle density determines all equilibrium properties, so the grand–potential functional can also be taken as a functional of the equilibrium particle density.

For definiteness, the statistical properties of an ensemble of non–interacting electrons are briefly reviewed. The energy eigenvalues ε_n are independent of the particle number, and the grand–potential follows by summing over the possible occupancies

of each eigenstate (0 or 1 for fermions) with the corresponding Boltzmann factor:

$$\Omega_0 = -\frac{1}{\beta}\sum_n \log\left[1 + e^{-\beta(\varepsilon_n - \mu)}\right] = U_0 - TS_0 - \mu N \quad . \tag{3.15}$$

The internal energy is simply $U_0 = \sum_n f_n\,\varepsilon_n$, the entropy of non–interacting fermions S_0 is given by

$$S_0 = -k_{\mathrm{B}}\sum_n \Big[f_n \log f_n + (1 - f_n)\log(1 - f_n)\Big] \;, \quad f_n = \frac{1}{e^{\beta(\varepsilon_n - \mu)} + 1} \tag{3.16}$$

and the occupation function f_n is now the Fermi–Dirac function. Introducing the DOS, Eq. 2.33, the grand–potential can be rewritten as

$$\Omega_0 = -\frac{1}{\beta}\int \mathrm{d}\varepsilon\, \log\left[1 + e^{-\beta(\varepsilon - \mu)}\right] n(\varepsilon) = -\int \mathrm{d}\varepsilon\, f(\varepsilon - \mu)\, N(\varepsilon) \tag{3.17}$$

integrating by parts and using the definition of the number of electrons at energy ε, Eq. 2.34. This formulation is the main ingredient in combining DFT and statistical mechanics, as will be seen in the following.

Generalising the previous expressions, Eqs. 2.4, 2.5 and 2.11,

$$\begin{aligned}\Omega[\hat{\rho}] \longrightarrow \Omega[n] &= F_{\mathrm{HK}}[n] + \int \mathrm{d}\vec{r}\, n(\vec{r})\Big(V_{\mathrm{ext}}(\vec{r}) - \mu\Big) \\ &= \Omega_0[n] - E_{\mathrm{H}}[n] + \Omega_{\mathrm{xc}}[n] - \int \mathrm{d}\vec{r}\, n(\vec{r})\, V_{\mathrm{xc}}(\vec{r})\end{aligned} \tag{3.18}$$

where the Hohenberg–Kohn energy functional becomes a Helmholtz free energy, and the remaining quantities are defined as before, with the new occupation function, and the appropriate double–counting terms. This result can also be shown directly by starting with Eq. 2.11, and allowing the occupation factors to vary, adding the correction term which keeps the functional variational [62].

The exchange–correlation contribution is now much more involved, but for the temperatures of relevance in typical condensed matter systems the main contribution is the correction to the specific heat. The LDA then generalises as

$$\Omega_{\mathrm{xc}}^{\mathrm{LDA}}[n] = E_{\mathrm{xc}}^{\mathrm{LDA}}[n] - \frac{T^2}{2}\int \mathrm{d}\vec{r}\, \gamma_{\mathrm{xc}}^{\mathrm{LDA}}(n(\vec{r})) \tag{3.19}$$

and a calculation of the specific heat coefficient for the homogeneous electron gas is provided in [63]. Working at finite temperature also eliminates many technical points which arise in the formal development of conventional DFT, concerning degenerate states.

Finite temperature DFT has not been pursued much in this form because the exchange–correlation functional does not capture the effects of thermally excited collective modes of the electronic system (phonons, magnons, ...). There are a few recent studies which attempt unified first–principles calculations of free energies of real materials [64], but most calculations still proceed by modelling thermal effects on top of conventional zero–temperature DFT calculations. A related question is what is the feedback between the thermally excited collective modes and the underlying electronic structure, or more generally what is the connection between the effective single–particle electronic structure and the ensemble averaged properties of a given system. An approximate way of addressing this matter will be described next.

3.2 The Coherent Potential Approximation

3.2.1 Types of disorder and the effective medium approach

The equilibrium properties of a physical system are given by an ensemble average over all possible configurations of the system. One can either take the brute force numerical approach, by sampling an appropriate set of configurations and estimating the partition function directly, or adopt some approximation scheme, to circumvent this calculation. The latter is the spirit of the effective medium approach: one constructs a fictitious system, the effective medium, which reproduces the average properties of the true physical system.

The problem can be further simplified if the collective modes of the system evolve on a much slower timescale than that pertaining to the electronic motions. This timescale separation is not as justifiable as the Born–Oppenheimer approximation, as there is no obvious small parameter to which one can refer. However it affords a very convenient picture: on the electronic timescale the collective DOFs are frozen in some configuration, and act as background to the electronic motion. The energy of each configuration then dictates the evolution of the system on the slower timescale. This is still a self–consistent problem, as the collective DOFs are themselves set up by the electrons.

One could consider different types of disorder, corresponding to the relevant DOFs (see Fig. 3.1):

- Chemical disorder: several kinds of atoms are present, not necessarily at the sites of a regular lattice;
- Positional disorder: there are atomic displacements from the sites of an ideal

periodic structure;

- Magnetic disorder: the orientations of the magnetic moments do not form an ordered structure.

Several types of disorder may be present at the same time, but from the computational point of view it will be very important if the locations of the atoms form a periodic structure or not, as in the former case the effective medium can be chosen to have the periodic properties of the underlying structure. For the formal developments such assumptions are not needed.

3.2.2 The Green function as a self–averaging operator

For a particular configuration of the system, the electronic properties can be computed through MST from the corresponding GF, as shown by Eq. 2.38. It then follows that the ensemble–averaged properties can be computed from the ensemble–averaged GF:

$$\langle G \rangle = G_0 + G_0 \langle \mathcal{T} \rangle G_0 \ , \quad \mathcal{T} = \sum_{ij} \tau_{ij} \tag{3.20}$$

as the reference GF is independent of the DOFs of the system.

If one instead refers the GF to the one corresponding to some effective medium $\widetilde{\mathcal{T}}$, independent of the local DOFs involved on the ensemble averaging, one obtains the modified Dyson equation which was already discussed in the subsection on the screening transformation:

$$\langle G \rangle = \widetilde{G} + \widetilde{G} \langle \Delta\mathcal{T} \rangle \widetilde{G} \ , \quad \widetilde{G} = G_0 + G_0 \widetilde{\mathcal{T}} G_0 \tag{3.21}$$

where $\langle \Delta\mathcal{T} \rangle$ is the averaged difference between the scattering properties of the real system and those of the effective medium. The prescription for setting up the effective medium is then simply to make this averaged difference as small as possible.

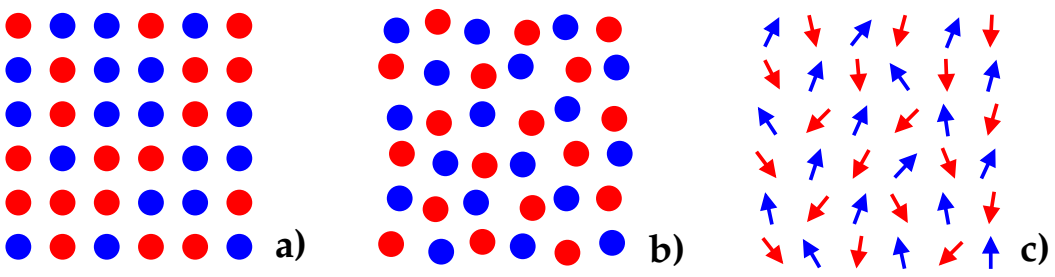

Figure 3.1: Some examples of disorder types: a) — Chemical disorder. b) — Positional disorder. c) — Magnetic disorder.

A self–consistent formulation is obtained by defining an approximation to the exact averaged difference, by algebraic or diagrammatic methods, and then to enforce $\langle\, \Delta\mathcal{T}\,\rangle_{\text{approx}} = 0$. This is the key idea for the CPA.

Turning around the initial definition of the GF as the resolvent of the Hamiltonian, one can introduce an effective Hamiltonian and an accompanying self–energy,

$$\langle\, G(z)\,\rangle^{-1} = z - \mathcal{H}_{\text{eff}}(z) = z - \mathcal{H}_0 - \Sigma(z) \quad . \tag{3.22}$$

These are non–Hermitian, satisfying instead the reality conditions $\mathcal{H}_{\text{eff}}(z^*) = \mathcal{H}_{\text{eff}}(z)^\dagger$ and $\Sigma(z^*) = \Sigma(z)^\dagger$, properties inherited from the GF. This also suggests the equivalence between a diagrammatic approach, where the focus is usually on approximating the self–energy, and the algebraic approach, where one seeks to approximate the scattering matrices.

3.2.3 The single–site coherent potential approximation

The single–site CPA is the equivalent in scattering theory of the mean–field theory in statistical mechanics: the problem is solved exactly for a single site, and the interactions with the rest of the system are described by an effective field. This was also done for the DFT problem: the Kohn–Sham scheme describes the motion of a single electron in a self–consistent field, so it is also a mean–field theory. The point here is that mean–field theories should not be regarded lightly.

Let ξ_i label the possible values at site i of the local DOF corresponding to the slowly evolving collective DOF, and suppose that it is possible to represent its effect by some local potential, $V_i(\xi_i)$, with associated single–site t–matrix $\underline{t}_i(\xi_i)$. The scattering properties of all sites other than i are determined by some effective t–matrices $\tilde{\underline{t}}_j$, whose construction is still to be specified. If one replaces $\tilde{\underline{t}}_i$ by $\underline{t}_i(\xi_i)$ at i, the site–diagonal SPO is immediately given by

$$\langle\, \underline{\tau}_{ii}\,\rangle_{\xi_i} = \widetilde{\underline{\tau}}_{ii} \left[\underline{1} + \left(\underline{t}_i^{-1}(\xi_i) - \tilde{\underline{t}}_i^{-1}\right) \widetilde{\underline{\tau}}_{ii}\right]^{-1} \tag{3.23}$$

by recalling Eq. 2.68. The host medium is the effective medium, whose properties are determined by the $\tilde{\underline{t}}_j$'s, and the respective SPO is denoted by $\widetilde{\underline{\tau}}_{ij}$. Introducing the impurity matrix

$$\underline{D}_i(\xi_i) = \left[\underline{1} + \left(\underline{t}_i^{-1}(\xi_i) - \tilde{\underline{t}}_i^{-1}\right) \widetilde{\underline{\tau}}_{ii}\right]^{-1} \tag{3.24}$$

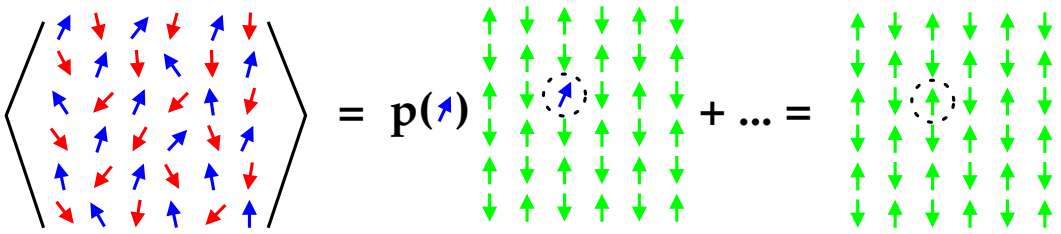

Figure 3.2: The construction of the partial averages. The statistical average is performed on all sites but one, indicated by the dotted circle. Then the averaging is completing by summing over the possible values of the DOF, weighted with the appropriate probability. Here this is illustrated for magnetic DOFs. The average magnetic moment on each site is represented by the green arrows, while the possible values of their orientations correspond to the red and blue arrows.

and the excess scattering matrix,

$$\underline{X}_i(\xi_i) = \left[\left(\underline{\tilde{t}}_i^{-1} - \underline{t}_i^{-1}(\xi_i) \right)^{-1} - \underline{\tilde{\tau}}_{ii} \right]^{-1} \tag{3.25}$$

the partially averaged SPO can be rewritten as

$$\langle\, \underline{\tau}_{ii} \,\rangle_{\xi_i} = \underline{\tilde{\tau}}_{ii}\, \underline{D}_i(\xi_i) = \underline{\tilde{\tau}}_{ii} + \underline{\tilde{\tau}}_{ii}\, \underline{X}_i(\xi_i)\, \underline{\tilde{\tau}}_{ii} \quad . \tag{3.26}$$

With the probability $P_i(\xi_i)$ of finding the value ξ_i for the local DOF, one can complete the averaging (see Fig. 3.2):

$$\langle\, \underline{\tau}_{ii} \,\rangle = \sum_{\xi_i} P_i(\xi_i)\, \langle\, \underline{\tau}_{ii} \,\rangle_{\xi_i} = \underline{\tilde{\tau}}_{ii} + \underline{\tilde{\tau}}_{ii} \left(\sum_{\xi_i} P_i(\xi_i)\, \underline{X}_i(\xi_i) \right) \underline{\tilde{\tau}}_{ii} \tag{3.27}$$

The effective medium is specified by requiring the vanishing of the average excess scattering:

$$\langle\, \underline{\tau}_{ii} \,\rangle = \underline{\tilde{\tau}}_{ii} \quad \Longrightarrow \quad \sum_{\xi_i} P_i(\xi_i)\, \underline{X}_i(\xi_i) = \underline{0} \tag{3.28}$$

which shows that this approximation is site–diagonal, and completes the single–site CPA prescription [45, 46, 65].

The site–diagonal GF follows without effort, from Eq. 2.57. For each value of ξ_i one can define the respective GF,

$$G_{ii}(\vec{r}, \vec{r}'; z, \xi_i) = Z_i(\vec{r}; z, \xi_i)\, \langle\, \underline{\tau}_{ii}(z) \,\rangle_{\xi_i}\, Z_i^{\times}(\vec{r}'; z, \xi_i) - Z_i(\vec{r}_<; z, \xi_i)\, J_i^{\times}(\vec{r}_>; z, \xi_i) \tag{3.29}$$

which require the scattering solutions corresponding to the potential $V_i(\xi_i)$, and the

respective partially averaged densities:

$$n_i(\vec{r};\xi_i) = -\frac{1}{\pi}\mathrm{Im\,Tr}\int_{\curvearrowright}\mathrm{d}z\, f(z)\, G_i i(\vec{r},\vec{r};z,\xi_i)\ , \quad \langle\, n_i(\vec{r})\,\rangle = \sum_{\xi_i} P_i(\xi_i)\, n_i(\vec{r};\xi_i) \tag{3.30}$$

and the fully averaged density is given. This is enough to close the loop from the effective potentials to the density and back, in self–consistent calculations.

The off–diagonal elements in site indices raise the question of how the single–site approximation applies:

$$G_{ij}(\vec{r},\vec{r}';z,\xi_i,\xi_j) = Z_i(\vec{r};z,\xi_i)\,\langle\, \underline{\tau}_{ij}(z)\,\rangle_{\xi_i,\xi_j}\, Z_j^{\times}(\vec{r}';z,\xi_j) \tag{3.31}$$

and the answer consistent with the single–site approximation is to ignore correlations between ξ_i and ξ_j [46]. Then

$$\langle\, \underline{\tau}_{ij}\,\rangle_{\xi_i,\xi_j} = \widetilde{\underline{\tau}}_{ij} + \widetilde{\underline{\tau}}_{ii}\,\underline{X}_i(\xi_i)\,\widetilde{\underline{\tau}}_{ij} + \widetilde{\underline{\tau}}_{ij}\,\underline{X}_j(\xi_j)\,\widetilde{\underline{\tau}}_{jj} + \widetilde{\underline{\tau}}_{ii}\,\underline{X}_i(\xi_i)\,\widetilde{\underline{\tau}}_{ij}\,\underline{X}_j(\xi_j)\,\widetilde{\underline{\tau}}_{jj} \tag{3.32}$$

which obeys

$$\sum_{\xi_i,\xi_j} P_i(\xi_i)\, P_j(\xi_j)\,\langle\, \underline{\tau}_{ij}\,\rangle_{\xi_i,\xi_j} = \widetilde{\underline{\tau}}_{ij} \tag{3.33}$$

so that no conditional probabilities apply. It is evident that each excess scattering matrix can appear at most once in the expression for $\langle\, \underline{\tau}_{ij}\,\rangle_{\xi_i,\xi_j}$, if the off–diagonal elements of the SPO are to be regained by completing the average in the single–site approximation, as shown.

To derive the electronic grand–potential in a concise form an expression for the integrated DOS in terms of the GF is needed. This was already provided in Eq. 2.35, Lloyd's formula. In the present context this specialises to

$$\langle\, N(\varepsilon)\,\rangle = N_0(\varepsilon) + \frac{1}{\pi}\mathrm{Im\,Tr}\,\log\underline{\widetilde{\tau}}(\varepsilon) + \frac{1}{\pi}\mathrm{Im\,Tr}\sum_{i,\xi_i} P_i(\xi_i)\log\underline{D}_i(\varepsilon;\xi_i)\ . \tag{3.34}$$

The first logarithm is a matrix logarithm in both site and angular momentum labels, while the rest are just matrix logarithms in angular momentum labels. The site–off–diagonal excess scattering contributions are neglected, and one can show that this is consistent with the variational requirement

$$\frac{\delta\langle\, N(\varepsilon)\,\rangle}{\delta\widetilde{\underline{t}}_i^{-1}(\varepsilon)} = \underline{0} \implies \sum_{\xi_i} P_i(\xi_i)\,\underline{X}_i(\xi_i) = \underline{0} \tag{3.35}$$

after some lengthy linear algebra, and is an alternative way of deriving the CPA.

The grand–potential then follows from Eqs. 3.17 and 3.18:

$$\langle\,\Omega\,\rangle = \langle\,\Omega_0[n]\,\rangle + \langle\,\Delta\Omega\,\rangle = -\int\!\mathrm{d}\varepsilon\, f(\varepsilon)\,\langle\,N(\varepsilon)\,\rangle + \langle\,\Delta\Omega\,\rangle \tag{3.36}$$

and the correction term $\langle\,\Delta\Omega\,\rangle$ results from the averaged double–counting terms:

$$\begin{aligned}\langle\,\Delta\Omega\,\rangle = &-\frac{1}{2}\sum_{i,\xi_i} P_i(\xi_i)\int\!\mathrm{d}\vec{r}\int\!\mathrm{d}\vec{r}\,'\, n_i(\vec{r},\xi_i)\,\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r}\,'|}\,n_i(\vec{r}\,',\xi_i)\\ &-\frac{1}{2}\sum_{i,j\neq i}\int\!\mathrm{d}\vec{r}\int\!\mathrm{d}\vec{r}\,'\,\langle\,n_i(\vec{r})\,\rangle\,\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r}\,'+\vec{R}_{ij}|}\,\langle\,n_j(\vec{r}\,')\,\rangle\\ &+\sum_{i,\xi_i} P_i(\xi_i)\left(\Omega_{\mathrm{xc}}[n_i(\xi_i)] - \int\!\mathrm{d}\vec{r}\, n_i(\vec{r},\xi_i)\, V_{\mathrm{xc}}(\vec{r},\xi_i)\right)\quad .\end{aligned} \tag{3.37}$$

A detailed derivation is given in Ref. [66] for magnetic systems, together with a discussion of all the assumptions involved in arriving at the final result. It is also shown that the averaged grand–potential is stationary with respect to changes in the densities,

$$\frac{\delta\langle\,\Omega\,\rangle}{\delta n_i(\vec{r},\xi_i)} = 0 \tag{3.38}$$

which is analogous to the stationary condition from convential DFT (see Eq. 2.5).

The single–site CPA has been applied to both model and first–principles calculations, the main applications being in the theory of random alloys [46] and magnetic systems [66]. Its dynamical extension was shown to be equivalent to Dynamical Mean Field Theory [67], which is used when strong electronic correlations are present. The main drawback is the single–site approximation, which was recently lifted, and will be briefly discussed next.

3.2.4 The non–local coherent potential approximation

The single–site approximation is the weak point of the CPA. This was realised quite early, and attempts have been made at generalisations since at least three decades ago. None were found to be completely satisfactory, as the CPA should satisfy some key requirements, which are obeyed by the single–site approximation:

- It should produce an analytical GF;
- The effective medium should preserve translational invariance, if present.

These points were addressed in recent years by the Non–Local CPA (NLCPA) scheme [68]. The main hurdle in lifting the single–site approximation is the ex-

ponential increase in the number of configurations which need to be considered to construct the effective medium. For an N–site extension of the CPA and for a binary alloy (two types of atoms), 2^N configurations exist (the vast majority of which might be equivalent or irrelevant to the statistical average). Techniques like importance sampling or Monte Carlo might ameliorate this point.

Chapter 4

Magnetism

In this chapter the basic concepts in magnetism are introduced, and a survey of the most common methods employed in first–principles calculations is given. The first section concerns the description and calculation of magnetic properties using the conventional ground state DFT. Then the effects of temperature are discussed.

4.1 Magnetism at zero temperature

4.1.1 Itinerant versus localised description

It is useful to begin by recalling the magnetic properties of isolated atoms, as dictated by Hund's rules. These are, in the Russell–Saunders scheme:

- First rule: the state with maximal value of the total spin angular momentum is the one with lowest energy. If each electron has spin angular momentum $\hat{S}_i$, then the total spin angular momentum is $\hat{S} = \sum_i \hat{S}_i$;
- Second rule: the state with maximal value of the orbital angular momentum which is compatible with the first rule has lowest energy. If each electron has orbital angular momentum $\hat{L}_i$, then in total we have $\hat{L} = \sum_i \hat{L}_i$;
- Third rule: the energy is minimised by the state which is compatible with the first and second rules, and has the value of the total angular momentum $\hat{J} = \hat{L} + \hat{S}$ which minimises the spin–orbit interaction, $E_{\mathrm{SO}} \approx \lambda \hat{L} \cdot \hat{S}$.

and their application is straighforward to atoms, where one deals with discrete energy levels. When several atoms are brought together to form a solid, however, the discrete nature of the energy spectrum is lost, as states belonging to different atoms hybridise, and electrons can hop from one atom to the next. If some states

are mostly localised around each atom, one could think of them as atomic–like, and try to apply Hund's rules. If on the other hand they are delocalised, there are no grounds to justify their usage.

As a mock example, consider bulk Fe: it takes on the bcc crystal structure at zero temperature, and the valence states are formed from the $3d^6 4s^2$ electrons. The $4s$ wavefunctions extend well beyond half the nearest–neighbour separation, and one could expect the corresponding electrons to be delocalised. The $3d$ wavefunctions, however, have shorter range, and so are candidates to a careful application of Hund's rules. The second rule does not apply, as the cubic environment around each atom will quench the orbital angular momentum. The third rule is of no concern here, and applying the first rule would suggest a total spin moment of $M_S = 3d^5_\uparrow - 3d^1_\downarrow = 4\mu_\mathrm{B}$. The experimental value is $M_S \approx 2.2\mu_\mathrm{B}$, which besides being half the naive estimate, it is also non–integer. The first goal of an electronic theory of magnetism is to explain the origin of the experimental value, and the nature of the magnetic moment.

The main idea is the balance between lowering the kinetic energy by delocalising the electrons across the solid, and lowering the exchange energy by spin polarisation. This argument was formulated mathematically by Stoner [69, 70], and can be roughly outlined as follows: let $n(\varepsilon)$ denote the DOS of the unpolarised electron system, which is the equivalent for a system of itinerant electrons of the discrete electronic spectrum of an atom. Splitting the up and down spin states by a small amount Δ leads to a small spin polarisation, given by

$$M = N_\uparrow - N_\downarrow = \int_{-\infty}^{\varepsilon_\mathrm{F}} \mathrm{d}\varepsilon \left[n(\varepsilon - \Delta/2) - n(\varepsilon + \Delta/2) \right] \approx n(\varepsilon_\mathrm{F})\, \Delta \ . \tag{4.1}$$

The total energy increases by the change in the kinetic energy, and decreases by the change in exchange energy, modelled by the Stoner parameter I,

$$E(M) - E(0) \approx \frac{1}{2} \left[n(\varepsilon_\mathrm{F}) \Delta^2 - I M^2 \right] = \frac{1 - I n(\varepsilon_\mathrm{F})}{2 n(\varepsilon_\mathrm{F})} M^2 = \frac{M^2}{2\chi} \tag{4.2}$$

and the enhanced spin susceptibility χ was identified. The unpolarised state becomes unstable when $I n(\varepsilon_\mathrm{F}) > 1$, which is the Stoner criterion for the appearance of a spontaneous magnetisation in an itinerant electron system. These ingredients are included in the DFT total energy functional, in a parameter–free way, and so allow first–principles calculations of materials–specific properties.

In insulators electrons are localised, and integer spin moments are often found, so the localised description applies naturally. The question then remains of whether it is meaningful to identify local magnetic moments in metals. Many experiments

on metallic systems can be interpreted by making use of the Ising and Heisenberg models, which conceptually require integer spin moments, while an interpretation in terms of the magnetism of delocalised electrons is much more cumbersome or even inconsistent with experimental data. Local magnetic moments are thus an emergent property of the electronic structure, set up by the correlated electronic motions. This picture is captured in reasonable detail by DFT at zero temperature, but a proper description of metallic magnetism at finite temperatures must incorporate these localised DOFs in the theory explicitly, as will be shown.

4.1.2 The rigid spin approximation and the magnetic force theorem

In Multiple Scattering Theory space is partitioned into cells surrounding atoms, and the Kohn–Sham potentials are partitioned in the same way; for simplicity the following discussing is given in terms of the ASA, the same approximation used for the calculations. In the ASA the potentials are assumed spherical, $V_i(\vec{r}) = V_i(r)$ and $\vec{B}_i(\vec{r}) = B_i(r)\,\hat{e}_i$. The orientation of the local magnetic field is given by the unit vector $\hat{e}_i$, and is constant throughout the atomic sphere. The single–site t–matrix is obtained by solving the associated Dirac equation with the magnetic field pointing along the z direction, and then rotated to the desired orientation:

$$\underline{t}_i(\hat{e}_i) = \underline{\mathcal{R}}(\hat{e}_i)\,\underline{t}_i(\hat{e}_z)\,\underline{\mathcal{R}}^{\dagger}(\hat{e}_i) \quad . \tag{4.3}$$

The magnetisation on cell i is given from Eq. 2.69 as

$$\vec{m}_i(\vec{r}) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\beta\vec{\Sigma}\, Z_i(\vec{r};\varepsilon,\hat{e}_i)\,\underline{\tau}_{ii}(\varepsilon,\{\hat{e}\})\,Z_i^{\times}(\vec{r};\varepsilon,\hat{e}_i) + \ldots \tag{4.4}$$

This expression shows that the magnetisation on cell i is not necessarily collinear with the orientation of the local magnetic potential $\hat{e}_i$, as the SPO depends on the orientations of all sites.

Assume now that the system is locked into some magnetic configuration specified by a set $\{\hat{e}\}$, and SCF potentials are computed. Focusing on site i, and keeping $B_i(r)$ fixed, vary the orientation of the local magnetic field. The total spin moment is

$$\vec{M}_i(\hat{e}_i) = \int \mathrm{d}\vec{r}\,\vec{m}_i(\vec{r}) = M_{\parallel}(\hat{e}_i)\,\hat{e}_i + \vec{M}_{\perp}(\hat{e}_i)\ , \quad \vec{M}_{\perp}(\hat{e}_i)\cdot\hat{e}_i = 0 \tag{4.5}$$

which is decomposed into a parallel and a perpendicular component to $\hat{e}_i$. If one had introduced a constraining field, as in Eq. 2.26, one could force $M_{\perp}(\hat{e}_i) = 0$; it is more instructive not to do so.

This procedure suggests a computational test for the presence of a local moment

on site i: $M_\perp(\hat{e}_i) \ll M_\parallel(\hat{e}_i)$ and $M_\parallel \approx$ constant. The physical interpretation of these conditions is that the magnetic moment at site i is mainly set up by the intra–atomic exchange interaction, in analogy with Hund's first rule, and is little affected by the magnetic configuration of the system. When this is so, one can speak of a good local moment, and justify the Rigid Spin Approximation (RSA), which is encoded in Eq. 4.3. Good local moments arise from d or f–electrons, which are fairly well localised; the delocalised sp–electrons are also spin–polarised, but their degree of polarisation will depend strongly on the magnetic configuration, as they are spread out over many sites. This intuitive picture helps explaining the dependence of the magnitude of the magnetic moment on the magnetic state. Normally non–magnetic atoms can also be polarised by their magnetic neighbours.

These ideas can be cast on a more formal footing by invoking the magnetic force theorem [25]. The orientations of the local magnetic fields can be taken as external parameters, and so the first order change in the total energy is just the change in the band energy,

$$E[\{\hat{e}+\delta\hat{e}\}] - E[\{\hat{e}\}] = \sum_i \frac{\delta E}{\delta \hat{e}_i} \cdot \delta\hat{e}_i \approx \sum_{i,n} \frac{\delta \omega_n}{\delta \hat{e}_i} \cdot \delta \hat{e}_i \tag{4.6}$$

starting from a reference configuration, using Eq. 2.18 and keeping the potentials fixed. This is a delicate approximation, as discussed by Bruno [25]; the main observation is that the contribution to the single–particle energy coming from letting the potentials relax self–consistently cancels out the contribution from the double–counting terms. Energy differences between different magnetic configurations can thus be estimated in a non–self–consistent way, and under some assumptions the magnon spectrum and the ordering temperature can also be computed [25].

4.1.3 Spin spirals or frozen magnons

An alternative and complementary technique to the magnetic force theorem is given by the generalised Bloch theorem [52], which allows calculations of special magnetic states known as spin spirals, without the use of large supercells. Spin spirals or frozen magnons are magnetic structures described by the following expression:

$$\vec{M}_i(\vec{q}) = M(\vec{q}) \left(\cos(\vec{q}\cdot\vec{R}_i + \phi)\sin\theta\,\hat{n}_1 + \sin(\vec{q}\cdot\vec{R}_i+\phi)\sin\theta\,\hat{n}_2 + \cos\theta\,\hat{n}_3 \right) \tag{4.7}$$

for vectors $\vec{R}_i$ in a Bravais lattices, and one magnetic atom per unit cell (the generalisation to more than one atom is simple). Here the unit vectors $\hat{n}_1$, $\hat{n}_2$ and $\hat{n}_3$ are the basis of a right–handed orthogonal coordinate system. The polar angle

θ is the cone angle, and ϕ is the azimuthal angle. The magnetic moments are of the same length $M(\vec{q})$ on all sites, and their orientation precesses with wavevector $\vec{q}$ about $\hat{n}_3$. Simple magnetic states, such as ferromagnetic and different kinds of antiferromagnetic states, can also be described by this expression (see Fig. 4.1).

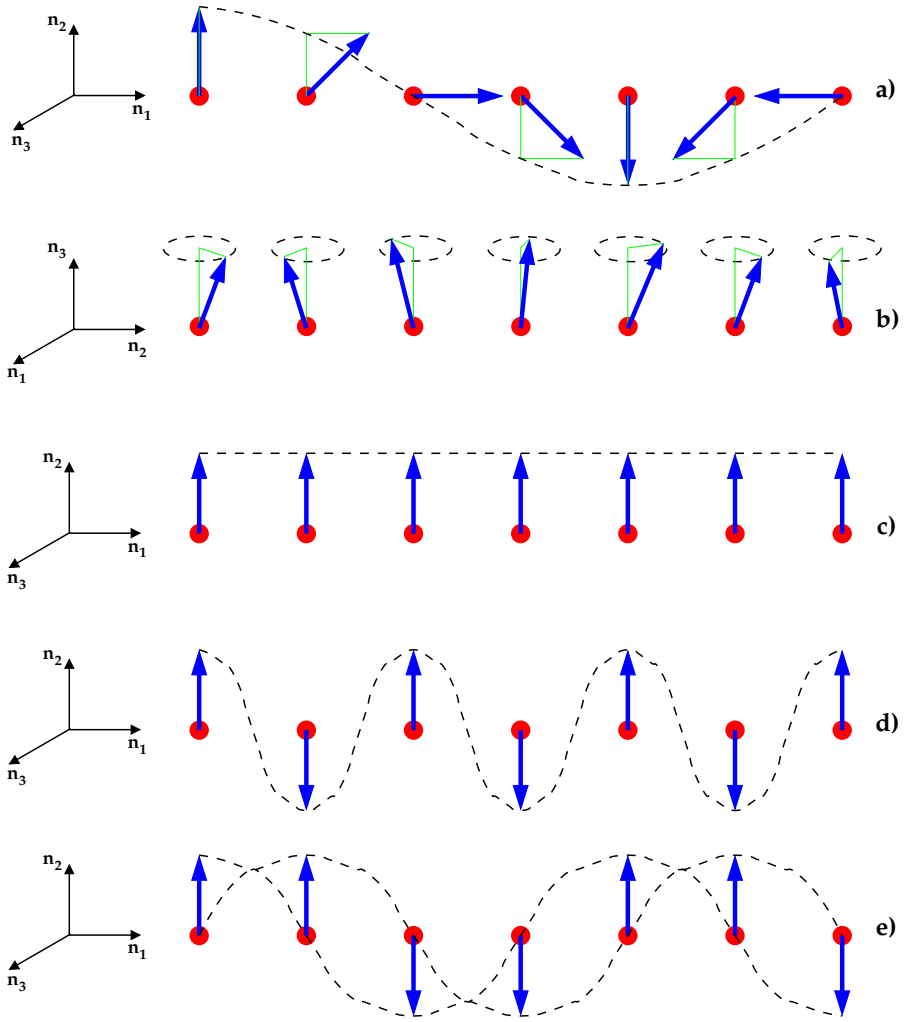

Figure 4.1: Some examples of spin spirals, from Eq. 4.7, on a linear chain of lattice constant a: a) — Flat spiral ($\theta = \pi/2$). b) — Conical spiral ($\theta \neq \pi/2$). c) — Ferromagnetic state ($q = 0$). d) — Antiferromagnetic state ($q = \pi/a$). e) — Up–up–down–down state. This has to be described by splitting the chain into two subchains of lattice constant $2a$, and then imposing the same spin spiral on each subchain.

The following derivation, appropriate for the MST–KKR formalism, follows Ref. [71]. If the SOC is left out of the Kohn–Sham Hamiltonian, the magnetic properties become independent of the overall orientation of the moments with respect to the

underlying lattice, and the azimuthal angle becomes redundant in a system with one atom per unit cell. One can thus choose the standard Cartesian coordinate axes for the unit vectors defining the orientation of the spin spiral.

The single–site t–matrices now transform as (RSA)

$$\underline{t}_i(\vec{q}) = \underline{\mathcal{R}}_i(\vec{q})\,\underline{t}_0(\hat{e}_z)\,\underline{\mathcal{R}}_i^{\dagger}(\vec{q}) \;, \quad \underline{\mathcal{R}}_i(\vec{q}) = \begin{pmatrix} e^{\mathrm{i}\vec{q}\cdot\vec{R}_i/2} & 0 \\ 0 & e^{-\mathrm{i}\vec{q}\cdot\vec{R}_i/2} \end{pmatrix} \underline{\mathcal{R}}(\theta,\phi) \tag{4.8}$$

where the operator $\underline{\mathcal{R}}_i(\vec{q})$ acts only on the spin components, in two stages: first the moment on site 0 is rotated away from the z direction by the rotation matrix $\underline{\mathcal{R}}(\theta,\phi)$; then it is rotated about the z axis by an angle $\vec{q}\cdot\vec{R}_i$ (the factor of $1/2$ comes from the spin matrix), and we obtain the single–site t–matrix at site i. The last rotation is equivalent to a translation in each spin component, as each exponential is in the form of the translation operator: each spin component will obey Bloch's theorem, but with different wavevectors [52].

The SPO follows immediately:

$$\underline{\mathcal{R}}_i(\vec{q})^{\dagger}\,\underline{\tau}_{ij}(\vec{q})\,\underline{\mathcal{R}}_j(\vec{q}) = \underline{t}_0\left[\delta_{ij} + \sum_{k\neq i}\left(\underline{\mathcal{R}}_i(\vec{q})^{\dagger}\,\underline{G}_{0,ik}\,\underline{\mathcal{R}}_k(\vec{q})\right)\left(\underline{\mathcal{R}}_k(\vec{q})^{\dagger}\,\underline{\tau}_{kj}\,\underline{\mathcal{R}}_j(\vec{q})\right)\right] \tag{4.9}$$

and can be obtained from a Fourier transform, as before (Eq. 2.64):

$$\underline{\mathcal{R}}_i(\vec{q})^{\dagger}\,\underline{\tau}_{ij}(\vec{q})\,\underline{\mathcal{R}}_j(\vec{q}) = \frac{1}{V_{\mathrm{BZ}}}\int \mathrm{d}\vec{k}\, e^{-\mathrm{i}\vec{k}\cdot\vec{R}_{ij}}\left[\underline{t}_0^{-1} - \underline{G}_0(\vec{k};\vec{q})\right]^{-1} \;. \tag{4.10}$$

All the electronic structure properties can then be computed as previously explained. The inversion is of a matrix in angular momentum but not site labels; this also shows that the calculation is no more demanding than a conventional ferromagnetic one.

As was already remarked before, orienting the t–matrices or the local magnetic fields along some prescribed directions gives no guarantee that the total spin moment will point in the same direction. This means that in general spin spiral calculations require constraining fields. For the special case of $\theta = \pi/2$, so–called flat spirals, it can be shown that no constraining fields are needed. The intuitive explanation is that in this case the net magnetic field on any site coming from all other sites is parallel or antiparallel to the assumed orientation of the moment on that site, and so no transverse component to the spin moment should be generated.

4.1.4 Model–based approaches using total energy calculations

The approaches described previously are completely first–principles, and provide useful information on their own. Physical insight can however be gained by trying to understand the results in terms of an appropriate spin model. For quantitative comparisons, the model should be able to reproduce the total energies and other magnetic properties to a reasonable degree. One can either choose a minimal model, or fit all available data in a systematic way. Then there is also the matter of whether to choose a quantum (spin operators) or a classical (unit vectors) spin model. The following discussion is made in terms of classical spin models; similar terms apply for quantum spin models, but may depend on the size of the spin (for instance certain terms are zero for spin–1/2 models, due to the properties of the Pauli matrices).

A subtle point which comes into question when modelling itinerant systems is the nature of the spin associated with each atoms. The only quantum spin operator which can be readily associated with a system of itinerant electrons is spin–1/2; if and how these combine to form a larger total spin operator is unclear, and as already discussed the spin moments are not necessarily integer. This explains the compromise in using classical unit vectors: the magnetic moment enters these models as a classical vector, of given length and well–defined orientation.

The ingredients usually found in classical spin models are:

- Pair interactions: $\sum_{ij} J_{ij}^{(n)} (\hat{e}_i \cdot \hat{e}_j)^n$; $n = 1$ gives the usual bilinear Heisenberg interactions, and $n = 2$ corresponds to biquadratic interactions;

- Multi–spin interactions; the most common is the four–spin term, which according to the restrictions to the site indices can represent a two, three or four–site interaction: $\sum_{ijkl} Q_{ijkl}[(\hat{e}_i \cdot \hat{e}_j)(\hat{e}_k \cdot \hat{e}_l) + (\hat{e}_i \cdot \hat{e}_k)(\hat{e}_j \cdot \hat{e}_l) - (\hat{e}_i \cdot \hat{e}_l)(\hat{e}_j \cdot \hat{e}_k)]$;

- Cost of changing the size of the magnetic moments from their equilibrium values: $(M - M_0)^2/2\chi_{\mathrm{L}}$; this is a form similar to the result obtained from Stoner theory, but the equilibrium value of the moment, M_0, can be non–zero. The denominator can be interpreted as a longitudinal susceptibility; if this is small then the moment will always be $M \approx M_0$, a good local moment;

- Anisotropic terms: these are discussed in detail below.

All these parameters will inevitably depend on the magnetic configuration, and so realistic models range from very simple (simple physical interpretation, bad fit) to very complex (the converse). This is the intrinsic problem in itinerant electron magnetism: environmental effects are very important for the electronic properties,

and in turn affect the magnetic properties. Examples of this approach are given in Refs. [72–77].

The systematic fitting procedure is now described. The energy of a magnetic configuration is expanded in contributions from all sites ($i \neq j \neq k \neq \ldots$):

$$E(\{\hat{e}\}) = E^{(0)} + \sum_i E_i^{(1)}(\hat{e}_i) + \sum_{i,j} E_{ij}^{(2)}(\hat{e}_i, \hat{e}_j) + \sum_{i,j,k} E_{ijk}^{(3)}(\hat{e}_i, \hat{e}_j, \hat{e}_k) + \ldots \quad (4.11)$$

and then a complete orthonormal set of functions of the orientations is introduced. The simplest choice is to use spherical harmonics (say real, with a suitable normalisation), and the corresponding expansion is

$$E_i^{(1)}(\hat{e}_i) = \sum_L E_{iL}^{(1)}\, Y_L(\hat{e}_i)\ , \quad E_{ij}^{(2)}(\hat{e}_i, \hat{e}_j) = \sum_{L,L'} E_{iL,jL'}^{(2)}\, Y_L(\hat{e}_i)\, Y_{L'}(\hat{e}_j)\ , \quad \ldots \quad (4.12)$$

If there is no magnetic anisotropy (no SOC taken into consideration), the energy is invariant under global rotations of the moments. This sets the single–site term to zero, and restricts the angular momentum sums in the other terms. This kind of expansion is known as the Spin Cluster Expansion [59, 60], which is the generalisation to magnetic systems of the Cluster Expansion used in alloy theory [57, 58]. Products of spherical harmonics can be transformed into familiar forms for the interaction energy. A few examples will be given later.

4.1.5 Thin films, surfaces and interfaces

In ordered bulk materials, there is a small number of inequivalent atoms, those in the unit cell. There is three–dimensional translational invariance, and often each atomic position possesses high symmetry. All these factors enormously reduce the number of parameters necessary to describe bulk properties.

When a surface or interface is present, such simplicity is lost. If the defect is taken as a perturbation on the bulk properties, the perturbed region typically extends over many bulk unit cells. There may also be qualitatively different properties: materials which are non–magnetic in the bulk might become magnetic at surfaces, due to the reduced coordination and the resulting loss of hybridisation, so that the Stoner criterion might be satisfied. Even for bulk magnetic materials, the changes in the electronic structure at surfaces or interfaces might also change the nature of the magnetic interactions, stabilising a magnetic state different from the bulk one.

Heterogeneous materials can also be grown by modern experimental techniques, and stacks of different magnetic materials in atomic contact through an interface,

or interacting via a non–magnetic separator, are of fundamental technological importance, besides remaining challenging for theoretical descriptions.

An extreme case is that of single atomic monolayers of a magnetic species grown on top of non–magnetic substrates. Magnetism in less than three dimensions not always leads to magnetic ordering, as stated by the Mermin–Wagner theorem [78]. The hybridisation with the substrate may also play a crucial role in determining the magnetic properties and the magnetic structure of the monolayer. As this is the simplest magnetic thin film one can consider, it will be analysed in detail in Chapter 5, to demonstrate the new theory being presented in this thesis, and to quantitatively illustrate the previous comments.

4.1.6 Relativistic effects and anisotropy

Magnetic anisotropy derives from the relativistic properties of electrons, through the SOC. It is useful to decompose the total magnetic energy into an isotropic and an anisotropic contribution, defined as the remainder. In this way, the anisotropic part is seen to be from one to three orders of magnitude smaller than the isotropic part, which explains why under certain circumstances it can be neglected in the calculations. However it has a qualitative impact: it defines the orientation of the magnetic moments in real space, and can also lift degeneracies between magnetic states, for instance favouring chiral states.

The two main forms of anisotropy considered are:

- On–site anisotropy: this refers to the anisotropic one–site term in the magnetic energy. It obeys the local point group symmetry [79, 80]. The lowest order term is the uniaxial anisotropy, which is very important in layered systems, and vanishes in cubic systems:

$$\sum_i \sum_{\alpha\beta} K_{i,\alpha\beta}\, \hat{e}_{i\alpha}\, \hat{e}_{i\beta} = \sum_i \sum_{n=1}^{3} K_{i,n}\, (\hat{e}_i \cdot \hat{u}_{i,n})^2 \;, \quad \alpha, \beta = x, y, z \tag{4.13}$$

  where $K_{i,n}$ are the eigenvalues and $\hat{u}_{i,n}$ the associated eigenvectors of the 3×3 matrix in cartesian components, on each site. The eigenvectors determine the directions along which the energy is minimum (easy axes) or maximum (hard axes);

- Pair or exchange anisotropy: anisotropic pair interactions. Each pair interac-

tion matrix has a well–known decomposition:

$$\sum_{ij}\sum_{\alpha\beta} J_{ij,\alpha\beta}\,\hat{e}_{i\alpha}\,\hat{e}_{i\beta} \;\longrightarrow\; \underline{J}_{ij} = \underline{J}^{\mathrm{I}}_{ij} + \underline{J}^{\mathrm{S}}_{ij} + \underline{J}^{\mathrm{A}}_{ij} \;. \tag{4.14}$$

The isotropic part reduces to the usual Heisenberg pair interaction:

$$\underline{J}^{\mathrm{I}}_{ij} = \frac{1}{3}\,(\mathrm{Tr}\,\underline{J}_{ij})\,\mathcal{I} \;\longrightarrow\; \frac{1}{3}\,(\mathrm{Tr}\,\underline{J}_{ij})\,\hat{e}_i\cdot\hat{e}_j \tag{4.15}$$

and is usually the dominant part. The antisymmetric part is rewritten as a Dzyaloshinskii–Moriya term [9, 10]:

$$\underline{J}^{\mathrm{A}}_{ij} = \frac{1}{2}\,(\underline{J}_{ij} - \underline{J}^{T}_{ij}) \;\longrightarrow\; \vec{D}_{ij}\cdot(\hat{e}_i\times\hat{e}_j) \tag{4.16}$$

and T denotes the transpose of the matrix. This interaction lifts the chiral degeneracy; taking a spin spiral as an example, the cross–product will be sensitive to the clockwise or anticlockwise precession of the orientations about $\hat{n}_3$. This term is only present if allowed by symmetry, see Refs. [10, 81]. The symmetric part can be decomposed as the on–site anisotropy:

$$\underline{J}^{\mathrm{S}}_{ij} = \frac{1}{2}\,(\underline{J}_{ij} + \underline{J}^{T}_{ij}) - \underline{J}^{\mathrm{I}}_{ij} \;\longrightarrow\; \sum_{n=1}^{3} J^{\mathrm{S}}_{ij,n}\,(\hat{e}_i\cdot\hat{u}_{ij,n})(\hat{e}_j\cdot\hat{u}_{ij,n}) \;. \tag{4.17}$$

It is sometimes called pseudo–dipolar anisotropy, by comparison with the dipole–dipole interaction, the coarse–grained version of the magnetostatic Breit interaction, Eq. 2.17:

$$\sum_{i\neq j}\left[\frac{3}{R^3_{ij}}\,(\vec{M}_i\cdot\vec{M}_j) - \frac{1}{R^5_{ij}}\,(\vec{M}_i\cdot\vec{R}_{ij})(\vec{M}_j\cdot\vec{R}_{ij})\right] \;. \tag{4.18}$$

These concepts will be of use when interpreting the results of the calculations for the magnetic monolayers, in Chapter 5.

4.2 Magnetism at finite temperature

4.2.1 Collective versus particle–hole excitations

The thermal properties of a magnetic system are determined by the kind of excited states which can be thermally populated, i.e., those roughly within $k_{\mathrm{B}}T$ of the ground state. The two main types of excitations are:

- Particle–hole excitations: an electron is removed from the occupied majority spin states to the unoccupied minority spin states. This is a longitudinal spin fluctuation, as the size of the magnetic moment is reduced. This process fits naturally within the Stoner picture;

- Collective excitations, or magnons: the orientations of the magnetic moments deviate from their equilibrium directions, in a coherent way. This transverse spin fluctuation is usually pictured as a spin spiral with a small cone angle, which is why these are also known as frozen magnons. They arise naturally when working with models of localised spins.

Typically the dominant contribution will come from magnons, as their minimum excitation energy is either zero (Goldstone mode) or the magnetic anisotropy energy (which is usually small). Particle–hole excitations make a sizable contribution only when transitions between electronic states of the same symmetry are considered; as these are usually separated in energy by the exchange splitting, they should only become important at high temperatures. In general both kinds of excitations might be present and mutually affect each other.

Taking a ferromagnetic system as an example, the temperature–dependent magnetic susceptibility very often follows the Curie–Weiss (CW) law above the Curie temperature T_c,

$$\chi(T) \equiv \left.\frac{\mathrm{d}M}{\mathrm{d}H}\right|_{H=0} = \frac{C}{T - T_c}\,, \quad C = \frac{\mu_\mathrm{B}^2\, q_c(q_c+2)}{3k_\mathrm{B}} \tag{4.19}$$

where q_c is known as the number of magnetic carriers per atom, which is an effective quantity. These can be compared with the value of the magnetic moment at zero temperature, $M_S = \mu_\mathrm{B}\, q_s$, and the ratio q_c/q_s [82] indicates whether a magnetic material is a good local moment system, $q_c/q_s \approx 1$, or a weakly ferromagnetic system, $q_c/q_s \gg 1$ (i.e. small q_s). Some examples are given in Table 4.1. EuO is an insulator, and the local moment corresponds to a spin of $S = 7/2$. $ZrZn_2$ is a weak itinerant ferromagnet; it is clear no local moment can be identified. Fe and Ni are in–between; Fe seems to fit the local moment criterion, but it would correspond to a non–integer spin. For Ni there is no simple interpretation.

The CW law is the high–temperature classical limit of the fluctuation–dissipation theorem, which applies when the thermal energy is much larger than the characteristic energies of the collective modes [1]. It then comes as no surprise that the Stoner model does not reproduce this behaviour, as $k_\mathrm{B}T \ll \varepsilon_F$ in most systems, and so the puzzle remains as to why the CW law is so common; the answer comes from

	q_c	q_s	q_c/q_s	T_c (K)
EuO	7.00	7.00	1.00	77
Fe	2.33	2.22	1.05	1043
Ni	0.89	0.61	1.25	631
$ZrZn_2$	0.86	0.16	5.38	26

Table 4.1: Some examples of q_c and q_s, from Wohlfarth [82].

the nature of thermal fluctuations about the equilibrium value of the magnetisation, as is explained by Spin Fluctuation Theory, outlined below.

4.2.2 Spin fluctuation theory

Fluctuations of the thermodynamic quantities are very important when determining the thermal properties of a physical system, in particular magnetic properties. Here a short exposition of the most basic spin fluctuation theory is given, for illustration purposes. The most general version, that of Moriya, is explained in detail in his book, Ref. [7]. This restricted description follows the treatment in [3], and should apply to weakly ferromagnetic systems, such as $ZrZn_2$.

The finite temperature properties of the Stoner model can be computed by standard statistical mechanics, but for the purpose of this discussion only the main results will be needed [3]. The Landau expansion of the free energy of an isotropic ferromagnetic Stoner system is

$$\mathcal{F} = \frac{A}{2}M^2 + \frac{B}{4}M^4 - MH = \frac{1}{2\chi_0}\left(\frac{T^2}{T_c^2} - 1\right)M^2 + \frac{1}{4\chi_0 M_0^2}M^4 - MH \qquad (4.20)$$

where M_0 is the value of the magnetisation at $T = 0$, and χ_0 sets the scale for the susceptibility. The equilibrium value of the magnetisation comes from the conditions $\mathrm{d}\mathcal{F}/\mathrm{d}M = 0$ and $\mathrm{d}^2\mathcal{F}/\mathrm{d}M^2 > 0$, and with $H = 0$

$$T > T_c: \quad M = 0\,; \quad T < T_c: \quad \frac{M^2}{M_0^2} = 1 - \frac{T^2}{T_c^2} \qquad (4.21)$$

with the inverse susceptibility

$$\frac{\chi_0}{\chi} = \frac{3M^2}{M_0^2} + \frac{T^2}{T_c^2} - 1 \qquad (4.22)$$

that does not follow the CW law. This is due to the neglect of fluctuations in the order parameter. The Curie temperature entering these expressions is given from a

more detailed analysis as [3]

$$k_{\mathrm{B}}^2 T_c^2 = \varepsilon_F^2 \Big(In(\varepsilon_{\mathrm{F}}) - 1 \Big) \tag{4.23}$$

which shows that it is of the order of the Fermi degeneracy temperature, much too high to explain the experimental values.

The order parameter is given by $\langle \vec{M} \rangle = M \, \hat{e}_z$. Consider now the addition of a thermal fluctuation, such that $\langle \vec{m} \rangle = 0$. Replacing

$$M^2 = \langle \vec{M}^2 \rangle \;\longrightarrow\; \langle (\vec{M} + \vec{m})^2 \rangle = M^2 + \langle m_\parallel^2 \rangle + 2 \langle m_\perp^2 \rangle \tag{4.24}$$

and the fluctuation is decomposed into a component parallel to the order parameter and two transverse ones, taken to be equal by symmetry. The other replacement is

$$M^4 \;\longrightarrow\; M^4 + M^2 \Big(6 \langle m_\parallel^2 \rangle + 4 \langle m_\perp^2 \rangle \Big) + 3 \langle m_\parallel^2 \rangle^2 + 4 \langle m_\parallel^2 \rangle \langle m_\perp^2 \rangle + 8 \langle m_\perp^2 \rangle^2 \tag{4.25}$$

and the modified free energy is

$$\mathcal{F} = \left[\frac{A}{2} + \frac{B}{4} \Big(6 \langle m_\parallel^2 \rangle + 4 \langle m_\perp^2 \rangle \Big) \right] M^2 + \frac{B}{4} M^4 - MH + \dots \tag{4.26}$$

The magnetisation is reduced by the fluctuations

$$\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}M} = 0 \;\;\Longrightarrow\;\; M^2 = -\frac{A}{B} - 3 \langle m_\parallel^2 \rangle - 2 \langle m_\perp^2 \rangle \approx M_0^2 - 3 \langle m_\parallel^2 \rangle - 2 \langle m_\perp^2 \rangle \tag{4.27}$$

where $-\frac{A}{B} \approx M_0^2$ is valid if the new Curie temperature is much smaller than the previous Stoner value, i.e., if the fluctuations are strong enough. It is computed from the inverse susceptibility in the paramagnetic state, where $M = 0$ and $\langle m_\parallel^2 \rangle = \langle m_\perp^2 \rangle = \langle m^2 \rangle$,

$$\frac{\mathrm{d}^2\mathcal{F}}{\mathrm{d}M^2} = \frac{1}{\chi} = A + 5B \langle m^2 \rangle = 0 \;\;\Longrightarrow\;\; \langle m_c^2 \rangle = \frac{M_0^2}{5} \;, \;\; T = T_c \tag{4.28}$$

which establishes the value of the fluctuations at T_c, Moriya's formula.

The temperature dependence of the fluctuations is still unknown, as their origin was not specified, and they were taken to be uncorrelated with the bulk order parameter. Assuming a linear isotropic behaviour (which can be loosely justified through the fluctuation–dissipation theorem, Eq. 3.14)

$$\langle m_\parallel^2 \rangle = \langle m_\perp^2 \rangle = \langle m^2 \rangle \approx \frac{M_0^2}{5} \frac{T}{T_c} \;\;\Longrightarrow\;\; \frac{M^2}{M_0^2} = 1 - \frac{T}{T_c} \tag{4.29}$$

and the total magnetisation, including the fluctuations, becomes

$$\langle\,(\vec{M}+\vec{m})^2\,\rangle = M^2 + 3\langle\,m^2\,\rangle = \begin{cases} \left(1-\frac{2}{5}\frac{T}{T_c}\right)M_0^2 & , \quad T<T_c \\ \frac{3}{5}\frac{T}{T_c}M_0^2 & , \quad T>T_c \end{cases} \tag{4.30}$$

so that it decreases as the temperature goes up to T_c, and then increases again. This also restores the CW law for the susceptibility, as now the main temperature dependence comes from the amplitude of the thermal fluctuations.

To close, one should point out that the present exposition, based on Landau theory augmented with spin fluctuations, still does not reproduce many experimental facts, such as the $T^{3/2}$ decay of the magnetisation in bulk systems, which is obtained when the fluctuations are magnons, or the critical behaviour near T_c [1]. These are well–known limitations of this kind of approach. The important concept that this brief outline was meant to demonstrate is that only with an adequate treatment of the fluctuations in the order parameters will the expected physical behaviour emerge from the statistical mechanics. The weak points in this treatment were the assumed temperature dependence of the fluctuations, and their independence of the average magnetic moment; more sophisticated treatments based on model Hamiltonians can be considered [7], and connected to first–principles calculations, as explained before.

4.2.3 Disordered local magnetic moments

If the system of interest is a metallic magnet in which local moments are present, longitudinal and transverse excitations are not independent of each other. As previously discussed, the size of the local moment will depend to some extent on the overall magnetic state of the system; if the magnetisation is being reduced by transverse fluctuations (magnon creation) then the size of the moments will also change, which is a longitudinal fluctuation. The electronic structure will also change: it will evolve from the sharp DOS corresponding to the completely ordered state at zero temperature to the broadened DOS corresponding to the disordered paramagnetic state. Here a crucial insight is that the paramagnetic state corresponds to the absence of long–range order of the orientations of the local moments, but not to the vanishing of the local moment size on each atom; this is the key idea in the theory of the Disordered Local Moment (DLM) state, which will be developed in MST language in the next chapter.

4.2.4 Other ways of modelling temperature–dependent effects

So far two ways of incorporating thermal effects into the theory of magnetism were mentioned. Spin Fluctuation Theory as presented was model–based, but one could conceivably derive all the parameters from electronic structure calculations, as explained by Moriya [7]. The DLM approach will be introduced in the next chapter couched in the MST–KKR formalism. Other approaches which remain parameter–free are those which compute the dynamical magnetic susceptibility from DFT (by linear response or explicit time dependence, e.g. [83]), and spin dynamics incorporated into or derived from electronic structure codes [84–86] (the magnetic equivalent of Carr–Parrinello molecular dynamics).

On the other hand, if the zero temperature DFT calculations were mapped to a model, as described before, then the thermodynamic properties of this model can be computed, and should qualitatively and sometimes even quantitatively describe the true magnetic properties of the system at finite temperature. The most commonly encountered methods are Monte Carlo based, such as spin dynamics [86] and, on a larger scale, micromagnetic simulations [22].

Chapter 5

Disordered Local Moments and Linear Response Theory

In this chapter the statistical mechanics of local magnetic moments is combined with finite temperature DFT, thus including in the first–principles treatment the effects of the slowly evolving magnetic DOFs, and healing some of the failures of finite temperature DFT. In the first part, the best single–site mean–field approximation (in a variational sense) is derived, and extended to be of a general form in its angular dependence. This scheme is then combined with DFT.

Then a linear response theory of the magnetic susceptibility is presented, based on the new theory. This gives direct access to the spin–spin correlation function, $S^{(2)}$, and so to the magnetic interactions in the system. The limiting cases of complete magnetic order and complete magnetic disorder are analysed, and the mean–field estimate of the ordering temperature is given. To conclude, an extension is briefly outlined, the Onsager cavity field construction, which can be used to correct some failures of the mean–field theory.

5.1 Disordered Local Moments and DFT

5.1.1 Introduction: variational approach to the Heisenberg model

To illustrate the statistical mechanics, the variational method is applied to the classical Heisenberg model, with local external fields $\vec{H}_i$ to break the symmetry. Choosing as reference Hamiltonian that of non–interacting moments coupled to

effective local fields, $\vec{h}_i$, which will be variational parameters, we have:

$$\mathcal{H} = -\frac{1}{2}\sum_{i,j\neq i} J_{ij}\,\hat{e}_i\cdot\hat{e}_j - \sum_i \vec{H}_i\cdot\hat{e}_i\ ,\quad \mathcal{H}_0 = -\sum_i \vec{h}_i\cdot\hat{e}_i\ . \tag{5.1}$$

Due to the single–site nature of the approximation, all quantities factor:

$$\mathcal{Z}_0 = \prod_i \mathcal{Z}_{0,i} = \prod_i \int \mathrm{d}\hat{e}_i\, e^{\beta\vec{h}_i\cdot\hat{e}_i} = (4\pi)^N \prod_i \frac{\sinh\beta h_i}{\beta h_i}\ . \tag{5.2}$$

The local order parameters are defined as

$$\vec{m}_i = \int \mathrm{d}\hat{e}_i\,\frac{e^{\beta\vec{h}_i\cdot\hat{e}_i}}{\mathcal{Z}_{0,i}}\,\hat{e}_i = \int \mathrm{d}\hat{e}_i\, P_i(\hat{e}_i)\,\hat{e}_i = L(\beta h_i)\,\hat{h}_i \approx \frac{\beta}{3}\,\vec{h}_i\ \ (\beta h_i \ll 1) \tag{5.3}$$

with the Langevin function $L(x) = \coth x - 1/x$. The free energy gives

$$\mathcal{F}_0 = -\frac{1}{\beta}\log\mathcal{Z}_0 = -\frac{1}{\beta}\sum_i \log\int \mathrm{d}\hat{e}_i\, e^{\beta\vec{h}_i\cdot\hat{e}_i} \implies -\frac{\partial\mathcal{F}_0}{\partial\vec{h}_i} = \vec{m}_i + \sum_{j\neq i}\vec{m}_j\cdot\frac{\partial\vec{h}_j}{\partial\vec{h}_i} \tag{5.4}$$

and the second term on the last right hand side accounts for a possible non–linear dependence of the vector parameters.

Minimising the variational inequality, Eq. 3.11,

$$\mathcal{F} \leq \mathcal{F}_1 = \mathcal{F}_0 + \langle\,\mathcal{H}-\mathcal{H}_0\,\rangle\ ,\quad \frac{\partial\mathcal{F}_1}{\partial\vec{h}_i} = 0 \implies \vec{h}_i = \vec{H}_i + \sum_{j\neq i} J_{ij}\,\vec{m}_j \tag{5.5}$$

which shows that the effective field is just the Weiss mean–field, and so as expected we regain the conventional mean–field approximation to the Heisenberg model. This corresponds to a non–linear set of coupled equations, through the local order parameters $\vec{m}_i$. The variational free energy is

$$\mathcal{F}_1(\beta;\{\vec{H}\}) = -\frac{1}{\beta}\sum_i \log\int \mathrm{d}\hat{e}_i\, e^{\beta\vec{h}_i\cdot\hat{e}_i} + \frac{1}{2}\sum_{i,j\neq i} J_{ij}\,\vec{m}_i\cdot\vec{m}_j \implies -\frac{\partial\mathcal{F}_1}{\partial\vec{H}_i} = \vec{m}_i \tag{5.6}$$

and the second term is a double–counting correction, akin to those that arise in DFT: the sum of single–site contributions counts the pair interaction twice. The last equation shows that the approximate free energy yields the correct expression for the local order parameter.

The magnetic susceptibility is

$$\frac{\partial \vec{m}_i}{\partial \vec{H}_j} = \frac{\partial \vec{m}_i}{\partial \vec{h}_i} \left(\delta_{ij} + \sum_{k \neq i} J_{ik} \frac{\partial \vec{m}_k}{\partial \vec{H}_j} \right) \iff \chi_{ij} = \chi_{0,i}\, \delta_{ij} + \chi_{0,i} \sum_{k \neq i} J_{ik}\, \chi_{kj} \tag{5.7}$$

and can formally be obtained by inverting a matrix in site indices,

$$\chi = \chi_0 + \chi_0\, J\, \chi \iff \chi^{-1} = \chi_0^{-1} - J \ . \tag{5.8}$$

This leads to the mean–field expression for the transition temperature from the paramagnetic state ($\vec{m}_i = 0, \forall i$) into a magnetically ordered state,

$$\chi_{0,i} \approx \frac{\beta}{3} \ , \ \ \det \chi^{-1} = 0 \implies \prod_n \Big(3k_{\mathrm{B}}T - \lambda_n \Big) = 0 \implies T_c = \frac{\lambda_{\max}}{3k_{\mathrm{B}}} \tag{5.9}$$

where λ_n are the eigenvalues of the matrix J, and $\lambda_{\max}$ is the largest. The associated eigenvector contains the information about the orientations of the moments on each site. These equations are usually Fourier transformed to exploit the translational invariance of the paramagnetic state, but that is not necessary for the general considerations given here.

The expression for T_c is independent of the dimensionality of the lattice, and so mean–field theory always predicts a magnetic phase transition for the classical Heisenberg model, which is incorrect if the lattice is 1D or 2D, as shown by the Mermin–Wagner theorem [78]. Furthermore, this expression for the magnetic susceptibility also violates the fluctuation–dissipation theorem, Eq. 3.14:

$$\chi_{ij} = \frac{\partial \vec{m}_i}{\partial \vec{H}_j} \neq \beta \Big(\langle\, \hat{e}_i\, \hat{e}_j \,\rangle - \langle\, \hat{e}_i \,\rangle \langle\, \hat{e}_j \,\rangle \Big) = \beta \Big(\langle\, \hat{e}_i^2 \,\rangle - \langle\, \hat{e}_i \,\rangle^2 \Big)\, \delta_{ij} \tag{5.10}$$

as the mean–field theory is only a theory for the order parameter, and not for the fluctuations. These statements should be kept in mind for later applications.

5.1.2 The best single–site approximation to a spin Hamiltonian

As the energetics of the spin configurations will be computed from DFT, their functional form is unknown, and so it is necessary to develop the theory for an arbitrary spin Hamiltonian. Let $\mathcal{H}$ represent the unknown spin Hamiltonian, and write the single–site approximation in terms of unspecified variational functions

$\mathcal{H}_0 = \sum_i g_i(\hat{e}_i)$. Carrying out a functional minimisation of Eq. 3.11 yields [66]

$$\frac{\delta \mathcal{F}_1}{\delta g_i(\hat{e}_i)} = 0 \implies \langle\, \mathcal{H} \,\rangle_{\hat{e}_i} - \langle\, \mathcal{H} \,\rangle = g_i(\hat{e}_i) - \langle\, g_i \,\rangle \;, \quad P_i(\hat{e}_i) = \frac{e^{-\beta g_i(\hat{e}_i)}}{\int \mathrm{d}\hat{e}_i \, e^{-\beta g_i(\hat{e}_i)}} \;. \tag{5.11}$$

$$\langle\, A \,\rangle_{\hat{e}_i} = \prod_{j \neq i} \int \mathrm{d}\hat{e}_j \, P_j(\hat{e}_j) \, A(\{\hat{e}\}) \;, \quad \langle\, A \,\rangle = \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \langle\, A \,\rangle_{\hat{e}_i} \tag{5.12}$$

This shows that the crucial quantity is then the partial average of the unknown Hamiltonian, $\langle\, \mathcal{H} \,\rangle_{\hat{e}_i}$: this means to average over all sites except site i, the full average being regained by performing the last average over the values of the DOF on site i. The main goal of the theory will be to compute this quantity from DFT.

Writing the Hamiltonian as a sum of contributions coming from different sites, $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 + \mathcal{H}_3 + \ldots$ $(i \neq j \neq k \neq \ldots)$,

$$\mathcal{H}_1 = \sum_i f_i^{(1)}(\hat{e}_i) \;, \quad \mathcal{H}_2 = \sum_{ij} f_{ij}^{(2)}(\hat{e}_i, \hat{e}_j) \;, \quad \mathcal{H}_3 = \sum_{ijk} f_{ijk}^{(3)}(\hat{e}_i, \hat{e}_j, \hat{e}_k) \;, \quad \ldots \tag{5.13}$$

we obtain the variational functions as a generalisation of the Weiss fields (care must be taken to collect all contributions, as shown for the two–site term):

$$\begin{aligned} g_i(\hat{e}_i) &= f_i^{(1)}(\hat{e}_i) + \sum_j \int \mathrm{d}\hat{e}_j \, P_j(\hat{e}_j) \left(f_{ij}^{(2)}(\hat{e}_i, \hat{e}_j) + f_{ji}^{(2)}(\hat{e}_j, \hat{e}_i) \right) + \cdots \\ &= -\vec{H}_i \cdot \hat{e}_i - \sum_j J_{ij} \, \hat{e}_i \cdot \vec{m}_j \quad \text{(Heisenberg model)} \end{aligned} \tag{5.14}$$

illustrated by comparing with the result obtained using the Heisenberg model, Eq. 5.1. The general variational free energy, with multiple–counting corrections, is

$$\begin{aligned} \mathcal{F}_1 &= -\frac{1}{\beta} \sum_i \log \int \mathrm{d}\hat{e}_i \, e^{-\beta g_i(\hat{e}_i)} - \sum_{n=2}^{n_{\max}} (n-1) \langle\, \mathcal{H}_n \,\rangle \\ &= \langle\, \mathcal{H} \,\rangle + k_{\mathrm{B}} T \sum_i \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \log P_i(\hat{e}_i) = \langle\, \mathcal{H} \,\rangle - T S_0 \end{aligned} \tag{5.15}$$

where terms up to some $n_{\max}$ number of coupled sites are considered. The second line shows that the approximate free energy is the sum of the internal energy, computed from the full Hamiltonian with the variational single–site probability functions, and the corresponding entropy.

To make these equations more transparent, and to connect with the results of the previous section, a complete set of angular functions must be introduced. For conve-

nience, let us expand the angular functions in renormalised *real* spherical harmonics, $\mathcal{Y}_L(\hat{e})$, with $L = (\ell, m)$, as detailed in Appendix A.

A general angular function is expanded as

$$f(\hat{e}) = f_0 + \sum_{L>0} f_L \, \mathcal{Y}_L(\hat{e}) \; , \quad f_0 = \frac{1}{4\pi} \int \mathrm{d}\hat{e} \, f(\hat{e}) \; , \quad f_L = \frac{2\ell+1}{4\pi} \int \mathrm{d}\hat{e} \, \mathcal{Y}_L(\hat{e}) \, f(\hat{e}) \tag{5.16}$$

The Hamiltonians are written as ($i \neq j \neq k \neq \cdots$, all L's > 0)

$$\begin{aligned} \mathcal{H} &= c + \sum_{iL} f_{iL}^{(1)} \, \mathcal{Y}_L(\hat{e}_i) + \sum_{iL,jL'} f_{iL,jL'}^{(2)} \, \mathcal{Y}_L(\hat{e}_i) \, \mathcal{Y}_{L'}(\hat{e}_j) + \ldots \\ \mathcal{H}_0 &= c_0 + \sum_{iL} g_{iL} \, \mathcal{Y}_L(\hat{e}_i) \; . \end{aligned} \tag{5.17}$$

The constants c and c_0 are not very relevant at present, but will be needed later.

The other advantage is that the functional minimisation becomes a conventional differentiation with respect to the parameters g_{iL}. Using

$$\frac{\partial P_j(\hat{e}_j)}{\partial g_{iL}} = -\beta P_j(\hat{e}_j) \sum_{L'} \frac{\partial g_{jL'}}{\partial g_{iL}} \left(\mathcal{Y}_{L'}(\hat{e}_j) - \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \mathcal{Y}_{L'}(\hat{e}_j) \right) \tag{5.18}$$

and the previous results we arrive at $c_0 = c$ and

$$\begin{aligned} g_{iL}(\{m\}) = f_{iL}^{(1)} &+ \sum_{jL'} \left(f_{iL,jL'}^{(2)} + f_{jL',iL}^{(2)} \right) m_{jL'} \\ &+ \sum_{jL',kL_1} \left(f_{iL,jL',kL_1}^{(3)} + f_{jL',iL,kL_1}^{(3)} + f_{jL',kL_1,iL}^{(3)} \right) m_{jL'} \, m_{kL_1} + \ldots \end{aligned} \tag{5.19}$$

Now the expectation values of the spherical harmonics replace the usual vector order parameters:

$$m_{iL} = \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \mathcal{Y}_L(\hat{e}_i) \; , \quad -1 \leq m_{iL} \leq 1 \tag{5.20}$$

which shows the convenience of the present choice of normalisation, and for definiteness the free energy is

$$\mathcal{F}_1 = -\frac{1}{\beta} \sum_i \log \int \mathrm{d}\hat{e}_i \, e^{-\beta \sum_L g_{iL} \mathcal{Y}_L(\hat{e}_i)} - \sum_{iL,jL'} f_{iL,jL'}^{(2)} \, m_{iL} \, m_{jL'} - \ldots \tag{5.21}$$

The external fields are defined as conjugate to the local DOFs. As these are now represented by the spherical harmonics, the coupling to the external fields is extended: $-\vec{H}_i \cdot \hat{e}_i \longrightarrow -\sum_L H_{iL} \, \mathcal{Y}_L(\hat{e}_i)$, which leads to a new magnetic susceptibility,

$$-\frac{\partial \mathcal{F}_1}{\partial H_{iL}} = m_{iL} \ , \quad \chi_{ij} = \frac{\partial m_{iL}}{\partial H_{jL'}} = \chi_{0,i} \left[\delta_{ij} + \sum_k S^{(2)}_{ik} \, \chi_{kj} \right] \tag{5.22}$$

with the on–site susceptibility given by (Eq. 5.18)

$$\chi_{0,iLL'} = \beta \left[\int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \mathcal{Y}_L(\hat{e}_i) \, \mathcal{Y}_{L'}(\hat{e}_i) - m_{iL} \, m_{iL'} \right] \xrightarrow[P_i(\hat{e}_i)=\frac{1}{4\pi}]{} \frac{\beta}{2\ell+1} \, \delta_{LL'} \tag{5.23}$$

and the generalised spin–spin correlation function, or $S^{(2)}$,

$$\begin{aligned} S^{(2)}_{iL,jL'} &= f^{(2)}_{iL,jL'} + f^{(2)}_{jL',iL} + \sum_{kL_1} \Big(f^{(3)}_{iL,jL',kL_1} + f^{(3)}_{jL',iL,kL_1} \\ &\quad + f^{(3)}_{iL,kL_1,jL'} + f^{(3)}_{jL',kL_1,iL} + f^{(3)}_{kL_1,iL,jL'} + f^{(3)}_{kL_1,jL',iL} \Big) \, m_{kL_1} + \ldots \\ &= f^{(2)}_{iL,jL'} + f^{(2)}_{jL',iL} + \sum_{kL_1} S^{(3)}_{iL,jL',kL_1} \, m_{kL_1} \end{aligned} \tag{5.24}$$

which shows how multi–spin interactions should contribute to the susceptibility, and generates a chain of correlation functions. This quantity plays the same role as the exchange parameters J_{ij} in the classical Heisenberg model, and can be seem to be completely symmetric in its labels. The on–site terms $f^{(1)}_{iL}$ in the spin Hamiltonian are handled by the on-site susceptibility.

It is also helpful to find the connection between different couplings used in spin models and the new formulation in terms of angular functions. Consider a model Hamiltonian, with $i \neq j \neq k \neq l$,

$$\begin{aligned} \mathcal{H} &= \sum_i K_i \, (\hat{e}_i \cdot \hat{n}_i)^2 + \sum_{i\alpha,j\beta} J_{i\alpha,j\beta} \, \hat{e}_{i\alpha} \, \hat{e}_{j\beta} + \sum_{i,j} B_{ij} \, (\hat{e}_i \cdot \hat{e}_j)^2 \\ &\quad + \sum_{i,j,k,l} Q_{ijkl} \left[(\hat{e}_i \cdot \hat{e}_j)(\hat{e}_k \cdot \hat{e}_l) - (\hat{e}_i \cdot \hat{e}_k)(\hat{e}_j \cdot \hat{e}_l) + (\hat{e}_i \cdot \hat{e}_l)(\hat{e}_j \cdot \hat{e}_k) \right] \ . \end{aligned} \tag{5.25}$$

The terms included are, in order: an on–site uniaxial magnetic anisotropy; anisotropic bilinear exchange interactions; isotropic biquadratic interactions; and the symmetrised four–spin coupling.

Making use of the sum rule that turns spherical harmonics into Legendre polynomials,

$$\sum_m \mathcal{Y}^m_\ell(\hat{e}_i) \, \mathcal{Y}^m_\ell(\hat{e}_j) = P_\ell(\hat{e}_i \cdot \hat{e}_j) \ , \quad P_0(x) = 1 \ , \quad P_1(x) = x \ , \quad P_2(x) = \frac{3x^2 - 1}{2} \tag{5.26}$$

this Hamiltonian maps onto the following interaction functions (omitting a con-

stant):

$$f^{(1)}_{i,2m} = \frac{2K_i}{3}\,\mathcal{Y}^m_2(\hat{n}_i)\ ,\quad f^{(2)}_{i1m,j1m'} = J_{im,jm'}\ ,\quad f^{(2)}_{i2m,j2m'} = \frac{2B_{ij}}{3}\,\delta_{mm'}$$
$$f^{(4)}_{i1m,j1m',k1m_1,l1m_2} = Q_{ijkl}\left[\delta_{mm'}\delta_{m_1m_2} - \delta_{mm_1}\delta_{m'm_2} + \delta_{mm_2}\delta_{m'm_1}\right] \tag{5.27}$$

and in principle these contributions can be extracted from the mean–field approximation functions $g_i(\hat{e}_i)$ or from the $S^{(2)}$.

The usefulness of this approach should be now clear. Although the statistical mechanics dictated by the single–site approximation might be qualitatively wrong, it provides an interpretative tool for connecting the electronic structure and magnetism. The coefficients of the variational functions and their dependence on the order parameters supply a mean of identifying the dominant spin couplings in the system, and the mean–field behaviour can also be a useful interpolation for systems of intermediate dimensionality (such as magnetic films a few atoms thick, which lie between 2D and 3D). It now remains to derive an approximation to compute the partial averages from DFT, which is explained in the next subsection.

5.1.3 Merging DFT and variational statistical mechanics

The previous statistical mechanical formalism relies on the identification of good local moments associated with some sites in the lattice. These are assumed to behave like classical unit vectors evolving under some unknown spin Hamiltonian dictated by the electronic structure, as explained in Chapter 3. There it was also pointed out that other atoms might become magnetised, without developing a local moment. These cannot be incorporated directly in the statistical mechanics, unless the formulation is extended to handle induced moments. Such an extension will not be considered here.

Taking as the slow DOF the orientation $\hat{e}_i$ of the local moment at site i, one can invoke the CPA to handle the construction of the required partial average, $\langle\,\mathcal{H}\,\rangle_{\hat{e}_i}$, as explained in Chapter 2. From the general expression Eq. 3.36, we have for the averaged grand potential $\langle\,\Omega\,\rangle = \langle\,\Omega_0\,\rangle + \langle\,\Delta\Omega\,\rangle$:

$$\langle\,\Omega_0\,\rangle = -\int\mathrm{d}\varepsilon\, f(\varepsilon)\, N_0(\varepsilon) - \frac{1}{\pi}\mathrm{Im\,Tr}\int\mathrm{d}\varepsilon\, f(\varepsilon)\,\log\underline{\underline{\widetilde{\tau}}}(\varepsilon)$$
$$-\frac{1}{\pi}\mathrm{Im\,Tr}\sum_i\int\mathrm{d}\varepsilon\, f(\varepsilon)\int\mathrm{d}\hat{e}_i\, P_i(\hat{e}_i)\log\underline{\underline{D}}_i(\varepsilon;\hat{e}_i) \tag{5.28}$$

(the double underline on the second term signifies a matrix in site and angular

momentum labels), with the impurity and excess scattering matrices,

$$\underline{D}_i(\varepsilon;\hat{e}_i) = \underline{1} + \underline{X}_i(\varepsilon;\hat{e}_i)\,\underline{\widetilde{\tau}}_{ii}(\varepsilon) \tag{5.29}$$

$$\underline{X}_i(\varepsilon;\hat{e}_i) = \left[\left(\underline{\tilde{t}}_i^{-1}(\varepsilon) - \underline{t}_i^{-1}(\varepsilon;\hat{e}_i)\right)^{-1} - \underline{\widetilde{\tau}}_{ii}(\varepsilon)\right]^{-1} \tag{5.30}$$

and the CPA condition

$$\int \mathrm{d}\hat{e}_i\, P_i(\hat{e}_i)\,\underline{X}_i(\varepsilon;\hat{e}_i) = \underline{0}\ . \tag{5.31}$$

The effective medium is now magnetic, and is required to reproduce the averaged magnetic properties of the system. The double–counting terms $\langle\,\Delta\Omega\,\rangle$ are given by the magnetic version of Eq. 3.37.

The averaged grand potential $\langle\,\Omega\,\rangle$ plays the same role as the averaged spin Hamiltonian $\langle\,\mathcal{H}\,\rangle$ from the previous section. The variational free energy then becomes

$$\mathcal{F}_1 = \langle\,\Omega\,\rangle + \frac{1}{\beta}\sum_i \int \mathrm{d}\hat{e}_i\, P_i(\hat{e}_i)\log P_i(\hat{e}_i) = \langle\,\Omega\,\rangle - TS_0 \tag{5.32}$$

and the variational functions $g_i(\hat{e}_i)$ that define the single–site probabilities $P_i(\hat{e}_i)$ (Eq. 5.11) must be specified:

$$\frac{\delta\mathcal{F}_1}{\delta g_i(\hat{e}_i)} = 0 \implies \langle\,\Omega\,\rangle_{\hat{e}_i} - \langle\,\Omega\,\rangle = g_i(\hat{e}_i) - \langle\,g_i\,\rangle \tag{5.33}$$

and observing that the averaged grand potential is stationary with respect to changes in the particle (and magnetisation) densities, Eq. 3.38, while the single–particle contribution to the grand potential, $\langle\,\Omega_0\,\rangle$, is stationary with respect to changes in the effective medium, Eq. 3.35. These steps have to be taken more carefully because one is not dealing with a fixed spin Hamiltonian: the effective medium will adjust to changes in all quantities, as the itinerant electrons will relax, and so these variations have to be considered in detail (see Ref. [66]).

The variational functions are then given by the following expression:

$$g_i(\hat{e}_i) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\log \underline{D}_i(\varepsilon;\hat{e}_i) + \Delta g_i(\hat{e}_i) \tag{5.34}$$

which can be conveniently expanded in spherical harmonics (see Eq. 5.16)

$$g_i(\hat{e}_i) = \sum_L g_{iL}\,\mathcal{Y}_L(\hat{e}_i)\ ,\quad g_{iL} = \frac{2\ell+1}{4\pi}\int \mathrm{d}\hat{e}_i\,\mathcal{Y}_L(\hat{e}_i)\,g_i(\hat{e}_i) = \frac{\partial\langle\,\Omega\,\rangle}{\partial m_{iL}}\ \ (L>0) \tag{5.35}$$

and the $L=0$ term can be disregarded, as it will not contribute to define the prob-

abilities $P_i(\hat{e}_i)$ that specify the effective medium and the approximate free energy. The double–counting correction is given by

$$\begin{aligned}\Delta g_i(\hat{e}_i) = &-\frac{1}{2}\int \mathrm{d}\vec{r}\int \mathrm{d}\vec{r}'\, n_i(\vec{r},\hat{e}_i)\,\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r}'|}\,n_i(\vec{r}',\hat{e}_i)\\ &-\sum_{j\neq i}\int \mathrm{d}\vec{r}\int \mathrm{d}\vec{r}'\, n_i(\vec{r},\hat{e}_i)\,\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r}'+\vec{R}_{ij}|}\,\langle\, n_j(\vec{r}')\,\rangle\\ &+\Omega_{\mathrm{xc}}[n_i(\hat{e}_i);\vec{m}_i(\hat{e}_i)] - \int \mathrm{d}\vec{r}\, n_i(\vec{r},\hat{e}_i)\, V_{\mathrm{xc}}(\vec{r},\hat{e}_i) - \int \mathrm{d}\vec{r}\,\vec{m}_i(\vec{r},\hat{e}_i)\cdot\vec{B}_{\mathrm{xc}}(\vec{r},\hat{e}_i)\end{aligned} \tag{5.36}$$

This term is usually neglected; this can sometimes be justified, but a critical assessment is always necessary [66]. These double–counting terms vanish if the particle and magnetisation densities are independent of the orientations of the local moments; changes in the local electronic structure driven by hybridisation with the rest of the system as the orientation of the moment varies ensures that this will never happen, and the importance of these effects should be reflected in the importance of these correction terms.

The formal prescription connecting DFT and the statistical mechanics of the local moments through the CPA is now complete, and describes the Disordered Local Moment (DLM) theory [26, 32, 33, 66, 87–90]. The picture is that of magnetic impurities described by the single–site t–matrices $\underline{t}_i(\hat{e}_i)$ embedded in an effective magnetic medium, set up by the $\underline{\tilde{t}}_i$'s. The CPA provides the partially averaged particle and magnetisation densities, $n_i(\vec{r},\hat{e}_i)$ and $\vec{m}_i(\vec{r},\hat{e}_i)$, and the corresponding potentials are obtained from the Poisson equation. The practical implementation of this scheme, however, must overcome a few obstacles, which are now analysed.

5.1.4 Practical aspects of the Disordered Local Moment theory

The main problem in the implementation of the theory is to define consistently the local moment and its orientation. In principle a constraining field must be included, to ensure that the partially averaged electronic structure produces a magnetic moment along the desired direction $\hat{e}_i$, see Eq. 2.26. As explained in Chapter 3, the magnetic moment will have the following form,

$$\vec{M}_i(\hat{e}_i) = \int \mathrm{d}\vec{r}\,\vec{m}_i(\vec{r},\hat{e}_i) = M_{\parallel}(\hat{e}_i)\,\hat{e}_i + \vec{M}_{\perp}(\hat{e}_i)\;,\quad \vec{M}_{\perp}(\hat{e}_i)\cdot\hat{e}_i = 0 \tag{5.37}$$

and the local moment condition being $M_{\perp} \ll M_{\parallel}$. The effect of the constraining field is to suppress $M_{\perp}$; in principle a different constraining field should be computed

for each required orientation of the local moment.

On the other hand, the following approximations will be used: the potentials are assumed to be spherically symmetric, the LDA is the chosen approximation to the exchange–correlation functional, and the magnetisation direction is assumed to be uniform inside each sphere (either MT or ASA constructions). The LDA ensures that the magnetic potential is collinear with the magnetisation direction, and the component of the magnetisation direction which is transverse to the local moment orientation is thus identically compensated point by point by the constraining field [85]. The following expressions should then apply:

$$\begin{aligned} n_i(\vec{r},\hat{e}_i) &= -\frac{1}{\pi}\mathrm{Im\,Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\langle\, G_{ii}(\vec{r},\vec{r};\varepsilon)\,\rangle_{\hat{e}_i} \\ \vec{m}_i(\vec{r},\hat{e}_i) &= -\frac{1}{\pi}\mathrm{Im\,Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\hat{e}_i\,(\beta\vec{\Sigma}\cdot\hat{e}_i)\,\langle\, G_{ii}(\vec{r},\vec{r};\varepsilon)\,\rangle_{\hat{e}_i} \end{aligned} \tag{5.38}$$

and the associated potentials $V_i(r,\hat{e}_i)$ and $B_i(r,\hat{e}_i)$ are defined as usual through the densities (the β in the last expression is one of the 4×4 Dirac matrices in relativistic theory, see Eq. 2.12).

The possible dependence of the length of the local moment on the orientation is handled by writing all quantities in a global cartesian frame,

$$\begin{aligned} \hat{e}_i = \sum_m \mathcal{Y}_1^m(\hat{e}_i)\,\hat{e}_m &= \mathcal{Y}_1^1(\hat{e}_i)\,\hat{e}_x + \mathcal{Y}_1^{-1}(\hat{e}_i)\,\hat{e}_y + \mathcal{Y}_1^0(\hat{e}_i)\,\hat{e}_z \\ &= \cos\phi_i \sin\theta_i\,\hat{e}_x + \sin\phi_i\sin\theta_i\,\hat{e}_y + \cos\theta_i\,\hat{e}_z \end{aligned} \tag{5.39}$$

and the local moment at site i becomes

$$\vec{M}_i(\hat{e}_i) = \sum_\alpha \sum_L M_{iL}^\alpha\, \mathcal{Y}_L(\hat{e}_i)\,\hat{e}_\alpha \quad (m=1,2,3 \;\longleftrightarrow\; \alpha = x,y,z) \tag{5.40}$$

by resumming the spherical harmonics expansion. The averaged magnetic moment vector then follows simply

$$\langle\, \vec{M}_i \,\rangle = \sum_\alpha \sum_L M_{iL}^\alpha\, \langle\, \mathcal{Y}_L(\hat{e}_i)\,\rangle\, \hat{e}_\alpha = \sum_\alpha \sum_L M_{iL}^\alpha\, m_{iL}\,\hat{e}_\alpha \tag{5.41}$$

using the generalised order parameters from Eq. 5.20. As an example, consider

$$\vec{M}_i(\hat{e}_i) = \big(M_0 + M_1\cos\theta_i\big)\,\hat{e}_i \implies \langle\, \vec{M}_i \,\rangle = \big(M_0\langle\,\cos\theta_i\,\rangle + M_1\langle\,\cos^2\theta_i\,\rangle\big)\,\hat{e}_z \tag{5.42}$$

assuming that the order parameter develops along the z–direction. Non–rigid be-

haviour of the magnetic moment will then naturally introduce the generalised order parameters.

Another practical question concerns the number of orientations $\hat{e}_i$ needed to compute the angular integrals. A systematic approach to the angular integrations is given by the Lebedev–Laikov meshes [91], which ensure numerical integration of spherical harmonics to machine precision up to a desired angular momentum cutoff. The quantities which have the strongest angular dependence are the probability distributions $P_i(\hat{e}_i)$, which are sharply peaked when the local order parameter is sizable. The densities, the potentials and the single–site functions $g_i(\hat{e}_i)$ can be expanded in spherical harmonics up to a fairly small cutoff, say $\ell_{\max} = 2$ or 3, which can be different for different quantities, and chosen to ensure that the neglected coefficients in the expansion are small. If orientations which are not part of the mesh are desired, an interpolation scheme is immediately provided by the truncated sum in spherical harmonics.

To conclude this section, the limiting case of the alloy analogy is reviewed. The magnetic system is pictured as a binary alloy of magnetic atoms pointing parallel ($\uparrow$) or antiparallel ($\downarrow$) to the local order parameter, which is equivalent to an Ising model [66]. This kind of problem is straighforward to solve in any KKR code which has a basic CPA implementation, and self–consistent potentials are readily available. A non–self–consistent evaluation of the several quantities in the theory so far would then be possible, using the four potentials obtained from the alloy calculation:

$$V_i(r, \hat{e}_i) = \frac{V_i(r,\uparrow) + V_i(r,\downarrow)}{2} + \frac{V_i(r,\uparrow) - V_i(r,\downarrow)}{2} (\hat{e}_i \cdot \hat{n}_i) \tag{5.43}$$

$$B_i(r, \hat{e}_i) = \frac{B_i(r,\uparrow) + B_i(r,\downarrow)}{2} + \frac{B_i(r,\uparrow) - B_i(r,\downarrow)}{2} (\hat{e}_i \cdot \hat{n}_i) \tag{5.44}$$

This goes one step beyond the RSA explained in Chapter 3, and incorporates the longitudinal effects of the magnetic effective medium.

To set up the magnetic alloy, the 'concentrations' of magnetic atoms parallel and antiparallel to the local order parameter on each site must be specified. A way of doing this is by resorting to the Langevin model in Eq. 5.3. Choosing the z–axis to be parallel to the local order parameter, the probability of finding the magnetic atom parallel or antiparallel to the order parameter is given by

$$P_\uparrow = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\frac{\pi}{2}} \mathrm{d}\theta \, P(\theta, \phi) = \frac{e^{\beta h} - 1}{e^{\beta h} - e^{-\beta h}} \quad , \quad P_\downarrow = 1 - P_\uparrow \quad , \quad \beta h = L^{-1}(m) \tag{5.45}$$

and the ratio of the Weiss field to temperature, βh, is given by inverting the Langevin function for the corresponding value of the order parameter, m. To illustrate the

m	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$L^{-1}(m)$	0.000	0.302	0.615	0.953	1.336	1.797	2.401	3.304	4.998
$P_\uparrow$	0.50	0.57	0.65	0.72	0.79	0.86	0.92	0.97	0.99
$P_\downarrow$	0.50	0.43	0.35	0.28	0.21	0.14	0.08	0.03	0.01

Table 5.1: The correspondence between the Langevin order parameter and the concentrations in the magnetic alloy analogy. See text for details.

correspondence, a few selected values are given in Table 5.1. When the order parameter is larger than, say, 0.8, the magnetic alloy becomes very dilute, and it makes more sense to think of 'impurities' (antiparallel magnetic atoms) in a 'host' (parallel magnetic atoms).

5.2 Linear response and the magnetic susceptibility

5.2.1 Linear response to a perturbation in a reference state

In linear response theory, the quantity of interest is the susceptibility which measures the proportionality between a change in an observable and the change in the external field that couples to it: the magnetic susceptibility is the change in the magnetisation $\vec{m}(\vec{r})$ of the system brought about by a change in the external applied magnetic field, $\vec{H}(\vec{r}')$. From Eq. 2.69 and the properties of Green functions it follows that the magnetisation response is coupled in general to the charge density response, through the changes in the effective potential:

$$\frac{\delta n(\vec{r})}{\delta \vec{H}(\vec{r}')} \equiv \chi^n(\vec{r},\vec{r}') = \int \mathrm{d}\vec{r}_1 \left(\chi_0^{nn}(\vec{r},\vec{r}_1) \frac{\delta V_{\mathrm{eff}}(\vec{r}_1)}{\delta \vec{H}(\vec{r}')} + \chi_0^{nm}(\vec{r},\vec{r}_1) \frac{\delta \vec{B}_{\mathrm{eff}}(\vec{r}_1)}{\delta \vec{H}(\vec{r}')} \right) \tag{5.46}$$

$$\frac{\delta \vec{m}(\vec{r})}{\delta \vec{H}(\vec{r}')} \equiv \chi^m(\vec{r},\vec{r}') = \int \mathrm{d}\vec{r}_1 \left(\chi_0^{mn}(\vec{r},\vec{r}_1) \frac{\delta V_{\mathrm{eff}}(\vec{r}_1)}{\delta \vec{H}(\vec{r}')} + \chi_0^{mm}(\vec{r},\vec{r}_1) \frac{\delta \vec{B}_{\mathrm{eff}}(\vec{r}_1)}{\delta \vec{H}(\vec{r}')} \right) \tag{5.47}$$

with the Kohn–Sham susceptibilities (the cyclic property of the trace was used)

$$\chi_0^{nn}(\vec{r},\vec{r}_1) = \frac{1}{\pi} \mathrm{Im}\,\mathrm{Tr} \int \mathrm{d}\varepsilon\, f(\varepsilon)\, G(\vec{r}_1,\vec{r};\varepsilon)\, G(\vec{r},\vec{r}_1;\varepsilon) \tag{5.48}$$

$$\chi_0^{mm}(\vec{r},\vec{r}_1) = \frac{1}{\pi} \mathrm{Im}\,\mathrm{Tr} \int \mathrm{d}\varepsilon\, f(\varepsilon)\, \beta\vec{\Sigma}\, G(\vec{r}_1,\vec{r};\varepsilon)\, \beta\vec{\Sigma}\, G(\vec{r},\vec{r}_1;\varepsilon) \tag{5.49}$$

$$\chi_0^{nm}(\vec{r},\vec{r}_1) = \frac{1}{\pi} \mathrm{Im}\,\mathrm{Tr} \int \mathrm{d}\varepsilon\, f(\varepsilon)\, G(\vec{r}_1,\vec{r};\varepsilon)\, \beta\vec{\Sigma}\, G(\vec{r},\vec{r}_1;\varepsilon) \tag{5.50}$$

$$\chi_0^{mn}(\vec{r},\vec{r}_1) = \frac{1}{\pi} \mathrm{Im}\,\mathrm{Tr} \int \mathrm{d}\varepsilon\, f(\varepsilon)\, \beta\vec{\Sigma}\, G(\vec{r}_1,\vec{r};\varepsilon)\, G(\vec{r},\vec{r}_1;\varepsilon) \tag{5.51}$$

and a possible change in the Fermi level should be considered [92]. This leads to two coupled integral equations for the susceptibilities, where the coordinate dependence is omitted for simplicity, and the appropriate integrations implied:

$$\chi^n = \chi_0^{nn}\,\mathcal{K}^{nn}\,\chi^n + \chi_0^{nn}\,\mathcal{K}^{nm}\,\chi^m + \chi_0^{nm}\,\mathcal{K}^{mn}\,\chi^n + \chi_0^{nm}\,\mathcal{K}^{mm}\,\chi^m \tag{5.52}$$

$$\chi^m = \chi_0^{mm} + \chi_0^{mm}\,\mathcal{K}^{mm}\,\chi^m + \chi_0^{mm}\,\mathcal{K}^{mn}\,\chi^n + \chi_0^{mn}\,\mathcal{K}^{nm}\,\chi^m + \chi_0^{mn}\,\mathcal{K}^{nn}\,\chi^n \quad . \tag{5.53}$$

The magnetic susceptibility χ^m begins with the bare magnetic susceptibility χ_0^{mm}, but the charge susceptibility χ^n is driven exclusively by feedback effects. The four kernels are given by

$$\begin{aligned} \mathcal{K}^{nn}(\vec{r},\vec{r}') &= \frac{\delta V_{\text{eff}}(\vec{r})}{\delta n(\vec{r}')} \;, \quad \mathcal{K}^{nm}(\vec{r},\vec{r}') = \frac{\delta V_{\text{eff}}(\vec{r})}{\delta \vec{m}(\vec{r}')} \\ \mathcal{K}^{mn}(\vec{r},\vec{r}') &= \frac{\delta \vec{B}_{\text{eff}}(\vec{r})}{\delta n(\vec{r}')} \;, \quad \mathcal{K}^{mm}(\vec{r},\vec{r}') = \frac{\delta \vec{B}_{\text{eff}}(\vec{r})}{\delta \vec{m}(\vec{r}')} \quad . \end{aligned} \tag{5.54}$$

These expressions hint at the level of complexity required in a full linear response calculation of the non–uniform static susceptibility; if the dynamical response is considered the problem is computationally even harder.

The simplifications usually considered are in the spirit of the adiabatic approximation, the same argument that was used to identify the local moments as slowly evolving DOFs. In the LDA and with the external field transverse to the magnetisation, its effect will be to induce a deviation of the orientation of the magnetisation away from its equilibrium direction. As the effective magnetic field points in the same direction as the magnetisation, only one term is necessary to describe this transverse response (the magnitude of the magnetisation is kept fixed)

$$\mathcal{K}^{mm}(\vec{r},\vec{r}') = \frac{\delta \vec{B}_{\text{eff}}(\vec{r})}{\delta \vec{m}(\vec{r}')} = \frac{B_{\text{xc}}(\vec{r})}{m(\vec{r})}\,\delta(\vec{r}-\vec{r}') \;, \quad \frac{\delta}{\delta \vec{m}} = \frac{1}{m}\frac{\delta}{\delta \hat{m}} \quad \text{(transverse only)} \tag{5.55}$$

and a more amenable integral equation must be solved,

$$\chi^m(\vec{r},\vec{r}') = \chi_0^{mm}(\vec{r},\vec{r}') + \int \mathrm{d}\vec{r}_1\, \chi_0^{mm}(\vec{r},\vec{r}_1)\,\frac{B_{\text{xc}}(\vec{r}_1)}{m(\vec{r}_1)}\,\chi^m(\vec{r}_1,\vec{r}') \tag{5.56}$$

which has the form of a bare Kohn–Sham susceptibility with an enhancement factor due to exchange and correlation effects. This is the typical ansatz for calculations of the transverse magnetic susceptibility, which circumvents the need to include the response from the charge density. It is expected to be a good approximation for local moment systems, as longitudinal fluctuations of the magnetisation are energetically

much more costly than the transverse ones. The theory to be developed next has a different starting point, and will combine the same ingredients in an alternative way, as the slow local moment DOFs were already extracted in the DLM picture.

5.2.2 The spin–spin correlation function and the magnetic susceptibility

The conventional Zeeman coupling of the local moment to the external field is given by, after Eq. 5.40,

$$\vec{H}_i \cdot \vec{M}_i(\hat{e}_i) = \sum_{\alpha L} H_{i\alpha}\, M^{\alpha}_{iL}\, \mathcal{Y}_L(\hat{e}_i) = \sum_{L} H_{iL}\, \mathcal{Y}_L(\hat{e}_i) \tag{5.57}$$

so the following correspondence applies:

$$\frac{\partial}{\partial H_{i\alpha}} = \sum_{L} M^{\alpha}_{iL}\, \frac{\partial}{\partial H_{iL}} \quad . \tag{5.58}$$

The response of the average local moment on site i, $\langle\, \vec{M}_i \,\rangle$, to a change in the external magnetic field on site j, $\vec{H}_j$, follows from Eq. 5.41:

$$\begin{aligned} \frac{\partial \langle\, \vec{M}_i \,\rangle}{\partial \vec{H}_j} &= \sum_{\alpha\beta} \sum_{LL'} \left(\frac{\partial M^{\alpha}_{iL}}{\partial H_{L'}}\, m_{iL}\, M^{\beta}_{jL'} + M^{\alpha}_{iL}\, M^{\beta}_{jL'}\, \frac{\partial m_{iL}}{\partial H_{jL'}} \right) \hat{e}_\alpha \otimes \hat{e}_\beta \\ &= \sum_{\alpha\beta} \left(\chi^{\mathrm{L}}_{i\alpha,j\beta} + \chi^{\mathrm{T}}_{i\alpha,j\beta} \right) \hat{e}_\alpha \otimes \hat{e}_\beta \end{aligned} \tag{5.59}$$

where the susceptibility separates into a longitudinal and a transverse response. As explained before, the longitudinal response couples to the charge density response, and describes high energy contributions, while the transverse part is associated with deviations in the orientations of the magnetic moments, which are the low–energy excitations in local moment systems. In the following only the transverse part of the magnetic susceptibility will be considered, and the double–counting corrections will be ignored. Under these approximations, the variational free energy is of the form (note a sign change)

$$\mathcal{F}_1 = \widetilde{\Omega} - \frac{1}{\beta} \sum_{iL} \log \int \mathrm{d}\hat{e}_i\, e^{\beta (g_{iL}(\{m\}) + H_{iL})\, \mathcal{Y}_L(\hat{e}_i)} \quad (L \geq 0) \tag{5.60}$$

with the grand–canonical potential for the reference system

$$\widetilde{\Omega} = -\int \mathrm{d}\varepsilon\, f(\varepsilon)\, N_0(\varepsilon) - \frac{1}{\pi} \mathrm{Im}\, \mathrm{Tr} \int \mathrm{d}\varepsilon\, f(\varepsilon)\, \log \underline{\underline{\widetilde{\tau}}}(\varepsilon) \tag{5.61}$$

and the single–site coefficients (sign change), which depend on the magnetic state through the effective medium,

$$g_{iL}(\{m\}) = \frac{2\ell+1}{4\pi}\int \mathrm{d}\hat{e}_i\, \mathcal{Y}_L(\hat{e}_i)\, \frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\log \underline{D}_i(\varepsilon;\hat{e}_i) \quad . \tag{5.62}$$

It is then straightforward to show that the free energy is stationary to changes in the order parameters m_{iL}, and that these are generated by the corresponding conjugate external fields H_{iL}:

$$\frac{\partial \mathcal{F}_1}{\partial m_{iL}} = 0\ , \quad -\frac{\partial \mathcal{F}_1}{\partial H_{iL}} = m_{iL} \tag{5.63}$$

using a sum rule for the spherical harmonics and the stationarity of the free energy to changes in the effective medium.

To compute the transverse magnetic susceptibility the following auxiliary susceptibility has to be evaluated (see Eqs. 5.22 and 5.23):

$$\frac{\partial m_{iL}}{\partial H_{jL'}} \equiv \chi_{iL,jL'} = \sum_{L_1} \chi_{0,iLL_1}\left[\delta_{ij}\delta_{L_1L'} + \sum_{kL_2} S^{(2)}_{iL_1,kL_2}\,\chi_{kL_2,jL'}\right] \tag{5.64}$$

with the spin–spin correlation function, $S^{(2)}$, now computed from the electronic structure,

$$S^{(2)}_{iL,jL'} \equiv \frac{\partial g_{iL}}{\partial m_{jL'}} = -\frac{\partial^2 \langle\, \Omega\, \rangle}{\partial m_{iL}\partial m_{jL'}} \quad . \tag{5.65}$$

In the spirit of the magnetic force theorem, Eq. 4.6, which is compatible with the current approximations, the free energy difference between magnetic states characterised by two different sets of order parameters is expressed as

$$\Delta\mathcal{F}_1 \approx \frac{1}{2}\sum_{iL,jL'} \delta m_{iL}\,\chi^{-1}_{iL,jL'}\,\delta m_{jL'} = \frac{1}{2}\sum_{iL,jL'} \delta m_{iL}\left[\chi^{-1}_{0,iLL'}\delta_{ij} - S^{(2)}_{iL,jL'}\right]\delta m_{jL'} \tag{5.66}$$

This shows that the auxiliary susceptibility is a useful concept by itself.

To compute the $S^{(2)}$, one must find the response of the effective magnetic medium to a change in the order parameters, caused by the change in the external fields, which is obtained from the CPA equation that defines the effective scattering matrices, Eq. 5.31:

$$\int \mathrm{d}\hat{e}_i \left(\frac{\partial P_i(\hat{e}_i)}{\partial H_{jL'}}\,\underline{X}_i(\varepsilon;\hat{e}_i) + P_i(\hat{e}_i)\,\frac{\partial \underline{X}_i(\varepsilon;\hat{e}_i)}{\partial H_{jL'}}\right) = \underline{0} \tag{5.67}$$

yielding, after some linear algebra and use of the CPA condition to integrate out

two terms,

$$\frac{\partial \underline{\tilde{t}}_i^{-1}(\varepsilon)}{\partial H_{jL'}} = \beta \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \underline{X}_i(\varepsilon;\hat{e}_i) \sum_L \Big(\mathcal{Y}_L(\hat{e}_i) - m_{iL} \Big) \Big[\delta_{ij}\delta_{LL'} + \frac{\partial g_{iL}}{\partial H_{jL'}} \Big]$$
$$+ \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \underline{X}_i(\varepsilon;\hat{e}_i) \sum_{k\neq i} \underline{\widetilde{\tau}}_{ik}(\varepsilon) \, \frac{\partial \underline{\tilde{t}}_k^{-1}(\varepsilon)}{\partial H_{jL'}} \, \underline{\widetilde{\tau}}_{ki}(\varepsilon) \, \underline{X}_i(\varepsilon;\hat{e}_i) \quad . \tag{5.68}$$

Identifying

$$\delta_{ij}\delta_{LL'} + \frac{\partial g_{iL}}{\partial H_{jL'}} = \sum_{L_1} \chi^{-1}_{0,iLL_1} \chi_{iL_1,jL'} \quad \text{and} \quad \frac{\partial \underline{\tilde{t}}_i^{-1}(\varepsilon)}{\partial H_{jL'}} = \sum_{kL_1} \frac{\partial \underline{\tilde{t}}_i^{-1}(\varepsilon)}{\partial m_{kL_1}} \, \chi_{kL_1,jL'} \tag{5.69}$$

and introducing the bare and full vertex response, respectively,

$$\underline{\Lambda}^{(1)}_{iL}(\varepsilon) = \beta \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \underline{X}_i(\varepsilon;\hat{e}_i) \sum_{L_1} \Big(\mathcal{Y}_{L_1}(\hat{e}_i) - m_{iL_1} \Big) \, \chi^{-1}_{0,iL_1L}$$
$$= \frac{2\ell+1}{4\pi} \int \mathrm{d}\hat{e}_i \, \underline{X}_i(\varepsilon;\hat{e}_i) \, \mathcal{Y}_L(\hat{e}_i) \quad , \quad \underline{\Lambda}^{(2)}_{i,jL}(\varepsilon) = \frac{\partial \underline{\tilde{t}}_i^{-1}(\varepsilon)}{\partial m_{jL}} \tag{5.70}$$

a closed equation is obtained

$$\underline{\Lambda}^{(2)}_{i,jL}(\varepsilon) = \underline{\Lambda}^{(1)}_{iL}(\varepsilon) \, \delta_{ij} + \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \underline{X}_i(\varepsilon;\hat{e}_i) \sum_{k\neq i} \underline{\widetilde{\tau}}_{ik}(\varepsilon) \, \underline{\Lambda}^{(2)}_{k,jL}(\varepsilon) \, \underline{\widetilde{\tau}}_{ki}(\varepsilon) \, \underline{X}_i(\varepsilon;\hat{e}_i) \tag{5.71}$$

which also supplies the $S^{(2)}$, from Eq. 5.62 and differentiating,

$$S^{(2)}_{iL,jL'} = -\frac{1}{\pi} \mathrm{Im\,Tr} \int \mathrm{d}\varepsilon \, f(\varepsilon) \, \underline{\Lambda}^{(1)}_{iL}(\varepsilon) \sum_{k\neq i} \underline{\widetilde{\tau}}_{ik}(\varepsilon) \, \underline{\Lambda}^{(2)}_{k,jL'}(\varepsilon) \, \underline{\widetilde{\tau}}_{ki}(\varepsilon)$$
$$= -\frac{1}{\pi} \mathrm{Im\,Tr} \int \mathrm{d}\varepsilon \, f(\varepsilon) \, \underline{\Lambda}^{(1)}_{iL}(\varepsilon) \, \underline{\widetilde{\tau}}_{ij}(\varepsilon) \, \underline{\Lambda}^{(1)}_{jL'}(\varepsilon) \, \underline{\widetilde{\tau}}_{ji}(\varepsilon) + \ldots \tag{5.72}$$

The structure of this $S^{(2)}$ derived from linear response of the effective medium is more complex than the transverse susceptibility introduced in the previous section, see Eq. 5.56. In that equation the full susceptibility could be related to a bare susceptibility which is the product of two GFs. The full expression for the $S^{(2)}$, and so for the DLM transverse susceptibility, also starts with two GFs, the SPOs, as indicated in the second line of Eq. 5.72, but then higher order terms also contribute in an averaged way, as seen from the equation for the full vertex, Eq. 5.71.

The $S^{(2)}$ and the DLM state do not depend explicitly on temperature, but only on the values of the order parameters. To see this, note that the CPA condition,

Eq. 5.31, can be rewritten as

$$\underline{0} = \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \underline{X}_i(\varepsilon;\hat{e}_i) = \sum_L \int \mathrm{d}\hat{e}_i \, P_i(\hat{e}_i) \, \mathcal{Y}_L(\hat{e}_i) \, \frac{2\ell+1}{4\pi} \int \mathrm{d}\hat{e}'_i \, \underline{X}_i(\varepsilon;\hat{e}'_i) \, \mathcal{Y}_L(\hat{e}'_i)$$
$$= \sum_L m_{iL} \, \frac{2\ell+1}{4\pi} \int \mathrm{d}\hat{e}_i \, \underline{X}_i(\varepsilon;\hat{e}_i) \, \mathcal{Y}_L(\hat{e}_i) = \sum_L m_{iL} \, \underline{\Lambda}^{(1)}_{iL}(\varepsilon) \tag{5.73}$$

so that specifying the order parameters also specifies the electronic structure. This means that the poor description of the temperature dependence of the magnetic properties in the single–site approximation does not affect the construction of the effective medium.

The physical transverse magnetic susceptibility is then just

$$\chi^{\mathrm{T}}_{i\alpha,j\beta} = \sum_{LL'} M^{\alpha}_{iL} \, \chi_{iL,jL'} \, M^{\beta}_{jL'} \tag{5.74}$$

which reverts to the conventional expression using only the $\ell = 1$, 3×3 block if the local moments are rigid, $\vec{M}_i(\hat{e}_i) = M_i \, \hat{e}_i$. Non–Heisenberg–like interactions are still included in general even in this case, as the susceptibility is given by the usual inverse,

$$\chi_{iL,jL'} = \left[\chi_0^{-1} - S^{(2)} \right]^{-1}_{iL,jL'} \quad . \tag{5.75}$$

5.2.3 Complete magnetic order and complete magnetic disorder

In the paramagnetic state all order parameters are zero (except the trivial one for $L = 0$, which is always 1). By symmetry the local moments behave as rigid, as the effective medium is now paramagnetic, and so there is no favoured direction (barring on–site anisotropies, which are small and will be neglected for the moment). Eq. 5.73 then becomes

$$\underline{\Lambda}^{(1)}_{i0}(\varepsilon) = \underline{0} \iff \frac{1}{4\pi} \int \mathrm{d}\hat{e}_i \, \underline{X}_i(\varepsilon;\hat{e}_i) = \underline{0} \iff P_i(\hat{e}_i) = \frac{1}{4\pi} \quad . \tag{5.76}$$

If the moments are arranged in a periodic lattice, there is translational invariance and one can Fourier transform the equations,

$$\chi^{-1}_{LL'}(\vec{q}) = \chi^{-1}_{0,LL'} - S^{(2)}_{LL'}(\vec{q}) = (2\ell+1) \, k_{\mathrm{B}} T \, \delta_{LL'} - S^{(2)}_{LL'}(\vec{q}) \tag{5.77}$$

using the limit indicated in Eq. 5.23. The mean–field transition temperature can be extracted from this equation, by setting it to zero, and the periodicity of the magnetic state is given by $\vec{q}$.

The same t–matrices applies on all sites, and so the $S^{(2)}$ becomes

$$S^{(2)}_{LL'}(\vec{q}) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\underline{\Lambda}^{(1)}_{L}(\varepsilon)\int\frac{\mathrm{d}\vec{k}}{V_{\mathrm{BZ}}}\,\Delta\underline{\widetilde{\tau}}(\vec{k}+\vec{q};\varepsilon)\,\underline{\Lambda}^{(1)}_{L'}(\varepsilon)\,\Delta\underline{\widetilde{\tau}}(\vec{k};\varepsilon)+\ldots$$
$$= -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\underline{\Lambda}^{(1)}_{L}(\varepsilon)\int\frac{\mathrm{d}\vec{k}}{V_{\mathrm{BZ}}}\,\Delta\underline{\widetilde{\tau}}(\vec{k}+\vec{q};\varepsilon)\,\underline{\Lambda}^{(2)}_{L'}(\vec{q};\varepsilon)\,\Delta\underline{\widetilde{\tau}}(\vec{k};\varepsilon) \quad (5.78)$$

with

$$\Delta\underline{\widetilde{\tau}}(\vec{k};\varepsilon) = \underline{\widetilde{\tau}}(\vec{k};\varepsilon) - \int\frac{\mathrm{d}\vec{k}}{V_{\mathrm{BZ}}}\,\underline{\widetilde{\tau}}(\vec{k};\varepsilon)\ . \quad (5.79)$$

The integration over the Brillouin zone of the SPOs will favour those vectors $\vec{q}$ for which there is strong nesting of the electronic states of the effective medium; some states which might nest otherwise may give a much weaker contribution, if the broadening introduced by the statistical average is significant. These nesting conditions are the explanation for the Ruderman–Kittel–Kasuya–Yoshida (RKKY) interaction in metallic systems [93–95], which is driven by Fermi surface nesting.

This expression for the $S^{(2)}$ also applies if the reference state is ferromagnetic, as translational invariance is preserved, with the appropriate effective medium t–matrices being used. The on–site susceptibility then has its general form from Eq. 5.23, and the susceptibility is given as before. It can be used to detect an instability of the ferromagnetic state against some modulation $\vec{q}$, as the ferromagnetic order develops.

Taking now the limit of complete magnetic order, we have

$$P_i(\hat{e}_i) \longrightarrow \delta(\hat{e}_i - \hat{n}_i)\ ,\ \ \underline{\tilde{t}}_i(\varepsilon) \longrightarrow \underline{t}_i(\varepsilon;\hat{n}_i) \quad (5.80)$$

where $\hat{n}_i$ denote the orientations of the magnetic moments in the completely ordered state. In this limit, the S2 becomes just a pair interaction:

$$\underline{X}_i(\varepsilon;\hat{e}_i) = \left[\left(\underline{t}_i(\varepsilon;\hat{n}_i) - \underline{t}_i^{-1}(\varepsilon;\hat{e}_i)\right)^{-1} - \underline{\tau}_{ii}(\{\hat{n}\};\varepsilon)\right]^{-1}\ ,\ \ \underline{\Lambda}^{(2)}_{i,jL}(\varepsilon) = \underline{\Lambda}^{(1)}_{iL}(\varepsilon)\,\delta_{ij} \quad (5.81)$$

as the second term in Eq. 5.71 vanishes. This shows that the thermal effects are twofold: the electronic structure is altered, as the sharp electronic states corresponding to complete magnetic order become more or less broadened, as described by the effective medium; and the pair interactions become renormalised by the average fluctuations from all other sites.

These concepts generalise naturally to a system which can be described by a finite number of sublattices. The interface between two magnetic systems, or a magnetic surface, however, have to be handled in a different way, as one or two ideally semi–

infinite systems have to be described. The DLM state must then obey the bulk magnetic boundary conditions, and the magnetic susceptibility has to be written in terms of the bulk susceptibilities, in a similar way to the procedure leading to Eq. 2.67.

5.2.4 Short–range correlations and the Onsager cavity field

The main drawback of the mean–field theory just described is the failure to incorporate short–range correlations, which leads to quantitative and qualitative discrepancies. To conclude this chapter on the theoretical aspects it is worth discussing a modification of the theory which amelliorates some of these faults. It is called the Onsager cavity field construction, going back to Lars Onsager pioneering work on dielectrics [96], and was applied to magnetism by Brout and Thomas [97].

The argument behind the Onsager cavity field construction is fairly simple to formulate: the effective field acting on the local moment at a given site must not include the interaction of the moment with itself; this generates a reaction field which is always parallel to the orientation of that moment, and so does not contribute to the statistical averages, and to the local order parameter. The simple Heisenberg model, Eq. 5.1, will be used to briefly illustrate this concept.

The reaction field is to be subtracted from the effective field:

$$\vec{h}_i = \vec{H}_i + \sum_{j \neq i} J_{ij}\, \vec{m}_j - \lambda_i\, \vec{m}_i \tag{5.82}$$

and the magnetic susceptibility is modified accordingly (the possible dependence of the reaction field parameter λ_i on the external fields is neglected)

$$\chi_{ij} = \chi_{0,i}\, \delta_{ij} + \chi_{0,i} \sum_{k \neq i} J_{ik}\, \chi_{kj} - \lambda_i\, \chi_{ij} \quad . \tag{5.83}$$

One can now reinstate the diagonal part of the fluctuation–dissipation theorem, Eq. 3.14,

$$\chi_{ii} \equiv \chi_{0,i} \quad \Longrightarrow \quad \lambda_i = \chi_{0,i}^{-1} \sum_{k \neq i} J_{ik}\, \chi_{ki} \tag{5.84}$$

so that the Onsager reaction field parameter λ_i is temperature dependent, and must be calculated self–consistently with the rest of the statistical quantities.

This simple correction scheme has been used successfully in the DLM theory [88] and in the order–disorder theory of alloys [98], as well as in handling strong electronic correlations and magnetism in the Hubbard model [99]. A model of classical spin fluctuations applicable for itinerant electron systems was discussed in the same way

[100]. For classical unit vectors it is equivalent to the spherical model [101], and it suppresses the mean–field magnetic phase transition in 2D systems, in the absence of anisotropy, thus restoring the Mermin–Wagner theorem [78]. In first–principles approaches which compute the transition temperature from the Fourier transform of the exchange interaction parameters there is a similar treatment known as the Random Phase Approximation [72].

Chapter 6

Magnetic monolayers on non–magnetic substrates

In this chapter the formalism previously introduced is applied to the problem of the properties of magnetic monolayers (MLs) supported on non–magnetic substrates. A brief overview of some theoretical and experimental results is given first, followed by calculations for a system which is well characterised theoretically and experimentally, the Mn ML on the W(001) surface. Then original theoretical results are given for Mn MLs on the (111) surfaces of Pd, Pt, Ag and Au. This is a systematic study of the influence of the substrate on the properties of the ML, and was spurred by experimental investigations of Mn on Ag(111). These calculations were already published [12].

6.1 Complex magnetism in supported monolayers

Modern experimental techniques developed over the past decades, such as molecular beam epitaxy (MBE), magnetron sputtering, ion beam sputtering, pulsed laser deposition (PLD) or metal organic chemical vapor deposition (MOCVD), afford unprecedented control in the growth of artificial magnetic structures. A range of characterisation techniques is also available, from electron and X–ray diffraction to scanning probes, such as scanning tunneling microscopy (STM). Neutron scattering is also becoming more frequent, such as polarised neutron reflectometry. The experimental progress has not been completely followed by the necessary theoretical developments, and the field of first–principles and micromagnetic modelling of the magnetic properties of nano and heterostructures is a very active one.

The reason why theoretical progress has been slower is that the loss of symmetry

and the intrinsic heterogeneity of most systems of interest make any analysis much more complicated, and even simple mean–field theories can become quite involved. A notorious example that has spread controversy and years of theoretical work consists of a few layers of Fe deposited on Cu(001) [89, 102–105]. This is the thin film version of the controversies over the magnetic ground state of fcc Fe, and is still on–going work. To gain insight into the several mechanisms driving the magnetic properties of these systems it is convenient to consider the simplest case possible: a single atomic layer of magnetic material deposited on a non–magnetic substrate. This is the best approximation available to a true two–dimensional system, and raises fundamental questions. Very recently the Mermin–Wagner theorem was extended to models which might be applicable to itinerant electron systems [106], and rules out magnetic ordering for continuous symmetry (like vector models) in the absence of anisotropy, as the original result [78].

A brief selection of first–principles results concerning magnetic MLs, often accompanying experimental findings, is now given. Fe on $4d$ and $5d$ substrates is sometimes no longer ferromagnetic (FM) [107]. FM Fe on W(110) has a chiral spin–wave spectrum [108], the domain walls separating magnetic domains are also chiral [109], and becomes antiferromagnetic (AFM) on the W(001) surface [11]. Fe on Ir(111) displays a very complicated magnetic structure [110], which has been interpreted as the magnetic analogue of a skyrmion lattice [111]. Mn exhibits a spin spiral state on W(001) and W(110) [112, 113], a triangular AFM state on Ag(111) [114], and was predicted theoretically to display a magnetic state corresponding to a superposition of three different wavevectors of the same symmetry on Cu(111) [115]. For the test calculations Mn_1/W(001) was chosen, and for the original results Mn_1/Ag(111) and the related Pd, Pt and Au substrates were analysed.

6.2 A test case: Mn monolayer on W(001)

6.2.1 Understanding the system

The Mn_1/W(001) system consists of a single Mn ML on the (001) surface of bcc W. The Mn atoms form a square lattice (taken in the xy–plane), matching the in–plane lattice constant of W (a = 3.165 Å), and relaxing towards the substrate, in comparison with the ideal W—W interlayer separation along the z–direction (this is not determined experimentally, but first–principles calculations give values in the 10–15 % range [116, 117], depending also on the assumed magnetic state). The experimental finding is that the magnetic state is a cycloidal spin–spiral ($\vec{q}\cdot\hat{n}_3 = 0$ and $\theta = \pi/2$, see Fig. 4.1, top), with a well–defined rotational sense [112].

To aid the understanding of the theoretical results, the simplest model with the necessary ingredients to generate such a magnetic state is now explained. Firstly, for a spin–spiral to be energetically favourable, at least two different isotropic exchange interactions must be present[1], which will be taken to be J_1 for nearest–neighbours (nn) in the Mn square lattice, and J_2 for next–nearest–neighbours (nnn). A schematic of this geometry is given in Fig. 6.1, left. Let us compute the energy of a generic spin spiral, according to Eq. 4.7 (the length of the spin moment is assumed independent of the wavevector, and can be absorbed into the exchange interactions):

$$\vec{m}_i(\vec{q}) = \Big(\cos(\vec{q}\cdot\vec{R}_i+\phi)\sin\theta\,\hat{n}_1 + \sin(\vec{q}\cdot\vec{R}_i+\phi)\sin\theta\,\hat{n}_2 + \cos\theta\,\hat{n}_3\Big)\ . \tag{6.1}$$

The energy due to isotropic pair interactions per atom is given by

$$E_{\rm iso}(\vec{q};\theta) = -\frac{1}{2N}\sum_{ij} J_{ij}\,\vec{m}_i(\vec{q})\cdot\vec{m}_j(\vec{q}) = \frac{J(0)-J(\vec{q})}{2}\sin^2\theta - \frac{J(0)}{2} \tag{6.2}$$

and the lattice Fourier sum is ($\vec{q}=\pi/a\,(q_x,q_y)$, with $q_x,q_y\in[-1,1]$)

$$J(\vec{q}) = \sum_j J_{0j}\cos(\vec{q}\cdot\vec{R}_{0j}) = 2J_1\left[\cos(\pi q_x)+\cos(\pi q_y)\right] + 4J_2\cos(\pi q_x)\cos(\pi q_y) \tag{6.3}$$

The extremal values of this energy are given by the conditions

$$\sin(\pi q_x)\Big(1+\frac{2J_2}{J_1}\cos(\pi q_y)\Big) = 0 \tag{6.4}$$

$$\sin(\pi q_y)\Big(1+\frac{2J_2}{J_1}\cos(\pi q_x)\Big) = 0 \tag{6.5}$$

$$\big(J(0)-J(\vec{q})\big)\sin 2\theta = 0 \tag{6.6}$$

The cone angle θ can take on the values 0 or $\pi/2$, and only in case of accidental degeneracy on any other values, as shown by the last equation. Flat spirals are then the norm[2].

The first two equations show that there are simple magnetic states corresponding to $(q_x,q_y) = (0,0),\ (\pm1,0),\ (0,\pm1)$ and $(\pm1,\pm1)$, which are FM, four symmetry–equivalent row–wise AFM, and four checkerboard AFM. Setting the term in brackets on the first two equations to zero gives the condition for the possibility of a spin spiral state: $-2J_2 \le J_1 \le 2J_2$. This implies that the nn interaction must be weak

[1] Or an appropriate anisotropic interaction, as will be shown below.

[2] $\theta = 0$ gives a FM arrangement, which is equivalent to a flat spiral with $\vec{q}=\vec{0}$.

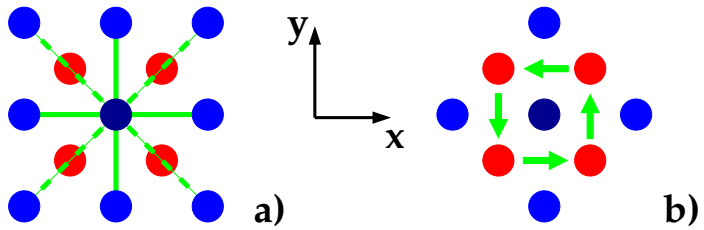

Figure 6.1: Geometry for the Mn ML supported on the W(001) substrate. Mn atoms are in blue, and W atoms in red. a) — nn interactions J_1 (solid green line) and nnn interactions (dashed green lines) between the central atom and the first and second lattice shells. b) — DM vectors between the central Mn atom and the four nearest neighbours.

enough compared to the nnn interaction. The spiral wavevector is then along one of the diagonals of the square lattice.

As expected, the isotropic interactions do not dictate the real–space orientation of the magnetic state, which is given by the unit vectors $\hat{n}_1$, $\hat{n}_2$ and $\hat{n}_3$. For this one must consider anisotropic interactions, see Chapter 3. The unidirectional Dzyaloshinsky–Moriya (DM) interaction [9, 10] can also fix the orientation of the magnetic moments, and is now analysed. Take the following form for the anisotropic nn pair interaction:

$$\underline{J}_{ij} = \begin{pmatrix} J_1 & D_z & -D_y \\ -D_z & J_1 & D_x \\ D_y & -D_x & J_1 \end{pmatrix} \ , \ \underline{J}_{ji} = \underline{J}_{ij}^T \tag{6.7}$$

and the overall symmetry of the pair interaction is reflected in the second equality. The non–zero off–diagonal coefficients of this matrix form an antisymmetric matrix, and results in the following coupling:

$$\vec{m}_i \cdot \underline{J}_{ij} \cdot \vec{m}_j = J_1 \, \vec{m}_i \cdot \vec{m}_j + \vec{D}_{ij} \cdot \left(\vec{m}_i \times \vec{m}_j \right) \ , \ \ \vec{D}_{ij} = (D_x, D_y, D_z) \ , \ \ \vec{D}_{ji} = -\vec{D}_{ij} \tag{6.8}$$

which introduces the DM vector $\vec{D}_{ij}$. The non–zero coefficients of this vector are dictated by the local symmetry of the pair of interacting sites [81]; for the current case, the square lattice supported by the non–magnetic substrate, $D_z = 0$, and the vector is in–plane and perpendicular to the line connecting i and j, with $|D_x| = |D_y| = D$ (see Fig. 6.1, right). Evaluating the cross product with the spin spiral form in Eq. 6.1, only one term gives a non–zero contribution to the lattice Fourier

sum (more details in Appendix B):

$$E_{\mathrm{DM}}(\vec{q};\theta;\hat{n}_3) = -\frac{1}{2N}\sum_{ij}\vec{D}_{ij}\cdot\left(\vec{m}_i\times\vec{m}_j\right) = -\frac{1}{2N}\sum_{ij}\vec{D}_{ij}\cdot\hat{n}_3\,\sin(\vec{q}\cdot\vec{R}_{ij})\sin^2\theta \quad (6.9)$$

A non–vanishing contribution to the energy requires a non–zero dot product, so the spin spiral will have a definite orientation in real space, according to the non–vanishing elements of the DM vector. The meaning of 'unidirectional' is now evident: the DM energy is not invariant under the $\vec{q}\to-\vec{q}$ transformation, only if $\hat{n}_3\to-\hat{n}_3$ also applies at the same time. This is why the DM interaction defines the rotational sense of a spin spiral, which is given by the pair $(\vec{q},\hat{n}_3)$.

In the present case this form specialises to

$$\begin{aligned} E_{\mathrm{DM}}(\vec{q};\theta;\hat{n}_3) &= D\left[(\hat{e}_x\cdot\hat{n}_3)\sin(\pi q_y) - (\hat{e}_y\cdot\hat{n}_3)\sin(\pi q_x)\right]\sin^2\theta && (6.10)\\ &= D\left[\cos\alpha\sin(\pi q_y) - \sin\alpha\sin(\pi q_x)\right]\sin^2\theta && (6.11)\end{aligned}$$

by writing $\hat{n}_3 = \cos\alpha\,\hat{e}_x + \sin\alpha\,\hat{e}_y$, and for collinear states $\sin(\pi q_x) = \sin(\pi q_y) = 0$, which shows that the DM energy is zero, as the cross product $\vec{m}_i\times\vec{m}_j$ is zero. In general the DM interaction can always lower the energy, by lifting the degeneracy coming from the isotropic energy, $E_{\mathrm{iso}}(\vec{q};\theta) = E_{\mathrm{iso}}(-\vec{q};\theta)$.

The stationary points of the energy come from (ignoring the cone angle)

$$\begin{aligned} \sin(\pi q_x)\Big(1+\frac{2J_2}{J_1}\cos(\pi q_y)\Big) - \frac{D}{J_1}\sin\alpha\cos(\pi q_x) = 0\\ \sin(\pi q_y)\Big(1+\frac{2J_2}{J_1}\cos(\pi q_x)\Big) + \frac{D}{J_1}\cos\alpha\cos(\pi q_y) = 0 \end{aligned} \quad (6.12)$$

$$\cos\alpha = \pm\frac{\sin(\pi q_y)}{\sqrt{\sin^2(\pi q_x)+\sin^2(\pi q_y)}}\,,\quad \sin\alpha = \mp\frac{\sin(\pi q_x)}{\sqrt{\sin^2(\pi q_x)+\sin^2(\pi q_y)}} \quad (6.13)$$

and the last two equations determine the two stationary directions of $\hat{n}_3$, and come from differentiating the energy with respect to the angle α. These are shown to depend on the modulation vector of the spin spiral, $\vec{q}$. One maximises the DM energy, while the other minimises it, while both directions are degenerate for the

isotropic energy. The spiral wavevector is then determined from the two equations

$$\sin(\pi q_x)\left[1+\frac{2J_2}{J_1}\cos(\pi q_y)\pm\frac{D}{J_1}\frac{\cos(\pi q_x)}{\sqrt{\sin^2(\pi q_x)+\sin^2(\pi q_y)}}\right]=0$$
$$\sin(\pi q_y)\left[1+\frac{2J_2}{J_1}\cos(\pi q_x)\pm\frac{D}{J_1}\frac{\cos(\pi q_y)}{\sqrt{\sin^2(\pi q_x)+\sin^2(\pi q_y)}}\right]=0 \qquad (6.14)$$

There are spin spiral solutions with $\vec{q}$ along the nn or nnn directions. The DM can also stabilise a spin–spiral by itself, as can be checked by setting $J_2 = 0$. From the equations for the angle α, it also follows that $\hat{n}_3 \cdot \vec{q} = 0$ in all cases.

Extending the discussion to include the uniaxial anisotropies would lead to unnecessary complication at this point. A general consequence of these anisotropies is that the energy can be lowered by deviating the orientations of the moments from the ideal spin spiral arrangement to align slightly with the anisotropy easy axes (this is sometimes called 'bunching' [8]). This also provides a mechanism to stabilise conical spin spirals: with $\hat{n}_3$ pointing along an easy axis, the cone angle can deviate from $\theta = \pi/2$. If different kinds of anisotropies are present, competition between them might occur. This is analogous to the competition between isotropic interactions which leads to the formation of non–collinear magnetic states, as explained above. A more detailed discussion is given in Appendix B.

Armed with these insights as to how the different ingredients combine to determine the magnetic state, we can now analyse the results of first–principles calculations of the spin–spin correlation function, and compare with published theoretical and experimental results.

6.2.2 Linear response from the paramagnetic state

The $Mn_1/W(001)$ self–consistent calculations were done using the Budapest SKKR code [47], in the ASA with $\ell_{\max} = 3$. A semicircular contour in the upper half of the complex plane with an asymetric distribution of 16 or 24 points was employed for the energy integrations, and 78 points in the irreducible wedge of the Brillouin zone were taken, which amounts to 624 points in the full zone, according to the C_{4v} symmetry of the system. The interfacial region was taken to comprise of eight layers of W, the Mn layer, and three layers of empty spheres modelling the decay towards the vacuum, and the matching was done to the semi–infinite W substrate and to the semi–infinite vacuum. The Mn layer was relaxed 10.8% towards W, in comparison with the ideal W–W interlayer separation, as in Ref. [116].

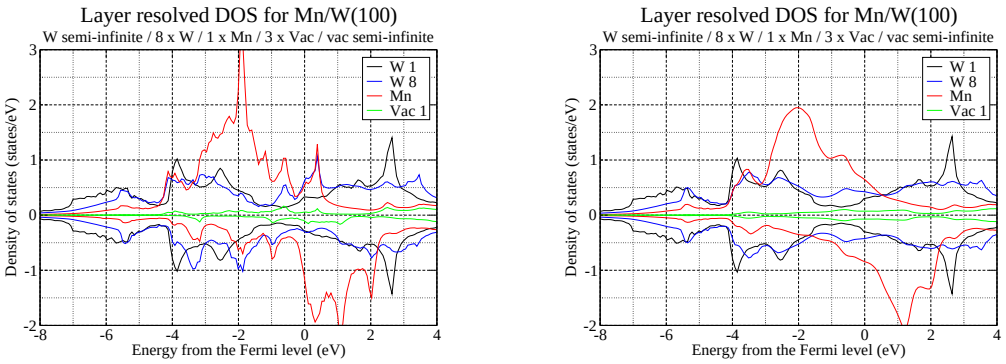

Figure 6.2: Layer–resolved DOS for ferromagnetic (left) and DLM (right) states of $Mn_1/W(001)$. The majority spin DOS is plotted as positive values, while the minority spin DOS is given as negative. W 1 is the substrate layer farthest from the Mn ML, while W 8 is in atomic contact with it. Vac 1 is the vacuum layer next to the Mn ML. In the DLM calculation there is no net spin polarisation; the Mn DOS is then computed from the partially averaged GF, to allow a comparison with the respective FM DOS.

Let us first compare the electronic structure and properties in the FM and DLM states. In the DLM state only Mn has a spin moment, of $3.10\mu_B$, which gives a net vector spin moment of zero upon averaging with $P(\hat{e}) = 1/4\pi$. This increases to $3.15\,\mu_B$ in the FM state, and the W layers magnetise in an alternating fashion: the layer next to Mn aligns antiparallel to the Mn moment, with $0.32\,\mu_B$, and successive layers align antiparallel to each other, with the induced magnetisation quickly decreasing to zero towards the bulk. The actual value of the spin moment depends on the value chosen for the relaxation towards the substrate, and different methods given different equilibrium geometries [116–118]. Our calculated values agree with those given in a recent paper by Shick [116] for the FM state.

The calculated layer resolved DOS, see Fig. 6.2, show that the sharp structure of the DOS in the FM state is broadened due to the magnetic disorder in the DLM state. To highlight the effect of hybridisation between the Mn and the nearest W layer, the DOS for the W layer farthest away from Mn is also given for comparison (along with the DOS in the first layer of vacuum). It can be seen that there are states in the W layer next to Mn which are not present in the bulk around the Fermi level. These are spin–polarised in the FM state, but there is no spin splitting in the DLM state, as W does not have a local magnetic moment. The Mn DOS remains spin–polarised in the DLM state, which demonstrates the presence of the local moment.

The $S^{(2)}$ calculations were done using only the $\ell = 1$ block of the general expres-

sion, Eq. 5.78, in the DLM state. This reverts back to the familiar vector notation. For the susceptibility, Eq. 5.77 becomes

$$\chi^{-1}_{\alpha\beta}(\vec{q}) = 3\,k_{\mathrm{B}}T\,\delta_{\alpha\beta} - S^{(2)}_{\alpha\beta}(\vec{q}) \quad , \quad (\alpha,\beta = x,y,z) \tag{6.15}$$

and the appropriate $S^{(2)}$ is

$$S^{(2)}_{\alpha\beta}(\vec{q}) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\underline{\Lambda}^{(1)}_{\alpha}(\varepsilon)\int \frac{\mathrm{d}\vec{k}}{V_{\mathrm{BZ}}}\,\Delta\underline{\widetilde{\tau}}(\vec{k}+\vec{q};\varepsilon)\,\underline{\Lambda}^{(1)}_{\beta}(\varepsilon)\,\Delta\underline{\widetilde{\tau}}(\vec{k};\varepsilon) \tag{6.16}$$

and only the bare vertex is considered in the following, except when otherwise noted. Calculations done using the full expression show only small numerical differences, and no change in the qualitative behaviour. A dense equidistant mesh (40×40 points) was used for the BZ integration, and the modulation vectors $\vec{q}$ were chosen from that mesh.

As explained in Appendix B, it is useful to decompose the 3×3 tensor in isotropic, traceless symmetric, and antisymmetric parts:

$$\begin{aligned} S^{(2)}_{\alpha\beta}(\vec{q}) &= S^{(2)}(\vec{q})\,\delta_{\alpha\beta} + S^{(2),\mathrm{S}}_{\alpha\beta}(\vec{q}) + S^{(2),\mathrm{A}}_{\alpha\beta}(\vec{q}) \;, \quad S^{(2)}(\vec{q}) = \frac{1}{3}\sum_{\alpha} S^{(2)}_{\alpha\alpha}(\vec{q}) \\ S^{(2),\mathrm{S}}_{\alpha\beta}(\vec{q}) &= \frac{S^{(2)}_{\alpha\beta}(\vec{q}) + S^{(2)}_{\beta\alpha}(\vec{q})}{2} - S^{(2)}(\vec{q})\,\delta_{\alpha\beta} \;, \quad S^{(2),\mathrm{A}}_{\alpha\beta}(\vec{q}) = \frac{S^{(2)}_{\alpha\beta}(\vec{q}) - S^{(2)}_{\beta\alpha}(\vec{q})}{2} \end{aligned} \tag{6.17}$$

To identify the main contributions, some representative plots are given in Fig. 6.3, left. The wavevectors are taken along straight lines in the BZ, with the endpoints indicated on the plot. The isotropic part is by far the dominant contribution, as expected, and already indicates that a spin spiral state would be more energetically favourable than the FM state $\vec{q} = (0,0)$. The checkerboard AFM, $\vec{q} = (1,1)\pi/a$, is very unfavourable, while the row–wise AFM, $\vec{q} = (1,0)\pi/a$, is not much lower in energy than the FM state, as well as all spin spirals on the line connecting those magnetic states. The result using the full vertex is also given, and as stated before the differences are small, for this system.

A fit to the J_1–J_2 model for the isotropic interactions, Eq. 6.3, is also given. The structure near the centre of the BZ, $\vec{q} = \vec{0}$, is not captured in this simple model. This is due to the long ranged nature of the interactions. To help quantify this, consider the quantity

$$J_0 = \sum_{n=1}^{\infty} z_n\,J_n \equiv S^{(2)}(\vec{0}) \tag{6.18}$$

which is the total isotropic interaction for the FM case. Here n labels the n–th

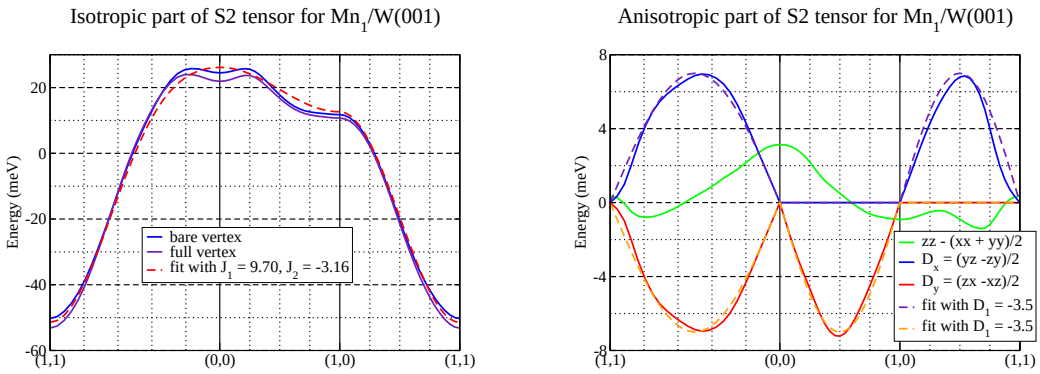

Figure 6.3: The most important components of $\underline{S}^{(2)}$ (see Eq. 6.17) for $Mn_1/W(001)$, along special symmetry directions in the BZ. The wavevector is given in units of π/a, and the fit uses the $S^{(2)}$ computed with the full vertex. The fit in the left panel is to Eq. 6.3, and in the right panel to Eq. 6.19.

shell of atoms, z_n the number of atoms in that shell, and J_n the value of the pair interaction between the atom at the origin and any atom on the n–th shell. The pair interactions are long–ranged and of varying sign, so this is an alternating series, thus defining a reliable truncation has to be done with care.

By inverse Fourier transforming (as a dense equidistant set of $\vec{q}$'s is available), these parameters can be determined without fitting. Taking the $S^{(2)}$ computed with the bare vertex as an example, we obtain $J_1 = 9.79$ meV and $J_2 = -2.89$ meV. The simple fit using Eq. 6.3 gives $J_1 = 9.70$ meV and $J_2 = -3.16$ meV, and is displayed in Fig. 6.3, left. The agreement is fair. However, $S^{(2)}(\vec{0}) = 24.5$ meV, while $4J_1 + 4J_2 = 27.6$ from the same data. The remaining 3 meV is the true value of the sum of all the pair interactions, of alternating sign, and it is very difficult to reproduce this number using a finite truncation and a fitting procedure. Returning to the previous model based solely on isotropic interactions, no spin spiral is predicted, as can be observed from the fitting curve and from $J_1 > 2|J_2|$. This contrasts with the untruncated first–principles results, which already indicate a spin spiral with $\vec{q} \approx (0.2, 0.2)\pi/a$, only slightly higher in energy than all other wavevectors with $q \approx 0.25\pi/a$ around the BZ centre.

The most important anisotropic contributions to the $S^{(2)}$ are shown in Fig. 6.3, right. By symmetry, the xz and yz elements of the uniaxial anisotropy are zero, as are the xy elements of the unidirectional anistropy, which would be the z–component of the DM vector. This is in agreement with the discussion in Ref. [81]. There is a weak in–plane uniaxial anisotropy, corresponding to an xy term, which is not shown, and the dominant uniaxial anisotropy is the out of plane element, which was

represented in Fig. 6.3 with the average of the xx and yy elements subtracted out. The form of the DM vector, for nn and nnn, is

$$\vec{D}_1(\vec{q}) = 2D_1\big[\sin(\pi q_x)\,\hat{e}_y - \sin(\pi q_y)\,\hat{e}_x\big] \tag{6.19}$$

$$\vec{D}_2(\vec{q}) = 4D_2\big[\sin(\pi q_x)\cos(\pi q_y)\,\hat{e}_y - \cos(\pi q_x)\sin(\pi q_y)\,\hat{e}_x\big] \tag{6.20}$$

On the path in reciprocal space with $(q_x, q_y) = (q, q)$, $D_x(\vec{q}) = -D_y(\vec{q})$, while for $(q_x, q_y) = (q, 0)$ we have $D_x(\vec{q}) = 0$ and for $(q_x, q_y) = (1, q)$ it follows that $D_y(\vec{q}) = 0$. All these relations hold for a general form of the DM vector, and are a consequence of the C_{4v} symmetry of the surface [81]. It can also be seen from Fig. 6.3, right that the nn approximation to the DM vector is well–justified. Taking the bare or the full vertex makes little difference for the actual numbers for the anisotropy.

Ignoring the uniaxial anisotropy for the moment, one can solve Eqs. 6.14 for the three symmetry directions shown in the plot, using the three fitted parameters, J_1, J_2 and D_1. As $D_1 < 0$, the solution for the minimum is obtained by choosing the minus sign in those equations (cf. Eqs. 6.11 and 6.13). Two solutions are almost degenerate minima of the model energy, $q = (0.17, 0.17)$ and $q = (0.25, 0.00)$, and interestingly are in approximate correspondence with the ring of almost degenerate maxima found for the calculated isotropic part of $S^{(2)}$. The uniaxial anisotropy energy is also about the same for both spin spirals, as can be seen from Fig. 6.3, right. The spiral vector might be slightly changed when this additional anisotropy energy is considered, but no qualitative change is expected. The absolute minimum is obtained for $q \approx (0.17, 0.17)$, which corresponds to a real–space period of 21 Å, or about four lattice constants of bulk W, and corresponds to a cycloidal spin spiral ($\hat{n}_3 \cdot \vec{q} = 0$, see Fig. 4.1, top).

To conclude, we refer back to Ref. [112], which reports on the experimental discovery of the magnetic state of $Mn_1/W(001)$, along with first–principles calculations. The magnetic ground state is a cycloidal spin spiral, with period of 22 Å, propagating along the diagonals of the square lattice. The first–principles calculations were done using the generalised Bloch theorem for spin spirals, and are in very good agreement with experiment. As was demonstrated by the previous discussion, our linear response theory of the instabilities of the DLM state finds essentially the same results, but from a very different starting point.

6.3 Mn monolayers on d^9 and d^{10} substrates

6.3.1 Introduction to the experimental results for $Mn_1/Ag(111)$

The inspiration for the theoretical investigation to be described in the following sections came from an experimental 'detective's work' on triangular Mn islands grown on the hexagonal (111) surface of fcc Ag [114]. Keeping the orientation of the surface fixed, there are four different types of islands, as depicted in Fig. 6.4. According to how the islands stack on the substrate, they can be called fcc (if there is no subsurface atom underneath the Mn atoms, Fig. 6.4, left) or hcp (if Mn atoms stack on top of the subsurface atoms, Fig. 6.4, right). The triangles can form in two different ways, and are creatively called 'up' and 'down' triangles. The islands can be considered as assembled from the building blocks highlighted by the green lines; fcc and hcp islands then also differ by the relative position of the surface atom with respect to these losenges.

The magnetic structure of the Mn islands was investigated using spin–polarised STM. It is fairly easy to establish that the magnetic state corresponds to triangular AFM, with 120° angles between neighbouring magnetic moments. The real detective work was in establishing the relative orientations of the moments in different types of islands. The moments in the fcc islands appear to be rotated by 30° or 60° from the moments in the hcp islands (this is part of the subtlety of interpreting the STM contrast). It was found that up and down islands display identical behaviour, for the same type of stacking.

First–principles calculations for this system predict that the magnetic ground-state should be row–wise AFM, in contradiction with experiment. Simple row by row AFM is predicted in Ref. [119], while Ref. [120] predicts double rows of FM spins antiparallel to each other. There is no quantitative explanation for this discrepancy so far.

6.3.2 Reference states for $Mn_1/X(111)$, X = Pd, Pt, Ag and Au

To investigate the disagreement between theory and experiment, a comparative study was made by calculating the properties of the Mn ML on the Ag substrate with those corresponding to the Pd, Pt and Au substrates, which are all fcc, and neighbours on the periodic table. The geometry used in our calculations was as follows. The interface region was composed of eight layers of the transition metal substrate, the Mn ML, and three layers of empty spheres. This was then matched to the semi–infinite bulk substrate and to the semi–infinite vacuum region.

No attempt was made at determining the equilibrium geometries for the four

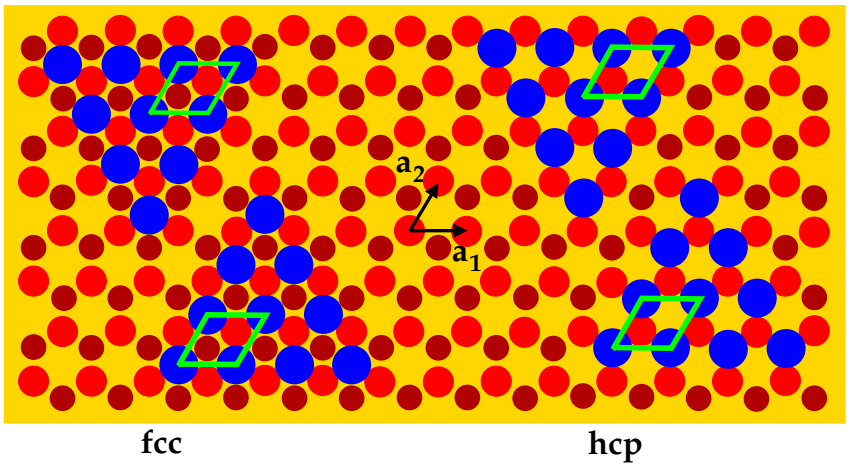

Figure 6.4: The geometry of the Mn triangular islands grown on an fcc(111) substrate. The Mn adatoms are in blue. The surface layer is represented by the bright red circles, while the subsurface layer is depicted with dark red circles. In the fcc stacking, the Mn adatoms stack on the hollow sites without a subsurface atom underneath (left), while for the hcp stacking, they stack on the hollow site directly on top of the subsurface atoms (right). A possible choice for the basis vectors spanning the hexagonal layers is also shown.

systems studied. However, the influence of different interlayer spacings between the magnetic ML and the substrate was investigated for the cases of Mn/Ag and Mn/Au, and it was found that there was no qualitative change in the magnetic properties, up to 15% inward relaxation. The dependence of the magnetic properties on the in–plane lattice constant was addressed in a previous study [121], and for the range relevant to our systems the same conclusion can be made. On the other hand, the relaxed geometrical parameters depend on the magnetic state and other approximations, such as the exchange–correlation functional and the handling of the charge density, so a definite statement is precluded. We thus progress with a simple model of the geometric structure, which we now describe.

To aid comparison, the average of the experimental lattice constants for Pd and Pt was used for both Pd and Pt (a = 3.905 Å), and likewise for Ag and Au (a = 4.085 Å). This means that the in–plane lattice constant for the Mn ML is the same for the Pd and Pt substrates, and also for the Ag and Au substrates. To estimate the inward relaxation, it was assumed that Mn grows on Cu(111) without any significant relaxation, thus defining the Mn ML hard sphere radius. By considering how these hard spheres stack on top of the larger ones corresponding to the substrate, the following relation was obtained for the ratio between the estimated and the ideal distance between planes: $d/d_0 = \sqrt{3/8\,(1 + a_{\mathrm{Cu}}/a_{\mathrm{X}})^2 - 1/2}$. We then round the

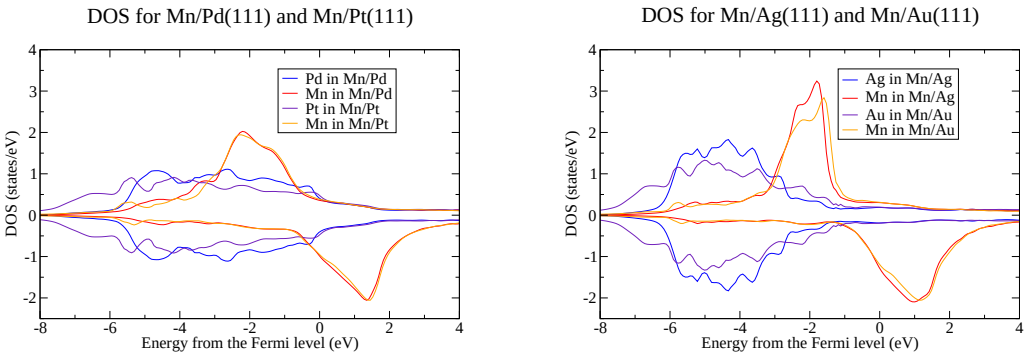

Figure 6.5: Layer–resolved DOS for Mn1/X(111), with X = Pd, Pt, Ag and Au. The majority spin DOS is plotted as positive values, while the minority spin DOS is given as negative. Only the DOS for the Mn and the nearest substrate layer is shown. In the DLM calculation there is no net spin polarisation; the Mn DOS is then computed from the partially averaged GF, to show the local spin splitting of the states; by averaging over all orientations of the local moment with $P_i(\hat{e}_i) = 1/4\pi$ no net spin polarisation remains. The substrate layers possess no local moments, and so no local spin splitting can be detected.

results and use 5% inward relaxation for the Pd and Pt substrates, and 10% for Ag and Au.

The influence of the substrate on the electronic structure of the Mn ML can be readily understood from the layer–resolved DOS, in the DLM state, shown in Fig. 6.5. The Mn minority states are just slightly affected by changing the substrate, as can be seen be comparing all four cases. The majority states, however, show a marked difference. For the Ag and Au substrates these states are fairly narrow, while for the Pd and Pt substrates they are much wider in energy, and show strong signs of hybridisation. This can be explained by the position of the substrate d–states. Ag and Au have filled d–bands, lower in energy than the Mn majority states, and so there is little hybridisation. On the other hand, Pd and Pt have partially filled d–bands, which extend up to the Fermi energy, so there is much stronger hybridisation with the Mn majority states. The strength of the hybridisation with the substrate also affects the size of the DLM spin moments: for Mn/Ag and Mn/Au it is about 3.9 μ_B, in agreement with previous calculations [119], while for Mn/Pd and Mn/Pt it drops to about 3.6 μ_B.

6.3.3 Magnetic interactions from the DLM state of Mn_1/X(111)

We now analyse the magnetic correlations in the DLM state of Mn_1/X(111), extracting the relevant information from the $S^{(2)}$, calculated according to Eq. 6.16.

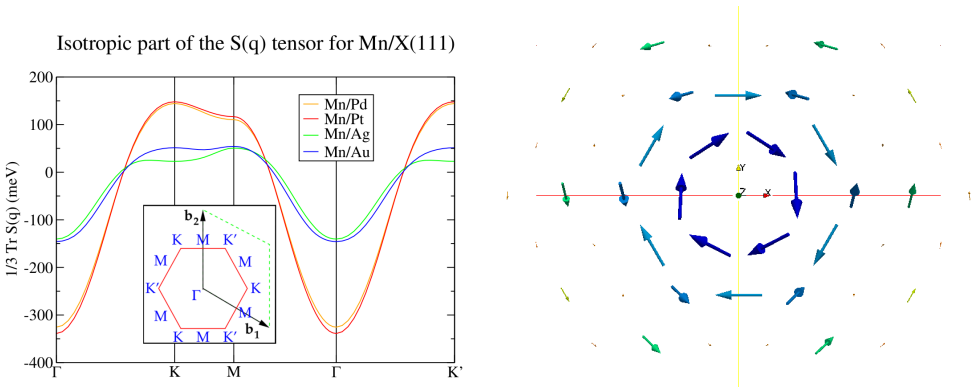

Figure 6.6: Left: isotropic part of the $S^{(2)}$ tensor, plotted along the high–symmetry lines of the hexagonal BZ. The symmetry points are shown on the inset, and correspond to simple magnetic states (see Fig. 6.7). Right: DM vectors in real space, computed by inverse Fourier transforming the first–principles data, plotted for all pairs formed by some atom and its neighbours in concentric shells. See Eq. 6.17 for the decomposition of the data into isotropic and anisotropic quantities.

3721 equidistant points in the 2D BZ were used. The isotropic part of the tensor is again the dominant part, and is given in Fig. 6.6, left. The discussion of the J_1–J_2 model on the hexagonal lattice is slightly more involved than on the square lattice, and so is given in detail in Appendix C. The main ideas will be quoted from there as needed.

For the isotropic part of the interaction, the J_1–J_2 model predicts extrema at Γ, M, K and K' (see inset in Fig. 6.6, left), as well as one on the Γ–K line. This is the behaviour seen in Mn/Ag and Mn/Au. For both cases the favoured state is row–wise AFM (Fig. 6.7, left), but Mn/Au shows an approximate degeneracy with the 120° AFM state (K or K' points). On the other hand, Mn/Pd and Mn/Pt show

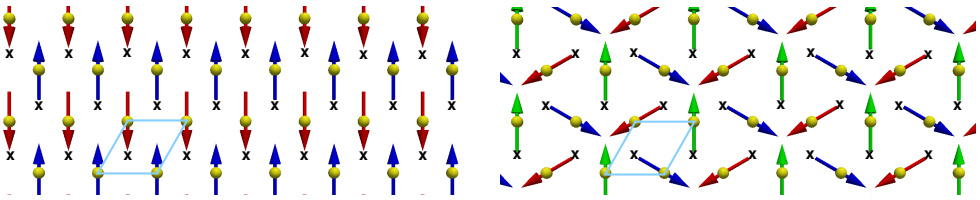

Figure 6.7: Simple magnetic states on the hexagonal ML. Arrows indicate the orientations of the Mn spins, while the crosses highlight one of the possibilities for the position of the substrate atoms. Left: row–wise AFM state, associated with the M point. Right: 120° AFM state, associated with the K or K' points. The rotational sense will depend on the specification of $\hat{n}_3$ (consult Appendix C for details). The chemical unit cell is highlighted in light blue.

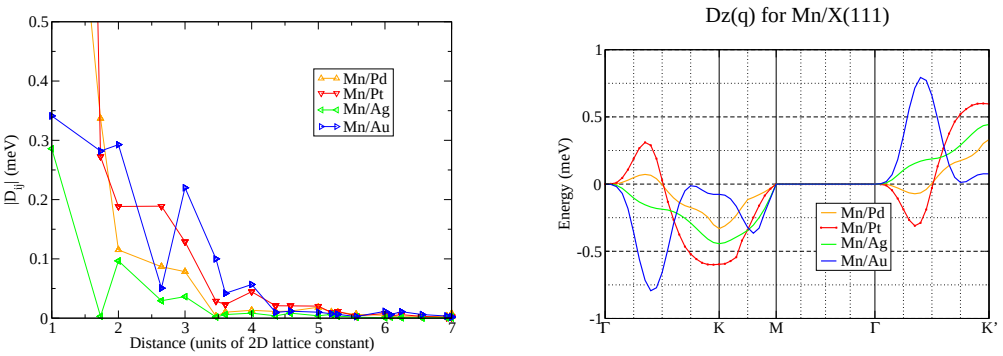

Figure 6.8: Left: length of the DM vectors in real space as a function of distance. The nn DM vector is much larger than all the others for Mn/Pd ($D = 1.2$ meV) and Mn/Pt ($D = 4.3$ meV). Right: the Fourier transformed z–component of the DM interaction. Departure from a simple sine–like form indicates that many shells of atoms are important in determining this particular unidirectional contribution, for all four substrates. See Appendix C for more details.

almost identical energy dispersions, favouring the 120° AFM state (Fig. 6.7, right). An extension of the J_1–J_2 model to include the third shell of neighbours shows that this additional interaction can modify considerably the magnetic phase diagram, see for instance Ref. [107]. Relying on the Fourier transformed quantities, as it is being done here, circumvents the need to determine the most important interactions and the model which best represents the data.

The anisotropic part of the tensor is dominated as before by the DM–type interaction. The nn DM vector is large for the Pd ($D = 1.2$ meV) and Pt ($D = 4.3$ meV) substrates, and relatively weak for the Ag and Au substrates ($D = 0.3$ meV). The length of the DM vector is plotted as a function of distance in Fig. 6.8, left. It can be seen that only the nn vector is greatly enhanced by the Pd and Pt substrates, and that there are substantial contributions up to the fifth or sixth shell of atoms. This was already illustrated for Mn/Au in Fig. 6.6, right.

As explained in Appendix C, the 120° AFM state acquires a well–defined rotational sense, i.e., $\hat{n}_3$, if the z–component of the DM vector is non–zero. This is the case for all four systems considered, as can be seen in Fig. 6.8, right. The symmetry considerations described in Appendix C are general, and can be checked in this plot. As there is a mirror plane on all the Γ–M lines, $D_z(\vec{q}) = 0$. The dispersion on the Γ–K lines is related to that on the Γ–K' lines by simply reversing the sign of $D_z(\vec{q})$. Competing contributions from different shells of atoms make $D_z(\vec{q})$ relatively small for Mn/Au at the K and K' points, in comparison with all other systems. Up to an

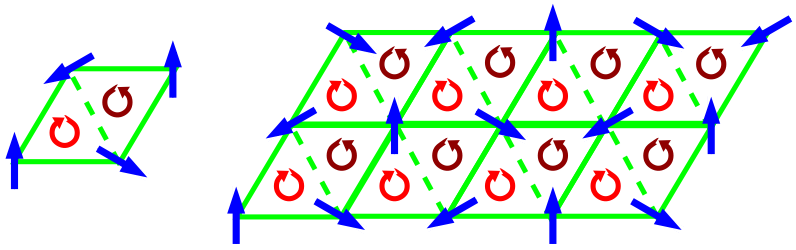

Figure 6.9: Left: a single rhombus with the two chiralities marked, according to Eq. 6.23 (see accompanying text also). Right: chirality pattern generated by repetition of the basic tile.

arbitrary global phase angle, the spins follow the expressions ($\vec{R}_i = i_1\vec{a}_1 + i_2\vec{a}_2$):

$$K \equiv \left(\frac{2}{3}, \frac{1}{3}\right): \quad \vec{m}_i = \cos\frac{2\pi}{3}(2i_1 + i_2)\,\hat{e}_x + \sin\frac{2\pi}{3}(2i_1 + i_2)\,\hat{e}_y \tag{6.21}$$

$$K' \equiv \left(\frac{1}{3}, \frac{2}{3}\right): \quad \vec{m}_i = \cos\frac{2\pi}{3}(i_1 + 2i_2)\,\hat{e}_x - \sin\frac{2\pi}{3}(i_1 + 2i_2)\,\hat{e}_y \tag{6.22}$$

These generate exactly the same magnetic pattern, and correspond to Fig. 6.7, right, with $\hat{e}_z$ pointing towards the reader (after an overall rotation of 90°). The weak in–plane anisotropy will then favour orientations either along the nn directions or perpendicular to them; this was not extracted from the first–principles data, as the numbers were below the numerical uncertainty.

One can also associate a chirality pattern with the chemical unit cell, see Fig. 6.9, left. Consider the single rhombus depicted on the left, split into two triangles. Travelling anticlockwise on each of those triangles, define the chirality vector as

$$\vec{\xi} = \vec{m}_1 \times \vec{m}_2 + \vec{m}_2 \times \vec{m}_3 + \vec{m}_3 \times \vec{m}_1 \tag{6.23}$$

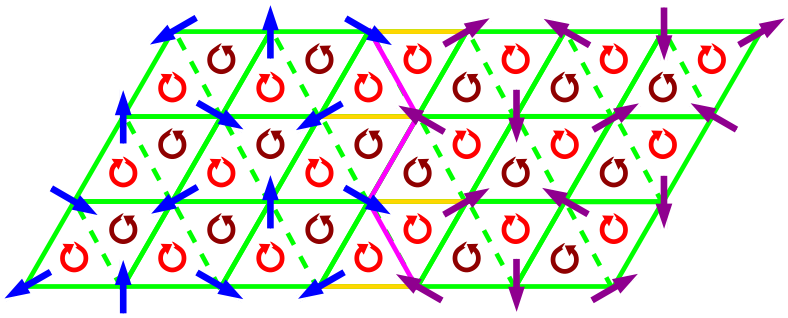

Figure 6.10: A possible interface between two chirality patterns. The bond marked yellow loses energy, while the bond in magenta gains energy.

where the numbers label the magnetic atoms in the order they are encountered while travelling around each triangle. An anti–clockwise symbol is assigned if $\vec{\xi} \cdot \hat{e}_z > 0$, and vice–versa. The resulting chirality pattern is shown on Fig. 6.9, right. The total or net chirality is zero, as might be expected.

Each rhombus can be assigned one of two chirality patterns, giving rise to two types of rhombus. Suppose that two magnetic domains are constructed out of each type of rhombus, see Fig. 6.10. If these are brought into contact, there will be a competition between the loss in isotropic exchange energy at the interface, and the gain in anisotropy energy by preserving the chirality pattern on each domain, and by aligning the moments along the easy axes. The sharp domain wall depicted here is purely schematic; in reality this might deform into some kind of Bloch or Neél wall, spanning more than two atoms.

The idea is to illustrate how the magnetic moments on different domains might appear at an angle to each other, from the point of view of an STM experiment, as in Ref. [114]. The colours chosen for the clockwise and anti–clockwise rotations of the moments on each triangles are not arbitrary (compare with the colours in Fig. 6.4): the conjecture is that the same chirality applies when the magnetic atoms surround a substrate atom, correlating with the alternating sign of D_z as one goes around the central atom; the position of the substrate atoms in fcc and hcp islands is related by a mirror symmetry along a nn direction, reversing the sign of D_z in all pairs (see Appendix C). Then each type of island, fcc or hcp, would have a different chirality pattern. Thus when going from an fcc island (Fig. 6.4, left) to an hcp island (Fig. 6.4, right), there would be a domain wall between these two chirality patterns just described.

6.3.4 Role of the substrate: band filling and spin–orbit coupling

A simple picture can be invoked to synthesise these calculations: on the Ag and Au substrates, the magnetic properties of the Mn ML are quite similar to those of a conceptual free standing or unsupported ML, as found before in other electronic structure calculations [119]. This explains the similarity between the $S^{(2)}$'s for the Ag and Au substrates. The Pd and Pt substrates then represent a strong enhancement of the nn interactions due to hybridisation with the substrate, which stabilises the 120° AFM state over the row–wise AFM state found for the Ag and Au substrates.

It can be seen that the anisotropic effects scale with the atomic number of the substrate, due to the SOC effect, as expected. The DM vectors on the first shell, however, show a marked difference: they are much larger in the Pd and Pt sub-

strates than in the Ag and Au substrates, regardless of the similarity between atomic numbers. This reveals another ingredient to a large anisotropic effect: strong hybridisation with the substrate. This is why the first shell DM vector in Mn/Pd is three times larger than the one in Mn/Au, despite the differences in atomic number, and so in SOC strength. This implies that a careful choice of substrate might yield similar anisotropic properties using lighter elements, thus dispensing the use of their heavier counterparts, which may be rarer and more expensive.

6.3.5 Comments on the Mn_1/Ag(111) system

We suggest two explanations for the disagreement between theory and experiment. Both are based on the delicate balance found between the two types of AFM:

- There may be some buried Mn atoms underneath the studied Mn triangular islands. This would enhance the hybridisation to the substrate, which has a very strong influence on the strength of the nn couplings, as found for the Pd and Pt substrates, and thus favours the 120° AFM state.

- There may be an inadequate treatment of the exchange–correlation effects in the Mn ML. Although this might be a small correction, the Mn_1/Ag(111) system, as described by the LDA, has a small energy difference between the two types of AFM states just discussed. The electronic structure changes that may be brought about by an improved treatment might be enough to tilt the preference from the row–wise to the 120° AFM state.

The experimental data on Mn_1/Ag(111) also suggests that similarly oriented magnetic islands with respect to the substrate have different magnetic domains, and so different magnetic anisotropies. We propose that this might be due to two different chirality patterns of the 120° AFM state, as explained in detail previously.

Chapter 7

Other applications and preliminary results

In this chapter other work and preliminary data is collected and explained. Firstly, the results of a theoretical investigation on the magnetic properties of MnSb are discussed. These are part of a join theoretical–experimental effort at the University of Warwick to characterise epitaxially stabilised MnSb polymorphs, particular the half–metallic properties of the zinc–blende structure.

Next a more detailed analysis of the properties of the intermediate magnetisation states between the FM and the DLM state is given for hcp and fcc Co. The feedback between the degree of magnetic order and the length of the local spin moment is exposed, and the applicability of an interpretation in terms of a conventional Heisenberg model is discussed.

To conclude we present data on the calculated first–principles mean–field magnetisation profile of an FeRh thin film. This will illustrate that the problem of finding the equilibrium magnetisation profile for an inhomogeneous system solves itself on the course of the self–consistent iterations, consistent with the supplied type of magnetic order — FM, AFM, etc.

7.1 Half–metallicity in zinc–blende MnSb

The half–metallic (HM) state (the majority spin DOS is metallic, and the minority spin DOS has a gap around the Fermi energy) was first predicted in electronic structure calculations for NiMnSb [122]. Since the original finding many other materials were predicted to be HM, including other Heusler alloys [123]. One quantity

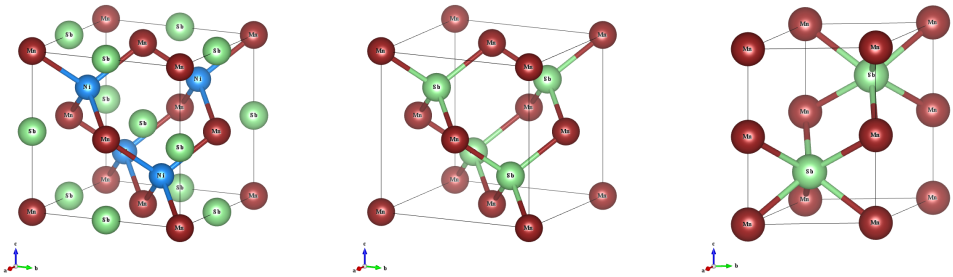

Figure 7.1: Crystal structure of the half–Heusler NiMnSb (left), the zb MnSb (middle), and the double hcp MnSb (right). Mn atoms are in dark red, Sb atoms in light green, and Ni atoms in light blue.

of interest is the intrinsic spin polarisation at the Fermi level,

$$p(\varepsilon_F) = \frac{n_\uparrow(\varepsilon_F) - n_\downarrow(\varepsilon_F)}{n_\uparrow(\varepsilon_F) + n_\downarrow(\varepsilon_F)} \tag{7.1}$$

with $n_\uparrow(\varepsilon_F)$ the DOS for the majority spin channel and $n_\downarrow(\varepsilon_F)$ the DOS for the minority spin channel. The predicted 100% spin polarisation at the Fermi energy has proved very hard to detect experimentally in many of those systems: extrinsic effects such as impurities and the difference between the bulk and the surface or interface properties, where measurements are usually made, but also intrinsic effects, such as temperature or strong correlation effects, conspire to cloud the data, or to destroy the complete polarisation [123]. A tell–tale sign of HM in electronic structure calculations is an integer value of the spin moment in otherwise metallic systems, for which this is very rarely the case.

The experimental interest in possibly HM MnSb concerns spintronics applications, where this material could play the role of a spin injector. It is also compatible with common semiconductors such as GaAs, forming clean and sharp interfaces [13]. The zinc–blende (zb) MnSb was predicted theoretically to be HM, but also that it should be unstable to a tetragonal distortion, while the true equilibrium structure is of NiAs type, or double hcp, and the HM is lost [124].

The structures of zb MnSb and cubic NiMnSb are closely related (see Fig. 7.1). The HM and the minority spin gap were shown to originate from the tetrahedral bonding of Mn with its neighbours. In NiMnSb Mn bonds tetrahedrally with Ni (while Sb forms octahedra around Mn), while in MnSb Sb plays the role of Ni in NiMnSb. For comparison the stable structure of MnSb is also shown on the same figure; no tetrahedral bonding occurs, and no HM is present.

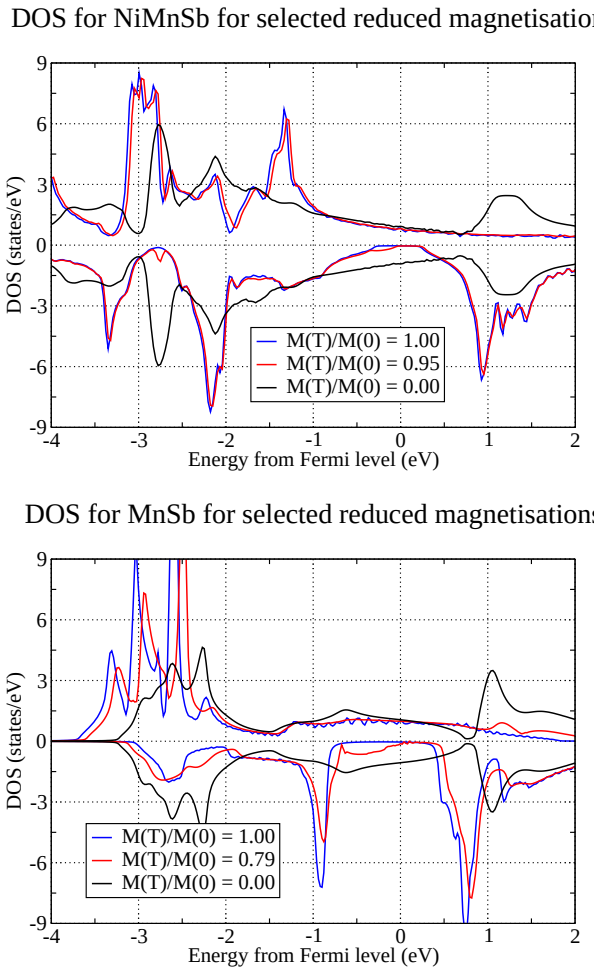

Figure 7.2: Comparison between the total DOS for NiMnSb (top) and zb MnSb (bottom) for selected reduced magnetisations (see text). The majority spin DOS is plotted as positive values, and the minority spin DOS as negative values.

The KKR calculations were done using the experimental lattice parameters determined for the MnSb polymorphs grown on GaAs. For zb MnSb $a = 6.50$ Å, while for the dhcp MnSb $a = 4.12$ Å, $c = 5.77$ Å. The alloy analogy described in Chapter 4 was used to investigate the evolution of the electronic structure as the relative magnetisation is reduced. In this model the atoms with a local moment (which are taken to be only Mn), form a binary alloy, of concentration $c_\uparrow$ for Mn atoms parallel to the magnetisation direction, and of concentration $c_\downarrow$ for Mn atoms antiparallel to it, with $c_\uparrow + c_\downarrow = 1$. The reduced magnetisation, $M(T)/M(0)$, is then obtained by taking the ratio of the total magnetic moment for each alloy composition with the value obtained for the FM case, $c_\uparrow = 1$; the temperature T is unknown. For

comparison with Ref. [14], calculations for NiMnSb were also performed, using their value of $a = 5.9$ Å.

Some results for the DOS are shown in Fig. 7.2. The top panel pertains to NiMnSb. The FM minority DOS ($M(T)/M(0) = 1$) shows a small gap around the Fermi energy, E_g, of about 0.5 eV. The empty Mn minority d states can be seen 1 eV above the Fermi energy, ε_F, while the majority d states are between 3 eV and 2 eV below ε_F, which an associated minority peak. These majority states are hybridised with the d states of Ni, which can be considered as being magnetised by Mn. The Mn spin moment then reduces from the ideal value of $4\,\mu_\mathrm{B}$ to $3.75\,\mu_\mathrm{B}$, while Ni picks up a spin moment of $0.25\,\mu_\mathrm{B}$, so that overall the spin moment is integer. The spin gap is filled with states very quickly as the magnetisation is reduced: at $M(T)/M(0) = 0.95$ these states are already at ε_F, and are the explanation for the fast drop in $p(\varepsilon_F)$, which is the parameter of interest for applications (see Fig. 7.3). The DOS for the DLM state is shown for completeness; no sign of the spin gap remains, by symmetry.

The bottom panel of Fig. 7.2 shows the DOS for zb MnSb. Let us consider the FM DOS first. The crucial difference is that the hybridisation is no longer with the d states of Ni, but with the p states of Sb. The conduction band is narrower than in NiMnSb, and the Mn d states more localised, as can be seen by the narrow peaks between 4 eV and 2 eV below ε_F (the majority states) and 0.5 eV above ε_F (the minority states). The peak about 1 eV below ε_F is from Sb; it polarises antiparallel to Mn, with a spin moment of $0.2\,\mu_\mathrm{B}$. Mn has a spin moment of $4.2\,\mu_\mathrm{B}$ in this system, so that the net spin moment is again $4\,\mu_\mathrm{B}$.

MnSb in the zb structure is much more robust against loss of complete spin polarisation at ε_F than NiMnSb. The large gap ($E_g \approx 1.1$ eV) and the absence of d states close to the Fermi level seem to be the key ingredients. As shown in Fig. 7.2, the states in the spin gap approach the Fermi level only for $M(T)/M(0) \approx 0.80$. The peak positions appear shifted due to the change in the Fermi energy between the two calculations.

Collecting $p(\varepsilon_F)$ for NiMnSb and MnSb as a function of the reduced magnetisation, we obtain Fig. 7.3, which should be compared with the plot given in Ref. [14] for NiMnSb. The behaviour of the polarisation at the Fermi level is easy to understand: it is saturated until some electronic states finally appear just below ε_F, and then it drops much faster than the total spin moment itself, as a small addition of states at the Fermi energy has little effect on the total spin moment, while causing the collapse of the polarisation.

These results are part of a joint paper with the experimentalists at the University

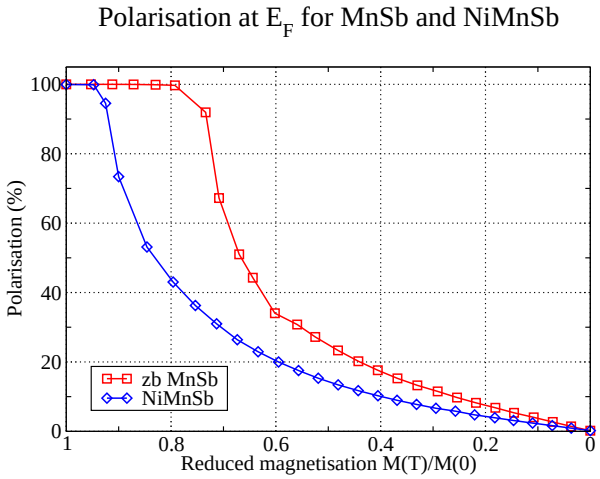

Figure 7.3: Polarisation at the Fermi energy as a function of the reduced magnetisation for NiMnSb and zb MnSb, after Eq. 7.1.

of Warwick who discovered the epitaxial polymorphs of MnSb, which has been submitted for publication, and form the theoretical background for the claim of HM and high spin polarisation at the Fermi level at high temperatures. This is due to the fact that $M(T)/M(0) \approx 0.80$ gives T not far from the Curie temperature, T_c. For NiMnSb, the experimental value is $T_c = 730$ K, while the mean–field estimate points to $T_c = 1290$ K, after Eq. 5.77. For MnSb the theory gives a similar value, $T_c = 1250$, so that a high experimental T_c is anticipated.

7.2 Bulk Co between the FM and the DLM state

In the previous section the simple alloy analogy was used to construct an effective medium mimicking the desired partially magnetised states. This is an interpolation scheme between the FM and the DLM states. The alloy analogy naturally reverts to the FM state (one component limit), and it also reproduces the DLM state correctly. The equivalence between the alloy analogy and the continuous distribution of local moment orientations in the DLM state was shown before [66], and is a consequence of isotropy of the paramagnetic state. The systems considered so far possessed good local moments (the size of the spin moment was fairly independent of the magnetic state), and so it is fair to enquire how does the theory perform, in an approximate way, for a system which does not satisfy the good local moment condition.

For this purpose bulk Co was chosen for the test calculations. Its zero temperature crystal structure is hcp, but the fcc structure has only slightly higher energy. The nn distance was taken to be 2.51 Å, which corresponds to the experimental $a_{\rm hcp}$

parameter. The hcp lattice was assumed to be ideal, $c/a = \sqrt{8/3}$, and for the fcc we have $a_{\mathrm{fcc}} = \sqrt{2}\, a_{\mathrm{hcp}}$, so that the nn distance is the same for both structures.

DOS for the FM and DLM states of hcp and fcc Co are shown in Fig. 7.4. The stability of FM hcp Co over FM fcc Co can be attributed to the details of the minority DOS around the Fermi level: there is a peak at ε_F for fcc Co, while for hcp Co ε_F is in a valley, which lowers the total energy. The spin moment is slightly lower for FM hcp Co, $1.58\,\mu_{\mathrm{B}}$, than for fcc Co, $1.63\,\mu_{\mathrm{B}}$. The difference is larger for the DLM states: $1.13\,\mu_{\mathrm{B}}$ for hcp Co and $0.94\,\mu_{\mathrm{B}}$ for fcc Co. This 60–70% reduction of the magnitude of the spin moment from the FM to the DLM state suggests that there is a feedback effect: the spin moment raises from its value on the DLM state by responding to the partial state of magnetisation, until it saturates when the FM state is reached.

We carry out the calculations for intermediate values of the order parameter by modifying the RSA picture, which should apply to good local moment systems. In the RSA we require the potential functions to be independent of the local moment orientations, except for the explicit orientation of the effective magnetic field:

$$V_i(\vec{r};\hat{e}_i) = V_i(\vec{r}) + B_i(\vec{r})\,\vec{\sigma}\cdot\hat{e}_i \quad . \tag{7.2}$$

This implies that the charge and magnetisation densities must also be independent of the orientations, in the way just described,

$$\begin{aligned} n_i(\vec{r};\hat{e}_i) \equiv n_i(\vec{r}) &= -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\hat{e}_i\, P_i(\hat{e}_i)\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\langle\, G_{ii}(\vec{r},\vec{r};\varepsilon)\,\rangle_{\hat{e}_i} \\ m_i(\vec{r};\hat{e}_i) \equiv m_i(\vec{r}) &= -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\hat{e}_i\, P_i(\hat{e}_i)\int \mathrm{d}\varepsilon\, f(\varepsilon)\,(\beta\vec{\Sigma}\cdot\hat{e}_i)\,\langle\, G_{ii}(\vec{r},\vec{r};\varepsilon)\,\rangle_{\hat{e}_i} \end{aligned} \tag{7.3}$$

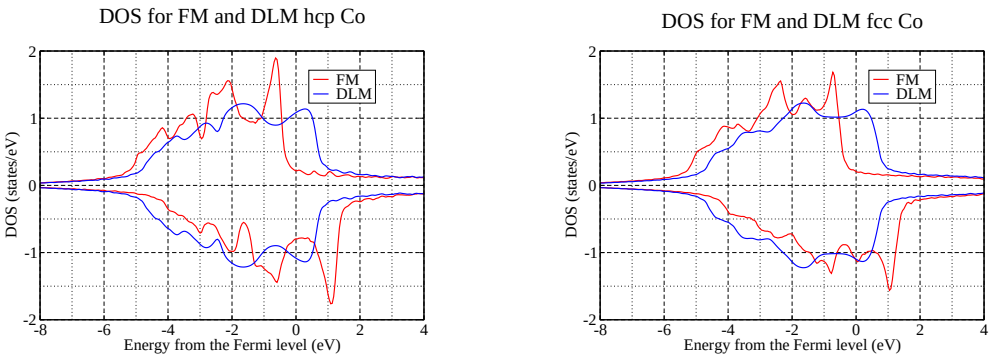

Figure 7.4: Total DOS for hcp and fcc Co in the FM and DLM states. The majority spin DOS is plotted as positive values, and the minority spin DOS as negative values.

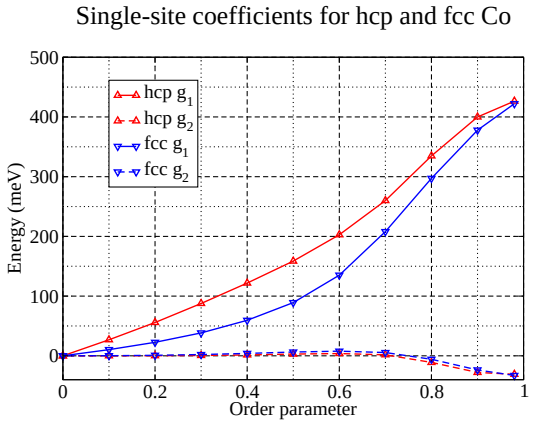

Figure 7.5: Expansion coefficients for the single–site functions for hcp and fcc Co (see Eq. 7.5). Linear dependence of g_1 on the order parameter corresponds to the mean–field approximation to the Heisenberg model, Eq. 5.5.

which gives $\vec{m}_i(\vec{r};\hat{e}_i) = m_i(\vec{r})\,\hat{e}_i$, and the connection between the densities and the potential functions is consistent with the LDA. This reproduces the averaged charge density, and so the correct number of valence electrons:

$$\langle\, n_i(\vec{r}\,)\,\rangle = \int \mathrm{d}\hat{e}_i\, P_i(\hat{e}_i)\, n_i(\vec{r};\hat{e}_i) = n_i(\vec{r}\,)\ , \quad Q_i = \int \mathrm{d}\vec{r}\, n_i(\vec{r}\,) \tag{7.4}$$

which motivates the choice of the averaging function, $P_i(\hat{e}_i)$.

For a FM order parameter, the partially averaged quantities depend only on the angle θ_i between the local moment orientation, $\hat{e}_i$ and the order parameter, $\langle\,\hat{e}_i\,\rangle$, which will be oriented along the z–direction in all sites (the magnetic anisotropy is very small in bulk Co, so it can be neglected here). The general spherical harmonic expansion used before then reduces to an expansion in Legendre polynomials in $\cos\theta_i$ (cf. Appendix A), and all the expressions simplify; the order parameter is then just a scalar, $\langle\,\cos\theta_i\,\rangle$.

The single–site functions (Eq. 5.35) that determine the probability distribution $P_i(\hat{e}_i)$ become

$$g_i(\hat{e}_i) = \sum_\ell g_{i\ell}\, P_\ell(\cos\theta_i) = g_{i0} + g_{i1}\cos\theta_i + g_{i2}\,\frac{3\cos^2\theta_i - 1}{2} + \ldots \tag{7.5}$$

(recall that constants do not affect the form of $P_i(\hat{e}_i)$).

One can also analyse the partially averaged single–site GF being used to construct the densities. This provides some information on how the magnetic moment changes

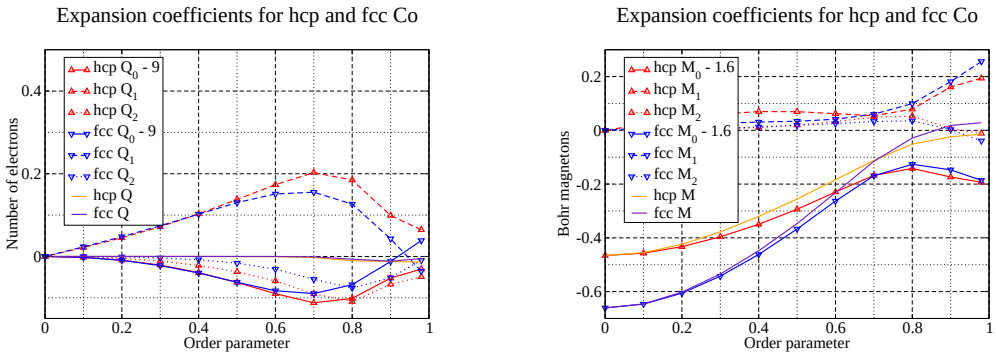

Figure 7.6: Expansion coefficients for the valence charge and spin moment for hcp and fcc Co (Eq. 7.8). See text for a detailed discussion.

as a function of its orientation with respect to the order parameter. The number of valence electrons for a given orientation is defined as

$$Q_i(\hat{e}_i) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\vec{r}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,\langle\, G_{ii}(\vec{r},\vec{r};\varepsilon)\,\rangle_{\hat{e}_i} \tag{7.6}$$

and the spin projection as

$$M_i(\hat{e}_i) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Tr}\int \mathrm{d}\vec{r}\int \mathrm{d}\varepsilon\, f(\varepsilon)\,(\beta\vec{\Sigma}\cdot\hat{e}_i)\,\langle\, G_{ii}(\vec{r},\vec{r};\varepsilon)\,\rangle_{\hat{e}_i} \tag{7.7}$$

which specialise in the current context to

$$Q_i(\hat{e}_i) = \sum_\ell Q_{i\ell}\,P_\ell(\cos\theta_i)\ ,\quad M_i(\hat{e}_i) = \sum_\ell M_{i\ell}\,P_\ell(\cos\theta_i)\ . \tag{7.8}$$

The calculations are performed by adjusting the inverse temperature β so that a desired value of the order parameter is obtained, after each self–consistent iteration.

In Fig. 7.5 we show our preliminary results for the first coefficients in the expansion of Eq. 7.5. The g_2 coefficient, corresponding to the second Legendre polynomial, is much smaller than g_1: this would suggest that a Weiss field–like description is valid, as done for the mean–field treatment of the simple Heisenberg model. However the dependence on the order parameter is not simply linear, as would be expected from Eq. 5.5. This also appears to be the case in preliminary calculations for bcc Fe, which is usually thought of as a good local moment system; they are not reported on this thesis. The nonlinearity illustrates the feedback effects between the electronic structure and the magnetic state; in this case it corresponds to the enhancement of the Co spin moment from its DLM value, and the knock–on effect on the Weiss

field (the coefficient g_1). A good discussion of how different approximations affect the magnetic moments for itinerant systems is given in Ref. [75], through the use of spin spirals.

The coefficients for the valence electron number, after Eq. 7.8, are given in Fig. 7.6, left. The orange and purple lines show the average number of electrons, which confirms that the correct valence is reproduced (9 was subtracted from the zeroth coefficient, so that only deviations are plotted). On the right panel of Fig. 7.6 the respective coefficients for the expansion of the spin moment are given. The orange and purple lines are again the averaged value; they follow the M_0 coefficient for most of the range of order parameters, which shows that the main effect is the enhancement of the length of the spin moment. Only when approaching the FM state do the other coefficients become important. In a good local moment system M_0 should be approximately constant over the entire range shown in the plot (size of the moment independent of the magnetic state), and the other coefficients should be much smaller than it, as was argued in Chapter 4 (see discussion of the RSA).

This approximate treatment of the full statistical mechanical problem suggests that the energetics in Co are ruled by the enhancement of the length of the spin moment, until the order parameter reaches 70–80%. Then the coefficient M_0 does not change much anymore, up to the FM state. It is tempting to label this behaviour as a crossover from Stoner–type behaviour, in which the main effect is the increase of the spin polarisation, to a Heisenberg–type behaviour, in which the transverse or orientational effects become important.

Two caveats apply: the effects of the approximation described in this section should be assessed, by comparison with less approximate schemes. The influence of the longitudinal fluctuations should also be considered separately: in the present scheme the length of the spin moments is not allowed to fluctuate, but is slaved to the magnetisation state, by minimisation of the DFT total energy.

7.3 The magnetisation profile of an FeRh thin film

We now apply the approach described in the previous section to an inhomogeneous system, an FeRh film embedded in bulk V. FeRh crystalises in the CsCl structure, while V adopts the bcc structure. The CsCl structure would revert to a bcc structure if both atoms were identical; the nn distances are very similar for both FeRh and V. For our setup we neglect the small mismatch and adopt an ideal bcc structure with $a = 3.02$ Å, the experimental value for V.

The film is spanned along the body diagonal, or (111) direction, and is taken to

be symmetric around its middle. The interfacial region consists of six layers of V, followed by eight Fe and Rh layers, an Fe layer which is the middle layer, and again eight Rh and Fe layers, finished by six V layers, for a total of 45 layers. Layers 1–6 are thus V, all the following odd numbers are Fe, while even numbers correspond to Rh, until layers 40–45, which are V again. Half of this interfacial region is depicted in Fig. 7.7. The preliminary calculations were done using $\ell_{\max} = 2$ instead of the standard value of 3 used in all previous calculations.

Bulk FeRh is well–known for a phase transition from its AFM ground state at low temperature to a FM state at high temperature [125]. The AFM state can be thought of as FM planes antiparallel to each other along the (111) direction, which is the justification for the choice of orientation in the film calculation. The FM and the AFM states can then be computed using the same structure, by just aligning consecutive Fe layers parallel or antiparallel to each other.

In the following the focus will be on the AFM configuration between the fully magnetised and the DLM state. The orientation of the magnetisation on each layer is fixed to be antiparallel between consecutive Fe layers, but the value of the order parameter is allowed to relax. This allows an approximate solution of the magnetic equation of state, the set of coupled equations given by Eqs. 5.19 and 5.20.

First some data pertaining to the fully magnetised FM and AFM configurations is briefly discussed. The Fe layers away from the interfaces with V have spin moments of $3.15\,\mu_B$ for the FM and $3.13\,\mu_B$ for the AFM case, while the Rh layers sandwiched by these Fe layers have an induced spin moment of $0.95\,\mu_B$ in the FM case only. The four Fe layers next to the two interfaces with V deviate from the bulk values. In the FM case, the layer in atomic contact with V shows a reduced spin moment of $2.44\,\mu_B$, which increases progressively to $2.98\,\mu_B$, $3.12\,\mu_B$ and $3.14\,\mu_B$. The Rh layers go from $0.62\,\mu_B$ to $0.88\,\mu_B$ and $0.95\,\mu_B$, again starting from the interface. In the AFM case the Fe spin moments go as $2.37\,\mu_B$, $2.94\,\mu_B$, $3.11\,\mu_B$ and $3.13\,\mu_B$. The Rh layers near the interface also display induced spin moments of $0.15\,\mu_B$, $0.04\,\mu_B$ and $0.01\,\mu_B$, due to the broken symmetry, which generates a net magnetic field on the Rh layers from the neighbouring antiparallel Fe layers, as their spin moments are different. The DLM Fe spin moments are essentially identical to the ones for the AFM setup, and there is no induced spin moment on any Rh layer in this case. The

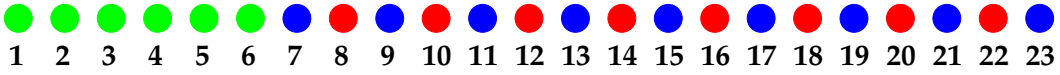

Figure 7.7: Left half of the interface region for the symmetric FeRh film embedded in V. Blue circles depict Fe layers, red circles Rh layers, and green circles V layers.

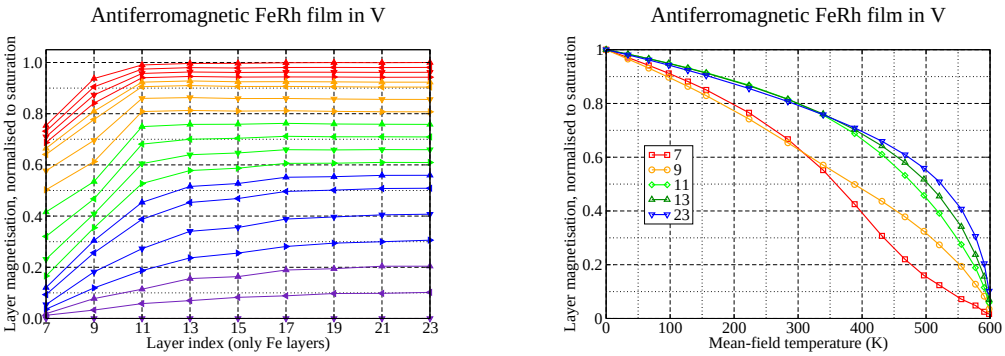

Figure 7.8: Left: average magnetisation on each Fe layer, normalised to the value of the central layer at $T = 0$ K. Only half of the layers is shown, as the film is symmetric. The central layer (23) was used as reference layer when performing the calculations. Right: the average magnetisation on selected layers, normalised to the value of each layer at $T = 0$ K, plotted versus the mean–field temperature derived from the central layer. For numbering of layers consult Fig. 7.7.

V layer next to the first Fe layer is polarised antiparallel to it, with $0.54\,\mu_B$ for the FM, $0.42\,\mu_B$ for the AFM case, and no spin moment for the DLM case. The other V layers also acquire a small spin polarisation, which quickly decays away from the interface.

To generate the magnetic state for an intermediate value of the order parameter, the value of the inverse temperature β is adjusted, such that the order parameter for the middle Fe layer (number 23) has a prescribed value. This is an effective constraint on the whole magnetic state of the system: the single–site functions for each layer, Eq. 5.35, which determine the respective probability functions $P_i(\hat{e}_i)$, and so the order parameter on each layer, relax with the electronic structure. Thus by the end of the self–consistent iterations, both the electronic structure and the magnetisation profile of the film are determined.

The magnetisation profile for the Fe layers in AFM FeRh film is shown in Fig. 7.8, left. The average magnetisation is given by the length of the averaged spin vector, and it is normalised to the value of the spin moment in the fully magnetised state for the central layer, number 23. Thus the topmost curve is just a plot of the values discussed before for the AFM state. It provides visual confirmation that bulk behaviour starts after the first three or four Fe layers. These interior layers start deviating from the bulk behaviour only when the normalised average magnetisation falls below 0.65, so that when it is down to 0.10 only two Fe layers on each side of the central Fe layer still appear to follow the bulk behaviour.

Fig. 7.8, right plots the average magnetisation, this time normalised to the fully

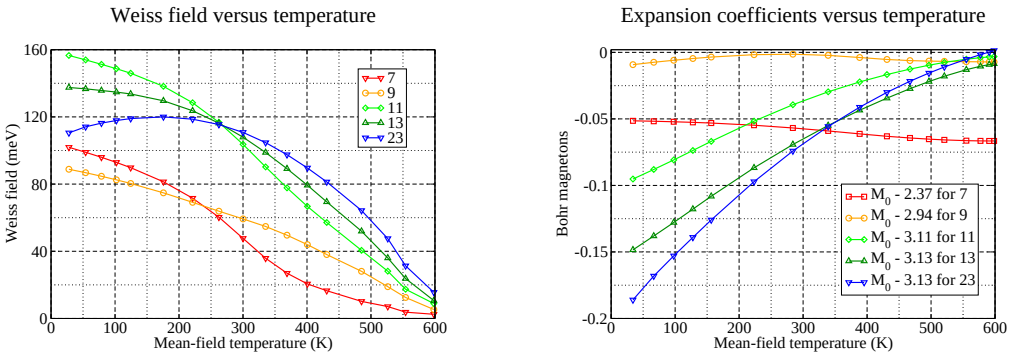

Figure 7.9: Left: Weiss field (the coefficient g_1, see Eq. 7.5) for selected layers. Right: the M_0 coefficient of the averaged spin moment, see Eq. 7.8. For numbering of layers consult Fig. 7.7.

magnetised AFM value of each layer, against the mean–field temperature extracted from constraining the order parameter on the central layer. This should be compared with the form expected from the Langevin function, $L(x) = \coth x - 1/x$. All the curves go to zero at about 600 K, which is the mean–field estimate for the transition temperature of the thin film. For reference, the experimental bulk transition temperature is $T = 670$ K, for the high temperature FM state [125]. The curve for layer 7, which is the Fe layer next to V, is similar to the mean–field solution in the presence of a finite external magnetic field. The role of the external field is played by the magnetisation of the interior of the FeRh film, which prevents the outermost Fe layers from becoming paramagnetic before the rest of the film does. Layer 9 displays similar behaviour. It is also possible that FM instead of AFM coupling between some of the layers would be energetically favourable; these calculations have not been performed so far.

The Weiss field acting on selected layers is plotted in Fig. 7.9, left. The middle layer shows a non–monotonic dependence on the mean–field temperature; this does not imply a non–monotonic dependence of the magnetisation, as can be confirmed in Fig. 7.8, right. As the Weiss field depends on the state of magnetisation of all layers (cf. Eq. 5.5, for instance), no simple interpretation appears possible. To illustrate the good local moment behaviour, Fig. 7.9, right displays the M_0 coefficient of the averaged spin moment, see Eq. 7.8. It should be noted that the range of values in this plot is much smaller than the total spin moments in the saturated AFM state, which were subtracted out. In layers 7 and 9 M_0 is almost constant, which was proposed before to be an indicator of good local moment behaviour. This is due to the presence of the interface, which reduces the hybridisation in those two layers.

The same physical mechanism applies for the magnetic monolayers discussed in the previous chapter; the reduced number of neighbours of the same kind of atom favours the intra–atomic exchange (Hund's rules) over the delocalisation energy gained by hybridisation with the neighbours.

The complex behaviour observed in this thin film is not unexpected: even a simple Heisenberg model with different coupling strengths between different layers would generate non–trivial behaviour [126]. The advantage of the present approach is that all quantities are computed on the same footing; the only input parameters are the structure and the desired type of magnetic order. For collinear magnetic states the only possible deviation from the initial magnetic order is by a spin flip; in the present context the Weiss field would have to change sign. If the electronic structure is such that instead of a sign change only a reduction of the value occurs, then the magnetic configuration is stationary (although it might not be the ground state). Direct comparisons would only be possible if the free energy of each configuration was available; this should include contributions from the other layers also, and so far it has not been implemented.

Chapter 8

Outlook

In Chapter 5 we presented a new theory of the magnetic properties of a heterogeneous system, firmly rooted in DFT, and thus first–principles. The GF formulation allowed for the natural incorporation of the statistical mechanics of the collective degrees of freedom, identified as local magnetic moments, and the main approximation was the single–site variational approach. The ensuing approximations (such as the form of the single–site potentials) were needed for computational reasons, and so can be lifted in due course. It was shown that an understanding of the magnetic properties could be given in terms of magnetic interactions between these local degrees of freedom, as extracted from the underlying electronic structure, using linear response to compute the spin–spin correlation function, or $S^{(2)}$.

There is still plenty of scope for extensions and generalisations of the theory. The single–site approximation can be lifted using the non–local coherent potential approximation (CPA), briefly outlined in Chapter 3; magnetic short–range order would then be naturally incorporated. A less demanding and also well–tried route is to introduce the Onsager cavity field, explained at the end of Chapter 5. Temperature plays a secondary role in the present theory: the prescription of the magnetic state through the values of the order parameters is the essential point. If the statistical mechanics of the longitudinal magnetic fluctuations are to be considered, then the mean–field temperature will take a leading role: it will control the equilibrium magnitude of the local moments, which will no longer be determined from the usual minimisation of the DFT total energy. As pointed out by Kakehashi [67], the conventional CPA used in this thesis is the static limit of dynamical mean–field theory: a dynamical extension of the theory can also be envisaged.

The usefulness of the $S^{(2)}$ concept was illustrated in Chapter 6. The paramagnetic disordered local moment (DLM) state was chosen as reference state, in contrast

to most other first–principles calculations, which are based on fully ordered magnetic states (simple or of spin spiral form). The latter approach can bias the results: magnetic interactions derived from the ferromagnetic (FM) state may be very different from those computed in some antiferromagnetic (AFM) state [77], or from the DLM state [127]. The pivotal role of the Dzyaloshinsky–Moriya interaction in establishing chiral magnetic states on the Mn monolayers was exposed [12]; this was due to the ability to handle the anisotropic effects generated by spin–orbit coupling on par with the conventional isotropic interactions. Simple models inspired by the information extracted from the first–principles calculations were also shown to be quite helpful. The material presented in Appendix B shows how the lattice Fourier transform of the $S^{(2)}$ can be used to compute the energetics of a perturbation with a spin spiral form. This is fairly general and should have wide applicability.

Chapter 7 focused on the connection between the electronic structure and the degree of magnetic order, as given by a simple interpolation version of the general theory developed in Chapter 5. The experimental finding by researchers at the University of Warwick of the zincblende MnSb in epitaxial films [13] led to an investigation of its predicted half–metallic properties. The spin polarisation at the Fermi level, $p(\varepsilon_F)$, remains saturated until the magnetisation is reduced to about 80% of its FM value; this is in contrast with the behaviour of $p(\varepsilon_F)$ in NiMnSb [14], for which the threshold is at 95%. Given the typical shape of the magnetisation curve as a function of temperature, this implies $p(\varepsilon_F) \approx 1$ for a much larger thermal range in MnSb than in NiMnSb, with potential benefits for spintronics applications.

The calculations for hcp and fcc Co were intended to test the limitations of the current approximate scheme. The main effect found was a reduction of the local moment from $1.6\,\mu_B$ in the FM state to about $1\,\mu_B$ in the DLM state. The linear dependence of the Weiss field on the order parameter expected from the mean–field theory of the simple Heisenberg model, Eq. 5.5, was not found: the magnetic interactions in the system depend on the magnitude of the local moment. Its angular dependence, measured with respect to the orientation of the FM magnetisation, revealed two different regimes: up to about 70% of the FM magnetisation the local moment is fairly independent of the orientation, but increases from its DLM value; then it becomes dependent on the orientation, but its magnitude increases only slightly, up to the FM value. This suggested an interpretation in terms of a crossover from a Stoner–type behaviour, where the dominant effect is the enhancement of the local moment length, to a Heisenberg–type behaviour, where the orientations of the local moments become the important parameters. This is a system for which it would be interesting to consider the effects of the thermally excited longitudinal

fluctuations, going beyond the current description.

The final example in Chapter 7 takes an FeRh film embedded in bulk V to demonstrate how the magnetic properties of a heterostructure can be computed, from the low temperature AFM state to the high temperature DLM state. The feedback between the electronic structure and the imposed magnetic state automatically determines the equilibrium magnetisation profile, during self–consistency, revealing the stabilisation of magnetism at the film interfaces by the bulk magnetisation of the film. It should be noted that even a simple mean–field theory based on the Heisenberg model becomes unwieldy, as different interaction parameters apply between different Fe layers; the present approach circumvents the need for such considerations. The actual equilibrium magnetic structure of the film was not determined, as only a simple imposed AFM order was considered; bulk FeRh is famous for its high–temperature transition from an AFM to an FM state [125], and how this picture is modified in a film of finite thickness will be left for future investigations.

The versatility of the present theory should be emphasised. Not only it allows detailed access to the magnetic interactions in a given system, through the $S^{(2)}$ calculations, but it can also solve an approximate magnetic equation of state, by imposing a given type of magnetic order, for a heterogeneous system. Both approaches are complementary; for instance, an instability of the magnetic reference state can be detected by the $S^{(2)}$'s at some intermediate stage between the fully ordered and the paramagnetic state. When there are many non–equivalent magnetic atoms, as in a heterostructure, the number of parameters provided by the $S^{(2)}$ might constitute too much information; here the equation of state approach allows for a simpler way of exploring the magnetic phases. Large systems can be considered, as the SKKR method scales linearly with system size (cf. Chapter 2).

Combined approaches with spin dynamics or Monte Carlo (MC) simulations are very appealing. These would allow a determination of the true magnetic ground state, without the need to impose its form, as done so far. It might also mitigate some of flaws of the single–site mean–field approximation: the single–site probability functions $P_i(\hat{e}_i)$ could be imported from, say, a MC simulation, and used in setting up the electronic structure. A realistic temperature dependence of the magnetic properties would also be obtained. The dependence of the magnetic interactions on the magnetic state would be available to the MC simulations; the most frequent approach is to compute the magnetic interactions once and for all from a chosen reference state and for some spin Hamiltonian, and assume they will be valid for the entire temperature range covered by the simulation [86].

To conclude, consider the wealth of experimental data available for magnetic

thin films and heterostructures [128], driven by fundamental as well as technological interest. Most of what is known about the finite temperature properties of such large systems is model based. First principles–theories of magnetism in these systems are still mostly restricted to conventional ground state DFT, or to fairly small system sizes. There are many opportunities for theoretical developments in this area, as well as fruitful collaborations with experimental groups. I hope the theoretical and computational developments presented in this thesis will make a useful contribute to address these questions, and spur new investigations.

Appendix A

Spherical harmonics

Spherical harmonics are eigenfunctions of the angular part of Laplace's equation in spherical coordinates:

$$\frac{1}{\sin^2\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\,\sin\theta\,\frac{\partial}{\partial\theta}+\frac{\partial^2}{\partial\phi^2}\right)Y_\ell^m(\theta,\phi)=-\ell(\ell+1)\,Y_\ell^m(\theta,\phi) \tag{A.1}$$

and obey the following relations,

$$Y_\ell^{m*}(\theta,\phi)=(-1)^m Y_\ell^{-m}(\theta,\phi)\ \ ,\ Y_\ell^{m*}(\pi-\theta,\phi+\pi)=(-1)^\ell Y_\ell^m(\theta,\phi) \tag{A.2}$$

They form a complete orthonormal basis set,

$$\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\mathrm{d}\theta\,\sin\theta\,Y_\ell^{m*}(\theta,\phi)\,Y_{\ell'}^{m'}(\theta,\phi)=\delta_{\ell\ell'}\delta_{mm'} \tag{A.3}$$

$$\sum_{\ell m}Y_\ell^{m*}(\theta,\phi)\,Y_\ell^m(\theta',\phi')=\delta(\cos\theta-\cos\theta')\delta(\phi-\phi') \tag{A.4}$$

and are given explicitly by the formula

$$Y_\ell^m(\theta,\phi)=(-1)^m\sqrt{\frac{2\ell+1}{4\pi}}\sqrt{\frac{(\ell-m)!}{(\ell+m)!}}\,P_\ell^m(\cos\theta)\,e^{\mathrm{i}m\phi} \tag{A.5}$$

which makes use of the associate Legendre polynomials, $P_\ell^m(x)$.

In the main body of the thesis the abbreviated notation $\hat{e}$ is used in place of the pair (θ,ϕ), as any unit vector is given by its spherical angles, and the two angular momentum labels are combined in a single one, $L=(\ell,m)$, which makes all

equations easier to write. A very useful result is the addition theorem,

$$\sum_m Y_\ell^{m*}(\hat{e}')\, Y_\ell^m(\hat{e}) = \frac{2\ell+1}{4\pi}\, P_\ell(\hat{e}\cdot\hat{e}') \tag{A.6}$$

The prefactor in front of the Legendre polynomial is inconvenient for some purposes, and will lead to a redefinition of the spherical harmonics below.

Real spherical harmonics can be obtained from the standard complex spherical harmonics by a simple unitary transformation, $\underline{U}$:

$$U_{LL'} = \delta_{\ell\ell'}\, U_{mm'} \;\;,\;\; U_{mm'} = \begin{cases} (-1)^m \frac{1}{\sqrt{2}}\, \delta_{m'm} + \frac{1}{\sqrt{2}}\, \delta_{m',-m} & (m>0) \\ 1 & (m=0) \\ \frac{\mathrm{i}}{\sqrt{2}}\, \delta_{m'm} + (-1)^m \frac{\mathrm{i}}{\sqrt{2}}\, \delta_{m',-m} & (m<0) \end{cases} \tag{A.7}$$

The spherical harmonics used in all the angular expansions in the main text are real renormalised spherical harmonics, obtained from the complex ones by:

$$\mathcal{Y}_L(\hat{e}) = \sum_{L'} W_{LL'} Y_{L'}(\hat{e}) \;\;,\;\; W_{LL'} = \sqrt{\frac{4\pi}{2\ell+1}}\, \delta_{\ell\ell'}\, U_{mm'} \tag{A.8}$$

The addition theorem now reads

$$\sum_m \mathcal{Y}_\ell^m(\hat{e}')\, \mathcal{Y}_\ell^m(\hat{e}) = P_\ell(\hat{e}\cdot\hat{e}') \tag{A.9}$$

and an extra factor has to be included in the orthonormality and completeness relations:

$$\frac{2\ell+1}{4\pi}\int \mathrm{d}\hat{e}\, \mathcal{Y}_L(\hat{e})\, \mathcal{Y}_{L'}(\hat{e}) = \delta_{LL'} \;\;,\;\; \sum_L \frac{2\ell+1}{4\pi}\, \mathcal{Y}_L(\hat{e})\mathcal{Y}_L(\hat{e}') = \delta(\hat{e}-\hat{e}') \tag{A.10}$$

Any angular function can be expanded in complex spherical harmonics as

$$f(\hat{e}) = \sum_L C_L\, Y_L(\hat{e}) \;\;,\;\; C_L = \int \mathrm{d}\hat{e}\, Y_L^*(\hat{e})\, f(\hat{e}) \tag{A.11}$$

and in terms of the renormalised ones as

$$f(\hat{e}) = \sum_L C_L\, \mathcal{Y}_L(\hat{e}) \;\;,\;\; C_L = \frac{2\ell+1}{4\pi}\int \mathrm{d}\hat{e}\, \mathcal{Y}_L(\hat{e})\, f(\hat{e}) \tag{A.12}$$

The first few renormalised spherical harmonics are given by:

$$
\begin{aligned}
&\mathcal{Y}_0^0 = 1 \\
&\mathcal{Y}_1^1 = \sin\theta\cos\phi\ ,\ \ \mathcal{Y}_1^0 = \cos\theta\ ,\ \ \mathcal{Y}_1^{-1} = \sin\theta\sin\phi \\
&\mathcal{Y}_2^2 = \frac{\sqrt{3}}{4}(1-\cos 2\theta)\cos 2\phi\ ,\ \ \mathcal{Y}_2^1 = \frac{\sqrt{3}}{2}\sin 2\theta\cos\phi\ ,\ \ \mathcal{Y}_2^0 = \frac{3}{4}\cos 2\theta - \frac{1}{4}\ , \\
&\mathcal{Y}_2^{-2} = \frac{\sqrt{3}}{4}(1-\cos 2\theta)\sin 2\phi\ ,\ \ \mathcal{Y}_2^{-1} = \frac{\sqrt{3}}{2}\sin 2\theta\sin\phi
\end{aligned}
\tag{A.13}
$$

Appendix B

Spin spirals and lattice sums

This appendix is dedicated to obtain the free energy of a spin spiral from a general quadratic form, which displays the translational invariance of the underlying Bravais lattice. For example, if one considers a perturbation of the paramagnetic or ferromagnetic state of the spin spiral form,

$$\delta\vec{m}_i = \delta m\big(\cos(\vec{q}\cdot\vec{R}_i+\phi)\sin\theta\,\hat{n}_1+\sin(\vec{q}\cdot\vec{R}_i+\phi)\sin\theta\,\hat{n}_2+\cos\theta\,\hat{n}_3\big) \tag{B.1}$$

the change in the free energy per site is given by (see Eq. 5.66)

$$\delta\mathcal{F} = \frac{1}{2N}\sum_{ij}\sum_{\alpha\beta}\delta m_{i\alpha}\left(\chi^{-1}_{0,i\alpha\beta}\,\delta_{ij}-S^{(2)}_{i\alpha,j\beta}\right)\delta m_{j\beta} = \delta\mathcal{F}^{(1)}+\delta\mathcal{F}^{(2)} \tag{B.2}$$

These relations hold:

$$\frac{1}{N}\sum_i\cos(\vec{q}\cdot\vec{R}_i)=\delta(\vec{q})\quad,\quad\frac{1}{N}\sum_i\sin(\vec{q}\cdot\vec{R}_i)=0 \tag{B.3}$$

where the delta function should read as a Kronecker delta, unless $N\to\infty$. Using elementary trigonometry the following identities can be derived:

$$\frac{1}{N}\sum_i\cos^2(\vec{q}\cdot\vec{R}_i+\phi)=\frac{1+\cos 2\phi\,\delta(\vec{q})}{2} \tag{B.4}$$

$$\frac{1}{N}\sum_i\sin^2(\vec{q}\cdot\vec{R}_i+\phi)=\frac{1-\cos 2\phi\,\delta(\vec{q})}{2} \tag{B.5}$$

$$\frac{1}{N}\sum_i\cos(\vec{q}\cdot\vec{R}_i+\phi)\sin(\vec{q}\cdot\vec{R}_i+\phi)=\frac{1}{2}\sin 2\phi\,\delta(\vec{q}) \tag{B.6}$$

$$(\hat{n}_1 \cdot \hat{e}_\alpha)(\hat{n}_1 \cdot \hat{e}_\beta) + (\hat{n}_2 \cdot \hat{e}_\alpha)(\hat{n}_2 \cdot \hat{e}_\beta) + (\hat{n}_3 \cdot \hat{e}_\alpha)(\hat{n}_3 \cdot \hat{e}_\beta) = \delta_{\alpha\beta} \tag{B.7}$$

This last expression is just the definition of the dot product between $\hat{e}_\alpha$ and $\hat{e}_\beta$ in the $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$ basis (which was introduced as a right–handed orthonormal system).

The on–site susceptibility is the same on all sites, according to the initial assumption, and so the first contribution to the free energy is, using vector–matrix notation,

$$\delta\mathcal{F}^{(1)} = \frac{(\delta m)^2}{4}\Big[\sin^2\theta\,\mathrm{Tr}\,\underline{\chi}_0^{-1} + (3\cos^2\theta - 1)\,\hat{n}_3 \cdot \underline{\chi}_0^{-1} \cdot \hat{n}_3\Big] \tag{B.8}$$

including the $\vec{q} = \vec{0}$ case, as then one can choose $\theta = 0$. Pulling out the isotropic part of the on–site susceptibility,

$$\chi_{0,\alpha\beta}^{-1} = \frac{1}{3}\mathrm{Tr}\,\underline{\chi}_0^{-1}\,\delta_{\alpha\beta} + \Big(\chi_{0,\alpha\beta}^{-1} - \frac{1}{3}\mathrm{Tr}\,\underline{\chi}_0^{-1}\,\delta_{\alpha\beta}\Big) = \chi_0^{-1}\,\delta_{\alpha\beta} + \widetilde{\chi}_{0,\alpha\beta}^{-1} \tag{B.9}$$

a separation between the isotropic and the anisotropic parts of the free energy is obtained,

$$\delta\mathcal{F}^{(1)} = \frac{(\delta m)^2}{2}\left[\chi_0^{-1} + \frac{3\cos^2\theta - 1}{2}\,\hat{n}_3 \cdot \underline{\widetilde{\chi}}_0^{-1} \cdot \hat{n}_3\right] \tag{B.10}$$

To exploit the translational invariance of the $S^{(2)}_{i\alpha,j\beta}$, i.e., the fact that it will depend only on $\vec{R}_{ij}$ and not on $\vec{R}_i$ and $\vec{R}_j$ separately, the following standard tricks will be used:

$$\cos(\vec{q}\cdot\vec{R}_j + \phi) = \cos(\vec{q}\cdot\vec{R}_{ij})\cos(\vec{q}\cdot\vec{R}_i + \phi) - \sin(\vec{q}\cdot\vec{R}_{ij})\sin(\vec{q}\cdot\vec{R}_i + \phi) \tag{B.11}$$

and

$$\sin(\vec{q}\cdot\vec{R}_j + \phi) = \cos(\vec{q}\cdot\vec{R}_{ij})\sin(\vec{q}\cdot\vec{R}_i + \phi) + \sin(\vec{q}\cdot\vec{R}_{ij})\cos(\vec{q}\cdot\vec{R}_i + \phi) \tag{B.12}$$

which introduce the cosine and sine lattice sums, and are related to the symmetric and antisymmetric parts of the tensor,

$$S^{(2),\mathrm{S}}_{\alpha\beta}(\vec{q}) = \sum_j \cos(\vec{q}\cdot\vec{R}_{ij})\,S^{(2)}_{i\alpha,j\beta} \quad , \quad S^{(2),\mathrm{S}}_{\beta\alpha}(\vec{q}) = S^{(2),\mathrm{S}}_{\alpha\beta}(\vec{q}) \tag{B.13}$$

$$S^{(2),\mathrm{A}}_{\alpha\beta}(\vec{q}) = \sum_j \sin(\vec{q}\cdot\vec{R}_{ij})\,S^{(2)}_{i\alpha,j\beta} \quad , \quad S^{(2),\mathrm{A}}_{\beta\alpha}(\vec{q}) = -S^{(2),\mathrm{A}}_{\alpha\beta}(\vec{q}) \tag{B.14}$$

As was done before, it is convenient to separate out the trace of the symmetric

tensor,

$$S^{(2),\mathrm{S}}_{\alpha\beta}(\vec{q}) = S^{(2)}(\vec{q})\,\delta_{\alpha\beta} + \widetilde{S}_{\alpha\beta}(\vec{q}) \tag{B.15}$$

The second contribution to the free energy reads

$$\begin{aligned}\delta\mathcal{F}^{(2)} = -\frac{(\delta m)^2}{4}\sum_{\alpha\beta}\Big[& S^{(2),\mathrm{S}}_{\alpha\beta}(\vec{q})\sin^2\theta\,\Big((\hat{n}_1\cdot\hat{e}_\alpha)(\hat{n}_1\cdot\hat{e}_\beta)+(\hat{n}_2\cdot\hat{e}_\alpha)(\hat{n}_2\cdot\hat{e}_\beta)\Big)\\ & + S^{(2),\mathrm{A}}_{\alpha\beta}(\vec{q})\sin^2\theta\,\Big((\hat{n}_1\cdot\hat{e}_\alpha)(\hat{n}_2\cdot\hat{e}_\beta)-(\hat{n}_2\cdot\hat{e}_\alpha)(\hat{n}_1\cdot\hat{e}_\beta)\Big)\\ & + 2\,S^{(2),\mathrm{S}}_{\alpha\beta}(\vec{0})\cos^2\theta\,(\hat{n}_3\cdot\hat{e}_\alpha)(\hat{n}_3\cdot\hat{e}_\beta)\Big]\end{aligned} \tag{B.16}$$

Recognising

$$(\hat{n}_1\cdot\hat{e}_\alpha)(\hat{n}_2\cdot\hat{e}_\beta)-(\hat{n}_2\cdot\hat{e}_\alpha)(\hat{n}_1\cdot\hat{e}_\beta) = (\hat{n}_1\times\hat{n}_2)\cdot(\hat{e}_\alpha\times\hat{e}_\beta) = \sum_\gamma \varepsilon_{\alpha\beta\gamma}\,\hat{e}_\gamma\cdot\hat{n}_3 \tag{B.17}$$

where $\varepsilon_{\alpha\beta\gamma}$ is the completely antisymmetric Levi–Civita symbol for the Euclidian metric, the second line of the previous equation can be rewritten as a DM term

$$\sum_{\alpha\beta\gamma}\varepsilon_{\alpha\beta\gamma}\,S^{(2),\mathrm{A}}_{\alpha\beta}(\vec{q})\,\hat{e}_\gamma\cdot\hat{n}_3 = \vec{D}(\vec{q})\cdot\hat{n}_3 \quad , \quad \vec{D}(-\vec{q}) = -\vec{D}(\vec{q}) \tag{B.18}$$

For $\vec{q} = \vec{0}$ or other collinear magnetic states the DM term is zero, as the cross product of the orientations of any pair of spins is zero.

Using vector–matrix notation the result can be written in a simple form,

$$\begin{aligned}\delta\mathcal{F}(\vec{q};\theta;\hat{n}_3) = & \frac{(\delta m)^2}{2}\Big(\chi_0^{-1} - S^{(2)}(\vec{0})\Big) + \frac{(\delta m)^2}{2}\sin^2\theta\,\Big(S^{(2)}(\vec{0}) - S^{(2)}(\vec{q})\Big)\\ & + \frac{(\delta m)^2}{2}\,\frac{3\cos^2\theta-1}{2}\,\hat{n}_3\cdot\Big(\underline{\widetilde{\chi}}_0^{-1} - \underline{\widetilde{S}}(\vec{0})\Big)\cdot\hat{n}_3\\ & + \frac{(\delta m)^2}{4}\sin^2\theta\,\Big[\hat{n}_3\cdot\Big(\underline{\widetilde{S}}(\vec{0}) - \underline{\widetilde{S}}(\vec{q})\Big)\cdot\hat{n}_3 - \vec{D}(\vec{q})\cdot\hat{n}_3\Big]\end{aligned} \tag{B.19}$$

The first line is the change in free energy from the isotropic terms, the second line is the uniaxial anisotropy of a FM perturbation, and the last line the unidirectional or DM and the uniaxial anisotropies for a spin spiral perturbation.

The equilibrium cone angle is derived from the stationarity conditions, and the anisotropies may stabilise a conical spiral.

Appendix C

The hexagonal lattice

The hexagonal lattice is taken to be spanned by the basis vectors $\vec{a}_1 = a\,(1,0)$ and $\vec{a}_2 = a\,(1/2, \sqrt{3}/2)$, so that $\vec{R}_i = i_1\,\vec{a}_1 + i_2\,\vec{a}_2$, with i_1 and i_2 integers. The basis vectors for the reciprocal lattice are then $\vec{b}_1 = 2\pi/a\,(1, -1/\sqrt{3})$ and $\vec{b}_2 = 2\pi/a\,(0, 2/\sqrt{3})$, and the modulation vector is written as $\vec{q} = q_1\,\vec{b}_1 + q_2\,\vec{b}_2$. The nn and nnn stars are given by the following (i_1, i_2) pairs:

- First star: (1,0), (-1,0), (0,1), (0,-1), (1,-1) and (-1,1);
- Second star: (1,1), (-1,-1), (2,-1), (-2,1), (1,-2) and (-1,2).

The corresponding geometry is given in Fig. C.1, for the case of the supported monolayer. There is a subtlety concerning the symmetry of this arrangement. It is C_{3v} and not C_{6v}: the rotational axis passing through the central atom is three–fold, not sixth–fold, as would apply if only one hexagonal layer would be present. This means that the nn and nnn atoms decompose into two sets. The atoms in each set relate to one another by a 120° rotation, while atoms from distinct sets are related by a single mirror plane. Two consecutive mirror transformations through two planes forming an angle of 60° are equivalent to a 120° rotation.

For the J_1–J_2 isotropic interaction, we have the following form:

$$\begin{aligned} J(\vec{q}) = {} & 2J_1\big[\cos 2\pi q_1 + \cos 2\pi q_2 + \cos 2\pi(q_1 - q_2)\big] \\ & + 2J_2\big[\cos 2\pi(q_1 + q_2) + \cos 2\pi(2q_1 - q_2) + \cos 2\pi(q_1 - 2q_2)\big] \end{aligned} \tag{C.1}$$

and the corresponding energy for a spin spiral is defined as usual,

$$E_{\mathrm{iso}}(\vec{q}\,;\theta) = \frac{J(0) - J(\vec{q})}{2}\sin^2\theta - \frac{J(0)}{2} \tag{C.2}$$

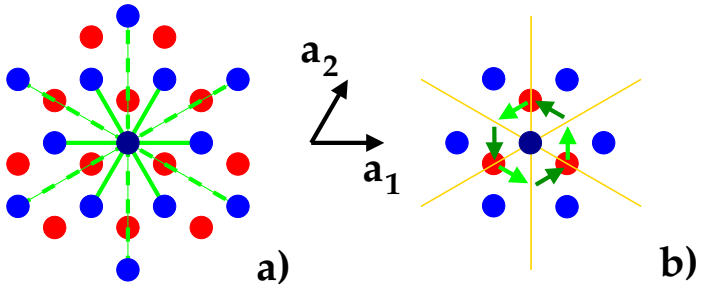

Figure C.1: Geometry for the Mn monolayer supported on the X(111) substrate. Mn atoms are in blue, and X atoms in red. a) — nn interactions J_1 (solid green line) and nnn interactions (dashed green lines) between the central atom and the first and second lattice shells. b) — DM vectors between the central Mn atom and the six nearest neighbours. Two sets of DM vectors exist, related by the mirror planes indicated in yellow. See the discussion in the text.

The stationarity conditions for q_1 and q_2 are then, after some trigonometry,

$$\begin{aligned}\sin\pi(2q_1-q_2)\Big[\cos\pi q_2+\frac{J_2}{J_1}\Big(\cos 3\pi q_2+2\cos\pi(2q_1-q_2)\Big)\Big]&=0\\ \sin\pi(2q_2-q_1)\Big[\cos\pi q_1+\frac{J_2}{J_1}\Big(\cos 3\pi q_1+2\cos\pi(2q_2-q_1)\Big)\Big]&=0\end{aligned}\qquad\text{(C.3)}$$

Setting $\sin\pi(2q_1-q_2)=0$ gives $q_2=2q_1+n$, and the second equation becomes

$$\sin 3\pi q_1\Big[\cos\pi q_1+\frac{3J_2}{J_1}\cos 3\pi q_1\Big]=0\qquad\text{(C.4)}$$

From $\sin 3\pi q_1=0$ we obtain the FM solution, $(q_1,q_2)=(0,0)$, and the 120° AFM solution, $(q_1,q_2)=(1/3,2/3)$ and the five other possibilities due to sixth–fold symmetry. Choosing instead the expression in square brackets to be zero, another equation arises:

$$\cos\pi q_1\Big[1-\frac{3J_2}{J_1}+\frac{6J_2}{J_1}\cos 2\pi q_1\Big]=0\qquad\text{(C.5)}$$

From $\cos\pi q_1=0$ we obtain the FM solution again, and the row–wise AFM solution, $(q_1,q_2)=(1/2,0)$ and the five other possibilities. Setting the term in brackets to zero gives the condition for a spin spiral solution,

$$-3J_2\le J_1\le 9J_2\qquad\text{(C.6)}$$

The corresponding wavevectors are of the form $(q, 2q)$ and equivalent lines related by 60° rotations[1].

Returning now to the symmetry of the supported monolayer, it is useful to recall that mirroring a vector $\vec{u}$ through a plane whose normal is given by $\hat{n}$ can be encoded in the following symmetric transformation P:

$$P\,\vec{u} = \vec{u} - 2(\vec{u}\cdot\hat{n})\,\hat{n} \implies P = 1 - 2\hat{n}\otimes\hat{n}\;,\quad P\,(P\,\vec{u}) = \vec{u} \implies P^{-1} = P \tag{C.7}$$

The three mirror planes for the C_{3v} symmetry have normals $\hat{n}_1 = (1/2, -\sqrt{3}/2, 0)$, $\hat{n}_2 = (0, 1, 0)$ and $\hat{n}_3 = (1/2, \sqrt{3}/2, 0)$, in cartesian components. The respective transformation matrices are thus

$$P_1 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}\;,\quad P_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\;,\quad P_3 = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{C.8}$$

It can be confirmed that two mirroring transformations by two adjacent mirror planes are equivalent to a rotation. For instance, $P_2\,(P_1\,\vec{u}) = \mathcal{R}_z(120°)\,\vec{u}$.

Take now the nn interaction between the central atom and its nn at $\vec{a}_1$ to be

$$J_{(1,0)} = \begin{pmatrix} J_\parallel & D_z & -D_y \\ -D_z & J_\parallel & D_x \\ D_y & -D_x & J_\perp \end{pmatrix} \tag{C.9}$$

with a uniaxial anisotropy given by the difference in the in–plane exchange $J_\parallel$ and the out–of–plane exchange $J_\perp$, and the form of the DM vector is to be specified. This identity must be satisfied:

$$J_{(-1,0)} = J^T_{(1,0)} = P_2\,J_{(1,0)}\,P_2 = \begin{pmatrix} J_\parallel & -D_z & D_y \\ D_z & J_\parallel & D_x \\ -D_y & -D_x & J_\perp \end{pmatrix} \implies D_x = 0 \tag{C.10}$$

The first equality is the overall symmetry of the pair interaction, $J_{ji} = J^T_{ij}$, while the second uses the symmetry of the system, as the atom at $-\vec{a}_1$ is the image of the one at $\vec{a}_1$ through the mirror transformation P_2. The form of the DM vector is then as sketched in Fig. C.1.

[1] In Ref. [119] it is incorrectly stated that these spin spirals should have wavevectors of the form $(q, 0)$ and symmetry equivalent lines. Their calculated energy dispersion for spin spirals shows a local minimum on the expected line, confirming the present theoretical analysis. This mistake reappears in some of the follow–up papers.

For the atom at $\vec{a}_2$ we have

$$J_{(0,1)}T = P_1\, J_{(1,0)}\, P_1 = \begin{pmatrix} J_\parallel & -D_z & -\frac{1}{2}D_y \\ D_z & J_\parallel & -\frac{\sqrt{3}}{2}D_y \\ \frac{1}{2}D_y & \frac{\sqrt{3}}{2}D_y & J_\perp \end{pmatrix} \tag{C.11}$$

The elements on the diagonal are unchanged by successive mirroring transformations (or rotations, for that matter). The in–plane components of the DM vector follow by successive 60° rotations, while the z–component alternates in sign as one goes along the atoms in the nn shell. This is colour–coded in Fig. C.1: the DM vectors are represented by their in–plane projection, and the sign of the z–component corresponds to the two shades of green. On the nnn shell the z–component is zero, as all pairs of atoms line in mirror planes [81]. This can be easily checked using the appropriate transformation matrices.

Relabelling $D_\parallel \equiv D_y$ and $D_\perp \equiv D_z$, the lattice sum for the nn DM vectors gives

$$\begin{aligned} \vec{D}(\vec{q}) = {} & D_\parallel \sqrt{3}\Big[\sin 2\pi(q_1 - q_2) - \sin 2\pi q_2\Big]\hat{e}_x \\ & + D_\parallel \Big[2\sin 2\pi q_1 + \sin 2\pi(q_1 - q_2) + \sin 2\pi q_2\Big]\hat{e}_y \\ & + 2D_\perp \Big[\sin 2\pi q_1 - \sin 2\pi(q_1 - q_2) - \sin 2\pi q_2\Big]\hat{e}_z \end{aligned} \tag{C.12}$$

For the 120° AFM state, there are two inequivalent modulation vectors, again stemming from the reduction of symmetry from C_{6v} to C_{3v}. For $(q_1, q_2) = (1/3, 2/3)$, $\vec{D}(\vec{q}) = 3\sqrt{3}D_\perp\,\hat{e}_z$, while for $(q_1, q_2) = (2/3, 1/3)$, $\vec{D}(\vec{q}) = -3\sqrt{3}D_\perp\,\hat{e}_z$. As the moments are noncollinear, there is a splitting of the energy if $D_\perp \neq 0$:

$$\delta\mathcal{F}_{\mathrm{DM}} = -\frac{(\delta m)^2}{2}\frac{3\sqrt{3}}{2}|D_\perp|\,, \quad \begin{cases} (q_1, q_2) = \left(\frac{1}{3}, \frac{2}{3}\right) & \Longrightarrow \hat{n}_3 = \frac{D_\perp}{|D_\perp|}\,\hat{e}_z \\ (q_1, q_2) = \left(\frac{2}{3}, \frac{1}{3}\right) & \Longrightarrow \hat{n}_3 = -\frac{D_\perp}{|D_\perp|}\,\hat{e}_z \end{cases} \tag{C.13}$$

where the orientation of $\hat{n}_3$ that minimises the DM free energy is indicated for both cases. This can be interpreted as a lifting of the chiral degeneracy of the 120° AFM state, as discussed in the main text.

Bibliography

[1] R. M. White, *Quantum Theory of Magnetism*, Springer Series in Solid-State Sciences (Springer-Verlag, Berlin Heidelberg, 2007), 3rd ed.

[2] P. Strange, *Relativistic quantum mechanics: with applications in condensed matter and atomic physics* (CUP, 1998).

[3] P. Mohn, *Magnetism in the Solid State*, Springer Series in Solid-State Sciences (Springer-Verlag, Berlin Heidelberg, 2006).

[4] R. Martin, *Electronic Structure — Basic Theory and Practical Methods* (CUP, 2004).

[5] J. Zabloudil, R. Hammerling, L. Szunyogh, and P. Weinberger, *Electron Scattering in Solid Matter* (Springer, 2005).

[6] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[7] T. Moriya, *Spin Fluctuations in Itinerant Electron Magnetism*, Springer Series in Solid-State Sciences (Springer-Verlag, Berlin Heidelberg, 1985).

[8] J. Kubler, *Theory of Itinerant Electron Magnetism*, International Series of Monographs on Physics (Oxford University Press, Oxford, 2000).

[9] I. Dzyaloshinksy, J. Phys. Chem. Solids **4**, 241 (1958).

[10] T. Moriya, Phys. Rev. **120**, 91 (1960).

[11] A. Kubetzka, P. Ferriani, M. Bode, S. Heinze, G. Bihlmayer, K. von Bergmann, O. Pietzsch, S. Blügel, and R. Wiesendanger, Phys. Rev. Lett. **94**, 087204 (2005).

[12] M. dos Santos Dias, J. B. Stauton, A. Deak, and L. Szunyogh, Phys. Rev. B **83**, 054435 (2011).

[13] S. A. Hatfield and G. R. Bell, Surf. Sci. **601**, 5368 (2007).

[14] M. Lezăić, P. Mavropoulos, J. Enkovaara, G. Bihlmayer, and S. Blügel, Phys. Rev. Lett. **97**, 026404 (2006).

[15] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[16] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many–Particle Systems* (Dover, 2003).

[17] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).

[18] W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).

[19] E. Engel and R. M. Dreizler, *Density Functional Theory — An Advanced Course* (Springer, 2011).

[20] W. Gordon, Z. Phys. **50**, 630 (1928).

[21] H. J. F. Jansen, Phys. Rev. B **38**, 8022 (1988).

[22] H. Kronmüller and M. Fähnle, *Micromagnetism and the Microstructure of Ferromagnetic Solids* (CUP, 2003).

[23] P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai, Phys. Rev. Lett. **53**, 2512 (1984).

[24] A. I. Liechtenstein, M. I. Katnelson, and V. A. Gubanov, J. Phys. F **14**, L125 (1984).

[25] P. Bruno, Phys. Rev. Lett. **90**, 087205 (2003).

[26] J. B. Staunton, S. Ostanin, S. S. A. Razee, B. L. Gyorffy, L. Szunyogh, B. Ginatempo, and E. Bruno, Phys. Rev. Lett. **93**, 257204 (2004).

[27] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).

[28] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).

[29] I. Dabo, A. Ferretti, N. Poilvert, Y. L. Li, N. Marzari, and M. Cococcioni, Phys. Rev. B **82**, 115121 (2010).

[30] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[31] M. Lüders, A. Ernst, M. Däne, Z. Szotek, A. Svane, D. Ködderitzsch, W. Hergert, B. L. Györffy, and W. M. Temmerman, Phys. Rev. B **70**, 205109 (2005).

[32] I. D. Hughes, M. Däne, A. Ernst, W. Hergert, M. Lüders, J. Poulter, J. B. Staunton, A. Svane, Z. Szotek, and W. M. Temmerman, Nature **446**, 650 (2007).

[33] I. D. Hughes, M. Däne, A. Ernst, W. Hergert, M. Lüders, J. B. Staunton, Z. Szotek, and W. M. Temmerman, New Journal of Physics **10**, 063010 (2008).

[34] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).

[35] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein, J. Phys. — Condens. Matter **9**, 767 (1997).

[36] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).

[37] J. Minar, J. Phys. — Condens. Matter **23**, 253201 (2011).

[38] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).

[39] R. J. Bartlett and M. Musial, Rev. Mod. Phys. **79**, 291 (2007).

[40] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[41] J. Kolorenc and L. Mitas, Rep. Prog. Phys. **74**, 026502 (2011).

[42] J. Korringa, Physica **13**, 392 (1947).

[43] W. Kohn and N. Rostoker, Phys. Rev. **94**, 1111 (1954).

[44] P. Lloyd and P. V. Smith, Adv. Phys. **21**, 69 (1972).

[45] A. Gonis, *Green Functions for Ordered and Disordered Systems* (Elsevier, 1992).

[46] J. S. Faulkner and G. M. Stocks, Phys. Rev. B **21**, 3222 (1980).

[47] L. Szunyogh, B. Újfalussy, P. Weinberger, and J. Kollar, Phys. Rev. B **49**, 2721 (1994).

[48] R. Zeller, P. H. Dederichs, B. Újfalussy, L. Szunyogh, and P. Weinberger, Phys. Rev. B **52**, 8807 (1995).

[49] E. M. Godfrin, J. Phys. Condens. Matter **3**, 7843 (1991).

[50] B. Lazarovits, L. Szunyogh, and P. Weinberger, Phys. Rev. B **65**, 104441 (2003).

[51] B. Lazarovits, L. Szunyogh, and P. Weinberger, Phys. Rev. B **67**, 024415 (2003).

[52] L. M. Sandratskii, Adv. Phys. **47**, 91 (1998).

[53] N. D. Mermin, Phys. Rev. **137**, A1441 (1965).

[54] R. P. Feynman, Phys. Rev. **97**, 660 (1955).

[55] R. Kubo, J. Phys. Soc. Japan. **17**, 1100 (1962).

[56] R.-J. Liu and T.-L. Chen, Phys. Rev. B **50**, 9169 (1994).

[57] J. M. Sanchez, F. Ducastelle, and D. Gratias, Physica **128**, 334 (1984).

[58] J. M. Sanchez, Phys. Rev. B **48**, 14013 (1993).

[59] R. Drautz and M. Fähnle, Phys. Rev. B **69**, 104404 (2004).

[60] R. Drautz and M. Fähnle, Phys. Rev. B **72**, 212405 (2005).

[61] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1* (Pergamon, 1980).

[62] M. Weinert and J. W. Davenport, Phys. Rev. B **45**, 13709 (1992).

[63] J. M. Luttinger, Phys. Rev. **119**, 1153 (1960).

[64] F. Körmann, A. Dick, B. Grabowski, B. Hallstedt, T. Hickel, and J. Neugebauer, Phys. Rev. B **78**, 033102 (2008).

[65] P. Soven, Phys. Rev. **156**, 809 (1967).

[66] B. L. Gyorffy, A. J. Pindor, J. Staunton, G. M. Stocks, and H. Winter, J. Phys. F — Met. Phys. **15**, 1337 (1985).

[67] Y. Kakehashi, Adv. Phys. **53**, 497 (2004).

[68] D. A. Rowlands, Rep. Prog. Phys. **72**, 086501 (2007).

[69] E. C. Stoner, Proc. Roy. Soc. London A **165**, 372 (1938).

[70] E. C. Stoner, Proc. Roy. Soc. London A **169**, 339 (1939).

[71] S. Mankovsky, G. H. Fecher, and H. Ebert, Phys. Rev. B **83**, 144401 (2011).

[72] M. Pajda, J. Kudrnovsky, I. Turek, V. Drchal, and P. Bruno, Phys. Rev. B **64**, 174402 (2001).

[73] O. N. Mryasov, Phase Transitions **78**, 197 (2005).

[74] O. N. Mryasov, U. Nowak, K. Y. Guslienko, and R. W. Chantrell, Europhys. Lett. **69**, 805 (2005).

[75] L. M. Sandratskii, R. Singer, and E. Sasioglu, Phys. Rev. B **76**, 184406 (2007).

[76] R. Wieser, E. Y. Vedmedenko, and R. Wiesendanger, Physical Review B **77**, 064410 (2008).

[77] A. Antal, B. Lazarovits, L. Udvardi, L. Szunyogh, B. Ujfalussy, and P. Weinberger, Phys. Rev. B **77**, 174429 (2008).

[78] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).

[79] H. B. Callen and E. Callen, J. Phys. Chem. Solids **27**, 1271 (1966).

[80] L. Szunyogh, B. Lazarovits, L. Udvardi, J. Jackson, and U. Nowak, Phys. Rev. B **79**, 020403 (2009).

[81] A. Crépieux and C. Lacroix, J. Magn. Magn. Mater. **182**, 341 (1998).

[82] E. P. Wohlfarth, J. Magn. Magn. Mater. **7**, 113 (1978).

[83] P. Buczek, A. Ernst, P. Bruno, and L. Sandratskii, Phys. Rev. Lett. **102**, 247206 (2009).

[84] V. I. Antropov, M. I. Katnelson, B. N. Harmon, M. van Schilfgaarde, and D. Kusnezov, Phys. Rev. B **54**, 1019 (1996).

[85] B. Újfalussy, B. Lazarovits, L. Szunyogh, G. M. Stocks, and P. Weinberger, Phys. Rev. B **70**, 100404 (2004).

[86] B. Skubic, J. Hellsvik, and O. Eriksson, J. Phys.: Condens. Matter **20**, 315203 (2008).

[87] J. Staunton, B. L. Gyorffy, G. M. Stocks, and J. Wadsworth, J. Phys. F — Met. Phys. **16**, 1761 (1986).

[88] J. B. Staunton and B. L. Gyorffy, Phys. Rev. Lett. **69**, 371 (1992).

[89] S. S. A. Razee, J. B. Staunton, L. Szunyogh, and B. L. Gyorffy, Phys. Rev. Lett. **88**, 147201 (2002).

[90] A. Buruzs, P. Weinberger, L. Szunyogh, L. Udvardi, P. I. Chleboun, A. M. Fischer, and J. B. Staunton, Phys. Rev. B **76**, 064417 (2007).

[91] V. I. Lebedev and D. N. Laikov, Doklady Mathematics **59**, 477 (199), http://www.ccl.net/cca/software/SOURCES/FORTRAN/Lebedev-Laikov-Grids/.

[92] S. Mankovsky and H. Ebert, Phys. Rev. B **74**, 054414 (2006).

[93] M. A. Ruderman and C. Kittel, Phys. Rev. **96**, 99 (1954).

[94] T. Kasuya, Prog. Theor. Phys. **16**, 45 (1956).

[95] K. Yosida, Phys. Rev. **106**, 893 (1957).

[96] L. Onsager, J. Am. Chem. Soc. **58**, 1486 (1936).

[97] R. Brout and H. Thomas, Physics (Long Island City, NY) **3**, 317 (1967).

[98] J. B. Staunton, D. D. Johnson, and F. J. Pinski, Phys. Rev. B **50**, 1450 (1994).

[99] M. Cyrot and H. Kaga, Phys. Rev. Lett. **77**, 5134 (1996).

[100] A. L. Shallcross, J. K. Glasbrenner, and K. D. Belashchenko, Phys. Rev. B **78**, 184419 (2008).

[101] D. E. Logan, Y. H. Szczech, and M. A. Tusch, Europhys. Lett. **30**, 307 (1995).

[102] J. Thomassen, F. May, B. Feldmann, M. Wuttig, and H. Ibach, Phys. Rev. Lett. **69**, 3831 (1992).

[103] D. Q. Li, M. Freitag, J. Pearson, Z. Q. Qiu, and S. D. Bader, Phys. Rev. Lett. **72**, 31112 (1994).

[104] T. Asada and S. Blügel, Phys. Rev. Lett. **79**, 507 (1997).

[105] H. L. Meyerheim, J.-M. Tonnerre, L. Sandratskii, H. C. N. Tolentino, M. Przybylski, Y. Gabi, F. Yildiz, X. L. Fu, E. Bontempi, S. Grenier, et al., Phys. Rev. Lett. **103**, 267202 (2009).

[106] D. Loss, F. L. Pedrocchi, and A. J. Leggett, Phys. Rev. Lett. **107**, 107201 (2011).

[107] B. Hardrat, A. Al-Zubi, P. Ferriani, S. Blügel, G. Bihlmayer, and H. S., Phys. Rev. B **79**, 094411 (2009).

[108] L. Udvardi and L. Szunyogh, Phys. Rev. Lett. **102**, 207204 (2009).

[109] S. Meckler, N. Mikuszeit, A. Pressler, E. Y. Vedmedenko, O. Pietzsch, and R. Wiesendanger, Phys. Rev. Lett. **103**, 157201 (2009).

[110] K. von Bergmann, S. Heinze, M. Bode, E. Y. Vedmedenko, G. Bihlmayer, S. Blügel, and R. Wiesendanger, Phys. Rev. Lett. **96**, 167203 (2006).

[111] S. Heinze, K. von Bergmann, M. Menzel, J. Brede, A. Kubetzka, R. Wiesendanger, G. Bihlmayer, and S. Blügel, Nature Phys. **7**, 713 (2011).

[112] P. Ferriani, K. von Bergmann, E. Y. Vedmedenko, S. Heinze, M. Bode, M. Heide, G. Bihlmayer, S. Blügel, and R. Wiesendanger, Phys. Rev. Lett. **101**, 027201 (2008).

[113] M. Bode, M. Heide, K. von Bergmann, P. Ferriani, S. Heinze, G. Bihlmayer, A. Kubetzka, O. Pietzsch, S. Blügel, and R. Wiesendanger, Nature **447**, 190 (2007).

[114] C. L. Gao, W. Wulfhekel, and J. Kirschner, Phys. Rev. Lett. **101**, 267205 (2008).

[115] P. Kurz, G. Bihlmayer, K. Hirai, and S. Blügel, Phys. Rev. Lett. **86**, 1106 (2001).

[116] A. B. Shick, F. Maca, M. Ondracek, O. N. Mryasov, and T. Jungwirth, Phys. Rev. B **78**, 054413 (2008).

[117] S. Dennler and J. Hafner, Phys. Rev. B **72**, 214413 (2005).

[118] P. Ferriani, S. Heinze, G. Bihlmayer, and S. Blügel, Phys. Rev. B **72**, 024452 (2005).

[119] S. Heinze, P. Kurz, D. Wortmann, G. Bihlmayer, and S. Blügel, App. Phys. A — Mater. **75**, 25 (2002).

[120] P. Kruger, M. Taguchi, and S. Meza-Aguilar, Phys. Rev. B **61**, 15277 (2000).

[121] D. Hobbs and J. Hafner, J. Phys. — Condens. Matter **12**, 7025 (2000).

[122] R. A. de Groot, F. M. Müller, P. van Engen, and K. J. H. Buschow, Phys. Rev. Lett. **50**, 2024 (1983).

[123] M. I. Katsnelson, V. I. Irkhin, L. Chioncel, A. I. Lichtenstein, and R. A. de Groot, Rev. Mod. Phys. **80**, 315 (2008).

[124] J.-C. Zheng and J. W. Davenport, Phys. Rev. B **69**, 144415 (2004).

[125] J. B. McKinnon, D. Melville, and E. W. Lee, J. Phys. C **1S**, 46 (1970).

[126] P. J. Jensen, D. Bennemann, and E. W. Lee, Surf. Sci. Rep. **61**, 129 (2006).

[127] S. Shallcross, A. E. Kissavos, V. Meded, and A. V. Ruban, Phys. Rev. B **72**, 104437 (2005).

[128] H. Zabel and S. D. Bader, *Magnetic Heterostructures* (Springer–Verlag, 2008).