Behandlung stark nichtkollinearer Magnetisierungsstrukturen mit der Spin-Cluster Entwicklung

Von der Fakultät Mathematik und Physik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von
Frank Dietermann
aus Schorndorf

Hauptberichter:	Prof. Dr. Manfred Fähnle
Mitberichter:	Prof. Dr. Christian Holm
Prüfungsvorsitz:	Prof. Dr. Jörg Wrachtrup
Tag der mündlichen Prüfung:	5.März 2012

Institut für theoretische und angewandte Physik der Universität Stuttgart
und Max-Planck Institut für intelligente Systeme
2012

Contents

List of Tables

List of Figures

Introduction

1. Background

Today, there is a multitude of different magnetic materials differing in composition and structure. Most surprisingly the different properties of these materials can be largely explained by the rather simple model of local magnetic moments which interact via exchange, dipole, and anisotropy interaction. Both exchange and dipole interaction favor particular relative angles between different magnetic moments, while the anisotropy favors particular angles with regards to symmetry axes of the crystal. This work concerns itself with the exchange interaction.

The exchange interaction is by far the strongest of the three, and in many cases favors either parallel or antiparallel alignment of the magnetic moments. Correspondingly, most materials show large regions compared to the interatomic distances where the magnetic moments are mostly aligned, with only small deviations from the favored alignment. Such configurations are proven to be described quite well qualitatively with the simple nearest-neighbour Heisenberg model by, e.g., micromagnetic simulations of domain walls, of magnetic vortices, etc. Especially micromagnetism has been developed into a well-honed tool which can be used by almost anyone with some background in engineering or physics.

However, during recent years the techniques to prepare and analyse magnetic samples have advanced greatly, especially for thin films. Some of these samples are no longer largely collinear, either because of ingenious patterning of the samples, or because the interactions in a material do not fit into the simple traditional picture of ferro- or antiferromagnetism. This requires extensions of the traditional Heisenberg exchange model. Most attempts at this problem use a more or less ad-hoc approach. The most intuitive, and historically first was to simply take interactions beyond the first neigbours into account. Some work groups have also considered interactions where the functional form is changed from Heisenberg's cosine, e.g., to a cosine squared. Some have also investigated interactions where one interaction is not limited to two atoms, e.g, terms of the form $(\mathbf{e}_i \cdot \mathbf{e}_j)(\mathbf{e}_k \cdot \mathbf{e}_l)$.

However, the justification why one specific extension of the Heisenberg model was chosen is quite often rather spurious. This becomes even worse if only specific symmetric configurations are considered - e.g., when someone investigates only spin spirals it becomes impossible to distinguish between a term $(\mathbf{e}_i \cdot \mathbf{e}_j)^2$ and a term of the form $(\mathbf{e}_i \cdot \mathbf{e}_j)(\mathbf{e}_k \cdot \mathbf{e}_l)$

if $\mathbf{r}_i - \mathbf{r}_j = \mathbf{r}_k - \mathbf{r}_l$.
The Spin Cluster Expansion, the method this work concerns itself solves this problem by employing a complete basis set which by definition encompasses all possible terms one may add to the classical Heisenberg model and thus allows - as long as the local magnetic moment picture is valid - comparable accuracy to ab-initio Density Functional Theory. The present author represents already the third “generation” working on this topic - the idea behind the SCE was due to R.Drautz [3] and the realization of this idea for bulk magnets was the work of R.Singer [1]. The goal of the present author was to expand the application range of the method, especially to the aforementioned thin films. This goal was not achieved due to technical difficulties. The present work has thus turned into a discussion of the details and intricacies of the process of modeling the magnetic exchange energy for bulk materials. These discussions should prove useful if the aforementioned technical difficulties can be removed, and for all other approaches to the modeling of the magnetic exchange energy in terms of local magnetic moments.

2. Structure of this work

This work focuses on the effects of symmetry in the framework of the Spin Cluster Expansion, and how these effects may be used. Because this requires detailed knowledge of the basis functions used in the Spin Cluster Expansion the work begins with a detailed presentation of the Spin Cluster Expansion, followed by a discussion of space group symmetry and specific symmetric configurations. The results are then used to develop methods to obtain information about the convergence limits of the Spin Cluster Expansion, i.e., how many basis functions of what type are needed to parametrize the energy of a specific system. The results are then verified by ab-initio Density Functional Theory calculations for bcc and fcc iron. The necessary tools for this verification are introduced in the last section of the theoretical part, and then results for bcc and fcc iron are presented.
This work also contains a thematically sorted index just before the bibliography.

3. Notation

Vectors are printed **bold**.
Matrices are printed $\mathsf{sans\ serif}$.
All summations are written down explicitly - no implicit summation.

Mathematical operators:

$\times$:	vector product *or* multi-line continuation of scalar product
$\cdot$:	scalar product
$\lfloor x \rfloor$:	nearest integer below x. E.g, $\lfloor -0.4 \rfloor = 1$
$\int f \,\mathrm{d}\Omega$:	integral of f over the whole space

Frequent scalar quantities:

E :	energy
e :	energy per atom
φ :	angle in the x-y-plane respective x-axis
ϑ:	angle respective z-axis
$Y^{lm}(\vartheta, \varphi)$:	spherical harmonic to orders l,m
N :	number of atoms within the considered system
a :	lattice constant
n :	charge density
ρ :	spin density matrix
$C^{LM}_{l_1 m_1 l_2 m_2}$:	Clebsch-Gordan-coefficent
$\begin{pmatrix} l_1 & l_2 & J \\ m_1 & m_2 & M \end{pmatrix}$:	3j-symbol

Frequent vector quantities:

(x, y, z) :	cartesian vector
$[hkl]$:	normalized vector in direction of the point identified by the Miller indices h,k,l of a cartesian unit cell
$\mathbf{r}$:	position vector
$\mathbf{r}_i$:	position vector of a specific atom i within a crystal
$\mathbf{e}_i$:	direction of the local magnetic moment
Ω :	list of all N directions of the magnetic moments in the system
$\mathbf{q}$:	wave vector of a spin-spiral configuration
$\mathbf{M}_i$:	magnetic moment of a specific atom i

1. Theory

This chapter details the theory of the Spin Cluster Expansion, with special emphasis on the symmetry of the lattice and on symmetric configurations of the magnetic moments. For readers with limited time at hand, a short summary is provided below which contains all essential information.

1.1. Summary

The Spin Cluster Expansion (SCE) relies on an adiabatic approximation which presumes that the magnetic energy of a system can be characterized by the orientations $\mathbf{e}_i$ of the local magnetic moments in this system. This energy $E(\{\mathbf{e}_i\})$ can be calculated by density functional theory for specific configurations of the $\{\mathbf{e}_i\}$, however, this requires a lot of numerical effort and does not provide an explicit, simple formula for calculating $E(\{\mathbf{e}_i\})$ for any configuration.
To resolve this difficulty, $E(\{\mathbf{e}_i\})$ is written as an expansion with respect to a set of complete basis functions, $E(\{\mathbf{e}_i\}) = \sum_\nu J_\nu \Phi_\nu$ where the index ν is a label to enumerate the basis functions of the complete set $\{\Phi\}$. The derivation of suitable basis functions and the treatment of various symmetries are detailed in Sections 1.3 through 1.5.
The basis functions $\mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ found are superpositions of products of the spherical harmonics, and can be characterized by a "cluster archetype" α_i - one set of n_α local moments or atoms which represents all equivalent sets under space group symmetry -, a "$\mathbf{l}_{\alpha i}$-tuple" - a set of n_α integer parameters representing different orders of spherical harmonics -, and an index $\mathbf{u}$ which enumerates the eigenvectors of a matrix arising due to the combination of $O(3)$ symmetry in the configuration space of the $\{\mathbf{e}_i\}$ and the rotational parts of the space group symmetry of the lattice. The energy can then be expanded as

$$E(\{\mathbf{e}_i\}) = \sum_{\alpha_i} \sum_{\mathbf{l}_{\alpha i}} \sum_{\mathbf{u}} J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}. \tag{1.1}$$

If the investigation is limited to specific symmetric configuration classes, the general basis set is overcomplete and one can use a symmetry-adapted, smaller basis set. For supercells this symmetry-adapted basis set consists of the cluster archetypes that fit within the supercell. For spin spiral configurations (frozen magnons), the symmetry-

adapted basis set is closely related to the extended Heisenberg model[1] , and for spin spirals of an arbitrary constant opening angle, or for spin spirals with varying but very small opening angles, the SCE and the extended Heisenberg model are equivalent.
To obtain the expansion coefficients $J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ of the energy of a system, a set of reference energies is calculated by ab-initio density functional theory and the values of the $J_{\alpha \mathbf{lu}}$ are obtained by a least-mean-squares fit of the SCE expansion (1.1) to the reference energies.

1.2. Adiabatic Approximation

Magnetism is a quantum mechanical effect of many-electron systems. Therefore, complete understanding and description of all aspects of magnetism has to be within the Hilbert space of the many-electron states.This is, however, notoriously difficult because of the complexity and size of the many-electron Hilbert space, the noncommutative algebra and the structure of the quantum mechanical equations. Therefore one tries - whenever possible - to find physical arguments to map the Hilbert space on some subspace of itself or on a space of other variables.
One way to do this is to make an “adiabatic approximation”, which in essence presumes that the system is always in its ground state with respect to some “slow” variables or observables. The state of the system shall then follow an adiabatic evolution, that is if the system starts in the ground state regarding specific expectation values of the slow observables, and moves slowly to a state with different expectation values of the slow observables, it ends in the corresponding ground state and not in some excited state. The validity and bounds of such assumptions are a rather interesting subject of ongoing research, especially for Hamiltonians without a gap in the eigenvalue spectrum [4, 5].
One adiabatic approximation in magnetism is to use the magnetic moments of individual atoms as the slow observables. This approach is in quite good qualitative agreement with experiments, as, for example, the Heisenberg model can be used to describe the characteristics of magnetic phase transitions, or the Landau-Lifshitz-Gilbert equation gives reasonable results for some types of magnetization dynamics. Therefore this adiabatic approximation is employed. As now the state of the system is uniquely identified by the magnetic moment configuration, the energy of the system may be expressed as function of the moment configuration.

One further step that is commonly used, e.g., when evaluating spin wave spectra [6, 7, 8], is to presume that only the configuration of the magnetic moment directions $\mathbf{e}_i = \mathbf{M}_i / |\mathbf{M}_i|$ are sufficient to determine the energy, and that the magnitudes of the magnetic moments $|\mathbf{M}_i|$ are slaved by these directions. In other words, the system is always in the ground state regarding the $\mathbf{e}_i$, which uniquely determines the corresponding $|\mathbf{M}_i|$.

[1] “Extended” in the sense that terms beyond nearest neighbours are considered, but the functional form $\mathbf{e}_i \cdot \mathbf{e}_j$ remains the same.

This is usually justified with the argument that *"The interatomic exchange parameters among MM's*[2] *(e.g., the J_{ij} parameters in a Heisenberg Hamiltonian) are small (<100 meV) compared to the characteristic electronic energies such as intraatomic exchange, bandwidth, etc. Thus the MM directions correspond to the slow degrees of freedom within the adiabatic approximation while the evolution of the magnitudes of the MM's is determined by the change of the electronic wave functions (fast degrees of freedom). Using the analogy with the Born-Oppenheimer approximation, the moment directions correspond to the slow nuclear motion, and the forces or torques governing their rotation arise from the rapid relaxation of the electronic system to the instantaneous directions (or positions) of the moments"* [9].

However, it has recently come to the attention of the present author that the intraatomic exchange is actually not significantly larger than the energy scale set by transversal fluctuations (see Tables A.1 and A.2 in Section A.7 for some calculations).

Sometimes one also finds the argument that *"The essence of these is that there is a separation between fast and slow motions in this problem, and attention to this fact can render it tractable. On a time scale, τ, long compared with the hopping time $\hbar/w$ ($\propto 10^{-15}$ s) where w is the relevant bandwidth, but short compared with an inverse spin-wave frequency $1/\omega_s$ ($\propto 10^{-13}$ s) electrons arrive at and leave a site with sufficient correlation between their spin orientations to leave the magnetisation, time averaged over τ, non-zero. This is the complex process of moment formation. Then, on the time scale of $1/\omega_s$, these moments can change their orientation, flip over for instance, as in thermal fluctuations described so well by the Heisenberg model. This is the slow motion."* [10], whereby the basic picture of moment formation has been proposed by J.Hubbard[11] as far as the author is aware.

However, the present author does not understand why this process shall preclude slow fluctuations in the absolute value of the localized magnetic moment, but permit slow fluctuations of the directions.

Such longitudinal fluctuations have already been investigated tentatively in, e.g., [12, 13], although Refs. [12, 13] do not allow for completely unrestricted, independent longitudinal fluctuations. The work [14] allowed for unrestricted and independent longitudinal fluctuations and obtained quite good agreement with experimental values for the Curie temperatures. Some ideas for an extension of the Spin Cluster Expansion formalism to an adiabatic approximation that takes independent longitudinal fluctuations into account and models the energy as a function of the $\mathbf{M}_i$ are presented in Section A.9.2.

Nevertheless, the work presented in the following was carried out under the assumption that the adiabatic approximation in terms of the directions holds. While the present author is unable to follow the reasoning behind this approximation as presented above, it does not necessarily indicate that this formulation of the adiabatic approximation is wrong. In fact, thermodynamical predictions employing this approximation and limiting

[2] Magnetic Moments

the configurations of the magnetic moments to spin spirals show quite good agreement with experimental findings, so one could a posterio justify the adiabatic approximation in terms of the directions.
Whatever the solution to this problem is, the arguments made and results achieved during this work should still prove useful, as the space of directions of the magnetic moments is a subspace of that of the magnetic moments, and therefore all insights gained within this subspace are also useful when discussing the space of magnetic moments. Furthermore, the SCE allows a considerable increase in accuracy regarding the description within the space of magnetic moment directions, which should allow to make theoretical predictions more precise, which in turn would make quantitative comparisons to experiments easier.

1.3. A Basis Set for Functions of the Magnetic Moment Directions

The goal of the Spin Cluster Expansion is to expand the energy of different magnetic states in terms of the directions of the local magnetic moments $\mathbf{e}_i$. These directions correspond to points on the surface of a sphere, and they can be parameterized in spherical coordinates by the angles ϑ_i, φ_i, which are limited to $\vartheta_i \in [0, \pi]$ and $\varphi_i \in [0, 2\pi]$. The results presented in this section have been obtained by R.Singer [1], R.Drautz and M.Fähnle[3] without any contribution by the author. Because the work by the author on symmetry relies on rather intricate details of the basis functions, the derivation of these basis functions is summarized here.
This derivation starts from a single atomic moment and then generalizes to countable numbers of magnetic moments. After this, the O(3) symmetry of a system without spin-orbit coupling has been taken into account[15] by R.Singer and M.Fähnle.

1.3.1. Local Space

The configuration space of the direction of one local magnetic moment is described by just two variables ϑ and φ or just one direction $\mathbf{e}$. The scalar product in the space of functions ("function space") of these two variables is defined as

$$\langle f|g\rangle = \frac{1}{4\pi}\int_S f^*(\Omega)g(\Omega)\,\mathrm{d}\Omega, \tag{1.2}$$

where $\Omega = (\vartheta, \varphi)$ is the configuration of the system, and the integral runs over all possible configurations (the surface of the sphere). A set of basis functions for the function space should be orthogonal and complete regarding this scalar product. One such orthogonal and complete basis set is the set of all scaled spherical harmonics $\{\sqrt{4\pi}Y^{lm}(\vartheta,\varphi)\}$, as

$$\begin{aligned}
&4\pi\langle Y^{lm}(\vartheta,\varphi),Y^{l'm'}(\vartheta,\varphi)\rangle\\
&=\int_S \left(Y^{lm}(\vartheta,\varphi)\right)^* Y^{l'm'}(\vartheta,\varphi)\,\mathrm{d}\Omega\\
&=\delta_{ll'}\delta_{mm'},\\
&\sum_{l=0}^{\infty}\sum_{m=-l}^{m=l}\left(Y^{lm}(\vartheta,\varphi)\right)^* Y^{lm}(\vartheta',\varphi')\\
&=\delta(\varphi-\varphi')\delta(\cos(\vartheta)-\cos(\vartheta'))\\
&=\delta(\mathbf{e}-\mathbf{e}').
\end{aligned} \tag{1.3}$$

1.3.2. Global Space

Presume a system formed by N atoms, each with a magnetic moment $\mathbf{m}$ and a corresponding direction $\mathbf{e}$. The distinguishable atoms are situated at positions $\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N$, and may be identified by their position. The magnetic moment directions can be described by numbers $\vartheta_1,\vartheta_2,\ldots,\vartheta_N$, and $\varphi_1,\ldots,\varphi_N$, or by N directions $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_N$.
The configuration Ω of this system shall describe which magnetic moment points in what direction. For this purpose, the configuration is written as, e.g., $\Omega=(\mathbf{e}_1^{\mathbf{r}_1},\mathbf{e}_2^{\mathbf{r}_2},\ldots,\mathbf{e}_N^{\mathbf{r}_N})$, where the lower index indicates specific values of the angles ϑ,φ and the upper index indicates which position (and correspondingly, atom) these values are associated with. $\Omega'=(\mathbf{e}_4^{\mathbf{r}_1},\mathbf{e}_N^{\mathbf{r}_2},\ldots,\mathbf{e}_6^{\mathbf{r}_N})$ would also be a valid, but different configuration of the system. To avoid ambiguity, the direction associated with position $\mathbf{r}_1$ shall always appear first in this list.
For example, consider two atoms at positions $\mathbf{r}_1$ and $\mathbf{r}_2$ and two directions, $\mathbf{e}_1=(1,0,0)$ and $\mathbf{e}_2=(0,0,1)$. The configuration could now be $\Omega=((1,0,0)^{\mathbf{r}_1},(0,0,1)^{\mathbf{r}_2})$ or $\Omega'=((0,0,1)^{\mathbf{r}_1},(1,0,0)^{\mathbf{r}_2})$. Both are different[3] configurations of the system.
The definition of the scalar product of the preceding section can be extended to the global function space as:

$$\begin{aligned}
\langle f|g\rangle&=\frac{1}{(4\pi)^N}\int_S f^*(\Omega)g(\Omega)\,\mathrm{d}\Omega=\\
&=\frac{1}{(4\pi)^N}\int_{S_A}\int_{S_B}\cdots\int_{S_\eta} f^*(\Omega)g(\Omega)\,\mathrm{d}\Omega_A\ldots\,\mathrm{d}\Omega_\eta.
\end{aligned} \tag{1.4}$$

[3] The seasoned condensed matter scientist may stall at this notion, because “only the relative angle between magnetic moments matters”. This is however a consequence of specific symmetries, which will be considered later. As a model system for the current considerations, take two different magnetic atoms on a surface with a magnetic field parallel to the surface.

The products of the basis functions of the local function space,

$$\Phi_{\mathbf{lm}}(\Omega) = \sqrt{4\pi}^N Y^{l_1m_1}(\mathbf{e}_\Box)Y^{l_2m_2}(\mathbf{e}_\Box)\dots Y^{l_Nm_N}(\mathbf{e}_\Box), \tag{1.5}$$

form orthogonal and complete basis functions for the global function space [3].
The $\mathbf{l}, \mathbf{m}$ on the left hand side and the boxes $\Box$ on the right hand side take a bit of explanation: The orders of the spherical harmonics are characterized by N values $l_1, l_2, \dots, l_N$, and further N values $m_1, m_2, \dots, m_N$. Each of these values has to be associated with one of the positions $\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N$, whereat the association of the m_i has to respect that $|m| \leq |l|$ for any given spherical harmonic. If the same numerical values $l_1, l_2, \dots, l_N$ are associated differently with the positions, the basis functions are different.
The argument $\mathbf{e}_\Box$ of a specific spherical harmonic is the numerical value that was associated by the configuration to the same atom as the orders l, m of this spherical harmonics. Consider the following example of two sites, $\mathbf{r}_1$ and $\mathbf{r}_2$, possible values $l_1 = 2, l_2 = 1$, and $m_1 = -1, m_2 = 0$. The configuration shall be $\Omega = (\mathbf{e}_1^{\mathbf{r}_1}, \mathbf{e}_2^{\mathbf{r}_2})$. The following four basis functions can be constructed:

$$\begin{aligned} F^1(\Omega) &= 4\pi Y^{20}(\mathbf{e}_1^{\mathbf{r}_1})Y^{1-1}(\mathbf{e}_2^{\mathbf{r}_2}), \\ F^2(\Omega) &= 4\pi Y^{2-1}(\mathbf{e}_1^{\mathbf{r}_1})Y^{10}(\mathbf{e}_2^{\mathbf{r}_2}), \\ F^3(\Omega) &= 4\pi Y^{10}(\mathbf{e}_1^{\mathbf{r}_1})Y^{2-1}(\mathbf{e}_2^{\mathbf{r}_2}), \\ F^4(\Omega) &= 4\pi Y^{1-1}(\mathbf{e}_1^{\mathbf{r}_1})Y^{20}(\mathbf{e}_2^{\mathbf{r}_2}). \end{aligned} \tag{1.6}$$

These are all different, orthogonal basis functions (the reader may work out the scalar product (1.4) between these functions). The notation has to reflect the association between the values $l_1, \dots, l_N$ and the positions $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$. In this work, the following convention is made: The values l_i are written in an "$\mathbf{l}$-tuple", which is a comma-separated list of numbers, e.g, $\mathbf{l} = (l_1^{\mathbf{r}_1}, l_2^{\mathbf{r}_2}, l_3^{\mathbf{r}_3}, \dots, l_N^{\mathbf{r}_N})$. To avoid ambiguity in the notation, each place in the $\mathbf{l}$-tuple is associated with a specific position. The value that comes first in the $\mathbf{l}$-tuple is associated with position $\mathbf{r}_1$, the second value with the position $\mathbf{r}_2$, etc. $\mathbf{l} = (l_3^{\mathbf{r}_1}, l_1^{\mathbf{r}_2}, \dots, l_2^{\mathbf{r}_N})$ is also a valid, but different $\mathbf{l}$-tuple. For the numbers m_i, the same convention is made, introducing a $\mathbf{m}$-tuple, e.g., $\mathbf{m} = (m_3^{\mathbf{r}_1}, m_5^{\mathbf{r}_2}, m_2^{\mathbf{r}_3}, \dots m_1^{\mathbf{r}_N})$.
The above examples are written in this notation as:

$$\begin{aligned} F^1(\Omega) &= \Phi_{(2^{\mathbf{r}_1},1^{\mathbf{r}_2})(0^{\mathbf{r}_1},-1^{\mathbf{r}_2})}(\Omega), \\ F^2(\Omega) &= \Phi_{(2^{\mathbf{r}_1},1^{\mathbf{r}_2})(-1^{\mathbf{r}_1},0^{\mathbf{r}_2})}(\Omega), \\ F^3(\Omega) &= \Phi_{(1^{\mathbf{r}_1},2^{\mathbf{r}_2})(0^{\mathbf{r}_1},-1^{\mathbf{r}_2})}(\Omega), \\ F^4(\Omega) &= \Phi_{(1^{\mathbf{r}_1},2^{\mathbf{r}_2})(-1^{\mathbf{r}_1},0^{\mathbf{r}_2})}(\Omega). \end{aligned} \tag{1.7}$$

1.3.3. Clusters

The $\mathbf{l}$-tuple contains in principle entries for all N atoms, however, these entries can be zero. As the spherical harmonic for $l = 0$, $\mathcal{Y}^{00}$ is a constant (equal to $1/\sqrt{4\pi}$) the spherical harmonics with associated $l = 0$ play no role in physical processes driven by changes in the $\mathbf{e}_i$ and do not need to be written explicitly in the basis functions and can be omitted from the notation. Then, a basis function for, e.g., $\Omega = (\mathbf{e}_1^{\mathbf{r}_1}, \mathbf{e}_2^{\mathbf{r}_2}, \mathbf{e}_3^{\mathbf{r}_3}, \mathbf{e}_4^{\mathbf{r}_4}, \ldots, \mathbf{e}_N^{\mathbf{r}_N})$ and, e.g., $\mathbf{l} = (l_1^{\mathbf{r}_1}, 0^{\mathbf{r}_2}, 0^{\mathbf{r}_3}, l_2^{\mathbf{r}_4}, \ldots, l_{n_\alpha}^{\mathbf{r}_\eta}, 0^{\mathbf{r}_{n_\alpha+1}}, \ldots, 0^{\mathbf{r}_N})$[4], would have the form

$$\begin{aligned}\Phi_{\alpha \mathbf{l}_\alpha \mathbf{m}_\alpha}(\Omega) = \sqrt{4\pi}^{n_\alpha} Y^{l_1 m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2 m_2}(\mathbf{e}_4^{\mathbf{r}_4}) \times \\ \times \ldots Y^{l_{n_\alpha} m_{n_\alpha}}(\mathbf{e}_{n_\alpha}^{\mathbf{r}_\eta}),\end{aligned} \tag{1.8}$$

where α on the left hand side shall contain the information at which n_α atoms the spherical harmonics are of an order unequal to zero. Its notation, and the meaning of the new $\mathbf{l}_\alpha, \mathbf{m}_\alpha$-tuples, will be explained shortly.
One rather important exception has to be made for the basis function where all elements of the $\mathbf{l}$-tuple are zero. This function is completely constant regardless of configuration

$$\Phi_{000} = \sqrt{4\pi}^N Y^{00}(\mathbf{e}^{\mathbf{r}_1}) \ldots Y^{00}(\mathbf{e}^{\mathbf{r}_N}) = \frac{\sqrt{4\pi}^N}{\sqrt{4\pi}^N} = 1. \tag{1.9}$$

This function naturally must remain within the basis set, as otherwise the basis set would be unable to describe constant functions.
α is denoted as a comma-separated list of positions, which shall always be in a specific sequence from $\mathbf{r}_1$ over $\mathbf{r}_2$ to $\mathbf{r}_N$, e.g. $\alpha = (\mathbf{r}_2, \mathbf{r}_5, \mathbf{r}_6, \mathbf{r}_7)$, which would mean in plain words "Only the spherical harmonics associated with positions $\mathbf{r}_2, \mathbf{r}_5, \mathbf{r}_6$ and $\mathbf{r}_7$ have orders $l^{\mathbf{r}_2}, l^{\mathbf{r}_5}, l^{\mathbf{r}_6}, l^{\mathbf{r}_7} \neq 0$". A specific sequence has to be chosen to avoid ambiguity. These atoms form a "cluster", which has a geometrical shape given by the n_α position vectors $\mathbf{r}_i$ as sketched in Figure 1.1.

The new $\mathbf{l}_\alpha$-tuple and $\mathbf{m}_\alpha$-tuple is again defined as a comma-separated list of values, which shall have only n_α entries and all entries of $\mathbf{l}_\alpha$ are larger than zero. To lighten the notation a bit, the association of the values l_i to the atoms shall be solely given by the position in this list: The first entry is assigned to the first position in the list α, the last entry to the last position in the list α, etc.
As an example, consider a three-atomic system with atoms at positions $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$, and a configuration $\Omega = (\mathbf{e}_1^{\mathbf{r}_1}, \mathbf{e}_2^{\mathbf{r}_2}, \mathbf{e}_3^{\mathbf{r}_3})$. Some examples for basis functions (not related to Figure 1.1) are then:

$$\begin{aligned}
\Phi_{(\mathbf{r}_1),(2)(1)} &= \sqrt{4\pi} Y^{21}(\mathbf{e}_1^{\mathbf{r}_1}), \\
\Phi_{(\mathbf{r}_1,\mathbf{r}_2),(1,3)(-1,1)} &= 4\pi Y^{1-1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{31}(\mathbf{e}_2^{\mathbf{r}_2}), \\
\Phi_{(\mathbf{r}_2,\mathbf{r}_3),(1,3)(-1,1)} &= 4\pi Y^{1-1}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{31}(\mathbf{e}_3^{\mathbf{r}_3}), \\
\Phi_{(\mathbf{r}_1,\mathbf{r}_2),(3,1)(-1,1)} &= 4\pi Y^{3-1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{11}(\mathbf{e}_2^{\mathbf{r}_2}), \\
\Phi_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3),(2,1,1)(-1,1,1)} &= \sqrt{4\pi}^3 Y^{2-1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{11}(\mathbf{e}_2^{\mathbf{r}_2}) \times \\
&\quad \times Y^{11}(\mathbf{e}_3^{\mathbf{r}_3}).
\end{aligned} \tag{1.10}$$

[4] $\eta \geq n_\alpha$ stands for the n_αth position where l is unqueal to zero.

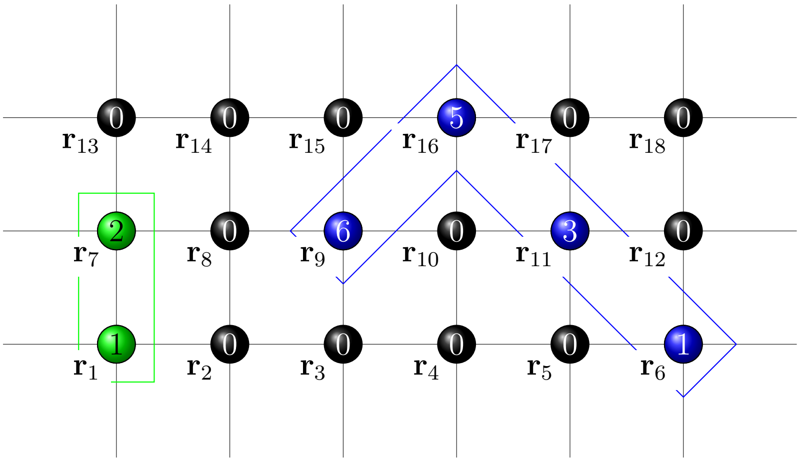

Figure 1.1.: Two-dimensional examples of clusters. The colour and the boxes identify different possible clusters, the numbers identify the corresponding l-values. All atoms at positions that do not belong to a cluster have the l value 0. For the green cluster: $\alpha = (\mathbf{r}_1, \mathbf{r}_7)$, $\mathbf{l}_\alpha = (1, 2)$; for the blue cluster: $\alpha = (\mathbf{r}_6, \mathbf{r}_9, \mathbf{r}_{11}, \mathbf{r}_{16})$, $\mathbf{l}_\alpha = (1, 6, 3, 5)$. The m_i have been omitted in the interest of brevity.

Some examples for *wrong* notation, or easy mistakes, are:

$$\begin{aligned}
\Phi_{(\mathbf{r}_1,\mathbf{r}_2),(1,0)(-1,0)} &\rightarrow \text{value of 0 in the l-tuple,} \\
\Phi_{(\mathbf{r}_3,\mathbf{r}_1),(1,3)(-1,1)} &\rightarrow \begin{array}{l}\text{the sequence of positions}\\ \text{in the cluster is wrong}\end{array}, \\
\Phi_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3),(2,1,1)(-1,1,2)} &\rightarrow |m^{\mathbf{r}_3}| = 2 > |l^{\mathbf{r}_3}| = 1\ .
\end{aligned} \tag{1.11}$$

Because the clusters are just convenient representations of the $\mathbf{l}$-tuple, the functions $\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha}$ are orthogonal in α, α' regarding the scalar product (1.4), and the completeness relation now requires a sum over all possible α:

$$\begin{aligned}
\langle \Phi_{\alpha'\mathbf{l}'_\alpha\mathbf{m}'_\alpha} | \Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha} \rangle &= \delta_{\alpha',\alpha}\delta_{\mathbf{l}'_\alpha,\mathbf{l}_\alpha}\delta_{\mathbf{m}'_\alpha,\mathbf{m}_\alpha}, \\
\sum_\alpha \sum_{\mathbf{l}_\alpha} \sum_{\mathbf{m}_\alpha} \left(\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha}(\tilde{\Omega})\right)^* \Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha}(\Omega) &= \delta(\tilde{\Omega} - \Omega) \\
&= \delta(\tilde{\mathbf{e}}^A - \mathbf{e}^A) \dots \delta(\tilde{\mathbf{e}}^N - \mathbf{e}^N).
\end{aligned} \tag{1.12}$$

$\mathbf{l}'_\alpha$ may refer to a different tuple to the same cluster α or to an l-tuple to another cluster α'. Therefore, it should technically be written as $\mathbf{l}'_{\alpha'}$, however this is very hard to read, e.g., in the Kronecker delta $\delta_{\mathbf{l}'_{\alpha'},\mathbf{l}_\alpha}$.

1.3.4. Symmetry Under Global Rotations of the Directions $\{\mathrm{e}_i\}$

Physical quantities have a well defined behaviour under rotations of the system investigated. E.g., the Hamiltonian of the condensed matter problem without spin orbit coupling transforms as a scalar, angular momenta or magnetic moments transform as vectors, etc. This behaviour has to be reproduced when the quantity is expanded in a basis set, which means that the basis used for the expansion has to have the same behaviour under rotations of the whole system. In the case of magnetic moment directions $\{\mathbf{e}_i\}$ as variables, these rotations of the whole system correspond to collective, simultaneous rotations of all $\mathbf{e}_i$.

The spherical harmonics form a basis set for the irreducible representations of the pure rotation group SO(3), that is under pure rotations of the arguments they have a well defined transformation behaviour:

$$Y^{lm}(\mathsf{R}^{-1}(\boldsymbol{\omega})\mathbf{e}) = \sum_{m'=-l}^{l} D^{(l)}_{m'm}(\boldsymbol{\omega}) Y^{lm'}(\mathbf{e}), \tag{1.13}$$

where $\boldsymbol{\omega}$ describes the rotation axis and angle, and the "Wigner D matrices" $D^{(l)}_{m'm}$ may be found in the literature, e.g., in [16]. This rotational behaviour can be directly compared to that of scalars, cartesian vectors, etc. to find the necessary basis set.
However, according to group theory, products of basis functions of irreducible representations of the same group do no longer form basis functions of irreducible representations of this group, but for reducible representations. This means that the rotational behaviour of the products of the spherical harmonics is not the same as in Eqn. (1.13). The basis functions of the irreducible representations can be calculated as linear combinations of the reducible basis functions. These basis functions of irreducible representations have again a well defined behaviour under rotations. This process has been carried out in [1, 15], and it was found that the linear combinations are only of basis functions $\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha}$ to the same α and $\mathbf{l}_\alpha$, but different $\mathbf{m}_\alpha$. One way to formulate the irreducible basis functions is

$$\mathcal{Y}^{l\mu}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\Omega) = \sum_{\mathbf{m}_\alpha} \mathrm{C}^{K_1}_{J_1,J_2}\mathrm{C}^{K_2}_{K_1,J_3}\ldots\mathrm{C}^{l\mu}_{K_{n_\alpha}-2,J_{n_\alpha}}\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha}(\Omega) \tag{1.14}$$

where the $\mathrm{C}^{K}_{J_i,J_j}$ are the Clebsch-Gordan coefficients, familiar from the coupling of two angular momenta. $\mathrm{C}^{K}_{J_i,J_j}$ is a shorthand for $\mathrm{C}^{k\mu}_{j_in_i,j_jn_j}$.
The j_i in the Clebsch-Gordan coefficient are the numerical values l_i out of the $\mathbf{l}_\alpha$-tuple, sorted in descending order so that $j_1 \geq j_2 \geq \ldots \geq j_{n_\alpha}$, and $j_1 = \max(\mathbf{l}_\alpha)$, $j_{n_\alpha} = \min(\mathbf{l}_\alpha)$.
The n_i are the m_i, which have to be chosen such that the pairs l_i, m_i remain intact.
In principle, the "angular momenta"[5] l_i may be coupled in any order to give the final

[5] The term is used due to the history of the Clebsch-Gordan coefficients. The discussion here is not about angular momenta, and just borrows from the calculus of angular momenta.

angular momentum l, and starting with the largest l_i is just a convention. This convention does not determine the intermediate angular momenta k_i, which remain a free parameter of the resulting basis functions $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$.
The k_i are bounded, because a Clebsch-Gordan coefficient $\mathrm{C}^{K}_{J_i,J_j}$ is only nonzero if the arguments j_i, j_j, k fulfill the requirement $j_i + j_j \geq k \geq |j_i - j_j|$. They are written in a "**k**-tuple", a comma-separated list of $n_\alpha - 2$ values: $\mathbf{k} = (k_1, k_2, \ldots, k_{n_\alpha-2})$. Contrary to the $\mathbf{l}_\alpha, \mathbf{m}_\alpha$-tuples, the k_i are *not* associated with specific positions. Instead, the sequence in the **k**-tuple gives the sequence of intermediate angular momenta for the successive couplings, with the first k_i associated with the first coupling, etc.
The intermediate μ_i are fixed for each given $\mathbf{m}_\alpha$, as a Clebsch-Gordan coefficient $\mathrm{C}^{k_i\mu_i}_{j_in_i,j_jn_j}$ is only nonzero if $n_i + n_j = \mu_i$. This means the first Clebsch-Gordan coefficient in (1.14) is only nonzero if $\mu_1 = n_1 + n_2$, and the the second Clebsch-Gordan coefficient is only nonzero if $\mu_2 = \mu_1 + n_3 = n_1 + n_2 + n_3$, etc.
The functions $\mathcal{Y}^{l\mu}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ are basis functions of the representation $D^{(l)}$, and transform under simultaneous rotations of all arguments as

$$\mathcal{Y}^{l\mu}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\mathsf{R}^{-1}(\boldsymbol{\omega})\{\mathbf{e}_i\}) = \sum_{\mu'=-l}^{l} D^{(l)}_{\mu'\mu}(\boldsymbol{\omega})\mathcal{Y}^{l\mu'}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\{\mathbf{e}_i\}). \tag{1.15}$$

The basis set for one specific physical quantity is given by the subset of $\{\mathcal{Y}^{l\mu}_{\alpha\mathbf{l}_\alpha\mathbf{k}}\}$ that has the same behaviour under rotations as this physical quantity. E.g., the basis set for scalars is $\{\mathcal{Y}^{00}_{\alpha\mathbf{l}_\alpha\mathbf{k}}\}$, and the basis set for vectors is $\{\mathcal{Y}^{11}_{\alpha\mathbf{l}_\alpha\mathbf{k}}\}$, $\{\mathcal{Y}^{10}_{\alpha\mathbf{l}_\alpha\mathbf{k}}\}$, $\{\mathcal{Y}^{1-1}_{\alpha\mathbf{l}_\alpha\mathbf{k}}\}$ (some details for cartesian vectors can be found in Section A.6).
The rest of this work concentrates on scalar functions, which means the relevant basis functions are $\mathcal{Y}^{00}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ (written as just $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ in the following).

$$\begin{aligned}
\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}} &= \sum_{\mathbf{m}_\alpha} \mathrm{C}^{K_1}_{J_1,J_2}\mathrm{C}^{K_2}_{K_1,J_3}\ldots\mathrm{C}^{K_{n_\alpha}-2}_{K_{n_\alpha}-3,J_{n-1}}\mathrm{C}^{00}_{K_{n_\alpha}-2,J_{n_\alpha}}\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha} \\
&= \sum_{\mathbf{m}_\alpha} \frac{(-)^{j_{n_\alpha}+n_{n_\alpha}}}{\sqrt{2j_{n_\alpha}+1}}\mathrm{C}^{K_1}_{J_1,J_2}\mathrm{C}^{K_2}_{K_1,J_3}\ldots\mathrm{C}^{l_{n_\alpha}-n_{n_\alpha}}_{K_{n_\alpha}-3,J_{n_\alpha-1}}\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha} \\
&:= \sum_{\mathbf{m}_\alpha} \begin{pmatrix} j_1 & \ldots & j_{n_\alpha} \\ n_1 & \ldots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha},
\end{aligned} \tag{1.16}$$

where the second step used the explicit formula[6] (A.14) for $\mathrm{C}^{00}_{K_{n_\alpha}-2,J_{n_\alpha}}$. The **k**-tuple now has only $n_\alpha - 3$ entries.
The coefficient defined by the last equation is a generalized version of Wigner's 3j-symbol,

[6] There is an error in the corresponding Eqn. (3.61) of[1], where the prefactor is written as $\frac{(-)^{l_{n_\alpha}-m_{n_\alpha}}}{\sqrt{2l_{n_\alpha}-m_{n_\alpha}}}$ instead of the correct $\frac{(-)^{j_{n_\alpha}-n_{n_\alpha}}}{\sqrt{2j_{n_\alpha}+1}}$. See also the next footnote.

which arises when three angular momenta couple to a resulting 0. It is thus called a "generalized Wigner coefficient"[7]. The coefficients $\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}}$ are only nonzero if

$$\begin{aligned} 2j_i - \sum_{j=1}^{n_\alpha} j_j &\leq 0 \;\; \forall i, \\ \text{and } k_{i-1} + j_{i+1} \geq k_i &\geq |k_{i-1} - j_{i+1}|\,, \\ \text{and } \sum n_i &= 0, \end{aligned} \tag{1.17}$$

where the second relation for the k_i does not hold for $i = 1$ and $i = n_\alpha - 2$. The first and last equation may be written in the l_i and the m_i as well, and therefore restrict the $\mathbf{l}_\alpha$ and $\mathbf{m}_\alpha$-tuple.

Investigating not only the pure rotations but the whole O(3) group (that is, SO(3) symmetry and the global inversion of all $\mathbf{e}_i$, which corresponds to symmetry under time inversion) leads to the further condition [15]:

$$\sum_{i=1}^{n_\alpha} l_i = \text{even, or equally } \sum_{i=1}^{n_\alpha} j_i = \text{even}, \tag{1.18}$$

which also means the basis functions are now real functions [15]. In case of $n_\alpha \leq 3$, no intermediate angular momenta k_i are necessary (for the formula, see the appendix, section A.2. For details, see [15].). For the case of $n_\alpha = 2$, the total angular momentum has to be the same for both spherical harmonics. For the case of $n_\alpha = 1$, the only basis function invariant under global rotations is proportional to $\mathcal{Y}^{00}$, however $l = 0$ was already excluded from the basis set.

These basis functions are complete for functions which transform as scalars under O(3), in the sense that any such function may be expanded in the $\{\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}\}$. The basis functions are also orthogonal regarding the scalar product (1.4):

$$\langle \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} | \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} \rangle = \delta_{\alpha,\alpha'} \delta_{\mathbf{l}_\alpha, \mathbf{l}'_\alpha} \delta_{\mathbf{k},\mathbf{k}'}. \tag{1.19}$$

Some examples to clarify the notation (for the formulas of the generalized Wigner coefficients, see Section A.2), using a generic configuration $\Omega = (\mathbf{e}_1^{\mathbf{r}_1}, \mathbf{e}_2^{\mathbf{r}_2}, \mathbf{e}_3^{\mathbf{r}_3}, \dots, \mathbf{e}_N^{\mathbf{r}_N})$:

$$\begin{aligned} \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2),(3,3)} &= \sum_{m_1} \begin{pmatrix} 3 & 3 \\ m_1 & m_2 \end{pmatrix} \Phi_{(\mathbf{r}_1,\mathbf{r}_2),(3,3),(m_1,m_2)} \\ &= \sum_{m_1} 4\pi \frac{(-)^{-m_1}}{\sqrt{7}} Y^{3m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{3-m_1}(\mathbf{e}_2^{\mathbf{r}_2}), \end{aligned}$$

[7] [1] writes these coefficients as $\begin{pmatrix} l_1 & \dots & l_{n_\alpha} \\ m_1 & \dots & m_{n_\alpha} \end{pmatrix}^{B_0}_{\mathbf{k}}$, where the B_0 indicates the sorting procedure. This has proven to be rather misleading, (see the prefactor error above), as the i of the l_i just differentiates numerical values, and their sequence in $\mathbf{l}_\alpha$ needs not be the descending sequence *chosen* in the Clebsch-Gordan coefficients.

$$
\begin{aligned}
\mathcal{Y}_{(\mathbf{r}_6,\mathbf{r}_7,\mathbf{r}_{12}),(1,2,1)} &= \sum_{m_1=-1}^{m_1=1}\sum_{m_2=-2}^{m_2=2}\begin{pmatrix} 2 & 1 & 1 \\ m_2 & m_1 & m_3 \end{pmatrix} \\
&\times \Phi_{(\mathbf{r}_6,\mathbf{r}_7,\mathbf{r}_{12})(1,2,1)(m_1,m_2,m_3)} \\
&= \sum_{m_1,m_2} \sqrt{4\pi}^3 \frac{(-)^{1-m_3}}{\sqrt{3}} \mathrm{C}^{1-m_3}_{2m_2,1m_1} \\
&\times Y^{1m_1}(\mathbf{e}_6^{\mathbf{r}_6}) Y^{2m_2}(\mathbf{e}_7^{\mathbf{r}_7}) Y^{1m_3}(\mathbf{e}_{11}^{\mathbf{r}_{12}}), \\
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_4,\mathbf{r}_6,\mathbf{r}_{13}),(3,2,1,2),(1)} &= \sum_{m_1,m_2,m_3} \begin{pmatrix} 3 & 2 & 2 & 1 \\ m_1 & m_2 & m_4 & m_3 \end{pmatrix}^{(1)} \\
&\times \Phi_{(\mathbf{r}_1,\mathbf{r}_4,\mathbf{r}_6,\mathbf{r}_{13}),(3,2,1,2),(m_1,m_2,m_3,m_3)} \\
= \sum_{m_1,m_2,m_3} & (4\pi)^2 \frac{(-)^{1-m_3}}{\sqrt{3}} \mathrm{C}^{1\mu}_{3m_1,2m_2} \mathrm{C}^{1-m_3}_{1\mu,2m_4} \\
&\times Y^{3m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{2m_2}(\mathbf{e}_4^{\mathbf{r}_4}) \\
&\times Y^{1m_3}(\mathbf{e}_6^{\mathbf{r}_6}) Y^{2m_4}(\mathbf{e}_{13}^{\mathbf{r}_{13}}).
\end{aligned}
\tag{1.20}
$$

The summations run only over $n_\alpha - 1$ elements of $\mathbf{m}$, because the condition $\sum m_i = 0$ of the generalized Wigner coefficients uniquely determines one m. E.g., for the three-atomic basis function $m_3 = -m_2 - m_1$.
Some examples for the orthogonality relation:

$$
\begin{aligned}
\langle \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2),(3,3)} | \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2),(3,3)} \rangle &= 1, \\
\langle \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2),(3,3)} | \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_3),(3,3)} \rangle &= 0, \\
\langle \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2),(3,3)} | \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2),(2,2)} \rangle &= 0, \\
\langle \mathcal{Y}_{(\mathbf{r}_6,\mathbf{r}_7,\mathbf{r}_{11}),(2,2,2)} | \mathcal{Y}_{(\mathbf{r}_6,\mathbf{r}_7,\mathbf{r}_{11}),(1,1,2)} \rangle &= 0, \\
\langle \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_4,\mathbf{r}_6,\mathbf{r}_{13}),(3,2,1,2),(1)} | \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_4,\mathbf{r}_6,\mathbf{r}_{15}),(3,2,1,2),(1)} \rangle &= 0.
\end{aligned}
\tag{1.21}
$$

1.4. Space Group Symmetry

The energy of a crystal is invariant under operations of the space group of the crystal lattice. Therefore, the basis functions should be chosen to be invariant under these operations. Similar to the rotational symmetry, the space group symmetry will induce linear dependencies between basis functions if it is not taken into account properly. This would mean the basis functions were no longer orthogonal, and numerical investigations are then ambiguous.
This section is based on work published already in [1]. This work is technically correct. However, some complications that arise due to the combination of the space group and

the SO(3) symmetry have been eventually overlooked, as [1] considers space group symmetry first. This problem was discovered, and largely solved, by R.Singer, who described this in some unpublished and rather terse notes which this work is based upon. These notes limited themselves to symmorphic space groups.
This chapter tries to give a comprehensive presentation, which also requires no differentiation between symmorphic and non-symmorphic space groups.

1.4.1. Some Group Theory

This section makes use of some concepts and theorems from group representation theory. These concepts are introduced below, after which the three necessary theorems are stated. For proofs of these theorems see, e.g., [17, 18][8].
A representation D^p of a group is defined as a mapping of the group elements to a set of d_p-dimensional matrices $\{\mathsf{D}^p\}$ that retains the group multiplication table under matrix multiplication. Two such representations are "equivalent" if there exists some linear transformation U between the matrices of the two representations.
The "projection operator" of a given representation[17, 18],

$$\hat{P}^p_{mn} = \frac{d_p}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \mathsf{D}^p_{mn} g \circ . \tag{1.22}$$

is an idempotent operator, $\hat{P}^p_{mn} = \left(\hat{P}^p_{mn}\right)^2$. The basis functions $\{\Psi\}$ of a given representation are the functions that transform by action of group elements exactly like the matrices of the representation:

$$g \circ \Psi_n = \sum_{m=1}^{d_p} \mathsf{D}^p_{mn}(g) \Psi_m, \tag{1.23}$$

where $g\circ$ indicates the action of the space group operation g to the function and p labels the representations. Note that one representation can have many different basis functions, because the definition of a basis function only requires it to transform in a specific way. The action of g on the SCE basis functions is detailed later in Section 1.4.3.
A rather simple but here very important case is the "identity" or "identical" representation D^1, which has dimension $d_1 = 1$, and the corresponding 1x1 matrix is $\mathsf{D}^1_{11} = 1$. Therefore, all basis functions of the identity representation are by definition invariant under application of any symmetry operation g.

[8] [17] limits itself unnecessarily to functions out of L^2, the space of quadratically integrable functions. Also, [17] defines the functions too narrowly as functions of variables out of R^3. According to [18], the theorems hold for any Hilbert space.

Only for the projection operator of the identical representation, the following relation holds for any $g_0 \in \mathcal{G}$:

$$\hat{P}^1_{11} g_0 \circ = \left(\sum_{g \in \mathcal{G}} g \circ \right) g_0 \circ = \sum_{g \in \mathcal{G}} (g \circ g_0) \circ = \sum_{\tilde{g} \in \mathcal{G}} \tilde{g} \circ = \hat{P}^1_{11}, \tag{1.24}$$

because a group must be closed under group multiplication $\circ$.

Theorem I: Any function Φ out of a Hilbert space may be expanded in basis functions of the inequivalent irreducible representations of a group:

$$\Phi = \sum_p \sum_{k=1}^{d_p} c^p_k \Psi^p_{k\Phi}, \tag{1.25}$$

where the required specific basis functions $\Psi^p_{k\Phi}$ out of all the possible basis functions Ψ^p_k for each representation depend on the function Φ.

Theorem II: All basis functions to different representations, and to different rows k of a specific representation are orthogonal

$$\langle \Psi^p_k | \Psi^{p'}_{k'} \rangle = \delta_{pp'} \delta_{kk'}, \tag{1.26}$$

where $\langle | \rangle$ is the scalar product of the Hilbert space. When a basis function $\tilde{\Psi}^p_k$ is a basis function to an equivalent representations $\tilde{D}^p$ of D^p, the right hand side of (1.26) needs not be unity.

Theorem III: The projection operator of a representation projects any function on the basis function of the representation, that is

$$\hat{P}^p_{mn} \Phi = c^p_n \Psi^p_{m\Phi}. \tag{1.27}$$

This has the corollary that all basis functions are eigenfunctions of the projection operator. and that the eigenfunctions of the projection operator are basis function of the given representation[9].

[9]Proof: Presume that there exists an eigenfunction f of $\hat{P}^p_{mn}$ that is not a basis function. Use the expansion (1.25) of f in equation (1.27). This gives $\hat{P}^p_{mn} f = c^p_n \Psi^p_{mf}$. However, because f is presumed to be an eigenfunction, $\hat{P}^p_{mn} f = af$. Comparison of the right hand sides contradicts the presumption that f is not a basis function.

1.4.2. Application to the SCE

As the energy function $E(\{\mathbf{e}_i\})$ is invariant under space group operation, it is a basis function of the identical representation. It may be expanded in the SCE basis functions (1.16),

$$E = \sum_{\alpha \mathbf{l}_\alpha \mathbf{k}} J_{\alpha \mathbf{l}_\alpha \mathbf{k}} \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}, \tag{1.28}$$

with interaction coefficients $J_{\alpha \mathbf{l}_\alpha \mathbf{k}} = \langle E | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle$.
Also, all SCE basis functions may be expanded in basis functions of the irreducible representations of the space group,

$$\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} = \sum_p \sum_{k=1}^{d_p} c^p_{k\alpha \mathbf{l}_\alpha \mathbf{k}} \Psi^p_{k\alpha \mathbf{l}_\alpha \mathbf{k}}. \tag{1.29}$$

The notation $\Psi^p_{k\alpha \mathbf{l}_\alpha \mathbf{k}}$ is to remind the reader that different SCE basis functions may need to be expanded by different basis functions for the space group.
As the energy function is a basis function of the identical representation of the space group, it follows from theorem II that:

$$\begin{aligned} J_{\alpha \mathbf{l}_\alpha \mathbf{k}} = \langle E | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle = \langle E \hat{P}^1_{11} | \sum_p \sum_{k=1}^{d_p} c^p_{k\alpha \mathbf{l}_\alpha \mathbf{k}} \Psi^p_{k\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle \\ = c^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}} \langle E | \Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle. \end{aligned} \tag{1.30}$$

The functions $\Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}}$ form a basis set for the energy, as

$$\begin{aligned} E = \hat{P}^1_{11} E = \sum_{\alpha \mathbf{l}_\alpha \mathbf{k}} J_{\alpha \mathbf{l}_\alpha \mathbf{k}} \hat{P}^1_{11} \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \\ \overset{(1.30),(1.27)}{=} \sum_{\alpha \mathbf{l}_\alpha \mathbf{k}} \left(c^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}} \right)^2 \langle E | \Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle \Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}}. \end{aligned} \tag{1.31}$$

Because the projection operator for D^1,

$$\hat{P}^1_{11} = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} g \circ, \tag{1.32}$$

is an idempotent operator, it has only eigenvalues 0 and 1. Eigenfunctions to the eigenvalue 0 are part of the kernel of $\hat{P}^1_{11}$, and they do not contribute to the energy expansion because the expansion coefficients $J_{\alpha \mathbf{l}_\alpha \mathbf{k}} = \langle E | \Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle = \langle E | \hat{P}^1_{11} | \Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle$ are zero if $\Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}}$ is an eigenfunction to the eigenvalue 0.
While the set of functions $\left\{ \Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}} \right\}$ for all $\alpha, \mathbf{l}_\alpha, \mathbf{k}$ form a basis set for the energy, they

are neither linear independent nor orthogonal, as will be shown in Section 1.4.6. However, one can construct a linear independent basis set for the eigenfunctions of $\hat{P}^1_{11}$ along the following arguments:
Any basis function Ψ^1_{1m} to the identical representation may be expanded in the SCE basis functions as the latter represent a complete basis:

$$\Psi^1_{1m} = \sum_{\alpha \mathbf{l}_\alpha \mathbf{k}} b^m_{\alpha \mathbf{l}_\alpha \mathbf{k}} \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}. \tag{1.33}$$

Per definition, the Ψ^1_{1m} are eigenfunctions to $\hat{P}^1_{11}$:

$$\hat{P}^1_{11} \Psi^1_{1m} = \Psi^1_{1m}. \tag{1.34}$$

Inserting Eqn. (1.33), and calculating the scalar product with $\mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'}$ yields

$$\langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \Psi^1_{1m} \rangle = \sum_{\alpha \mathbf{l}_\alpha \mathbf{k}} b^m_{\alpha \mathbf{l}_\alpha \mathbf{k}} \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle. \tag{1.35}$$

Only the eigenfunctions to the eigenvalue 1 are of interest. Therefore:

$$\begin{aligned} &\langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \Psi^1_{1m} \rangle = \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \Psi^1_{1m} \rangle = \\ &= \sum_{\alpha \mathbf{l} k} b^m_{\alpha \mathbf{l}_\alpha \mathbf{k}} \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle = \sum_{\alpha \mathbf{l}_\alpha \mathbf{k}} b^m_{\alpha \mathbf{l}_\alpha \mathbf{k}} \delta_{\alpha,\alpha'} \delta_{\mathbf{l}_\alpha, \mathbf{l}'_\alpha} \delta_{\mathbf{k},\mathbf{k}'}, \end{aligned} \tag{1.36}$$

where the orthogonality of the SCE basis functions was used. One can construct an eigenvalue problem for the coefficients $b^m_{\alpha \mathbf{l} \mathbf{k}}$ by subtracting the right hand side of Eqn. (1.36) from Eqn. (1.35):

$$\sum_{\alpha \mathbf{l}_\alpha \mathbf{k}} b^m_{\alpha \mathbf{l}_\alpha \mathbf{k}} \left[\langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle - \delta_{\alpha,\alpha'} \delta_{\mathbf{l}_\alpha, \mathbf{l}'_\alpha} \delta_{\mathbf{k},\mathbf{k}'} \right] = 0. \tag{1.37}$$

Solving this eigenvalue problem will yield a set of orthogonal, linear independent eigenvectors to the eigenvalue 1, which then form an orthogonal, linear independent basis set for the energy.

1.4.3. Action on the Basis Functions

Before the matrix elements of the projection operator regarding the SCE basis functions can be calculated, the effect of a space group symmetry operation acting on a SCE basis function must be defined. For the purpose of this work, a crystal shall consist of only one type of undistinguishable atoms, which are ordered according to a Bravais lattice.

Action on the $\Phi_{\mathbf{lm}}$

This discussion shall begin with the products $\Phi_{\mathbf{lm}}(\Omega)$, defined back in Eqn. (1.5), before the clusters were defined. There, N different atoms were situated at N different positions $(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$. The configuration was given by, e.g., $(\mathbf{e}_1^{\mathbf{r}_1}, \mathbf{e}_2^{\mathbf{r}_2}, \ldots)$, the assignment of the orders of the spherical harmonics by, e.g., $\mathbf{l} = (l_1^{\mathbf{r}_1}, l_2^{\mathbf{r}_2}, l_3^{\mathbf{r}_3}, \ldots)$ and $\mathbf{m} = (m_1^{\mathbf{r}_1}, m_2^{\mathbf{r}_2}, m_3^{\mathbf{r}_3}, \ldots)$. The basis function $\Phi_{\mathbf{lm}}$ to these examples is:

$$\Phi_{\mathbf{lm}}(\Omega) = \sqrt{4\pi}^N Y^{l_1^{\mathbf{r}_1} m_1^{\mathbf{r}_1}}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2^{\mathbf{r}_2} m_2^{\mathbf{r}_2}}(\mathbf{e}_2^{\mathbf{r}_2}) \ldots . \tag{1.38}$$

If the atoms are undistinguishable, and arranged in a symmetric way (for our purposes, on a Bravais lattice), there exist space group operations g which transform the points of the Bravais lattice, which for our purpose shall also be the positions of the undistinguishable atoms, $\mathbf{r}_1, \mathbf{r}_2, \ldots$, into each other. This will lead to a group action on the $\Phi_{\mathbf{lm}}$. Such a group action on a function maps one function f to a different function $\tilde{f} = g \circ f$ out of the same function space. The conventional way [19] to define a group action $g\circ$ is to define

$$(g \circ f)(x) = f(g^{-1} \circ x) \tag{1.39}$$

via a given group action $g\circ$ on the arguments x of the functions[10]. To find the modified function $(g \circ f)$ of the original argument x, one has to insert the modified argument $g^{-1} \circ x$ into the original function f. The resulting expression must then be transformed until it is a function of the original argument x, and this function can then be identified with $g \circ f$. From a practical point of view, this transformation is necessary because many useful properties of the functions (orthogonality, etc.) are only defined for functions of the same argument.

The argument of the $\Phi_{\mathbf{lm}}$ is the configuration, $\Omega = (\mathbf{e}_1^{\mathbf{r}_1}, \mathbf{e}_2^{\mathbf{r}_2}, \ldots)$. To each point $\mathbf{r}_i$ some direction $\mathbf{e}$ is assigned by the "assignment function" $\mathbf{e}(\mathbf{r}_i)$. Then, by Eqn. (1.39), $(g \circ \mathbf{e})(\mathbf{r}_i) = \mathbf{e}(g^{-1} \circ \mathbf{r}_i)$, which means that after the action of g, the orientation of the magnetic moment at the position $\mathbf{r}_i$ is the orientation which was originally assigned to the position $g^{-1} \circ \mathbf{r}_i$.

It is somewhat more useful within our notation to ask where the magnetic moment direction that was assigned to $\mathbf{r}_i$ is after the group action. That means solving the equation $(g \circ \mathbf{e})(\mathbf{r}_j) = \mathbf{e}(\mathbf{r}_i)$, or according to Eqn. (1.39), $\mathbf{e}(g^{-1}\mathbf{r}_j) = \mathbf{e}(\mathbf{r}_i)$. This is solved by $\mathbf{r}_j = g \circ \mathbf{r}_i$.

Therefore, the group action on the configuration can be written as $g \circ \Omega = (\mathbf{e}_1^{g \circ \mathbf{r}_1}, \mathbf{e}_2^{g \circ \mathbf{r}_2},$

[10] These arguments need not be real numbers, as suggested by the usage of x. [19] gives an example for a function on the vertices of a square. Also, the group action on the arguments and the group action on the function are, in principle, two different mappings - one maps an argument to another argument, and the other maps a function to another function. Many authors therefore resort to different symbols for the group action on the argument, the group action on the function, etc. The author thinks this is rather distracting, and prefers the use of parentheses where ambiguity might arise.

$\ldots$), which moves the numerical values $(\mathbf{e}_1, \mathbf{e}_2, \ldots)$ from their original positions to new positions in the crystal, where the new positions are given by the group action on the old positions.
With this group action on the argument of the $\Phi_{\mathbf{lm}}$, Eqn. (1.39) yields for the group action on the $\Phi_{\mathbf{lm}}$:

$$\begin{aligned}(g \circ \Phi_{\mathbf{lm}})(\Omega) &= \Phi_{\mathbf{lm}}(g^{-1} \circ \Omega) \\ (g \circ \Phi_{\mathbf{lm}})(\mathbf{e}_1^{\mathbf{r}_1}, \ldots) &= \Phi_{\mathbf{lm}}(\mathbf{e}_1^{g^{-1}\circ\mathbf{r}_1}, \mathbf{e}_2^{g^{-1}\circ\mathbf{r}_2}, \ldots) \\ &= \sqrt{4\pi}^N Y^{l_1^{\mathbf{r}_1} m_1^{\mathbf{r}_1}}(\mathbf{e}_i^{\mathbf{r}_1}) Y^{l_2^{\mathbf{r}_2} m_2^{\mathbf{r}_2}}(\mathbf{e}_j^{\mathbf{r}_2}) \\ &\times \ldots Y^{l_h^{g^{-1}\circ\mathbf{r}_1} m_h^{g^{-1}\circ\mathbf{r}_1}}(\mathbf{e}_1^{g^{-1}\circ\mathbf{r}_1}) \\ &\times Y^{l_k^{g\circ\mathbf{r}_2} m_k^{g\circ\mathbf{r}_2}}(\mathbf{e}_2^{g^{-1}\circ\mathbf{r}_2}),\end{aligned} \tag{1.40}$$

where $\mathbf{e}_i$ and $\mathbf{e}_j$ are the orientations of the magnetic moments that were associated with $g \circ \mathbf{r}_1, g \circ \mathbf{r}_2$ in the original configuration Ω, and are now associated with $\mathbf{r}_1, \mathbf{r}_2$.
l_h, m_h and l_k, m_k are the orders of the spherical harmonics which are associated with $g^{-1} \circ \mathbf{r}_1, g^{-1} \circ \mathbf{r}_2$ by the l-tuple $\mathbf{l}$.
The next step is to transform back to the original configuration, where $\mathbf{e}_1$ is associated with $\mathbf{r}_1$, $\mathbf{e}_2$ is associated with $\mathbf{r}_2$, etc., without changing the function on the right hand side. This means that the same numerical argument (e.g., $\mathbf{e}_1$) has to enter the spherical harmonic with the same orders (e.g., l_h, m_h). To do so, the numerical values l_1, m_1 need to be associated with $g \circ \mathbf{r}_1$, and the numerical values l_h, m_h with $\mathbf{r}_1$, etc. This can be achieved by applying the group action g to the $\mathbf{r}_i$ in the **l**- and **m**-tuple such that:

$$\begin{aligned} g \circ \mathbf{l} &:= (l_1^{g\circ\mathbf{r}_1}, l_2^{g\circ\mathbf{r}_2} \ldots), \\ g \circ \mathbf{m} &:= (m_1^{g\circ\mathbf{r}_1}, m_1^{g\circ\mathbf{r}_2} \ldots). \end{aligned} \tag{1.41}$$

With this definition,

$$\Phi_{g\circ\mathbf{l} g\circ\mathbf{m}}(\Omega) = \Phi_{\mathbf{lm}}(g^{-1} \circ \Omega), \tag{1.42}$$

And therefore, by Eqn. (1.39),

$$(g \circ \Phi_{\mathbf{lm}})(\Omega) = \Phi_{g\circ\mathbf{l} g\circ\mathbf{m}}(\Omega) \tag{1.43}$$

defines the group action on the $\Phi_{\mathbf{lm}}$ by the action on the tuples as given in (1.41).

Action on the $\Phi_{\alpha \mathbf{l}_\alpha \mathbf{m}_\alpha}$

The cluster α encoded the information which atoms were associated with a spherical harmonic with an order l unequal to zero. It was written as a comma-separated list of n_α positions, e.g., $\alpha = (\mathbf{r}_3, \mathbf{r}_5, \ldots,)$, with the convention that the positions always

appear in a specific sequence - $\mathbf{r}_1$ is always written before $\mathbf{r}_2$, etc.
The $\mathbf{l}_\alpha$-tuple ($\mathbf{m}_\alpha$-tuple) contained the information what the actual values of the l_i were at the postions listed in α, and was written as a comma-delimited list with n_α entries, e.g., $(l_1, l_2, , \ldots, l_{n_\alpha})$. If now the original $\mathbf{l}, \mathbf{m}$-tuples, which were defined for the whole system, change according to Eqn.(1.45), the orders unequal to zero will be moved to new positions given by $g \circ \mathbf{r}$. As the purpose of the cluster was to encode precisely this information, it has to change to

$$\alpha' = g \circ \alpha := \mathsf{S}(g \circ \mathbf{r}_3, g \circ \mathbf{r}_5, \ldots, g \circ \mathbf{r}_{\eta_\alpha}), \tag{1.44}$$

where S indicates a sorting or permutation of the $\mathbf{r}_i$ so that the smallest subscript of the $\mathbf{r}_i$ appears first. This has to be written explicitly, because the correct sequence (of $\mathbf{r}_1$ before $\mathbf{r}_2$, etc. in α) is not guaranteed after a symmetry operation has been applied[11].
The $\mathbf{l}_\alpha$-tuple is just a list of numerical values, which do not change by a symmetry operation. However, because the position in this list is tied to the spatial positions listed α, the same reordering as for the cluster has to be applied to sequence of these values in the $\mathbf{l}_\alpha$- and $\mathbf{m}_\alpha$-tuple of the cluster:

$$\begin{aligned} g \circ \mathbf{l}_\alpha &= \mathsf{S}(\mathbf{l}_\alpha), \\ g \circ \mathbf{m}_\alpha &= \mathsf{S}(\mathbf{m}_\alpha). \end{aligned} \tag{1.45}$$

With these definitions, the group action on the $\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha}$ is given by

$$\begin{aligned} (g \circ \Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha})(\Omega) &= \Phi_{g\circ\alpha g\circ\mathbf{l}_\alpha g\circ\mathbf{m}_\alpha}(\Omega) \\ &= \Phi_{g\circ\alpha\mathsf{S}(\mathbf{l}_\alpha)\mathsf{S}(\mathbf{m}_\alpha)}(\Omega). \end{aligned} \tag{1.46}$$

For an example, consider Figure 1.2. There, a molecule of six indistinguishable atoms forming a hexagon is depicted. The points of this hexagon can be transformed into each other by six counterclockwise rotations[12] through 0 degrees, 60 degrees, 120 degrees, etc. These rotations are labeled $C_6^0, C_6^1, C_6^2, \ldots, C_6^5$. The position vectors are $\mathbf{r}_1$ for the lower left atom, and $\mathbf{r}_2$ for the lower right atom, etc, proceeding counterclockwise around the hexagon.
This example uses an initial basis function $\Phi_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5),(3,1,2),(1,0,-1)}$, and an initial configuration (which one may write in two dimensions as just one angle, ϑ) of $\Omega = (\frac{\pi}{3}^{\mathbf{r}_1}, \frac{2\pi}{3}^{\mathbf{r}_2}, \pi^{\mathbf{r}_3}, \frac{4\pi}{3}^{\mathbf{r}_4}, \frac{5\pi}{3}^{\mathbf{r}_5}, 0^{\mathbf{r}_6})$ as depicted in the left column of Figure 1.2.
The first row of Figure 1.2 displays the transformation of the configuration, $g^{-1} \circ \Omega$

[11] One has to choose a specific sequence to ensure that there is only one way to write a basis function. With this convention, basis functions are equal if in the cluster and tuples the same values appear in the same sequence. Without this convention, one could write one basis functions as, e.g., $\Phi_{(\mathbf{r}_1,\mathbf{r}_2)(2,1)(-2,0)}$ or as $\Phi_{(\mathbf{r}_2,\mathbf{r}_1)(1,2)(0,-2)}$. This would complicate formulas and the numerical implementation.

[12] There are also symmetry operations that mirror the atoms, or invert the positions. These are omitted in the interest of brevity.

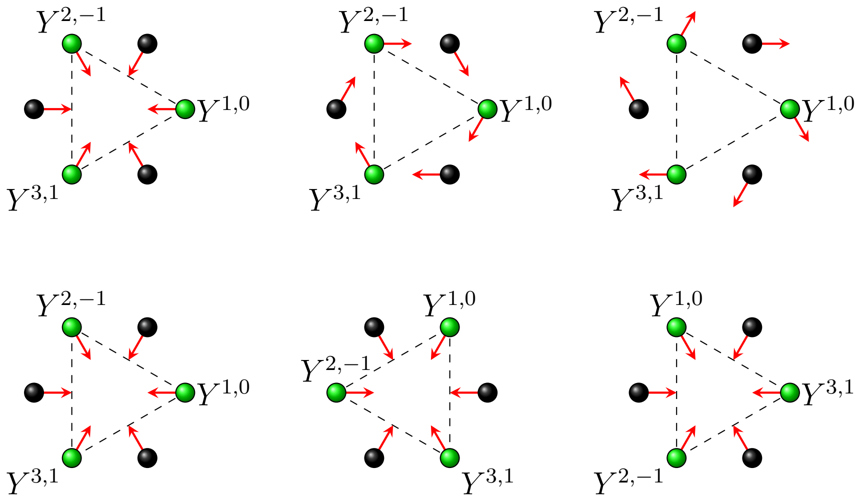

Figure 1.2.: Transformation of a basis function $\Phi_{\alpha\mathbf{lm}}$. Atoms in green are part of the current cluster α, and the associated spherical harmonics are written next to the atoms. Top row: $\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha}(g^{-1}\circ\Omega)$, Bottom row: $(g\circ\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}_\alpha})(\Omega)$, Left to right: C_6^0, C_6^1, C_6^2. Dashed lines exist just to guide the eye.

and the corresponding $\Phi_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5),(3,1,2),(1,0,-1)}\ (g^{-1}\circ\Omega)$, and the second row displays $(g\circ\Phi_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5),(3,1,2),(1,0,-1)})(\Omega)$. The following table contains $(g\circ\Phi_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5),(3,1,2),(1,0,-1)})(\Omega)$ for all symmetry operations:

Op	$g\circ\{\mathbf{r}_i\}$	basis function	position (Fig. 1.2)
C_6^0	$(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5)$	$\Phi_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5),(3,1,2),(1,0,-1)}$	upper left
C_6^1	$(\mathbf{r}_2,\mathbf{r}_4,\mathbf{r}_6)$	$\Phi_{(\mathbf{r}_2,\mathbf{r}_4,\mathbf{r}_6),(3,1,2),(1,0,-1)}$	upper center
C_6^2	$(\mathbf{r}_3,\mathbf{r}_5,\mathbf{r}_1)$	$\Phi_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5),(2,3,1),(-1,1,0)}$	upper right
C_6^3	$(\mathbf{r}_4,\mathbf{r}_6,\mathbf{r}_2)$	$\Phi_{(\mathbf{r}_2,\mathbf{r}_4,\mathbf{r}_6),(2,3,1),(-1,1,0)}$	lower left
C_6^4	$(\mathbf{r}_5,\mathbf{r}_1,\mathbf{r}_3)$	$\Phi_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_5),(1,2,3),(0,-1,1)}$	lower center
C_6^5	$(\mathbf{r}_6,\mathbf{r}_2,\mathbf{r}_4)$	$\Phi_{(\mathbf{r}_2,\mathbf{r}_4,\mathbf{r}_6),(1,2,3),(0,-1,1)}$	lower right

$g\circ\{\mathbf{r}_i\}$ lists the transformation of the initial positions in the cluster, before the necessary reordering S takes place.

Action on the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$

The symmetry operations do not affect the generalized Wigner coefficients and the $\mathbf{k}$-tuple, as only the sorted values of the $\mathbf{l}_\alpha$-tuple, $\mathbf{j}$ enter the coefficients. The $\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}}$

are therefore not changed if the sequence of the l_i changes. Therefore, the action of g on the O(3)-symmetrized basis functions is:

$$\begin{aligned} g \circ \mathcal{Y}_{\alpha \mathbf{l} \mathbf{k}} &= \sum_{\mathbf{m}} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} g \circ \Phi_{\alpha \mathbf{l}_\alpha \mathbf{m}_\alpha} \\ &= \sum_{\mathbf{m}} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \Phi_{g \circ \alpha \mathsf{S}(\mathbf{l}_\alpha) \mathsf{S}(\mathbf{m}_\alpha)}. \end{aligned} \tag{1.47}$$

1.4.4. Matrix Elements

The matrix elements regarding the $\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}$ are

$$\langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle. \tag{1.48}$$

With the definition of the $\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}$, Eqn. (1.16), and the definition of the projection operator this becomes:

$$\begin{aligned} &\langle \sum_{\mathbf{m}'_\alpha} \begin{pmatrix} j'_1 & \dots & j'_{n_\alpha} \\ n'_1 & \dots & n'_{n_\alpha} \end{pmatrix}^{\mathbf{k}'} \Phi_{\alpha' \mathbf{l}'_\alpha \mathbf{m}'_\alpha} \\ &| \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} g \circ | \\ &\sum_{\mathbf{m}_\alpha} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \Phi_{\alpha \mathbf{l}_\alpha \mathbf{m}_\alpha} \rangle \\ =& \frac{1}{|\mathcal{G}|} \sum_{\mathbf{m}'_\alpha \mathbf{m}_\alpha} \begin{pmatrix} j'_1 & \dots & j'_{n_\alpha} \\ n'_1 & \dots & n'_{n_\alpha} \end{pmatrix}^{\mathbf{k}'} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \\ &\times \sum_{g \in \mathcal{G}} \langle \Phi_{\alpha' \mathbf{l}'_\alpha \mathbf{m}'_\alpha} | \Phi_{g \circ \alpha g \circ \mathbf{l}_\alpha g \circ \mathbf{m}_\alpha} \rangle. \end{aligned} \tag{1.49}$$

Because the $\Phi_{\alpha \mathbf{l}_\alpha \mathbf{m}_\alpha}$ are orthogonal in $\alpha, \mathbf{l}_\alpha, \mathbf{m}_\alpha$ this matrix element is zero unless there is at least one group element $g_{\alpha \alpha' \mathbf{l}_\alpha \mathbf{l}'_\alpha}$ that transforms α to α' and $\mathbf{l}_\alpha$ to $\mathbf{l}'_\alpha$. Remember that $\hat{P}^1_{11} g_{\alpha \alpha' \mathbf{l}_\alpha \mathbf{l}'_\alpha} = \hat{P}^1_{11}$ according to Eqn. (1.24), and the matrix element may be written as:

$$\begin{aligned} \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle &= \delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}} \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} g_{\alpha \alpha' \mathbf{l}_\alpha \mathbf{l}'_\alpha} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle \\ &= \delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}} \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}} \rangle. \end{aligned} \tag{1.50}$$

The "orbit" $\mathcal{O}_{\alpha' \mathbf{l}'_\alpha}$ of a specific $\alpha' \mathbf{l}'_\alpha$ is the set of all clusters and l-tuples that can be created by some group operation acting on $\alpha' \mathbf{l}'_\alpha$. $\delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}}$ is one if $\alpha \mathbf{l}_\alpha$ is an element of

the orbit $\mathcal{O}_{\alpha'\mathbf{l}'_\alpha}$, and zero otherwise.
The scalar product forming the other part of the matrix elements can be analyzed a bit further:

$$\begin{aligned}
\langle \mathcal{Y}_{\alpha'\mathbf{l}'_\alpha\mathbf{k}'}|\hat{P}^1_{11}|\mathcal{Y}_{\alpha'\mathbf{l}'_\alpha\mathbf{k}}\rangle &= \frac{1}{|\mathcal{G}|}\langle \mathcal{Y}_{\alpha'\mathbf{l}'_\alpha\mathbf{k}'}|\sum_{g\in\mathcal{G}} g\circ|\mathcal{Y}_{\alpha'\mathbf{l}'_\alpha\mathbf{k}}\rangle \\
&\overset{(1.16)}{=} \frac{1}{|\mathcal{G}|}\sum_{\mathbf{m}'_\alpha\mathbf{m}_\alpha}\begin{pmatrix} j'_1 & \dots & j'_{n_\alpha} \\ n'_1 & \dots & n'_{n_\alpha}\end{pmatrix}^{\mathbf{k}'}\begin{pmatrix} j'_1 & \dots & j'_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha}\end{pmatrix}^{\mathbf{k}} \\
&\times \sum_{g\in\mathcal{G}}\langle \Phi_{\alpha'\mathbf{l}'_\alpha\mathbf{m}'_\alpha}|\Phi_{g\circ\alpha' g\circ\mathbf{l}'_\alpha g\circ\mathbf{m}_\alpha}\rangle.
\end{aligned} \tag{1.51}$$

The $\Phi_{\alpha\mathbf{l}_\alpha\mathbf{m}}$ are still orthogonal in $\alpha, \mathbf{l}_\alpha, \mathbf{m}$. This means that the scalar product is zero for any group element in the sum over the group elements that does not map α' to itself and $\mathbf{l}'_\alpha$ to itself. Therefore, the sum over all group elements may be replaced by the sum over the set of group elements that leave $\alpha', \mathbf{l}'_\alpha$ invariant. This set shall be called $\mathcal{G}_{\alpha'\mathbf{l}'_\alpha}$. In group theory, this is called the "stabilizer" or "isotropy group" of $\alpha'\mathbf{l}'_\alpha$. It is a subgroup[13] of $\mathcal{G}$. If the origin of the coordinate system is conveniently chosen to be the center of mass of α', all of these elements have to be rotations. The group has always at least one entry (the identity), but can have as many entries as there are rotations in $\mathcal{G}$. The projection operator of the identical representation of $\mathcal{G}_{\alpha'\mathbf{l}'_\alpha}$ is

$$\hat{P}^1_{\mathcal{G}_{\alpha'\mathbf{l}'_\alpha}} = \frac{1}{\left|\mathcal{G}_{\alpha'\mathbf{l}'_\alpha}\right|}\sum_{g\in\mathcal{G}_{\alpha'\mathbf{l}'_\alpha}} g\circ . \tag{1.52}$$

[13] $g\circ\alpha' g\circ\mathbf{l}'_\alpha = \alpha'\mathbf{l}'_\alpha \rightarrow g^{-1}\circ\alpha' g^{-1}\circ\mathbf{l}'_\alpha = \alpha'\mathbf{l}'_\alpha$, therefore the inverse element to g is also part of the stabilizer. Any product of two elements g_1, g_2 in the stabilizer is clearly also an element of the stabilizer. The set is therefore closed and contains an inverse to each element within. The identity is always part of the stabilizer.

With this definition, the matrix elements of $\hat{P}^1_{11}$, the projection operator of the whole group, may be written as

$$\begin{aligned}
&\langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle = \delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}} \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}} \rangle \\
&= \frac{\left|\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}\right|}{|\mathcal{G}|} \delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}} \langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}} | \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}} \rangle \\
&\overset{(1.16),(1.52)}{=} \frac{\delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}}}{|\mathcal{G}|} \sum_{\mathbf{m}'_\alpha \mathbf{m}_\alpha} \begin{pmatrix} j'_1 & \cdots & j'_{n_\alpha} \\ n'_1 & \ldots & n'_{n_\alpha} \end{pmatrix}^{\mathbf{k}'} \\
&\qquad \times \begin{pmatrix} j'_1 & \cdots & j'_{n_\alpha} \\ n_1 & \ldots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \sum_{g \in \mathcal{G}_{\alpha' \mathbf{l}'_\alpha}} \langle \Phi_{\alpha' \mathbf{l}'_\alpha \mathbf{m}'_\alpha} | \Phi_{g \circ \alpha' g \circ \mathbf{l}'_\alpha g \circ \mathbf{m}_\alpha} \rangle \\
&\qquad = \frac{\delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}}}{|\mathcal{G}|} \sum_{\mathbf{m}'_\alpha \mathbf{m}_\alpha} \begin{pmatrix} j'_1 & \cdots & j'_{n_\alpha} \\ n'_1 & \ldots & n'_{n_\alpha} \end{pmatrix}^{\mathbf{k}'} \\
&\qquad \times \begin{pmatrix} j'_1 & \cdots & j'_{n_\alpha} \\ n_1 & \ldots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \sum_{g \in \mathcal{G}_{\alpha' \mathbf{l}'_\alpha}} \delta_{\mathbf{m}'_\alpha, g \circ \mathbf{m}_\alpha},
\end{aligned} \tag{1.53}$$

where $\left|\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}\right|$ in the prefactor of the first step appears to compensate for the $1/\left|\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}\right|$ inherent in the definition of $\hat{P}^1_{\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}}$.
To discuss the eigenvectors, it is convenient to recast this equation as

$$\langle \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'} | \hat{P}^1_{11} | \mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} \rangle = \frac{\left|\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}\right|}{|\mathcal{G}|} \delta_{\alpha \mathbf{l}_\alpha, \mathcal{O}_{\alpha' \mathbf{l}'_\alpha}} A_{\mathbf{k}', \mathbf{k}}, \tag{1.54}$$

with

$$\begin{aligned}
A_{\mathbf{k}', \mathbf{k}} = \frac{1}{\left|\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}\right|} &\sum_{\mathbf{m}'_\alpha \mathbf{m}_\alpha} \begin{pmatrix} j'_1 & \cdots & j'_{n_\alpha} \\ n'_1 & \ldots & n'_{n_\alpha} \end{pmatrix}^{\mathbf{k}'} \\
\times &\begin{pmatrix} j'_1 & \cdots & j'_{n_\alpha} \\ n_1 & \ldots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \sum_{g \in \mathcal{G}_{\alpha' \mathbf{l}'_\alpha}} \delta_{\mathbf{m}'_\alpha, g \circ \mathbf{m}_\alpha}.
\end{aligned} \tag{1.55}$$

The seemingly odd choice of prefactors $1/\left|\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}\right|$ and $\left|\mathcal{G}_{\alpha' \mathbf{l}'_\alpha}\right| / |\mathcal{G}|$ will prove useful later.

1.4.5. Eigenvectors

The matrix of the projection operator may be written as a direct product of a matrix B with elements

$\mathsf{B}_{\alpha'\mathbf{l}'_\alpha,\alpha\mathbf{l}_\alpha} = \left|\mathcal{G}_{\alpha'\mathbf{l}'_\alpha}\right| / |\mathcal{G}|\, \delta_{\alpha\mathbf{l}_\alpha,\mathcal{O}_{\alpha'\mathbf{l}'_\alpha}}$ and a matrix A with elements $A_{\mathbf{k}',\mathbf{k}}$ given by Eqn. (1.55), as for each element of the matrix the following holds:

$$\langle \mathcal{Y}_{\alpha'\mathbf{l}'_\alpha\mathbf{k}'}|\hat{P}^1_{11}|\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}\rangle_{\alpha\alpha'\mathbf{l}_\alpha\mathbf{l}'_\alpha\mathbf{k}\mathbf{k}'} = \mathsf{B}_{\alpha'\mathbf{l}'_\alpha,\alpha\mathbf{l}_\alpha} A_{\mathbf{k}',\mathbf{k}}. \tag{1.56}$$

The matrix A is the same for each element of the orbit $\mathcal{O}_{\alpha'\mathbf{l}'_\alpha}$, and may be calculated for any representative of this orbit.
If eigenvectors $\mathbf{b}$ of B and eigenvectors $\mathbf{u}$ of A have been obtained, that is $\sum_{\mathbf{k}} \mathsf{A}_{\mathbf{k}',\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} = \lambda_A c^{\mathbf{u}}_{\mathbf{k}} \forall \mathbf{k}'$ and $\sum_{\alpha,\mathbf{l}_\alpha} \mathsf{B}_{\alpha'\mathbf{l}'_\alpha,\alpha\mathbf{l}_\alpha} b_{\alpha\mathbf{l}_\alpha} = \lambda_B b_{\alpha\mathbf{l}_\alpha} \forall \alpha', \mathbf{l}'_\alpha$, the components of the eigenvectors for the whole matrix of the projection operator, $b^m_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ may be constructed as $u_{\mathbf{k}} b_{\alpha\mathbf{l}_\alpha}$, because

$$\begin{aligned} \sum_{\alpha,\mathbf{l}_\alpha\mathbf{k}} \mathsf{B}_{\alpha'\mathbf{l}'_\alpha,\alpha\mathbf{l}_\alpha} \mathsf{A}_{\mathbf{k}',\mathbf{k}} u_{\mathbf{k}} b_{\alpha\mathbf{l}_\alpha} &= \left(\sum_{\alpha,\mathbf{l}_\alpha} \mathsf{B}_{\alpha'\mathbf{l}'_\alpha,\alpha\mathbf{l}_\alpha} b_{\alpha\mathbf{l}_\alpha}\right)\left(\sum_{\mathbf{k}} \mathsf{A}_{\mathbf{k}',\mathbf{k}} u_{\mathbf{k}}\right) \\ &= \lambda_A \lambda_B u_{\mathbf{k}} b_{\alpha\mathbf{l}_\alpha} \forall \alpha', \mathbf{l}'_\alpha, \mathbf{k}' \end{aligned} \tag{1.57}$$

The orthogonal eigenvectors $\mathbf{u}$ of $A_{\mathbf{k}',\mathbf{k}}$ are nontrivial and have to be determined numerically, and their components are denoted $c^{\mathbf{u}}_{\mathbf{k}}$. In the basis of the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$,

$$\mathbf{u} = \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}. \tag{1.58}$$

The eigenvectors of the matrix B can be calculated analytically. If the basis set is ordered such that basis functions with clusters and $\mathbf{l}_\alpha$-tuples in the same orbit $\mathcal{O}_{\alpha'\mathbf{l}'_\alpha}$ are next to each other, the matrix B is block diagonal and of the form

$$\mathsf{B} = \begin{pmatrix} \mathsf{B}_1 & 0 & \dots & 0 & 0 \\ 0 & \mathsf{B}_2 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & \mathsf{B}_{\mu-1} & 0 \\ 0 & 0 & \dots & 0 & \mathsf{B}_\mu \end{pmatrix}, \tag{1.59}$$

where the blocks represent different orbits, indexed by $1, \dots, \mu$, of the form

$$\mathsf{B}_i = \frac{\left|\mathcal{G}_{\alpha'_i\mathbf{l}'_{\alpha i}}\right|}{|\mathcal{G}|} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}, \tag{1.60}$$

where only the prefactor and the dimension of a block depend on the specific orbit $\mathcal{O}_{\alpha'_i\mathbf{l}'_{\alpha i}}$ that this block represents. The notation $\mathcal{O}_{\alpha'_i\mathbf{l}'_{\alpha i}}$ specifies an orbit by one (arbitrarily chosen) element $\alpha_i, \mathbf{l}_{\alpha i}$ of the orbit. This element is called a “cluster archetype”.

These blocks do *not* have dimension $|\mathcal{G}|$, but dimension $|\mathcal{G}| \,/\, \left|\mathcal{G}_{\alpha_i'\mathbf{l}_{\alpha i}'}\right|$, because the orbit has only this many different elements[14]. Each B_i has only one[15] eigenvector $\tilde{\mathbf{b}}_i = (1, \ldots, 1)$ to the eigenvalue 1. The dimension of this eigenvector is $|\mathcal{G}| \,/\, \left|\mathcal{G}_{\alpha_i'\mathbf{l}_{\alpha i}'}\right|$, and the basis of the B_i matrix are the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ where all $\alpha, \mathbf{l}_\alpha$ belong to one specific orbit i. All different, linear independent eigenvectors $\mathbf{b}_i$ of B can be constructed as $(\tilde{\mathbf{b}}_1, 0 \ldots 0)$, $(0, \tilde{\mathbf{b}}_2, \ldots, 0), \ldots, (0, 0 \ldots, \tilde{\mathbf{b}}_\mu)$, and may be written in the basis of all $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ as

$$\mathbf{b}_i = \sum_{\alpha,\mathbf{l}_\alpha \in \mathcal{O}_{\alpha_i\mathbf{l}_{\alpha i}}} 1 \cdot \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}} = \sum_{\alpha,\mathbf{l}_\alpha} \delta_{\alpha\mathbf{l}_\alpha, \mathcal{O}_{\alpha_i\mathbf{l}_{\alpha i}}} \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}. \tag{1.61}$$

There is one eigenvectors $\mathbf{b}_i$ of B to the eigenvalue 1 for each orbit $\mathcal{O}_{\alpha_i\mathbf{l}_{\alpha i}}$, which means that the label i of the orbit can be used to identify the eigenvector.
The eigenvectors Ψ^1_{1m} of the direct product matrix of B and A, written in the basis of the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$, are then:

$$\begin{aligned} \Psi^1_{1m} &= \sum_{\alpha,\mathbf{l}_\alpha} \delta_{\alpha\mathbf{l}_\alpha, \mathcal{O}_{\alpha_i\mathbf{l}_{\alpha i}}} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}} \\ &= \frac{1}{|\mathcal{G}_{\alpha_i\mathbf{l}_{\alpha i}}|} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \sum_{g \in \mathcal{G}} g \circ \mathcal{Y}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{k}} \end{aligned} \tag{1.62}$$

where the second step just used that one may create the orbit by applying the group operations. The prefactor arises because the sum runs over the whole group, whereby each actually different element of the orbit will be generated $|\mathcal{G}_{\alpha_i\mathbf{l}_{\alpha i}}|$ times. The index m of the Ψ^1_{1m} indicates both different orbits i and different $\mathbf{u}$. Therefore, the following symbol is defined:

$$\Psi^1_{1m} = \mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}, \tag{1.63}$$

to better reflect this and to retain visual consistency with the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$. Contrary to the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$, α_i and $\mathbf{l}_{\alpha_i}$ are representatives of a whole orbit i, and naturally, only one representative may be chosen for each orbit.
These basis functions are orthogonal (and therefore, linear independent),

$$\langle \mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} | \mathcal{Y}^{\mathcal{G}}_{\alpha_i'\mathbf{l}_{\alpha i}'\mathbf{u}'} \rangle = \delta_{\alpha_i\mathbf{l}_{\alpha i}, \alpha_i'\mathbf{l}_{\alpha i}'} \delta_{\mathbf{u},\mathbf{u}'}, \tag{1.64}$$

[14] This follows fromt the so-called orbit-stabilizer theorem. A proof is given in the Appendix, Section A.4. See also [20].

[15] Every B_i is an idempotent matrix, $\mathsf{B}_i\mathbf{x} = \mathsf{B}_i^2\mathbf{x}\forall\mathbf{x}$. This means it has only eigenvalues $\lambda_i = 0, 1$. As $\mathrm{Tr}(\mathsf{B}_\mathsf{i}) = \sum_\mathsf{i} \lambda_\mathsf{i} = 1$, it follows that there is only one eigenvalue 1, and all other eigenvalues are zero. This also means that $\mathsf{A}_{\mathbf{k}',\mathbf{k}}$ has also only eigenvalues 0 or 1, because the eigenvalues of the matrix of the whole projection operator, given by the product of the eigenvalues of B and A according to Eqn. (1.57), need to be 1 or 0.

where $\delta_{\alpha_i \mathbf{l}_{\alpha i}, \alpha_i' \mathbf{l}_{\alpha i}'}$ is zero if $\alpha_i \mathbf{l}_{\alpha i}$ and $\alpha_i' \mathbf{l}_{\alpha i}'$ represent different orbits, and one if not. The energy may be expanded as

$$E = \sum_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}, \tag{1.65}$$

where the sum runs only over $\alpha_i, \mathbf{l}_{\alpha_i}$ that represent different orbits.
Except the eigenvectors of A, and the prefactor $1/|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$, the procedure to calculate these functions is straightforward: Start with a given cluster "archetype" at one point, and then apply all symmetry operations, thus generating all possible "pictures" of this cluster throughout the crystal. Sum over the values of the $\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}$ of these symmetry pictures.

1.4.6. Why the Ψ^1_{1m} Are Not a Useful Basis Set

This subsection shall detail that the $\Psi^1_{1\alpha \mathbf{l}_\alpha \mathbf{k}}$ are neither orthogonal nor linearly independent, and therefore not very useful as a basis set. This is not important for the numerical investigations, but hopefully educational.
First, the definition of the new, orthogonal and linear independent basis functions may be written as

$$\begin{aligned} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} &= \frac{1}{|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \sum_{g \in \mathcal{G}} g \circ \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}} \\ &= \frac{|\mathcal{G}|}{|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \hat{P}^1_{11} \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}} \\ &\overset{(1.27)}{=} \frac{|\mathcal{G}|}{|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} c^1_{1\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}} \Psi^1_{1\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}. \end{aligned} \tag{1.66}$$

On the other hand, the basis function of the space group $\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}$ that a given $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ would be projected upon can be expanded in the $\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}$ as

$$\begin{aligned}
\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}} &= \sum_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \langle\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}|\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}\rangle\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \\
&\overset{(1.27)}{=} \sum_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \frac{1}{c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}}\langle\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}P^1_{11}|\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}\rangle\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \\
&= \frac{1}{c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}}\sum_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \langle\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}|\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}\rangle\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \qquad (1.67)\\
&\overset{(1.62)}{=} \frac{1}{c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}}\sum_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \delta_{\alpha\mathbf{l}_\alpha,\mathcal{O}_{\alpha_i\mathbf{l}_{\alpha i}}} c^{\mathbf{u}}_{\mathbf{k}}\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \\
&= \frac{1}{c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}}\sum_{\mathbf{u}} c^{\mathbf{u}}_{\mathbf{k}}\mathcal{Y}^{\mathcal{G}}_{\alpha\mathbf{l}_\alpha\mathbf{u}}
\end{aligned}$$

where the second to last step used the orthogonality of the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$, and the last step used that if due to the Kronecker delta the orbit has to contain $\alpha, \mathbf{l}_\alpha$, one might just as well use $\alpha, \mathbf{l}_\alpha$ as representative of this specific orbit.
Not all of the functions $\{\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}\}$ are orthogonal, as the scalar product of two such functions is given by

$$\langle\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}|\Psi^1_{1\alpha'\mathbf{l}'_\alpha\mathbf{k}'}\rangle = \frac{1}{c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}c^1_{1\alpha'\mathbf{l}'_\alpha\mathbf{k}'}}\delta_{\alpha\mathbf{l}_\alpha,\mathcal{O}_{\alpha'\mathbf{l}'_\alpha}}\sum_{\mathbf{u}} c^{\mathbf{u}}_{\mathbf{k}}c^{\mathbf{u}}_{\mathbf{k}'} \qquad (1.68)$$

where the orthogonality of the $\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}$ was used.
If one considers two rotationally invariant basis functions, $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ and $\mathcal{Y}_{\alpha'\mathbf{l}'_\alpha\mathbf{k}}$ to different elements $\alpha, \mathbf{l}_\alpha$ and $\alpha', \mathbf{l}'_\alpha$ out of one orbit and to equal $\mathbf{k}$, $\hat{P}^1_{11}$ projects these functions on proportional functions (see Eqn. (1.67)). Therefore, the functions $\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}$ and $\Psi^1_{1\alpha'\mathbf{l}'_\alpha\mathbf{k}}$, $\alpha', \mathbf{l}'_\alpha \in \mathcal{O}_{\alpha\mathbf{l}_\alpha}$ are linearly dependent.
Functions $\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}$ and $\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}'}$ to the same $\alpha, \mathbf{l}_\alpha$, but to different $\mathbf{k}, \mathbf{k}'$ can be linearly dependent as well. This is not evident from the formula above, however, one may insert Eqn. (1.66) into Eqn. (1.67) to give

$$\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}} = \frac{1}{c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}}\frac{|\mathcal{G}|}{|\mathcal{G}_{\alpha\mathbf{l}_\alpha}|}\sum_{\mathbf{k}'}\sum_{\mathbf{u}} c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}'}c^{\mathbf{u}}_{\mathbf{k}}c^{\mathbf{u}}_{\mathbf{k}'}\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}'} \qquad (1.69)$$

and therefore

$$c^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}\left(\frac{|\mathcal{G}_{\alpha\mathbf{l}_\alpha}|}{|\mathcal{G}|} - \sum_{\mathbf{u}}(c^{\mathbf{u}}_{\mathbf{k}})^2\right)\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}} = \sum_{\mathbf{k}'\neq\mathbf{k}}\sum_{\mathbf{u}} c^{\mathbf{u}}_{\mathbf{k}}c^{\mathbf{u}}_{\mathbf{k}'}\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}'}, \qquad (1.70)$$

which means at least one $\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}}$ may be written as a linear combination of all other $\Psi^1_{1\alpha\mathbf{l}_\alpha\mathbf{k}'}$, unless $c^{\mathbf{u}_0}_{\mathbf{k}'} \propto \delta_{\mathbf{k},\mathbf{k}'}$ for one specific eigenvector $\mathbf{u}_0$. $c^{\mathbf{u}_0}_{\mathbf{k}'} \propto \delta_{\mathbf{k},\mathbf{k}'}$ would mean that only one specific $\mathbf{k}$ contributes to the eigenvector $\mathbf{u}_0$, which also means that this $\mathbf{k}$ contributes to no other eigenvector, as the $\mathbf{u}$ are orthogonal. This is, however, a rare case.

1.5. Configurational symmetry

With reasonable computational effort, only small parts of the global configuration space may be probed with density functional theory. There are two classes of symmetric configurations that can be probed with numerical ease, which are called spin spirals and supercells.

Spin spirals are configurations that have the following symmetry:

$$\begin{aligned} \vartheta(\mathbf{T}+\mathbf{r}_0) &= \vartheta(\mathbf{r}_0), \\ \varphi(\mathbf{T}+\mathbf{r}_0) &= \varphi(\mathbf{r}_0) + \mathbf{q}\cdot\underbrace{(\mathbf{r}_i-\mathbf{r}_0)}_{:=\mathbf{T}}, \end{aligned} \tag{1.71}$$

where $\mathbf{r}_0$ indicates the position at the (arbitrarily chosen) origin of the coordinate system, and the relation holds for all elements of the translation subgroup with associated translation vectors $\mathbf{T}$. This means that all magnetic moment directions have the same angle regarding a specific, global axis and rotate coherently around this axis. The phase of the rotation at a specific atom is determined by the position vector of the atom. The symmetry operation underlying these configurations is a combination of a translation in the space of the position vectors $\mathbf{r}_i$, and a rotation in the space of the magnetic moment directions $\mathbf{e}_i$. By employing a modified Bloch theorem [21] these configurations can be easily calculated for any wave vector and opening angle.

Supercell configurations are configurations where

$$\begin{aligned} \vartheta\Big(\sum_{k=1}^{3} n_k\mathbf{T}_k + \mathbf{r}_i\Big) &= \vartheta(\mathbf{r}_i), n_k \in \mathbb{Z} \\ \varphi\Big(\sum_{k=1}^{3} n_k\mathbf{T}_k + \mathbf{r}_i\Big) &= \varphi(\mathbf{r}_i), n_k \in \mathbb{Z} \end{aligned} \tag{1.72}$$

holds for any position $\mathbf{r}_i$. The directions of the magnetic moments thus repeat exactly after any multiple of the three translations $\{\mathbf{T}_k\}$, which are chosen out of the subgroup of pure translations of the spacegroup. The spatial cell defined by these translation vectors $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$ is called the supercell.

The configuration space of a crystal which contains a very large, but countable number

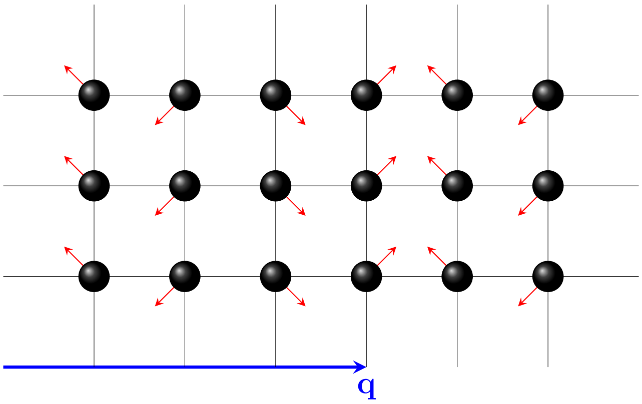

Figure 1.3.: Example of a spin-spiral structure. Wave vector $\mathbf{q} = 2\pi/a0.25(1\,0)$

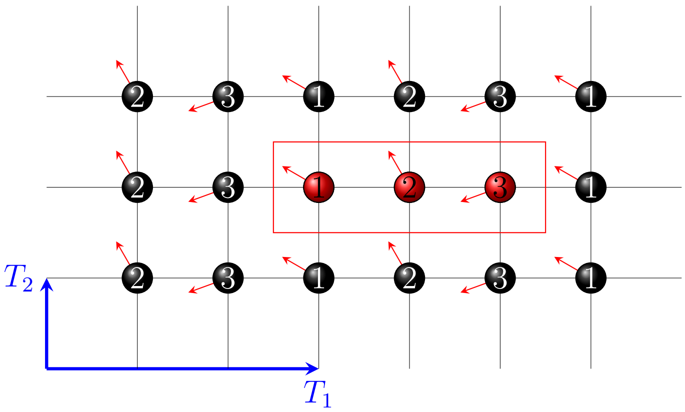

Figure 1.4.: Example of a two-dimensional supercell structure with translation vectors $\mathbf{T}_1 = a\,(3\,0)\,, \mathbf{T}_2 = a\,(0\,1)$. All atomic magnetic moments with equal numbers have the same directions $\mathbf{e}_i$

N_{at} of atoms is

$$K_{\mathrm{Cr}} = S_1^2 \otimes S_2^2 \otimes \ldots \otimes S_{N_{\mathrm{at}}}^2 := \left(S^2\right)^{N_{\mathrm{at}}}, \tag{1.73}$$

where S^2 indicates the surface of the 3-dimensional sphere.
The configuration space of the supercell is

$$K_{\mathrm{SC}} = S_1^2 \otimes S_2^2 \otimes \ldots \otimes S_{N_{\mathrm{SC}}}^2 := \left(S^2\right)^{N_{\mathrm{SC}}}, \tag{1.74}$$

with a finite number N_{SC} of atoms within the supercell. This number is limited to about 16-32 atoms for numerical investigations, because the computation time required for DFT calculations scales with the third power of N_{SC}.
For spin spiral configurations the configuration space is even smaller:

$$K_{\mathrm{SS}} = \underbrace{S^1}_{\text{opening angle}} \otimes \underbrace{E^3}_{\text{wave vector}}, \tag{1.75}$$

where S^1 is the surface of the two-dimensional sphere, i.e., a circle. The configuration space of the wave vector is, of course, actually the first Brillouin zone in a crystal. Obviously, for the rather small numbers N_{SC} accessible by DFT calculations,

$$\begin{aligned} K_{\mathrm{SC}} &\subset K_{\mathrm{Cr}}, \\ K_{\mathrm{SS}} &\subset K_{\mathrm{Cr}}, \\ K_{\mathrm{SS}} &\not\subset K_{\mathrm{SC}}. \end{aligned} \tag{1.76}$$

The central problem is that our goal is to obtain a description of $E(\{\mathbf{e}_i\})$ valid in K_{Cr}, while numerical results are only available for K_{SS} and K_{SC}, which are both of much smaller dimension than K_{Cr}.

1.5.1. Spin Spiral Symmetry

It is convenient to discuss spin spiral configurations within the framework of the rotationally symmetric $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$, and consider the space group symmetry later.
Assuming the configuration has a spin spiral symmetry, it can be shown after some algebra (presented in Section A.5.1) that

$$\begin{aligned} \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\Omega_{\mathrm{SS}}) = &\sum_{\tilde{l}=l_0,\mathrm{even}}^{l_f} P_{\tilde{l}}(\cos(\vartheta_0)) \sum_{\mathbf{m}^+} 2\cos(\mathbf{q}\cdot\mathbf{T}_{\mathbf{m}_+}) \\ &\times \begin{pmatrix} j_1 & \ldots & j_{n_\alpha} \\ n_1 & \ldots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} N_{\mathbf{l}_\alpha\mathbf{m}^+}, \end{aligned} \tag{1.77}$$

where the set of all possible $\mathbf{m}$ has been decomposed into two parts, $\{\mathbf{m}\}_+$ and $\{\mathbf{m}\}_-$, such as that to each $\mathbf{m}_+$ there is exactly one $\mathbf{m}_- = -\mathbf{m}_+$. For the vector $\mathbf{m}$ with

$m_i = 0 \forall i$ there is no such counterpart, and it is simply defined to be part of $\{\mathbf{m}\}_+$. $\mathbf{T}_{\mathbf{m}+} = \sum_{i=1}^{n_\alpha} m_i \mathbf{r}_i$ is a translation vector of the crystal lattice, as $m_i \in \mathbb{Z}$, and the $N_{\mathbf{l}_\alpha \mathbf{m}^+}$ are constant, real coefficients which depend on $\mathbf{l}_\alpha$ and $\mathbf{m}^+$. The maximum possible order of the Legendre polynomials $P_{\tilde{l}}(\cos(\vartheta_0))$ is given by the sum over the elements of the $\mathbf{l}$-tuple, $l_f = \sum_{i=1}^{n_\alpha} l_i$.
The basis function depends only on the opening angle ϑ_0 of the spin spiral, but not on the initial phase φ_0. This is because the basis functions are invariant under global rotations and therefore do not depend on the phase of the spin spiral.
As noted in the introduction to this topic, the configuration space is equivalent to $S^1 \otimes E^3$, and the Legendre polynomials form a basis set of the circle S^1 and the sines and cosines form a basis set for E^3.
The cosines are the inversion symmetric part of the plane-wave basis of E^3, reflecting that the transformation $\mathbf{q} \mapsto -\mathbf{q}$ is just a global phase difference of π, which does not change the $\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}$ because they are invariant under global rotations.
The even order of the Legendre polynomials is due to the fact that the original basis functions were invariant under collective inversion of all magnetic moments, which in the spin spiral case is equivalent to $\vartheta \mapsto \vartheta + \frac{\pi}{2}$ and $\mathbf{q} \mapsto -\mathbf{q}$. The effect on the wave vector is irrelevant (see above), and the symmetry under $\vartheta \mapsto \vartheta + \frac{\pi}{2}$ corresponds to symmetry under inversion on the interval $[-1, 1]$, which the Legendre polynomials are defined on, and this is only fulfilled by the even Legendre polynomials.
The above formula therefore represents the decomposition of the overcomplete SCE basis set towards a complete basis set for spin spirals, which consists of products of Legendre polynomials to even order and cosines.
Concerning space group symmetry, remember that the invariant space group basis functions are constructed from the $\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}$ by

$$\mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} = \frac{1}{|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \sum_{g \in \mathcal{G}} g \circ \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}. \tag{1.78}$$

The action of a group element on the $\mathcal{Y}^{SS}_{\alpha \mathbf{l}_\alpha \mathbf{k}}$ may be deduced (see Section 1.4.3) by the action of the inverse group element on the configuration, which is given by Eqn. (1.71):

$$\begin{aligned} \vartheta(\mathbf{r}_i) &= \vartheta(0) \forall i, \\ \varphi(\mathbf{r}_i) &= \varphi(0) + \mathbf{q} \cdot (\mathbf{r}_i - \mathbf{r}_0). \end{aligned} \tag{1.79}$$

$\vartheta_{\mathbf{r}_i}$ is the same for any $\mathbf{r}_i$, and therefore invariant under the action of all group elements. $\varphi_{\mathbf{r}_i}$ only depends on the difference between $\mathbf{r}_i$ and $\mathbf{r}_0$, so if a translation is applied to both position vectors, nothing changes and the configuration is therefore invariant under the translational parts of all group elements.
A rotation R^{-1} will change the direction of the difference vectors, which is equivalent to a rotation R of $\mathbf{q}$, as only the scalar products between $\mathbf{q}$ and the difference vectors

matters. Therefore, one may write the invariant space group basis functions as

$$\begin{aligned}\mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}(\Omega_{\mathrm{SS}}) = \frac{|T|}{|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|} &\sum_{\tilde{l}=l_0,\mathrm{even}}^{l_f} P^{\tilde{l}}(\cos(\vartheta_0)) \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \\ &\times \sum_{\mathbf{m}^+} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} N_{\mathbf{l}_{\alpha i}\mathbf{m}^+} \\ &\times \sum_{R\in\mathcal{G}_0} 2\cos((\mathsf{R}\mathbf{q})\cdot \mathbf{T}_{\mathbf{m}_+}).\end{aligned} \tag{1.80}$$

$\mathcal{G}_0$ is the point group of the space group, that is the group of all rotations that occur in some element of the space group, regardless of the translational part of that element. The prefactor $|T|$ represents the number of pure translations in $\mathcal{G}$.

1.5.2. Relation to the extended Heisenberg model

In the extended Heisenberg model, the cosines of the angles are used as basis functions:

$$\mathcal{F}^i_{\mathrm{H}} = \mathbf{e}_0 \cdot \mathbf{e}_i \tag{1.81}$$

for each atom i. The direction of the magnetic moment at sites i and 0 is for spin spirals given by

$$\mathbf{e}_i = \begin{pmatrix} \sin(\vartheta_0)\cos(\mathbf{q}\cdot\mathbf{r}_i) \\ \sin(\vartheta_0)\sin(\mathbf{q}\cdot\mathbf{r}_i) \\ \cos(\vartheta_0) \end{pmatrix}, \mathbf{e}_0 = \begin{pmatrix} \sin(\vartheta_0) \\ 0 \\ \cos(\vartheta_0) \end{pmatrix} \tag{1.82}$$

and the scalar product gives

$$\mathcal{F}^i_{\mathrm{H}} = \cos^2(\vartheta_0) + \sin^2(\vartheta_0)\cos(\mathbf{q}\cdot\mathbf{r}_i). \tag{1.83}$$

Note that the functional dependence on ϑ_0 is the same for all Heisenberg basis function. This means that within the extended Heisenberg models, the energy needs to have the same functional dependence on ϑ_0 as well. This prediction of a generalized Heisenberg model will be compared to density functional theory data later in Section 2.2.1.
For constant opening angles ϑ_0, these functions reduce to

$$\mathcal{F}^i_{\mathrm{H}} = a + b\cos(\mathbf{q}\cdot\mathbf{r}_i). \tag{1.84}$$

For small opening angles, the extended Heisenberg basis functions are approximately

$$\mathcal{F}^i_{\mathrm{H}} \approx (1-\vartheta_0^2) + \vartheta_0^2\cos(\mathbf{q}\cdot\mathbf{r}_i) \approx \vartheta_0^2(\cos(\mathbf{q}\cdot\mathbf{r}_i) - 1), \tag{1.85}$$

where the constant term 1 has been omitted in the last step.
The $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ for these configurations, as shown in Eqn. (1.77), have the same quadratic

behaviour in ϑ_0 for small ϑ_0, because all Legendre polynomials $P^l(x)$ to even order l are proportional to x^2 for small x. Furthermore, each cosine within Eqn. (1.77) may be replaced with the corresponding Heisenberg basis function, because the $\mathbf{T}_{\mathbf{m}_+}$ are lattice vectors. This means that for a set of spin-spiral configurations with one constant opening angle, or with various but small opening angles, the SCE basis functions and the extended Heisenberg model are equivalent, in the sense that one may be expressed as a linear combination of the other. Therefore, the results of any investigation that only uses spin-spiral configurations with one constant, or various, but small opening angles can be described within an extended Heisenberg model with the same accuracy as within a SCE - which explains the great success the model has enjoyed in describing spin spirals [22, 23, 24].

1.5.3. Supercells

One of the main results of this subsection will be that for supercell configurations, the basis functions to cluster archetypes within the supercell form a complete basis set. All other basis functions may be written as linear combinations of those within the supercell, for all configurations that have a common supercell symmetry. Naturally, this is not true for general configurations, however, as mentioned in the introduction to this section, our DFT code does not allow for investigating general configurations but only for supercells and spin spirals.

This subsection shall detail the problem by investigating the basis functions. What the results means for numerical investigations, and how the validity of a given SCE derived from spin spirals and supercells for general configurations can be checked is the topic of the next subsections and section.

It is convenient again to first discuss the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$, or specific "cluster pictures" in the sense of Section 1.4.5. These functions are defined by the cluster, which shall be $\alpha = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{n_\alpha})$, each position has a magnetic moment direction assigned by Ω, which shall be $(\mathbf{e}_1^{\mathbf{r}_1}, \mathbf{e}_2^{\mathbf{r}_2}, \ldots, \mathbf{e}_{n_\alpha}^{\mathbf{r}_{n_\alpha}})$. Also, each position is assigned a specific spherical harmonic by $\mathbf{l}_\alpha$, which shall be $(l_1, l_2, \ldots, l_{n_\alpha})$. The k-tuple shall be $(k_1, k_2, \ldots, k_{n_\alpha-3})$. This prototypical basis function is

$$\begin{aligned}\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\Omega) = \sqrt{4\pi}^{n_\alpha} \sum_{\mathbf{m}} \begin{pmatrix} j_1 & \ldots & j_{n_\alpha} \\ n_1 & \ldots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \\ \times Y^{l_1 m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2 m_2}(\mathbf{e}_2^{\mathbf{r}_2}) \ldots Y^{l_{n_\alpha} m_{n_\alpha}}(\mathbf{e}_{n_\alpha}^{\mathbf{r}_{n_\alpha}}).\end{aligned} \tag{1.86}$$

The supercell is formed by the positions inside the volume spanned by $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$ from one arbitrarily defined origin which shall have the position vector $(0, 0, 0)$.

A cluster picture can contain a position vector $\mathbf{r}_i$ that is outside of the supercell, that is $\mathbf{r}_i \cdot \mathbf{T}_k \geq |\mathbf{T}_k|^2$ for some $k \in (1, 2, 3)$. Because the numerical values of the magnetic moment directions repeat exactly after $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$ for all configurations with a common supercell

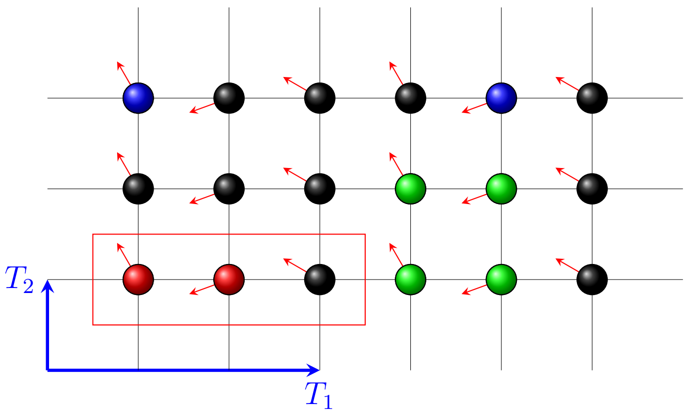

Figure 1.5.: Illustration of backfolding into the supercell. Red lines mark the boundary of the supercell. For all configurations that have this specific symmetry, the basis functions with some $\mathbf{l}_\alpha$ to the blue cluster are equal to basis functions with the same $\mathbf{l}_\alpha$ of the red cluster. Basis functions to the green cluster may be written as linear combinations of the basis functions to the red cluster.

symmetry, the direction associated with this position vector has for all configurations with a common supercell symmetry the same value as the direction associated with $\mathbf{r}_j = \mathbf{r}_i - \sum_k \lfloor \mathbf{r}_i \cdot \mathbf{T}_k / a^2 \rfloor \mathbf{T}_k$ within the supercell, where $\lfloor x \rfloor$ indicates the nearest lower integer to x and a is the lattice constant. A sketch is provided in Figure 1.5. When calculating the $\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}}(\Omega_{\mathrm{SC}})$ according to Eqn.(1.86) for different configurations with a common supercell symmetry, the position $\mathbf{r}_i$ could be replaced by the position $\mathbf{r}_j$ within the supercell, without changing the value of the basis function. This holds for all such supercell configurations. Correspondingly, the cluster $\alpha = (\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_{n_\alpha})$ could be just as well written as $\alpha' = (\mathbf{r}_1, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_{n_\alpha})$ for all such supercell configurations. This is referred to as "backfolding", because $\mathbf{r}_i$ is "folded" back into the supercell. If the position $\mathbf{r}_j$ was not already part of the original cluster α, the basis function is for all supercell configurations Ω_{SC} equal to the basis function to α' (with equal $\mathbf{l}_\alpha, \mathbf{u}$). An example is the blue cluster in Figure 1.5.

However, if the position $\mathbf{r}_j$ was already part of the original cluster α, or if several atoms are backfolded on the same $\mathbf{r}_j$, the new cluster α' would no longer contain only different positions. An example is the green cluster in Fig. 1.5. This means the product of all spherical harmonics ($\propto \Phi_{\alpha \mathbf{l}_\alpha \mathbf{m}_\alpha}$) contains two (or more) spherical harmonics which have for all such supercell configurations the same argument. Such products can be decomposed to a linear combination of spherical harmonics by the following formula:

$$
\begin{aligned}
Y^{l_i m_i}(\vartheta,\varphi) Y^{l_j m_j}(\vartheta,\varphi) = \sum_{l=|l_i-l_j|}^{l_i+l_j} & \left[\frac{(2l_i+1)(2l_j+1)}{4\pi(2l+1)}\right]^{\frac{1}{2}} \\
& \times \mathrm{C}^{l\mu}_{l_i m_i, l_j m_j} Y^{l\mu}(\vartheta,\varphi) \mathrm{C}^{l0}_{l_i 0, l_j 0},
\end{aligned}
\tag{1.87}
$$

with $\mu = m_i + m_j$.

Naturally, one may insert this equality into the definition of the basis function without changing the properties of the basis function, i.e., the result must still be a rotationally invariant function. A basis function which contains such a product of two spherical harmonics to the same argument just once can then be written as a linear combination of all basis functions to the cluster α' which contains $\mathbf{r}_j$ only once,

$$
\begin{aligned}
\mathcal{Y}_{(\mathbf{r}_1,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_{i-1},\mathbf{r}_i=\mathbf{r}_j+\mathbf{T},\mathbf{r}_{i+1},\ldots,\mathbf{r}_{n_\alpha}),\mathbf{l}_\alpha,\mathbf{k}}(\Omega_{\mathrm{SC}}) &= \sum_{\mathbf{l}'_\alpha \mathbf{k}'} c^{\mathbf{l}_\alpha \mathbf{k}}_{\mathbf{l}'_\alpha \mathbf{k}'} \mathcal{Y}_{\alpha' \mathbf{l}'_\alpha \mathbf{k}'}(\Omega_{\mathrm{SC}}) \\
\text{where } \alpha' &= (\mathbf{r}_1,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_{i-1},\mathbf{r}_{i+1},\ldots,\mathbf{r}_{n_\alpha-1}), \\
\mathbf{l}'_\alpha &= (l_1, l_2, \ldots, \tilde{l}_j, \ldots, l_{i-1}, l_{i+1}, \ldots, l_{n_\alpha}), \\
\text{and } \tilde{l}_j &\in [|l_i - l_j|, l_i + l_j], \\
\mathbf{T} &= \sum_{i=1}^{3} n_i \mathbf{T}_i, n_i \in \mathbb{Z}.
\end{aligned}
\tag{1.88}
$$

The restriction on the possible $\mathbf{l}'_\alpha$-tuples and the cluster α' follows from the orthogonality of the $\Phi_{\alpha \mathbf{l}_\alpha \mathbf{m}_\alpha}$. If the cluster α contains several products of spherical harmonics to the same argument, the process simply has to be repeated.

Because the whole set of basis functions was complete for all configurations, this also means that the set of all basis functions to clusters within a given supercell are complete for all configurations with a common supercell symmetry.

Detailed calculations of the $c^{\mathbf{l}_\alpha \mathbf{k}}_{\mathbf{l}'_\alpha \mathbf{k}'}$ for the subset of basis functions to clusters with up to four atoms, and $\mathbf{l}_\alpha$ already ordered as a decreasing sequence[16], $l_1 \geq l_2 \geq l_3 \geq \ldots \geq l_{n_\alpha}$, can be found in Section A.5.2. These results are summarized in the following list of

[16] This was presumed just to lighten the notation a bit. The general cases can be rather easily deduced from the calculations in Section A.5.2, but are difficult to express in our notation.

equations:

$$
\begin{aligned}
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2=\mathbf{r}_1+\mathbf{T})(l_1,l_1)} &= \text{const} \\
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3=\mathbf{r}_2+\mathbf{T})(l_1,l_2,l_3)} &\propto \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2)(l_1,l_1)} \\
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3=\mathbf{r}_1+\mathbf{T})(l_1,l_2,l_3)} &\propto \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2)(l_2,l_2)} \\
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2=\mathbf{r}_1+\mathbf{T},\mathbf{r}_3)(l_1,l_2,l_3)} &\propto \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2)(l_3,l_3)} \\
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2=\mathbf{r}_1+\mathbf{T},\mathbf{r}_3,\mathbf{r}_4)(l_1,l_2,l_3,l_4)(k)} &\propto \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_3,\mathbf{r}_4)(k,l_3,l_4)} \\
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4=\mathbf{r}_3+\mathbf{T})(l_1,l_2,l_3,l_4)(k)} &\propto \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)(l_1,l_2,k)} \\
\mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3=\mathbf{r}_2+\mathbf{T},\mathbf{r}_4)(l_1,l_2,l_3,l_4)(k)} &\propto \sum_{\tilde{l}=|l_2-l_3|}^{l_3+l_4} c_{\tilde{l}\mathbf{l}k} \\
&\times \mathcal{Y}_{(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_4)(l_1,\tilde{l},l_3)}.
\end{aligned}
\tag{1.89}
$$

Only for the last case is the backfolded basis function actually a superposition of the basis functions to clusters within the supercell. Also, the results for $n_\alpha = 2, 3$ hold for any order of the l_i in $\mathbf{l}_\alpha$. Unfortunately, this convenient result does not hold once space group symmetry is taken into account.
The invariant space group basis functions are calculated from the $\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}$ as

$$
\begin{aligned}
\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}(\Omega) &= \frac{1}{|\mathcal{G}_{\alpha\mathbf{l}_\alpha}|}\sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \sum_{g\in\mathcal{G}} g \circ \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\Omega) \\
&= \frac{1}{|\mathcal{G}_{\alpha\mathbf{l}_\alpha}|}\sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \sum_{t\in\mathcal{T}} \sum_{\tilde{g}} t \circ \left(\tilde{g} \circ \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\Omega)\right),
\end{aligned}
\tag{1.90}
$$

where in the second step the whole space group was divided into the translation subgroup $\mathcal{T}$ and $\tilde{g}$ which contains the rotations, glide planes and screw axes within the Wigner-Seitz cell. As for supercell configurations the directions repeat exactly after each multiple of $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$, it is only necessary to sum over the translations within the supercell:

$$
\begin{aligned}
\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}(\Omega_{\mathrm{SC}}) &= \frac{|T|}{|T|_{\mathrm{SC}}\,|\mathcal{G}_{\alpha\mathbf{l}_\alpha}|}\sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \times \\
&\times \sum_{t\in\mathcal{T}_{\mathrm{SC}}} \sum_{\tilde{g}} t \circ \left(\tilde{g} \circ \mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}}(\Omega_{\mathrm{SC}})\right),
\end{aligned}
\tag{1.91}
$$

wher $|T|_{\mathrm{SC}}$ is the number of translation within the supercell.
For the cluster pictures, basis functions to one cluster were always mapped to a superposition of basis functions of only one other cluster, and in many cases for small clusters (see above) only on one $\mathbf{l}'_\alpha$-tuple. However, the equivalent fact does not hold true for the cluster archetypes of the $\mathcal{Y}^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}$, basis functions to one archetype may be superpositions of basis functions to many archetypes.

For an example, consider Figure 1.5, and picture a supercell with $\mathbf{T}'_2 = 2\mathbf{T}_2$. If the blue cluster picture is rotated by 90 degrees, it is then mapped upon a constant, whereas the red cluster picture is not. Because the invariant space group basis functions $\mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$ are constructed as a superposition of all possible pictures of the cluster archetype, the basis function to a cluster archetype larger than the supercell will be a superposition of basis functions to several cluster archetypes within the supercell.
Note that if there is one group operation that can fit the cluster archetype into the supercell, the whole basis function can not be a linear combination of the basis functions to clusters within the supercell.

1.5.4. Consequences for numerical investigations

The goal is to obtain an expansion of the energy in terms of the basis functions:

$$E = \sum_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}} J_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}} \tag{1.92}$$

where the $J_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$ are constant coefficients. To obtain these coefficients, it is necessary to have some information about the physical system. This information can be gathered by probing a set of various configurations Ω^j, indexed by j, with density functional theory, which yields the set of energies $E^j(\Omega^j)$ associated with these configurations.
But as mentioned before, these configurations have to be supercells or spin spirals. For supercells it was shown in the preceding chapter that a basis with cluster archetypes within the supercell is complete. The energy of all such configurations can therefore be written as

$$E^{\mathrm{SC}} = \sum_{\alpha_i \in \mathrm{SC}, \mathbf{l}_{\alpha i}, \mathbf{u}} J^{\mathrm{SC}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}, \tag{1.93}$$

where the cluster archetypes of $\mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$ have to be chosen within the supercell.
The same basis functions also contribute to the expansion valid for all configurations (Eqn. (1.92)), but with different coefficients $J_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}} \neq J^{SC}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$. This is because supercell configurations map basis functions of cluster archetypes that are too large for the supercell to superpositions of basis functions within the supercell, and the coefficients $J^{\mathrm{SC}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$ obtained by supercell calculations will then be a superposition[17] of the true coefficient $J_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$ to the cluster archetypes which fit into the supercell, and the $\{J_{\alpha_j \mathbf{l}_{\alpha j}\mathbf{u}'}\}$ to the cluster archetypes $\{\alpha_j\}$ which do not fit into the supercell. However, if the SCE shall be valid not only for configurations of (or below) a specific supercell size, the general

[17] If the mapping of the basis functions to cluster archetypes α_j larger than the supercell to the basis functions to cluster archetypes within the supercell is given by

$$\mathcal{Y}^{\mathcal{G}}_{\alpha_j \mathbf{l}_{\alpha_j}\mathbf{u}'}(\Omega_{\mathrm{SC}}) = \sum_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}} c^{\alpha_j \mathbf{l}_{\alpha_j}\mathbf{u}'}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}(\Omega_{\mathrm{SC}}),$$

$J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ are needed.
The obvious idea is to use spin spirals to solve this problem, as they seem to represent changes of the magnetic moment directions on a longer length scale. This is, however, a fallacy, as spin spiral configurations are in principle *more* symmetric than supercell configurations. As was shown in Section 1.5.1, the full SCE basis set is overcomplete for spin spiral configurations, and may be written as as superposition of a symmetry adapted basis set consisting of cosines and Legendre polynomials. However, this basis set is adapted to specific configurations, and would be unable to describe general configurations.
Using the general SCE basis functions for spin spirals introduces many linear dependencies within these basis functions, and therefore does not allow a unique identification of expansion coefficients $J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$. These linear dependencies are also significantly harder to compute analytically (and therefore, to avoid) than in the case of supercells.
Intuitively, if the supercell investigated by density functional theory is "big enough", the $J^{\mathrm{SC}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ should be good approximations of the $J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$. The theoretical background of this intuiton is explored in Section 1.7, followed by means to determine what is actually "big enough" for a physical system.

1.6. The Energy per Atom, Numerical Implementation

So far, all formulas were derived for the energy $E(\Omega)$. This is an extensive quantity, and consequentially diverges for an infinite crystal. A finite, intensive quantity is the energy per atom, denoted $e(\Omega) = E(\Omega)/n_{\mathrm{At,Cry}}$. The number of atoms in a crystal, $n_{\mathrm{At,Cry}}$ is given by

$$n_{\mathrm{At,Cry}} = n_{\mathrm{bas}} \left|T\right| , \tag{1.94}$$

where n_{bas} is the number of atoms within the Wigner-Seitz cell, and $|T|$ the number of pure translations within the space group of the crystal.

---

it follows that

$$\begin{aligned} E(\Omega_{\mathrm{SC}}) &= \sum_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\Omega_{\mathrm{SC}}) \\ &= \sum_{\alpha_i \in \mathrm{SC}} \sum_{\mathbf{l}_{\alpha i} \mathbf{u}} \underbrace{\left(J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} + \sum_{\alpha_j \notin \mathrm{SC}} \sum_{\mathbf{l}_{\alpha j} \mathbf{u}'} J_{\alpha_j \mathbf{l}_{\alpha j} \mathbf{u}'} c^{\alpha_j \mathbf{l}_{\alpha j} \mathbf{u}'}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \right)}_{J^{\mathrm{SC}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}, \end{aligned}$$

where $\alpha \in \mathrm{SC}$ indicates the cluster archetypes that can fit into the supercell. This holds for all configurations that have one specific supercell symmetry.

If the energy is expanded in terms of the $\mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$,

$$E(\Omega) = \sum_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}, \tag{1.95}$$

it follows that

$$e(\Omega) = \sum_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \frac{1}{n_{\mathrm{bas}} |T|} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}. \tag{1.96}$$

For the actual numerical implementation of the SCE, only the energy per atom matters. Therefore, instead of the functions defined in Eqn. (1.91),

$$\begin{aligned} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\Omega_{\mathrm{SC}}) = & \frac{|T|}{|T|_{\mathrm{SC}}} |\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}| \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \\ & \times \sum_{t \in \mathcal{T}_{\mathrm{SC}}} \sum_{\tilde{g}} t \circ (\tilde{g} \circ \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}(\Omega_{\mathrm{SC}})) , \end{aligned} \tag{1.97}$$

the following functions are computed:

$$\begin{aligned} \gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\Omega_{\mathrm{SC}}) = & \frac{1}{|T|_{\mathrm{SC}}\, n_{\mathrm{bas}}} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \\ & \times \sum_{t \in \mathcal{T}_{\mathrm{SC}}} \sum_{\tilde{g}} t \circ (\tilde{g} \circ \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}(\Omega_{\mathrm{SC}})) . \end{aligned} \tag{1.98}$$

Comparison shows that $1/n_{\mathrm{bas}} |T|$, which will arise anyway when moving to the energy per atom, has been absorbed into the definition of the basis functions. The factor $1/ |\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$ from Eqn. (1.91) is omitted in this new definition, and therefore the expansion of the energy per atom becomes

$$e(\Omega) = \sum_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \underbrace{\frac{J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}}{|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|}}_{j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}} \gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} := \sum_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}. \tag{1.99}$$

Absorbing $|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$ into the definition of the interaction coefficients makes the program code slightly shorter, as there is no need to calculate $|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$ and to keep track of the result[18]. The drawback is that the $\gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ are just orthogonal and no longer orthonormal, however this has no practical consequences in the experience of the author. As the interaction coefficients are determined by a least mean square fit, a scaling of the basis functions just results in the inverse scaling of the interaction coefficients.

[18] Also, the author has to admit that for quite some time this factor was implemented incorrectly as $|\mathcal{G}_\alpha|$, the number of group elements that transforms α into itself (Eqn. (3.4.1) in [1], there called D).

One caveat is that if the absolute values of interaction coefficients to cluster archetypes with different $|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$ are to be compared, $|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$ has to be taken into account. For this case, a short standalone program exists that computes $|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$.
For spin spirals the program does not use Eqn. (1.77), but instead simply uses that spin spiral configurations (and therefore, the basis functions) are invariant under translations, therefore proportional to the number of translations, and evaluates

$$\gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\Omega_{\mathrm{SS}}) = \frac{1}{n_{\mathrm{bas}}} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \sum_{\tilde{g}} t \circ \left(\tilde{g} \circ \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}(\Omega^{\alpha}_{\mathrm{SS}})\right). \tag{1.100}$$

which, for the numerical implementation, is just Eqn. (1.98) with $T_{\mathrm{SC}} = 1$. The configuration $(\Omega^{\alpha}_{\mathrm{SS}})$ is calculated for each position in the cluster picture $\tilde{g} \circ \alpha$ according to Eqn. (1.71), $\vartheta(\mathbf{r}_i) = \vartheta_0$ and $\varphi(\mathbf{r}_i) = \varphi_0 + \mathbf{q} \cdot \mathbf{r}_i$.

1.7. Convergence of the Expansion

The sum over $\alpha_i, \mathbf{l}_{\alpha i}, \mathbf{u}$ in the expansion (1.99) necessarily needs to converge to some finite value e, for any configuration Ω. As the absolute values of the SCE basis functions do not converge for all configurations, this means the absolute values of the expansion coefficients need to converge. It is presumed that the absolute values of the spin cluster expansion coefficients converge

1. with regard to a maximum l-tuple entry $l_{\max}$:
$$|j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}| \begin{cases} < \epsilon & \text{if } \max(\mathbf{l}_{\alpha i}) > l_{\max} \\ \geq \epsilon & \text{if } \max(\mathbf{l}_{\alpha i}) \leq l_{\max} \end{cases} \tag{1.101}$$
This also limits the eigenvectors $\mathbf{u}$, as the eigenvectors $\mathbf{u}$ are constructed from the basis functions to the possible $\mathbf{k}$-tuples, whose entries are limited to $k_{\max} \leq 2l_{\max}$.

2. with regard to a maximum interaction length $r_{\max}$:
$$|j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}| \begin{cases} < \epsilon & \text{if } \exists i,j \in \alpha_i \text{ so that } |\mathbf{r}_i - \mathbf{r}_j| > r_{\max} \\ \geq \epsilon & \text{if } \nexists i,j \in \alpha_i \text{ so that } |\mathbf{r}_i - \mathbf{r}_j| > r_{\max} \end{cases} \tag{1.102}$$

3. with regard to a maximum number of atoms $n_{\max}$ in the cluster archetype:
$$|j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}| \begin{cases} < \epsilon & \text{if } n_\alpha > n_{\max} \\ \geq \epsilon & \text{if } n_\alpha \leq n_{\max} \end{cases} \tag{1.103}$$

Naturally, suitable values for ϵ are yet to be determined. Numerical tests for these criteria are detailed in the following subsections.
One very important consequence of is that once the limits $l_{\max}$, $\mathbf{r}_{\max}$ and $n_{\max}$ are established, the set of possible basis functions that fulfills these criteria is finite, and one has no longer to deal with an infinite set of possible interactions.

1.7.1. l Convergence

The $l_{\max}$ convergence limit can in principle be tested by calculating the values of the interaction coefficients $j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ for various $\max(\mathbf{l}_{\alpha i})$. These values should be less than ϵ if $\max(\mathbf{l}_{\alpha i})$ exceeds $l_{\max}$. For practical purposes, it is more convenient to define the limit from the accuracy of the SCE, in this case one would expect that only basis functions with $\max(\mathbf{l}_{\alpha i}) \leq l_{\max}$ are necessary to reach an accuracy below a predefined limit $\tilde{\epsilon}$.
$l_{\max}$ can only be investigated for spin spiral or supercell configurations, because no DFT data can be obtained for other configurations. Correspondingly, one has to take the effect of these configurational symmetries on the basis functions into account. This effect is that for spin spirals and supercells $\max(\mathbf{l}_{\alpha i})$ of the relevant clusters may exceed $l_{\max}$, because in both cases the basis functions are mapped on superpositions were sums over elements of the $\mathbf{l}_{\alpha i}$-tuple enter. However, both types of configurations may still be used to obtain estimates for the quantity $n_{\max} l_{\max}$.
If spin spiral configurations with a constant wave vector $\mathbf{q} = \mathbf{q}_0$ and various opening angles ϑ_0 are investigated, one can write the expansion Eqn. (1.99) as

$$e(\vartheta_0) = \sum_{\tilde{l}=\text{even}} j_{\tilde{l}}(\mathbf{q}_0) P_{\tilde{l}}(\cos(\vartheta_0)). \tag{1.104}$$

where Eqn. (1.80) was used, and all $\mathbf{q}_0$-dependent terms as well as all prefactors have been collected in the effective interaction coefficients $j_{\tilde{l}}(\mathbf{q}_0)$. $\tilde{l}$ may, in principle, assume all even integer values, as in Eqn. (1.80) the sum over one specific $\mathbf{l}_{\alpha i}$-tuple enters as the upper limit, and in Eqn. (1.99) one has to sum over all possible $\mathbf{l}_{\alpha i}$ tuples.
However, if the presumptions for convergence of the interaction coefficients hold, the only important basis functions are those with $\mathbf{l}_{\alpha i}$-tuples of dimension $n_{\max}$ or less, and with elements of $l_{\max}$ or less in the $\mathbf{l}_{\alpha i}$-tuple. Therefore, one may limit the sum over $\tilde{l}$ to values below or equal to $n_{\max} l_{\max}$, without loss of accuracy. The absolute values of the $j_{\tilde{l}}$ should also be negligible above $n_{\max} l_{\max}$.
Therefore, one can estimate the lower bound of the product $n_{\max} l_{\max}$ by calculating $e(\vartheta_0)$ for some fixed wave vector and fitting this curve with Legendre polynomials. The highest order of Legendre polynomials required to fit $e(\vartheta)$ with sufficient accuracy (which is yet to be determined) is then an estimate for $n_{\max} l_{\max}$.
This is only an estimate, because according to Eqn. (1.80) each basis function is mapped to a superposition of Legendre polynomials, and omitting the highest order of $\tilde{l}$ in this superposition does not have to carry the same penalty for the accuracy of the model as omitting the whole basis function would. To double-check the estimate, $l_{\max}$ may also be estimated from supercell calculations.
For supercell configurations, all basis functions to cluster archetypes larger than the supercell are mapped on superpositions of basis functions of the clusters within the supercell by the backfolding procedure. At the heart of the this mapping is Eqn. (1.87), which limits the "total angular momentum" $\sum_j l_j$ of basis function that contribute to

the superposition to values lower or equal than that of the original basis function.
The basis function with the largest possible "total angular momentum" and a significant interaction coefficient has the "total angular momentum" of $n_{\max}l_{\max}$ (see the above discussion for spin spirals). If this basis function is backfolded into the supercell, no basis function in the resulting superposition may have a "total angular momentum" greater than $n_{\max}l_{\max}$. Conversely, it should be possible to describe the energies of any supercell configuration with basis functions which have a total angular momentum of less or equal than $n_{\max}l_{\max}$.
Therefore, one may estimate the lower boundary of $n_{\max}l_{\max}$ by calculating the energy of various configurations for one specific supercell, and fit the resulting data with all cluster basis functions with $\mathbf{l}_{\alpha i}$-tuples up to a certain $\sum_j l_j$. The highest $\sum_j l_j$ needed is then an estimate for $n_{\max}l_{\max}$.

1.7.2. Spatial Convergence

One can not reliably test for $\mathbf{r}_{\max}$ by spin spirals, because the positions $\{\mathbf{r}_j\}_i$ of the cluster archetype enter the effective spin-spiral basis functions (Eqn. (1.80)) only in the form of $\mathbf{T}_{\mathbf{m}_+}\cdot\mathbf{q}$, where the translation $\mathbf{T}_{\mathbf{m}_+} = \sum_{j=1}^{n_\alpha} m_j\mathbf{r}_j$ depends on the positions $\{\mathbf{r}_j\}_i$ in the cluster archetype α_i and the elements $\{m_j\}$ of $\mathbf{m}_+$. This precludes a distinction between basis functions to larger values for the elements of $\mathbf{l}_{\alpha i}$ (and correspondingly, larger possible values for the $\{m_j\}$), and basis functions to larger position vectors $\{\mathbf{r}_j\}_i$ in the cluster archetype.
However, the $\mathbf{r}_{\max}$ convergence can be tested by calculating the energy difference due to a localized configuration change in supercells of increasing size along one direction, as shall be illustrated in the following. For example, use a "one-dimensional" supercell, where $\mathbf{T}_1$ and $\mathbf{T}_2$ are the smallest possible translations, and $\mathbf{T}_3$ is succesively increased. The ferromagnetic configuration serves as a reference and the magnetic moment at the origin of the supercell is flipped (sketched in Figure 1.6).
Consider the energy per atom, multiplied by the number of atoms N_{SC} in the supercell $E^{\mathrm{SC}} = e(\Omega)\cdot n_{\mathrm{bas}}\,|T|_{\mathrm{SC}}$. From the point of view of the Spin Cluster Expansion, this quantity is

$$\begin{aligned}
E^{\mathrm{SC}}(\Omega) &= n_{\mathrm{bas}}\,|T|_{\mathrm{SC}} \sum_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} j^{\mathrm{SC}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}\gamma^{\mathcal{G}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}}(\Omega) \\
&\overset{(1.98)}{=} \sum_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} j^{\mathrm{SC}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \frac{n_{\mathrm{bas}}\,|T|_{\mathrm{SC}}}{|T|_{\mathrm{SC}}\,n_{\mathrm{bas}}} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \\
&\quad\times \sum_{t\in\mathcal{T}_{\mathrm{SC}}}\sum_{\tilde{g}} t\circ\tilde{g}\circ\mathcal{Y}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{k}}(\Omega) \\
&= \sum_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} j^{\mathrm{SC}}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{u}} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \sum_{t\in\mathcal{T}_{\mathrm{SC}}}\sum_{\tilde{g}} t\circ\tilde{g}\circ\mathcal{Y}_{\alpha_i\mathbf{l}_{\alpha i}\mathbf{k}}(\Omega).
\end{aligned} \tag{1.105}$$

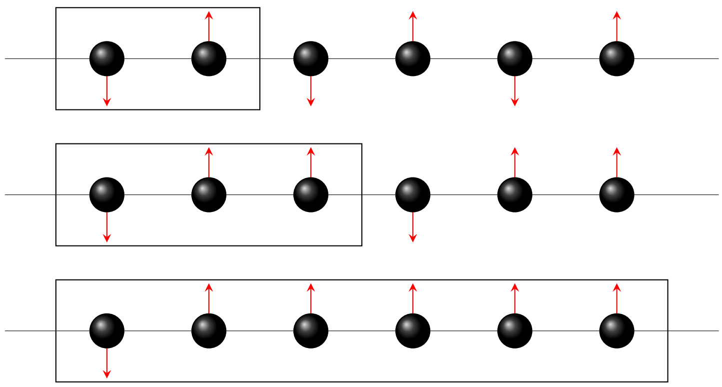

Figure 1.6.: Illustration for the numerical test for the maximum interaction range $r_{\max}$. The configuration change is localized at the origin of the supercell, and the calculation is carried out for increasingly larger supercells. Once the supercell becomes larger than $r_{\max}$, the quantity $e \cdot n_{\mathrm{SC}}$ will not change anymore.

The energy difference between two configurations Ω_1, Ω_2 is

$$\begin{aligned} E^{\mathrm{SC}}(\Omega_1) - E^{\mathrm{SC}}(\Omega_2) = & \sum_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} j^{\mathrm{SC}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \sum_{\mathbf{k}} c^{\mathbf{u}}_{\mathbf{k}} \\ & \times \sum_{t \in \mathcal{T}_{\mathrm{SC}}} \sum_{\tilde{g}} t \circ \tilde{g} \circ (\mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}(\Omega_1) - \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}(\Omega_2)). \end{aligned} \tag{1.106}$$

If the difference between configurations Ω_1 and Ω_2 is localized at position $\mathbf{r}_0$, naturally the values of $t \circ \tilde{g} \circ \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}(\Omega_1)$ and $t \circ \tilde{g} \circ \mathcal{Y}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{k}}(\Omega_2)$ only differ if the picture of the cluster archetye α_i, $t \circ \tilde{g} \circ \alpha_i$, contains $\mathbf{r}_0$ - after the backfolding into the supercell has been taken into account where necessary.
If the size of the supercell is now increased, the set of translations within the new supercell contains a new element. For the calculation of the basis functions one has to sum over all these elements, however, for the basis functions that fit within the old supercell no additional[19] pictures that contain $\mathbf{r}_0$ will be created, and therefore the difference between the values of these these basis functions for Ω_1 and Ω_2 does not change.
This means any change observed in the energy difference by DFT when moving from one supercell size to the other must be due to the backfolding (which changes the values of the effective interaction coefficients $j^{\mathrm{SC}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ for the basis functions which could fit into

[19]Due to the backfolding, it suffices to look at the pictures inside the supercell. If the supercell size is increased in direction of $\mathbf{T}_3$, the only new pictures possible are old pictures that are translated away from $\mathbf{r}_0$.

the old supercell), and to the "new" basis functions that just about fit into the new supercell.
But once the super cell size increases to larger than $r_{\max}$, only clusters with $j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} < \epsilon$ are backfolded, which by definition contribute nothing to the energy change, and all new basis functions that only fit into the new supercell must have negligible interaction coefficients as well because they must contain positions with distances larger than $r_{\max}$. Therefore the energy change has to converge for large supercells, and the size of that supercell where the energy is suitably converged is an estimate for $r_{\max}$.

1.8. How to Obtain the Expansion Coefficients

So far the framework of the SCE has been presented. This is a very general framework which can be used in principle to approximate any function over a configuration space of sphere surfaces. However, the SCE by itself provides no information about a physical system. This information has to be obtained by other means.
The tool of choice is ab-initio density functional theory (DFT), which is described in the following subsection. This allows to probe the ground state energy of a many electron system for various configurations of the magnetic moments, just as the adiabatic approximation in Section 1.2 was defined. A very short overview of DFT as it applies to noncollinear magnetism is given in the following subsection. A general overview of DFT can be found in almost all recent solid state theory text books, and a good, detailed introduction to noncollinear magnetism within DFT is given in [21].
The SCE is then used to interpolate between discrete points obtained by DFT by fitting the SCE to the DFT data points. Details of this fit are described in subsections 1.8.5 to 1.8.6.
This section contains no original work by the present author.

1.8.1. Density Functional Theory

Density functional theory starts from the quantum mechanical many-body problem defined by the Hamiltonian

$$\hat{H}_{\text{many-body}} = \sum_{i=1}^{n_e} \frac{\hat{p}_i^2}{2m_e} + \sum_{i=1}^{n_e} V_{\text{ext}}(\mathbf{r}_i) + \sum_{i<j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{1.107}$$

where the Born-Oppenheimer approximation was used, spin-orbit coupling has been neglected, n_e is the number of electrons, e the charge and m_e the mass of one electron. The external potential V_{ext} contains all potentials that do not depend on the electron wave function, e.g., the potential of the nuclei, external electric fields, etc.
Instead of the conventional search for the eigenstates belonging to this Hamiltonian,

the object of density functional theory is just the spin density matrix[20] $\rho_0(\mathbf{r})$ of the ground state. This is because it can be shown that there is a one-to-one relationship between this density and the ground state (this is the essence of the "first theorem of Hohenberg and Kohn", [25]) . If such a one-to-one relationship exists, one may calculate any operator expectation value as a functional of the ground state density.
To obtain this ground state density matrix, the following energy functional is minimized with regards to ρ:

$$\begin{aligned} E_{\mathrm{DFT}}[\rho] = T[\rho] + E_{\mathrm{xc}}[\rho] + & \underbrace{E_{\mathrm{ext}}[\rho]}_{=\int \mathrm{d}^3\mathbf{r} V(\mathbf{r})\rho(\mathbf{r})} \\ & + \frac{e^2}{2}\int \mathrm{d}^3\mathbf{r} \int \mathrm{d}^3\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \end{aligned} \tag{1.108}$$

where $n(\mathbf{r})$ is the charge density part of the spin density matrix.
There are two key features in this functional:

- First, the main part of the Coulomb energy is described by the "Hartree energy":

$$\begin{aligned} E_{\mathrm{Cou}} &= \left\langle \Psi \left| \sum_{i<j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right| \Psi \right\rangle \\ \to E_{\mathrm{Har}} &= \frac{e^2}{2}\int \mathrm{d}^3\mathbf{r} \int \mathrm{d}^3\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \end{aligned} \tag{1.109}$$

  where Ψ is the (unknown) many-body wave function.

- Second, the main contribution of the kinetic energy, is approximated as the kinetic energy $T[\rho]$ of a system of non-interacting particles. This model system is chosen such as to have the same ground state density as the actual, physical system.

$$\begin{aligned} E_{\mathrm{kin}}^{\mathrm{real}} &= \left\langle \Psi \left| \sum_i \frac{\hat{p}_i^2}{2m_e} \right| \Psi \right\rangle \\ \to E_{\mathrm{kin}}^{\mathrm{model}} &= \sum_{i,\mathrm{occ}} \left\langle \Phi_i^{\mathrm{mod}} \left| \frac{\hat{p}^2}{2m_e} \right| \Phi_i^{\mathrm{mod}} \right\rangle \\ &= \sum_{i,\mathrm{occ}} \epsilon_i - \int \mathrm{d}^3\mathbf{r}\ \mathrm{Tr}(\rho W_{\mathrm{eff}}), \end{aligned} \tag{1.110}$$

  where the Φ_i^{mod} are the eigenstates of the non-interacting particles in the model system, and the ϵ_i are the corresponding eigenvalues.

[20] The basis of this matrix are the spin eigenfunctions, $|\uparrow\rangle$ and $|\downarrow\rangle$. It may be decomposed into the charge density part and the spin density part according to $\rho = 0.5n(\mathbf{r})\mathbb{1} + 0.5\mathbf{m}(\mathbf{r})\cdot\sigma$, where $\sigma = (\sigma_x, \sigma_y, \sigma_z)$, with the well-known Pauli spin matrices $\{\sigma_i\}$.

> The potential W_{eff} of the model system can be calculated from the requirement that the ground state density matrix has to be equal to ρ_0 of the physical system.

The first theorem of Hohenberg and Kohn, requires that there is some functional that fulfills $E_{\text{xc}}[\rho_0] = E_{\text{Cou}}[\rho_0] - E_{\text{Har}} + E_{\text{kin}}^{\text{real}}[\rho_0] - E_{\text{kin}}^{\text{model}}[\rho_0]$ for the ground state density ρ_0. But this is no definition that can be used in a practical calculation, because the calculation of E_{xc} from this definition would require knowledge of the many-body wave function. Therefore, $E_{\text{xc}}[\rho]$ has to be approximated. This is done by investigating the defining equation for model systems where (good approximations of) the many-body wave function is known. For further details on how such approximations are obtained, see, e.g., [26].
$E_{\text{xc}}[\rho]$ contributes only a small part of the total energy of a state, however, large parts of the energy differences between different states (e.g., magnetic, nonmagnetic state) are contributed by this functional. Therefore, a good approximation of this functional is absolutely essential.
To actually find the ground state the functional (1.108) has to be minimized regarding ρ. This is accomplished by starting with an initial guess ρ_1 for the ground state density matrix ρ_0, and then solving a self-consistency problem where the residual is given by the difference between the current ρ_n and the $\tilde{\rho}_{n+1}$ that solves the model system given by

$$\hat{H}_{\text{model}} = \frac{\Delta^2}{2m_e} + W_{\text{eff}}[\rho_n](\mathbf{r}). \tag{1.111}$$

The self-consistency problem can be solved by various quasi-Newton mixing algorithms, e.g., [27]. In these methods, the actual ρ_{n+1} after one step is a mix of several of the previous steps $\rho_i, i \leq n$ and $\tilde{\rho}_{n+1}$.
If this cycle reaches self-consistency, that is $\rho_n - \rho_{n+1} < \epsilon$ for some suitably chosen ϵ, one has reached the minimum of the model system with regards to ρ, and therefore obtained the ground state density matrix ρ_0 for the physical system, from which the ground state energy may be calculated.

1.8.2. Constrained Density Functional Theory

The original formulation of DFT only allowed to investigate the global ground state of a system. To investigate other states, the problem is reformulated as the ground state of a limited class of density matrices which, e.g., have to correspond to one specific expectation value of an operator. This is accomplished by employing Lagrange multipliers [28]. This work is interested in states which correspond to specific magnetic moments $\mathbf{M}_i^0$ where i labels the atoms. These moments are prescribed at the start of a DFT calculation, and then the following energy functional is minimized [29]:

$$E_{\text{CDFT}}[\rho] = E_{\text{DFT}}[\rho] - \sum_i \mathbf{B}_i^L \cdot (\mathbf{M}_i - \mathbf{M}_i^0) \tag{1.112}$$

where the $\mathbf{B}_i^L$ are the Lagrange multipliers.
The resulting spin density matrix represents the density matrix with the lowest energy of all density matrices that fulfill the requirement that the magnetic moments $\mathbf{M}_i$ correspond in direction and absolute value to $\mathbf{M}_i^0$. If one is only interested in fixing the directions of the magnetic moments to some $\mathbf{e}_i^0$, the modified energy functional is

$$E_{\mathrm{CDFT}}[\rho] = E_{\mathrm{DFT}}[\rho] - \sum_i \mathbf{B}_i^L \cdot \underbrace{(\mathbf{e}_i^0 \times (\mathbf{M}_i \times \mathbf{e}_i^0))}_{\mathbf{M}_i^{\perp} - \mathbf{M}_i^{0\perp}}, \tag{1.113}$$

where $\mathbf{M}_i^{\perp}$ is the part of the magnetic moment orthogonal to $\mathbf{e}_i^0$, and the constraint $\mathbf{M}_i^{0\perp}$ is zero (no orthogonal moment).
To find the ground state, the energy functional is minimized regarding ρ and the $\mathbf{B}_i^L$. To minimize with regards to the $\mathbf{B}_i^L$, the change of $\mathbf{B}_i^L$ in one step of the self-consistency cycle is estimated as $\mathbf{M}_i^0 - \mathbf{M}_i$ for general constraints on the magnetic moments, and correspondingly as $\mathbf{M}_i^{\perp 0} - \mathbf{M}_i^{\perp} = -\mathbf{M}_i^{\perp}$ for transversal constraints. The minimization of $\mathbf{B}_i^L$ is then carried out by the same mixing methods mentioned in Section 1.8.1, concurrently with the minimization of ρ.

1.8.3. Key Features and Approximations in DFT Implementations

Basis Sets

The one-particle wave functions of the model system for the kinetic energy are always expanded in a complete basis set of E^3, the three dimensional euclidean space. Any complete basis set for E^3 may be chosen, and many different sets have been used in different DFT computer codes.
These basis functions are in many DFT codes partitioned into two parts: One part defined over positions far away from the nuclei, where the potential is essentially constant, and one part close to the nuclei where the $1/(\mathbf{r} - \mathbf{R}_{\mathrm{nuc}})$ potential of the nuclei diverges. For the part far away from the nuclei, functions ranging from plane waves to spherical waves to Hankel functions are used. The resulting methods are called “APW” (augmented plane wave), “ASW” (augmented spherical wave), or “LMTO” (linear muffin tin orbital) methods. The last name has historical reasons, the regions close to the nuclei are called “muffin tins” in all three methods. For the functions close to the nuclei, solutions to a Schrödinger equation with spherical potential are used. This is the augmentation referred to in the classification of the methods.

Atomic Sphere Approximation (ASA)

There are two unrelated parts that are sometimes summarized under this approximation: For the first approximation, the spherical “muffin tins” around the atoms are chosen such

that the total volume of all muffin tin spheres in one unit cell is equal to the volume of the unit cell. This means the muffin tin spheres overlap.
If now an integral for the expectation value of an operator runs over the whole space, it can be decomposed into an integral over the muffin tin spheres, and an integral over the interstitium. Because the integral over the muffin tins will cover the overlap regions twice (once for each muffin tin), one has to subtract from the whole integral the integral over the overlap region once. The approximation is now that the latter integral over the overlap region and the integral over the interstitium are presumed to cancel each other, and therefore the integral over the space within the muffin tins is a good approximation of the real integral.

$$\begin{aligned}\int_{E^3} \ldots \, \mathrm{d}^3\mathbf{r} = \int_{\mathrm{MT}} \ldots \, \mathrm{d}^3\mathbf{r} + \int_{\mathrm{Inter}} \ldots \, \mathrm{d}^3\mathbf{r} \\ - \int_{\mathrm{Overlap}} \ldots \, \mathrm{d}^3\mathbf{r} \approx \int_{MT} \ldots \, \mathrm{d}^3\mathbf{r}.\end{aligned} \tag{1.114}$$

The other approximation is that every quantity within the muffin tins, including the spin density, is presumed to be spherically symmetric. This can be motivated by the fact that the dominant part of the external potential is the potential of the nucleus, which is spherically symmetric. If the density is spherically symmetric, the Hartree energy will be spherically symmetric as well. The only sources of asymmetry are then external potentials or spin-orbit coupling.

Atomic Moment Approximation/Spin-ASA

Exchange-correlation functionals are usually derived for collinear systems, and therefore the off-diagonal parts of the spin density matrix are often not taken into account in a specific approximation of this functional. In the “local spin density approximation” of the exchange-correlation functional, only local effects are taken into account, and the exchange-correlation functional is approximated as

$$E_{\mathrm{xc}} = \int n(\mathbf{r}) e_{\mathrm{xc}} \left(n(\mathbf{r}), |\mathbf{m}(\mathbf{r})| \right) \, \mathrm{d}^3\mathbf{r}, \tag{1.115}$$

where e_{xc} is a function of the local charge density and magnetization. Therefore, the correct way to treat the noncollinearity would be to calculate the absolute value of the local magnetization at each point of the integration grid used for the exchange-correlation functional. However, because usually the magnetization is largely collinear in the muffin-tin region, $|\mathbf{m}(\mathbf{r})|$ is quite often supplanted by $\mathbf{m}(\mathbf{r}) \cdot \mathbf{e}_{\mathbf{M}}$, where $\mathbf{M}$ is the direction of the total magnetic moment in this sphere. This approximation has been coined “Atomic Moment Approximation” (AMA) [30] or “Spin-ASA” [2].

1.8.4. DFT Program Codes Used

This work covers two exemplary materials, iron in the bcc (α) phase, and iron in the fcc (γ) phase. Unless otherwise stated, the numerical results for bcc iron have been obtained by the present author, and the numerical results for fcc iron have been obtained by R.Singer during his PhD-thesis [1]. For bcc iron, a program code based upon the "Stuttgart TB-LMTO-ASA" code developed in the department Andersen at the Max Planck Institute for Solid State Research has been used. This code was heavily modified by various PhD students in the working group of Prof. M. Fähnle, but this modified code is currently not publicly available. It is simply called "LMTO" throughout this work, and employs a LMTO basis set, the Atomic Sphere Approximation and the Atomic Moment Approximation. The code branch of Andersen et al. is available to the public, but has since then evolved considerably from the common ancestor with the program used in this work.
For bcc iron, the effect of the basis set and the ASA approximation were controlled by calculations with "FLEUR", a program code maintained by the working group of Prof. Blügel at the Forschungszentrum Jülich. This code employs an augmented plane wave basis set and the AMA, but not the ASA. No significant differences were found.
The results for fcc iron have been obtained via a program code currently maintained by L.Sandratskii at the Theory Department of the Max Planck Institute for Microstructure Physics, which employs an augmented spherical wave basis set and neither the ASA nor the AMA. It is called "MASW" throughout this work. For the reasons behind this choice, see [1, 31].
A tabular overview:

Program	basis	ASA	AMA
LMTO	linear muffin tin orbitals	yes	yes
FLEUR	augmented plane waves	no	yes
MASW	augmented spherical waves	no	no

1.8.5. The Expansion Coefficients

Formally the expansion coefficients are defined as:

$$\begin{aligned} J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} &= \langle E | \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \rangle \\ &= \frac{1}{(4\pi)^N} \int_{S_1} \cdots \int_{S_N} E(\{\mathbf{e}_i\}) \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\{\mathbf{e}_i\}) \, \mathrm{d}\Omega_1 \ldots \, \mathrm{d}\Omega_N. \end{aligned} \tag{1.116}$$

It would be theoretically possible to tackle this integral over the phase space with multi-dimensional numerical integration methods, using DFT to obtain energies at the integration grid points. This is highly impractical due to the computational inefficiency of DFT, although such a calculation is possible for highly symmetric configurations (i.e.

small supercells), and an example of this is presented in Section 2.2.1. A much more practical approach is to preselect a group of n_{bas} possible expansion coefficients or basis functions, and then employ a least-mean square fit of n_{conf} reference energies $\{E^{\mathrm{DFT}}\}$ obtained by DFT against the predictions of the SCE expansion $\{E^{\mathrm{SCE}}\}$. This usually works reliably even with significantly less reference configurations than would be needed to solve the integral in Eqn. (1.116).
The standard form of the least-mean squares problem is

$$\mathbf{x}^t \left(\mathsf{A}^t\mathsf{A}\right)\mathbf{x} - 2\mathbf{b}^t\mathsf{A}\mathbf{x} := \min, \tag{1.117}$$

where in our case[21] $\mathbf{x}$ are the n_{bas} possible expansion coefficients considered, $\mathbf{b}$ contains the n_{conf} reference energies and A is the $n_{\mathrm{conf}} \times n_{\mathrm{bas}}$ matrix of the basis functions for the corresponding configurations. A least-mean squares fit is a very well known and common problem, therefore this work does include neither a detailed description nor an introduction to the methods used to solve this problem. The numerical library used to solve the problem is MINPACK, developed by the University of Chicago as operator of Argonne National Laboratory, and available from www.netlib.org/minpack .
To understand the following chapters on numerical results, the concepts of the residual sum of squares, cross validation, and the singular value decomposition are necessary. These are defined below.

Residual sum of squares

The residual sum of squares is defined as:

$$\mathrm{RSS} := \mathbf{x}_0^t \left(\mathsf{A}^t\mathsf{A}\right)\mathbf{x}_0 - 2\mathbf{b}^t\mathsf{A}\mathbf{x}_0 \tag{1.118}$$

that is, the remaining (residual) value of the cost function after minimization. If there are enough suitably chosen reference configurations this will also be a good measure of

[21] Written explicitly:

$$\begin{aligned}
\sum_{j=1}^{n_{\mathrm{conf}}} \left| \sum_{\alpha\mathbf{lu}} J_{\alpha\mathbf{lu}} \mathcal{Y}^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}(\Omega_j) - E^{\mathrm{DFT}}(\Omega_j) \right|^2 &: = \min \\
\sum_{j=1}^{n_{\mathrm{conf}}} \left| \sum_i A_{ji}x_i - b_j \right|^2 &: = \min \\
\sum_{j=1}^{n_{\mathrm{conf}}} \left[\left(\sum_i A_{ji}x_i \right)^2 - 2b_j \sum_i A_{ji}x_i + b_j^2 \right] &: = \min \\
\mathbf{x}^t \left(\mathsf{A}^t\mathsf{A}\right)\mathbf{x} - 2\mathbf{b}^t\mathsf{A}\mathbf{x} + \underbrace{\mathbf{b}^t\mathbf{b}}_{\mathrm{const}} &: = \min,
\end{aligned}$$

the expected error. To make this connection more apparent, the rest of this work uses the normalized residual sum of squares,

$$\begin{aligned} \mathrm{nRSS} :&= \frac{1}{n_{\mathrm{conf}} - n_{\mathrm{bas}}} \left[\mathbf{x}_0^t \left(\mathsf{A}^t \mathsf{A} \right) \mathbf{x}_0 - 2 \mathbf{b}^t \mathsf{A} \mathbf{x}_0 \right] \\ &= \frac{1}{n_{\mathrm{conf}} - n_{\mathrm{bas}}} \sum_{j=1}^{n_{\mathrm{conf}}} \left| E^{\mathrm{SCE}}(\Omega_j) - E^{\mathrm{DFT}}(\Omega_j) \right|^2 . \end{aligned} \tag{1.119}$$

The prefactor $1/(n_{\mathrm{conf}} - n_{\mathrm{bas}})$ instead of $1/n_{\mathrm{conf}}$ is due to the fact that as long as A has full rank, one can always fit n_{bas} interaction coefficients exactly.
The expected standard deviation of the SCE model to density functional theory is then $\sigma = \sqrt{\mathrm{nRSS}}$.

Cross validation

This refers to the simple method of partitioning the available reference energies either manually or automatically into a training and a testing set. The SCE model is fitted to the training set of reference energies, and then used to predict the energies of the testing set.
If there are sufficiently many and suitably chosen reference configurations in the training set, the nRSS of the training set and the normalized sum of squared prediction errors defined by

$$\mathrm{nPSS} := \frac{1}{n_{\mathrm{test}}} \sum_{j \in \{\mathrm{Test}\}} \left| E^{\mathrm{SCE}}(\Omega_j) - E^{\mathrm{DFT}}(\Omega_j) \right|^2 . \tag{1.120}$$

are expected to be approximately equal. If this is not the case, then this suggests that the training set has a strong sample bias, i.e., the solution to the minimization problem depends strongly on the current training set, which means that this subset of the reference data does not cover a sufficient part of the total phase space to obtain the true solution.
The SCE program code also calculates an “automated” nPSS (termed anPSS in the following), where this procedure is repeated several times for randomly chosen training and testing sets. The size of the training set is hereby constant at 90% of all configurations, and the size of the testing set is correspondingly fixed at 10% of all configurations[22].
The procedure is repeated at least 20 times and until the average of the anPSS over all past runs changes less than 0.1%.
If this average prediction error is significantly larger than the nRSS it is very easy to

[22] A variant of cross validation used for the classic cluster expansion of alloys, and presented in [1], is to use a testing set of size one, and repeat the calculations n_{conf} times (every configuration is in the testing set once). However, as n_{conf} is of the order of 10^3 for the SCE, the resulting quantity is effectively indistinguishable from the RSS.

find training sets with large sample bias, and this casts doubt on the sample bias of the complete reference set.
The term "nPSS" is henceforth used to refer to the quantity in Eqn. (1.120) for manually chosen testing sets, while the term "anPSS" refers to the result of the automated procedure of the preceding paragraph. If one manually divides the set of all reference configurations into a training and testing set, the anPSS is obtained from further subdividing the training set, but does not consider the energies of the testing set.

Singular value decomposition

This decomposition of the matrix A is defined as the right hand side of

$$\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^{\dagger}, \tag{1.121}$$

where Σ is a $n_{\text{conf}} \times n_{\text{bas}}$ rectangular diagonal matrix, U is a unitary $n_{\text{conf}} \times n_{\text{conf}}$ matrix and V is a unitary $n_{\text{bas}} \times n_{\text{bas}}$ matrix. The diagonal elements of Σ are called the "singular values" of A.
In the singular value decomposition, rank deficiency shows up as vanishing singular values, as U and V have full rank and therefore $\text{rk}(\mathsf{A}) = \text{rk}(\Sigma)$. The matrix A should have full rank, because $\text{rk}(\mathsf{A}) = \text{rk}\left(\mathsf{A}^t\mathsf{A}\right)$, and

$$\left(\mathsf{A}^t\mathsf{A}\right)_{mn} = \sum_{j=1}^{n_{\text{conf}}} \mathcal{Y}_m{}^{\mathcal{G}}(\Omega_j)\mathcal{Y}_n{}^{\mathcal{G}}(\Omega_j) \approx \delta_{mn} \tag{1.122}$$

should be guaranteed for a high enough number of suitably chosen configurations due to the orthogonality of the basis functions. If one finds vanishing singular values, then either the basis functions used are not orthogonal (e.g., due to overlooked fundamental symmetries) or the reference configurations used are not representative for the whole configuration space.

1.8.6. Identifying an Efficient Basis Set

Not all basis function within the convergency limits of the expansion are equally important for fitting the energy landscape. To obtain the best numerical efficiency of the resulting model one aims to use the least amount of computation time that still provides sufficient accuracy for the task the model shall be used for.
This requires an estimate how much time the calculation of the value of a specific basis function requires. This calculation requires the following steps:

Spherical harmonics: A basis function contains a product over n_α spherical harmonics. As one spherical harmonic is typically used in many basis functions, it is efficient to precompute the values of all spherical harmonics that occur in some basis function for each magnetic moment direction in a configuration, and then store these values in a lookup table. Then, obtaining the value of one specific spherical harmonic requires just one reference to this lookup table. Therefore, the computation time required for this part of the calculation of the basis functions scales with n_α.

Rotational symmetry: This involves the calculation of the generalized Wigner coefficients, which do not depend on the specific configuration and are therefore again precomputed and stored in a lookup table.
Furthermore, one has to sum over all possible $\mathbf{m}_\alpha$-tuples (see Eqn. (1.14)). For each $\mathbf{m}_\alpha$-tuple the program requires one reference to the lookup table for the Wigner coefficients, and n_α references to the lookup table for the spherical harmonics.
Each element m_j in the $\mathbf{m}_\alpha$ tuples has $2l_j + 1$ possible values, where l_j is the element of the $\mathbf{l}_\alpha$-tuple associated with the same position and spherical harmonic. Therefore, there are $\prod_{j=1}^{n_\alpha}(2l_j + 1)$ possible values of the $\mathbf{m}_\alpha$-tuple.
However, the generalized Wigner coefficients are only nonzero if $\sum_j m_j = 0$ holds. $\mathbf{m}_\alpha$-tuples that fulfill this condition are again precomputed, and a rough estimate for the remaining $\mathbf{m}_\alpha$-tupels is given by $\prod_{j=1}^{n_\alpha}(2l_j + 1)$, $l_j \neq \max(\mathbf{l}_{\alpha i})$, that is the product skips over the largest value of $\mathbf{l}_{\alpha i}$. Therefore, the computation time required for this part of the calculation of the value of the basis functions is expected to be roughly proportional to $\prod_{j=1}^{n_\alpha}(2l_j + 1)$, $l_j \neq \max(\mathbf{l}_{\alpha i})$.

Space group symmetry: The space group symmetry introduces a sum over the action of all group elements on a cluster archetype (see Eqn. (1.98)). Because the sum runs over all group elements, it has the same number of elements for all basis functions. Therefore, this sum has no impact on the relative computation time between basis functions[23].
Calculating the pictures generated by the group elements acting on the archetypes requires more time for cluster archetypes with more atoms. However, these pictures may be precomputed for any application where one is interested in many different configurations to one specific supercell symmetry, e.g., for a Monte Carlo simulation with a fixed supercell. Therefore, this part of the calculation is disregarded when estimating the computation time required.
Furthermore, the basis functions to $n_\alpha \geq 4$ contain a sum over the components of the eigenvector $\mathbf{u}$. These eigenvectors represent superpositions of basis functions to equal α_i, $\mathbf{l}_{\alpha i}$ with different $\mathbf{k}$. However, as the $\mathbf{k}$ only enter the generalized Wigner coefficients, which are configuration independent, one may precompute the necessary superpositions of the generalized Wigner coefficients, and store them in

[23]This part of the computation does scale with the number of atoms (translations) in the supercell, giving an overall linear scaling of the SCE with the number of atoms in the supercell.

a lookup table. Then the eigenvectors have no impact on the computation time.

Overall, the computation time t_{comp} of a basis function is estimated as

$$t_{\mathrm{comp}} \propto n_\alpha \prod_{j=1}^{n_\alpha} (2l_j + 1), \quad l_j \neq \max(\mathbf{l}_{\alpha i}). \tag{1.123}$$

Numerical tests for 1000 configurations to a 4-atomic supercell symmetry in bcc iron indicate this is a reasonable estimate, giving a proportionality factor of on average 7μs (on an AMD Opteron 2218 CPU), where the proportionality factor differs for individual basis functions by up to 3μs, which is probably due to the rough estimate for the number of $\mathbf{m}_\alpha$-tuples, compiler optimizations and the inevitable numerical noise inherent in measuring computation times.
To determine an efficient basis set, the change in the nRSS per computation time, $\Delta\mathrm{nRSS}/t_{\mathrm{comp}}$ associated with adding a specific basis function to a SCE model is compared to the same quantity for all other basis functions under consideration. The basis function that most efficiently decreases the nRSS is then added to the model, whereupon the process is repeated until some predefined limit for the nRSS or for the number of basis functions is reached.
ΔnRSS is calculated by fitting once without and once with this specific basis function. Because t_{comp} is roughly proportional to $n_\alpha \prod_{j=1}^{n_\alpha}(2l_j + 1), \quad l_j \neq \max(\mathbf{l}_\alpha)$ according to Eqn. (1.123), ΔnRSS is divided by this quantity to decide which basis function most efficiently decreases the nRSS.
This method has been coined "forward reduction" in [1], although there only the largest decrease in the nRSS decided which basis function should be added to the model, leading to a nRSS optimized with regards to the number of basis functions. However, the author is of the opinion that the number of basis functions is a not particularly useful quantity by itself.

2. Results For bcc Iron

The bcc crystal structure of iron is the iron crystal structure of lowest energy up to 1183 K and between 1663 K and the melting point of iron at 1812 K. The ground state is ferromagnetic, with magnetic moments of 2.26 μ_B per atom. This magnetic moment is mostly due to the six 3d-electrons, which contribute about 2.35 μ_B to the total moment, while the two 4s electrons contribute -0.015 μ_B (the negative sign indicates that the magnetic moment is antiparallel to the magnetic moment of the 3d electrons) and the 3p electrons contribute -0.05 μ_B. Even in the antiferromagnetic state, the local magnetic moments still have absolute values of 1.88 μ_B.

2.1. Density Functional Theory (DFT)

The expansion coefficients $j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ are obtained by a least-mean squares fit over a set of reference configurations, whose energies have been obtained by means of ab-initio density functional theory. While such a fit can sometimes act as a filter for random noise [32, 33], it will always reproduce any systematic errors inherent in the energies of the reference configurations. Any random noise inherent in the energies of the reference configurations also inherently limits the best achievable nRSS of the SCE model relative to these reference energies.
There are two sources of error in the DFT data, one pseudorandom due to numerical integration issues, and one systematic due to the necessary approximations to the exchange-correlation functional. These are detailed below.

2.1.1. k Point Convergence

The effective potential for the model system of DFT (Eqn. (1.108)) inherits the space group symmetry of the crystal. This means that the model wave functions Φ^{mod} are Bloch functions with a wave vector $\mathbf{k}$ out of the first Brillouin zone. To calculate expectation values of observables, it is necessary to integrate over the first Brillouin zone, using Fermi-Dirac statistics.
Any implementation of DFT in a computer program necessarily chooses a discrete and finite set of $\mathbf{k}_i$ out of the first Brillouin zone to represent the Brillouin zone. The integral over the first Brillouin zone then has to be approximated as a sum over the elements of $\{\mathbf{k}_i\}$. A large, well distributed set naturally gives a better description of the Brillouin

zone integrals at the cost of increased numerical effort.
One problem that arises when these integrals are computed is that if the Fermi surface is defined too narrowly, numerical oscillations occur during the search for the ground state density matrix, because eigenstates of the model system with eigenvalues close to the Fermi surface that were occupied in one step might not be occupied in the next step. To avoid this, the Fermi surface is usually smeared out to some degree, so that these states are partially occupied, and instead of oscillating between 0 and 1 the occupation $f(\epsilon)$ changes more smoothly between steps.
For the calculations presented here, a Gaussian smearing of the form

$$f(\epsilon) = \frac{1}{2}\left[1 - \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{e}^{-t^2}\,\mathrm{d}t\right], \qquad x = \frac{\epsilon - \epsilon_\mathrm{F}}{w} \tag{2.1}$$

is used, where ϵ_F is the Fermi energy and w is a suitably chosen parameter. A comparatively coarse sampling of the eigenvalues (corresponding to a small set $\{\mathbf{k}_i\}$) requires are greater degree of smearing to suppress oscillations than a very fine-grained sampling. On the other hand a very large smearing will misrepresent the Fermi surface, which tends to overestimate the energy.
Usually these errors are investigated by selecting a couple of configurations (e.g., the ferromagnetic and the antiferromagnetic state) and increasing the number of k-points until the difference between the resulting energies is small enough. As mentioned before, less smearing of the Fermi surface is required if the grid density is higher, and so the width w of the Gaussian is chosen such as to keep $(n_\mathbf{k})^{\frac{1}{3}} \cdot w$ constant where $n_\mathbf{k}$ is the current size of the set $\{\mathbf{k}_i\}$.
After such calculations, the parameters $n_\mathbf{k} = 33^3$ and $w = 0.2272727\ldots$ mRy ($(n_\mathbf{k})^{\frac{1}{3}} \cdot w = 7.5$mRy) were adopted. These parameters were used for the calculation of all reference energies for the SCE fit.
The estimated error was of the order of 2 μRy, and was estimated by calculating several canted configurations with different combinations of $n_\mathbf{k}$ and w, and observing the convergence of the difference of the energy relative to the ferromagnetic ground state energy calculated with the same $n_\mathbf{k}, w$.
However, for spin spirals another problem that to the knowledge of the author has so far not been discussed in the literature occurs: The modified translational symmetry of spin spirals corresponds to a shift of the elements of $\{\mathbf{k}_i\}$ by $\mathbf{q}$ for majority spin wave functions and $-\mathbf{q}$ for minority spin wave functions [21, 2]. However, the standard algorithms used in most DFT programs to choose an optimal set of $\{\mathbf{k}_i\}$ for the Brillouin zone integration (in the LMTO program, the algorithm of Moreno and Soler [34]) have not been adapted to address this issue - if such an adaption is possible.
A straightforward test for this effect is to calculate spin spirals with different $\mathbf{q}$-vectors but an opening angle of zero. Physically, this is just one state, the ferromagentic state. However, the different $\mathbf{q}$-vectors will introduce different shifts of the $\mathbf{k}$ vectors and induce small differences which are an estimate for the severity of the issue. Naturally. this

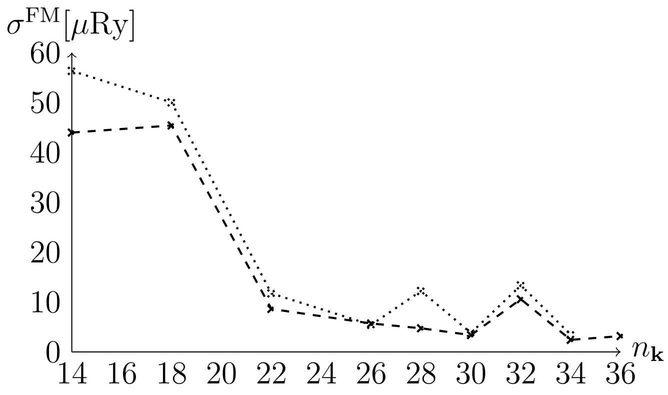

Figure 2.1.: k-point convergence and error estimate. Dashed curve: $(n_{\mathbf{k}})^{\frac{1}{3}} \cdot w = 7.5$mRy
Dotted curve: $(n_{\mathbf{k}})^{\frac{1}{3}} \cdot w = 100$mRy

effect is expected to be small when the set $\{\mathbf{k}_i\}$ covers the first Brillouin zone well.
To estimate this error, the following procedure is adopted:

1. A set $\{\mathbf{q}\}$ of 209 spin-spiral wave vectors is chosen, which covers the first Brillouin zone reasonably well. A DFT calculation for a spin-spiral with an opening angle of zero is done for each element of $\{\mathbf{q}\}$. The standard deviation of the resulting energies relative to the mean is used as an estimate of the error inherent in the calculations.

2. These calculations are repeated for larger and larger sets $\{\mathbf{k}_i\}$. The smearing w of the Fermi surface is again chosen such as to keep the product $(n_{\mathbf{k}})^{\frac{1}{3}} \cdot w$ constant.

3. This is in turn repeated for various choices of the product $(n_{\mathbf{k}})^{\frac{1}{3}} \cdot w$.

The result of this procedure is the standard deviation

$$\sigma^{\mathrm{FM}}(n_{\mathbf{k}}, w) := \frac{1}{\sqrt{n_{\mathbf{q}}}} \sqrt{\sum_{\mathbf{q}} \left(E^{\mathrm{FM}}(\mathbf{q})\right)^2 - \left(\sum_{\mathbf{q}} E^{\mathrm{FM}}(\mathbf{q})\right)^2} \tag{2.2}$$

where $n_{\mathbf{q}}$ is the size of the set $\{\mathbf{q}\}$. This is presented in Figure 2.1. Different choices of $(n_{\mathbf{k}})^{\frac{1}{3}} \cdot w$ do not have a large effect on σ. For the choice of parameters for the reference calculations, $n_{\mathbf{k}} = 33^3, w = 0.2272727\ldots$ mRy, the estimate for σ is 8.3 μRy. Therefore, this work uses 8.3μRy as an error estimate for the DFT date.
If supercell configurations are calculated, the volume of the real-space cell is increased, and the volume V_{rec} of the reciprocal space cell correspondingly decreases. The size of $\{\mathbf{k}_i\}$ for the supercell, $n_{\mathbf{k}}^{\mathrm{SC}}$ is chosen such that $n_{\mathbf{k}}^{\mathrm{SC}}/V_{\mathrm{rec}}$ remains approximately constant. This does not result in k-points that are equally distributed within the common volume

Functional	$E_{\mathrm{FM}} - E_{\mathrm{AFM}}$
Barth and Hedin (LSDA,1972)[35]	29.7 mRy
Perdew and Wang (LSDA,1986) [36]	31.2 mRy
Perdew and Wang (GGA,1986) [36]	33.6 mRy
Perdew, Burke and Enzenhofer (GGA,1996) [37]	34.4 mRy

Table 2.1.: Energy difference between ferromagnetic and antiferromagnetic configuration of bcc iron for various energy functionals at the experimental lattice constant

of two different reciprocal cells, just in the same k-point density. This has the effect of shifting the energies slightly (which is later corrected for), but a negligible effect on energy differences.

2.1.2. Density Functionals

A source of systematic error are the functionals used to approximate the exchange-correlation term in Eqn. (1.108). There have been many, increasingly sophisticated approaches to the problem of modeling the exchange-correlation functional [26], starting from the so-called "local spin-density functional approximation" (LSDA), which takes only the local spin density into account, over the "generalized gradient approximation" (GGA), which includes gradient terms of the spin density, to the so-called "meta-GGA". However, the generalization of the more sophisticated approaches to noncollinear magnetism is not always clear, and quite often noncollinear magnetism is not taken into account when the functionals are derived. Furthermore, functionals beyond the GGA were not implemented within the computer codes used for this work. Therefore the calculations presented for bcc iron were done within the LSDA, except some of the following model calculations within GGA.

Most functionals do reproduce the ground state magnetic moment rather well, but yield significantly varying results for, e.g., the difference between the energies of the ferromagnetic and the antiferromagnetic state, or the lattice constant. Some examples for the energy difference between the ferromagnetic and the antiferromagnetic state are given in Table 2.1. The differences between (modern) functionals are of the order of 1mRy. This is an unsatisfactory state of affairs, as this is much higher than the numerical error discussed above, and it will be shown below that the fit of the SCE to the DFT data also has a much higher accuracy. Thus, the SCE very likely describes the energy landscape of DFT more accurately than DFT describes the real, physical energy landscape.

For all calculations for bcc iron, unless otherwise stated, the LSDA-functional of Perdew and Wang was used. For all calculations for fcc iron, the GGA-functional of Perdew, Burke and Enzenhofer was used.

$\tilde{l}$	$\mathbf{q}_1$		$\mathbf{q}_2$	
	$j_{\tilde{l}}$ [mRy]	$\sqrt{\text{nRSS}}$ [μRy]	$J_{\tilde{l}}$ [mRy]	$\sqrt{\text{nRSS}}$ [μRy]
2	-19.022	998	-9.457	390
4	2.228	167	0.948	28
6	0.374	60	-0.085	12
8	0.128	32	0.042	10
10	0.108	17	0.031	6

Table 2.2.: Values of $j_{\tilde{l}}(\mathbf{q})$ in mRy, obtained by least mean square fit The columns marked $\sqrt{\text{nRSS}}$ give the square root of the nRSS that can be achieved if all Legendre polynomials up to the current $\tilde{l}$ are used.

2.2. SCE Parameters

One central assumption is that the spin cluster expansion converges with regard to the maximum order $l_{\max} = \max(\mathbf{l}_{\alpha i})$ and the interaction range $r_{\max} = \max(|\mathbf{r}_j - \mathbf{r}_k| \in \alpha_i)$. This section details numerical evidence that backs these assumptions.

2.2.1. l Convergence

Spin Spirals

It was shown in section 1.5.3 that for spin spirals with any wave vector $\mathbf{q}_0$, the dependence of the energy per atom on the opening angle for a fixed wave vector, $e(\vartheta_0)|_{\mathbf{q}}$ can be described by Legendre polynomials with orders up to at most $n_{\max} l_{\max}$, where $n_{\max}$ is the maximum number of atoms in the cluster. Therefore, the energy per atom was calculated for 15 different opening angles for each of the aformentioned 209 wave vectors, and the curve $e(\vartheta_0)|_{\mathbf{q}}$ was plotted for each wave vector. Two wave vectors were selected by visual inspection, and then $e(\vartheta_0)|_{\mathbf{q}}$ was fitted with Legendre polynomials of up to eighth order.
A plot of the curve $e(\vartheta_0)|_{\mathbf{q}}$ for the wave vectors $\mathbf{q}_1 = 2\pi/a(0.867,\ 0.867,\ 1.067)$ and $\mathbf{q}_2 = 2\pi/a(0.533,\ 0.933,\ 1.067)$ is shown in Figure 2.2, together with the prediction of the Heisenberg model (see Section 1.5.2) that $e(\vartheta_0)|_{\mathbf{q}} \propto \sin^2(\vartheta_0)$. A table of the effective interaction coefficients $j_{\tilde{l}}$ as defined in Eqn. (1.104), and the RSS achievable with increasing order of the Legendre polynomials is given in Table 2.2.

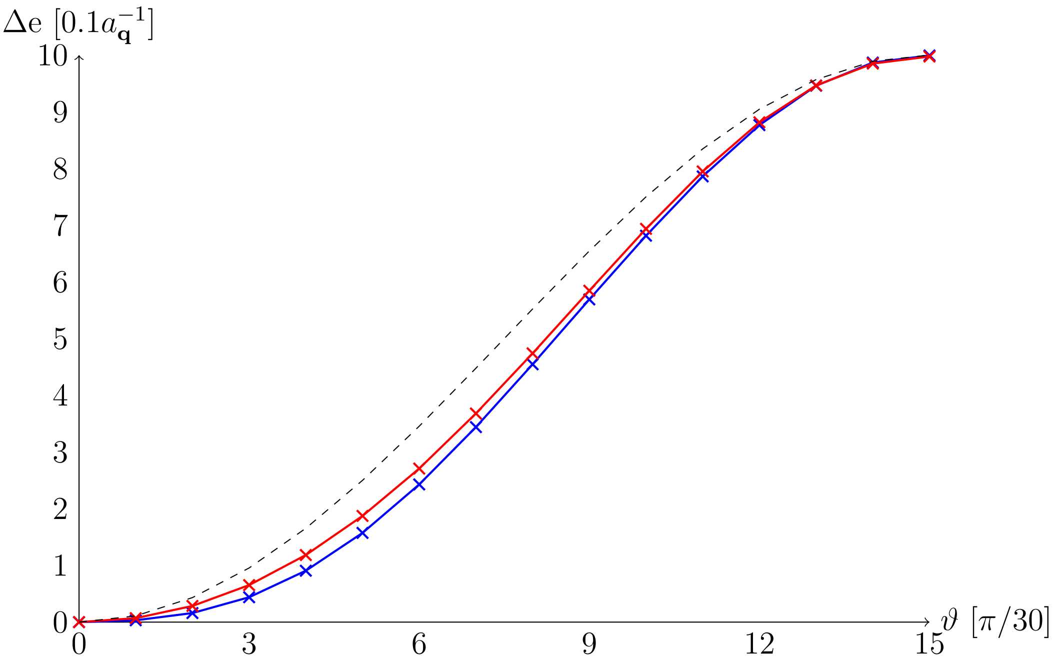

Figure 2.2.: $e(\vartheta_0)|_{\mathbf{q}}$ for $\mathbf{q}_1 = 2\pi/a(0.867, 0.867, 1.067)$ (blue) and $\mathbf{q}_2 = 2\pi/a(0.533, 0.933, 1.067)$ (red), each curve shifted so that $e(0)|_{\mathbf{q}} := 0$, and scaled so that $e(\frac{\pi}{2})|_{\mathbf{q}} := 1$. The scaling factor $a_{\mathbf{q}}$ necessary is $a_{\mathbf{q}_1}^{-1} = 26.4$ mRy, $a_{\mathbf{q}_2}^{-1} = 13.5$ mRy. The dashed line marks the prediction of the Heisenberg model that $e(\vartheta_0)|_{\mathbf{q}} \propto \sin^2(\vartheta_0)$

Two-atomic Supercells

It was also shown in section 1.5.3 that for all configurations any supercell symmetry the maximum "total angular momentum" l_{tot} of all significant basis functions in the expansion of the energy per atom for configurations that have such a supercell symmetry obeys

$$l_{\text{tot}} = \sum_j l_j \leq l_{\max} n_{\max}, \tag{2.3}$$

where l_j labels the elements of a l-tuple $\mathbf{l}_{\alpha i}$, and $n_{\max}$ is the maximum number of atoms in a cluster representative, and $l_{\max}$ is the maximum order of one element of the $\mathbf{l}_{\alpha i}$ tuple. Therefore, one can determine $l_{\max} n_{\max}$ by determining the significant basis functions for a specific supercell symmetry, and the corresponding largest l_{tot}.

First, l-convergence is tested for a two-atomic supercell. Such supercells can be constructed in several ways, e.g., by taking the smallest cubic supercell or by enlarging the primitive unit cell of the bcc lattice. These represent different classes of configurations, however, the theoretical predictions should hold true for any such class. For the calculations presented here, the smallest cubic supercell was used.

For all two-atomic supercells there is only one variable that identifies the configuration, the angle $\vartheta \in [0, \pi]$ between the two magnetic moment directions due to the global rotational symmetry. There is also only one possible cluster archetype $\alpha_1 = (\mathbf{r}_1, \mathbf{r}_2)$, containing in the case presented here the two atoms at positions[1] $\mathbf{r}_1 = O$ and $\mathbf{r}_2 = 0.5\sqrt{3}a[111]$, and the basis functions have the $\mathbf{l}$-tupel (l, l). Only basis functions with $2l \leq n_{\max} l_{\max}$ should be necessary to describe the energy.

Because the configuration space is very simple this is also an opportunity to test the fitting procedure against the direct calculation of the expansion coefficients as mentioned in section 1.8.5. The interaction coefficients for the energy are defined as

$$\begin{aligned} j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} &= \langle E | \gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}} \rangle \\ &\overset{\text{2-atomic}}{=} \frac{1}{(4\pi^2)} \int E(\mathbf{e}_1, \mathbf{e}_2) \gamma^{\mathcal{G}}_{(\mathbf{r}_1, \mathbf{r}_2),(l,l)}(\Omega)\, \mathrm{d}\Omega \\ &= \frac{1}{2} \int_0^\pi E(\vartheta) \gamma^{\mathcal{G}}_{(\mathbf{r}_1, \mathbf{r}_2),(l,l)}(\vartheta) \sin(\vartheta)\, \mathrm{d}\vartheta. \end{aligned} \tag{2.4}$$

where the last step made use of the rotational invariance of the system. The numerical integration was done by applying Simpson's rule over n_{conf} equidistant points in $[0, \pi]$. Calculations for $n_{\text{conf}} = 11, 101$ and 3143 are presented in Table 2.3. $|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$ has the same value of 12 for all two-atomic interactions which are considered here, so it does not need to be considered when comparing the absolute values.

It is possible to obtain reliable values for the interaction coefficients by the least mean

[1] To keep the notation brief, the author employs the Miller indices of the cubic supercell to describe directions in the crystal. E.g., [111] denotes the vector $1/\sqrt{3}(1, 1, 1)$, [100] the vector (1,0,0), etc.

$\mathbf{l}_{\alpha i}$	$n_{\mathrm{conf}} = 11$		$n_{\mathrm{conf}} = 101$	
	fit	integral	fit	integral
(1,1)	9.307	9.318	9.308	9.308
(2,2)	0.600	0.604	0.601	0.600
(3,3)	-0.172	-0.107	-0.173	-0.173
(4,4)	0.021	0.042	0.020	0.020
(5,5)	-0.014	0.256	-0.012	-0.010
(6,6)	-0.007	0.083	-0.006	-0.003
$\sqrt{\mathrm{nRSS}}$	12	601	12	51

Table 2.3.: Values of $j_{(\mathbf{r}_1,\mathbf{r}_2),(l,l)}$ in mRy for two-atomic supercell symmetry, obtained by least mean square fit versus numerical calculation of the defining integral. The last row gives the square root of the nRSS in μRy if all basis functions are used.

square fit for an even lower number of configurations - the first three interaction coefficients can be estimated quite well from just 5 configurations - if these are distributed evenly through the interval $\vartheta = [0, \pi)$, and not, e.g., bunched at $\vartheta = \pi/2$.

Larger supercells

To double-check the results from the two-atomic calculations presented above, and to illuminate the effects from the linear superpositions due to the backfolding of the basis functions as discussed in Section 1.5.4, additional calculations for a three-atomic supercell symmetry with translation vectors $\mathbf{T}_1 = 1.5\sqrt{3}a[1\bar{1}1]$, $\mathbf{T}_2 = 0.5\sqrt{3}a[\bar{1}11]$ and $\mathbf{T}_3 = 0.5\sqrt{3}a[11\bar{1}]$ were done. Furthermore, configurations to a four-atomic supercell symmetry with translation vectors $\mathbf{T}_1 = 2\sqrt{3}a[1\bar{1}1]$, and the same $\mathbf{T}_2$ and $\mathbf{T}_3$ as for the three-atomic supercell were investigated. The three-atomic supercell contains iron atoms at the positions $\mathbf{r}_1 = 0[1\bar{1}1]$, $\mathbf{r}_2 = 0.5\sqrt{3}a[1\bar{1}1]$ and $\mathbf{r}_3 = \sqrt{3}a[1\bar{1}1]$. The four-atomic supercell contains one more atom at position $\mathbf{r}_4 = 1.5\sqrt{3}a[1\bar{1}1]$.
The configurations to the three-atomic supercell symmetry can be described by the angles $\vartheta_2^{\mathbf{r}_2}$, $\vartheta_3^{\mathbf{r}_3}$ and $\varphi_3^{\mathbf{r}_3}$, setting $\vartheta_1^{\mathbf{r}_1}$, $\varphi_1^{\mathbf{r}_1}$ and $\vartheta_2^{\mathbf{r}_2}$ to zero, as there always exists a global rotation of all magnetic moment directions $\{\mathbf{e}_i\}$ that transforms any given configuration into a configuration of this type, and the energy is invariant under global rotations. Naturally, the configurations to the four-atomic supercell symmetry are described by the same variables plus $\vartheta_4^{\mathbf{r}_4}$, $\varphi_4^{\mathbf{r}_4}$.
Configurations to the three-atomic supercell symmetry were calculated for $\vartheta_2^{\mathbf{r}_2}, \vartheta_3^{\mathbf{r}_3} = 0, \pi/8, \ldots, \pi$ and $\varphi_3^{\mathbf{r}_3} = 0, \pi/8, \ldots, 2\pi$. This does not represent a homogeneous distribution of a given magnetic moment on the sphere surface, but should still be good enough to give a reliable estimate for $n_{\max} l_{\max}$. 1000 configurations to the four-atomic supercell

$\mathbf{l}_{\alpha i}$	$n_{\text{conf}} = 3143$		
	fit	integral	$\sqrt{\text{nRSS}}$
(1,1)	9.308	9.309	754
(2,2)	0.601	0.602	211
(3,3)	-0.173	-0.172	28
(4,4)	0.021	0.022	19
(5,5)	-0.012	-0.009	12
(6,6)	-0.006	-0.002	11
$\sqrt{\text{nRSS}}$	11	15	–

Table 2.4.: Same as Table 2.3. The additional last column gives the square root of the nRSS in μRy if all basis functions up to the current l are used in the least mean square fit.

symmetry were created by selecting the direction of each magnetic moment randomly and evenly distributed on the sphere surface by choosing φ randomly out of $[0, 2\pi)$ and $\cos(\vartheta)$ randomly out of $[-1, 1)$.

For the three-atomic supercell as defined three paragraphs above, there are only two possible cluster archetypes, $\alpha_1^3 = (\mathbf{r}_1, \mathbf{r}_2)$ with two atoms and $\alpha_2^3 = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ with three atoms (the upper index labels the supercell size). Depending on the $\mathbf{l}_{\alpha i}$ of specific basis functions, these basis functions may have different $|\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}|$. To allow for a comparison between the absolute values, Table 2.5 therefore lists the $j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ and the corresponding $J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ for the three-atomic supercell.

For the four-atomic supercell symmetry, there are two possible two-atomic cluster archetypes, $\alpha_1^4 = (\mathbf{r}_1, \mathbf{r}_2)$ and $\alpha_2^4 = (\mathbf{r}_1, \mathbf{r}_3)$. Furthermore, there are two different three-atomic cluster archetypes, $\alpha_3^4 = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \alpha_2^3$ and $\alpha_4^4 = (\mathbf{r}_1, \mathbf{r}_3, \mathbf{r}_4)$, and finally there is one four-atomic cluster archetype, $\alpha_5^4 = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)$. Together with the possible $\mathbf{l}_{\alpha i}$-tuples this results in more than a hundred possible basis functions to $l_{\text{tot}} \leq 10$, which can not be discussed here in the same detail as for the three-atomic supercell symmetry. Therefore, the forward reduction procedure described in Section 1.8.6 was employed to identify the most efficient basis functions in this set.

The result of this procedure is that there are no interactions to $l_{\text{tot}} > 8$ under the top 20 most efficient basis functions. Using these basis functions allows to fit the reference energies to the four-atomic supercell symmetry with a $\sqrt{\text{nRSS}}$ of 8.5μRy. Furthermore, there are no interactions with $l_{\text{tot}} > 6$ under the top 10 most efficient basis functions, and the corresponding fit leads to a $\sqrt{\text{nRSS}}$ of 12.5μRy.

To illustrate the effects of the backfolding of basis functions on the interaction coefficients obtained from calculations for specific supercell symmetries (see Section 1.5.3), Table 2.6 lists the interaction coefficients of basis functions that are part of the set of

cluster	$\mathbf{l}_{\alpha i}$	$j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$	$\lvert\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}\rvert$	$J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$	$\sqrt{\text{nRSS}}$	
α_1^3	(1,1)	8.4076	12	100.891	258	258
α_1^3	(2,2)	0.4603	12	5.524	128	128
α_1^3	(3,3)	0.0440	12	0.528	127	14
α_1^3	(4,4)	-0.0008	12	-0.010	127	5
α_1^3	(5,5)	-0.0063	12	-0.076	127	6
α_1^3	(6,6)	0.0041	12	0.049	127	5
α_2^3	(2,1,1)	0.3509	6	2.105	32	34
α_2^3	(2,2,2)	-0.0507	12	-0.608	19	22
α_2^3	(3,2,1)	-0.0613	6	-0.367	13	19
α_2^3	(3,3,2)	0.0183	6	0.110	13	7
α_2^3	(4,2,2)	0.0136	6	0.082	10	11
α_2^3	(4,3,1)	-0.0278	6	-0.167	7	8
α_2^3	(4,3,3)	-0.0009	6	-0.005	7	5
α_2^3	(4,4,4)	0.0064	12	0.077	6	6
α_2^3	(6,3,3)	-0.0048	6	-0.029	6	6
α_2^3	(6,4,2)	0.0078	6	0.047	5	6

Table 2.5.: Values of $j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ in mRy and $J_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ obtained by least mean square fit against 1377 configurations to a three-atomic supercell symmetry. The second last column gives the square root of the nRSS in μRy if all basis functions up to this row are used in the fit. The last column gives the same quantity if this basis function and all basis functions which lead to a larger reduction in the nRSS are used for the fit.

α_i	$\mathbf{l}_{\alpha i}$	$j^{2\text{at}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$	$j^{3\text{at}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$	$j^{4\text{at}}_{\alpha_i \mathbf{l}_{\alpha i}\mathbf{u}}$
α_1^3	(1,1)	9.308	8.4076	9.4168
α_1^3	(2,2)	0.601	0.4603	0.2495
α_1^3	(3,3)	-0.173	0.0440	0.0092
α_2^3	(2,1,1)	–	0.3509	0.2301
α_2^3	(2,2,2)	–	-0.0507	-0.0217
α_2^3	(3,2,1)	–	-0.0613	-0.0149
α_2^3	(3,3,2)	–	0.0183	-0.0222
α_2^3	(4,3,1)	–	-0.0278	-0.0182

Table 2.6.: Comparison of the values of common interaction coefficients for various supercell sizes. All values in mRy.

possible basis functions for at least the three-atomic and the four-atomic supercell, and are within the 20 most efficient basis functions of the four-atomic supercell.

2.2.2. Spatial Convergence

The assumption is that there exists a length $r_{\max}$ so that all expansion coefficients corresponding to cluster archetypes α_i with larger $\max(|\mathbf{r}_i - \mathbf{r}_j| \in \alpha_i)$ are effectively zero. It was shown in subsection 1.7.2 that an estimate of this value may be determined by the following procedure:

Start with a supercell of two atoms, one at the origin and one at $0.5\sqrt{3}[1\bar{1}1]$, with their magnetic moments oriented ferromagnetically. The energy per atom of this configuration is calculated by DFT. Then, the direction of the magnetic moment at the origin is flipped, and the difference in the energy per atom regarding the ferromagnetic configuration is calculated and scaled with the factor two, as there are two atoms in the supercell. This value is called Δe_2.

Now enlarge the supercell by one atom in the $[1\bar{1}1]$ direction. Again, only the magnetic moment of the iron atom at the origin is flipped, and the energy difference is calculated and scaled with the factor 3, giving Δe_3 . This number will be smaller, because the overall configuration is less canted as there is now a larger distance between canted moments. For some large enough supercell these numbers should converge to a finite value. The supercell size where this convergence is “good enough” is an estimate for $r_{\max}$.

Data for bcc iron is shown in Figure 2.3. The difference between Δe_9 and Δe_{10} is 147 μRy.

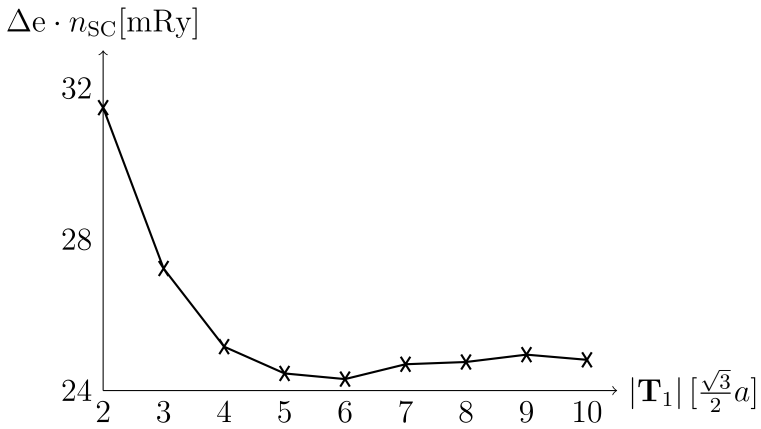

Figure 2.3.: Spatial convergence - energy per atom difference Δe, scaled with the number of atoms in the supercell (n_{SC}) against the size of the supercell.

2.2.3. Discussion

One can see from Table 2.6 that the backfolding of basis functions, as discussed in Section 1.5.4, has a considerable effect on the values of the $j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$, sometimes even changing the sign of the value of an interaction coefficient.
The absolute values of the interaction coefficients show a rather clear convergence behaviour with regards to higher values in $\mathbf{l}_{\alpha i}$. The convergence is monotonous for two-atomic supercells. For three-atomic supercells, there are some small inconsistencies, however, if one looks at the last column of Table 2.5, the basis functions where this is not monotonous have a very small effect on the nRSS, which means they are not important to reproduce the DFT data. Therefore they are probably very susceptible to small numerical effects.
The DFT calculations for the reference energies for the l-convergence were all calculations with a constant set $\{\mathbf{k}_i\}$, as each investigation was done for one specific supercell symmetry. Therefore the reference energies should be subject to the lower estimate of about 2-3 μRy for the pseudorandom numerical errors discussed in Section 2.1.1. Yet the best $\sqrt{\text{nRSS}}$ achieved (spin spirals, three-atomic supercell calculations) was about two times higher, and in case of the two-atomic supercell the best $\sqrt{\text{nRSS}}$ was even roughly four times higher than the estimated numerical error. While the author has no explanation for this, he considers such accuracy good enough for any practical purpose. Overall, using basis functions to $l_{\text{tot}} \leq 8$ should give a model that is accurate to 10-20 μRy.
The convergence with regards to the maximum interaction length $r_{\max}$ is, unfortunately, not as encouraging. Even moving from nine to ten atoms between two noncollinear

planes, the energy difference still changes by 14.7 μRy per atom. While this is of the same order of magnitude as the effects of the l-convergence, exhaustive DFT calculations on this length scale are completely infeasible. To investigate general configurations on this length scale, one would have to set up a supercell which has roughly this size in each direction. Such calculations would require far too much computation time[2]. They are also hindered by the fact that the mixing methods that solve the DFT self consistency problem quite often fail to find the minimum of the energy for very large supercells. Furthermore, the space of possible configurations gains two dimensions with each atom added to the supercell, which means sampling a representative part of the space of possible configurations quickly becomes impossible.

2.3. General SCE

2.3.1. Reference Configurations

One of the goals of this work was to construct a SCE valid for as many configurations as possible. Therefore, the reference configurations whose energies the SCE is fitted against should be as varied as possible. This is, however, limited by the available computational resources for DFT calculations, exacerbated by the bad O(3) scaling of DFT implementations with the number of atoms in the supercell, the principal problem of the number of dimensions of the configuration space increasing with O(2N-3) with the number of atoms in the supercell (the 3 is due to the symmetry under global rotations), and the not particularly robust mixing methods. A tabular overview over the configurations considered by the author is provided in Table 2.7.
A small stumbling block is that different supercell symmetries have slightly different values of, e.g., the ferromagnetic energy, because the set $\{\mathbf{k}_i\}$ for the Brillouin zone integration of DFT is chosen differently by the automatic algorithm implemented in the LMTO program. This has a negligible effect on the difference of the energy of a given configuration relative to the energy of the ferromagnetic configuration of the same supercell geometry. However, when the energies of configurations to different supercell symmetries are used together in the least-mean squares fit this effect can be considerable. Therefore, the energy of the ferromagnetic configuration for a given supercell symmetry is subtracted from all energies to the same supercell symmetry.
The primitive cell in Table 2.7 represents spin spiral configurations. For these configurations, 15 values for the opening angle ϑ_0 were chosen equidistantly out of $[\pi/30, \pi/2]$, and for each such opening angle energies to a common set of 208 wave vectors were

[2] The calculation for 10 atoms in the supercell took about 10 hours on the computers used for this work. A supercell of the same size in each direction would contain 1000 atoms, and the computational time required scales with the third power of the number of atoms. One calculation would therefore take about one millennium. If the mixing algorithm would work.

label	$\mathbf{T}_1$	$\mathbf{T}_2$	$\mathbf{T}_3$	$n_{\mathrm{At}}^{\mathrm{SC}}$
wav	$0.5\sqrt{3}a[1\bar{1}1]$	$0.5\sqrt{3}a[\bar{1}11]$	$0.5\sqrt{3}a[11\bar{1}]$	1
prim-2	$\sqrt{3}a[1\bar{1}1]$	$0.5\sqrt{3}a[\bar{1}11]$	$0.5\sqrt{3}a[11\bar{1}]$	2
prim-3	$1.5\sqrt{3}a[1\bar{1}1]$	$0.5\sqrt{3}a[\bar{1}11]$	$0.5\sqrt{3}a[11\bar{1}]$	3
prim-4	$2\sqrt{3}a[1\bar{1}1]$	$0.5\sqrt{3}a[\bar{1}11]$	$0.5\sqrt{3}a[11\bar{1}]$	4
cub-2	$2a[100]$	$a[001]$	$a[001]$	4
cub-2x2	$2a[100]$	$2a[001]$	$a[001]$	8
cub-2x2x2	$2a[100]$	$2a[001]$	$2a[001]$	16

Table 2.7.: Types of reference configurations used for the SCE. This table gives the translation vectors of the respective supercell and the number $n_{\mathrm{At}}^{\mathrm{SC}}$ of atoms within. Further information is presented in Table 2.8.

calculated. The wave vectors were chosen not only along the high-symmetry lines of the Brillouin zone, but also from within the Brillouin zone.
Calculations for supercells which are formed by elongating the one-atomic primitive cell of Table 2.7 along $[1\bar{1}1]$ were initially intended only for the l-convergence calculations described in Section 2.2.1. However, they have been included in the set of reference energies because the cubic supercells do not extend very far along $[1\bar{1}1]$, yet the author found many "important" two-atomic interactions along this direction in [31], and therefore wanted to investigate this in some more detail..
The mean difference of the energy per atom to the respective ferromagnetic ground state over all configurations is 9.391 mRy, with the median[3] at 7.066 mRy, and the average deviation from the mean difference of the energy per atom is 8.423 mRy. The largest difference of the energy per atom is 31.282 mRy and occurs for the antiferromagnetic configuration.

2.3.2. Considered Basis Functions

The main part of the set of considered basis functions is constructed from all cluster archetypes with up to $n_{\max} = 4$ atoms that fit in the 8-atomic supercell described in Table 2.7, with l-tuples that obey $l_{\mathrm{tot}} \leq 8$. There are 233 such interactions after all symmetries are taken into account, consisting of 12 basis functions to two-atomic cluster archetypes, 39 basis functions to three-atomic cluster archetypes and 182 basis functions to four-atomic cluster archetypes.
This set of possible interactions has two weaknesses: First, the cluster archetypes are rather short-ranged, and second, there are not many interactions along [111], however, the author found many important two-atomic interactions in this direction in [31]. To

[3] The value separating the higher half from the lower half of a set of values.

label	e_{FM}	n_{conf}	method
wav	31.41	3117	systematic (*)
prim-2	:=0.0	3143	systematic
prim-3	21.24	1377	systematic
prim-4	20.59	1000	random
cub-2	0.05	869	random
cub-2x2	2.58	770	random
cub-2x2x2	0.05	95	random

Table 2.8.: Types of reference configurations used for the SCE. This Table gives the energy per atom e_{FM} for the ferromagnetic state in μRy relative to the value for 2-atomic supercells, the number n_{conf} of DFT calculations for the respective supercell and the method used to generate the directions of the magnetic moments for each DFT calculation. A "systematic" method to generate the configurations indicates an equidistant set of values in $[0, \pi]$ for each ϑ and in $[0, 2\pi]$ for each φ. "random" indicate randomly chosen values in $[-1, 1]$ for $\cos(\vartheta)$ and in $[0, 2\pi]$ for φ. The primitive cell represents spin-spiral configurations, the spefic wave vectors and opening angles are given in the main text. The value for e_{FM} for the primitive cell represents the average over 208 wave vectors.

somewhat alleviate these weaknesses, the pair-interactions up to $|\mathbf{r}| \leq 3.84a$ with $l \leq 2$ were added, along with a three-atomic cluster archetype along [111] which contains the atom at the origin and the atoms at $0.5\sqrt{3}a[111]$, $\sqrt{3}a[111]$. The interaction coefficients of most of these additional basis functions are solely determined by energies to spin spiral configurations, because for the considered supercell configurations most of the additional basis functions are linear combinations of basis functions to cluster archetypes within the supercell (see Section 1.5.3). An exception are the basis functions to the additional three-atomic cluster archetype and the two closest additional two-atomic cluster archetypes along [111], where the interaction coefficients are also determined by the energies to configurations to the 16-atomic supercell symmetry and by the energies of (rather symmetric) configurations to the bar-shaped three- and four-atomic supercell symmetries obtained from elongating the primitive cell along $[1\bar{1}1]$.
Altogether, 283 basis functions are considered.

2.3.3. Results and Discussion For bcc Iron

Necessity of considering spin spirals and supercells: A fit solely against energies of spin spiral configurations or solely against energies of supercell configurations is unable to predict the energies of the other type of configuration with good accuracy. Obtaining a model only from spin spiral configurations predicts the supercell reference energies with a $\sqrt{\text{nPSS}}$ of about 0.67 mRy (for the definition of the nPSS see the discussion of cross validation in Section 1.8.6), and obtaining a model only from supercell configurations predicts the supercell configurations with a $\sqrt{\text{nPSS}}$ of about 0.44 mRy. The $\sqrt{\text{nRSS}}$ for the fitted energies per atom is in both cases far smaller, of the order of 0.08 mRy.

The values of the respective nPSS have been obtained by employing the forward reduction procedure as described in Section 1.8.6. This procedure tries to find the most computationally efficient fit for the corresponding training set, e.g., spin spirals, and does not optimize for predicting the respective testing set, e.g., supercells. Using too many basis functions can lead to overfitting, where, e.g., the numerical noise and the remaining effects of the different sets $\{\mathbf{k}_i\}$ used for the Brillouin zone integration during the DFT calculations inherent in the energies of the respective reference configurations are fitted, which can lead to comparatively large (up to 3 mRy) errors regarding the energies in the testing set. The author simply manually tracked the nPSS for the testing set during the forward reduction procedure and selected the smallest occuring value.

This is also necessary because if the model is obtained only from spin spirals the SCE basis set is overcomplete (see Section 1.5.1) and linear dependencies show up. To avoid such linear dependencies when the model is obtained from supercell configurations, all basis functions to cluster archetypes larger than the 8-atomic supercell (the “additional” basis functions from Section 2.3.2) were removed from the set of considered basis functions. Somewhat surprisingly, there is still one basis function that shows linear dependency if the model is obtained only from configurations to supercell symmetry. This linear dependency does not exist if all reference configurations including those to spin-spiral symmetry are used for the fit, and omitting the basis function in question has no measurable effect on the quality of the fit. Upon closer inspection, the linear dependency involves eight basis function with a very complex interplay of backfolding for the considered types of supercells and the relevant Clebsch-Gordan coefficients. The author did not further investigate this phenomena, as it does not occur for general configurationsw

Full model: Fitting all 283 considered basis functions against all reference configurations yields a $\sqrt{\text{nRSS}}$ of 50.3 μRy. This is of the same order of magnitude as the accuracy expected from the l-convergence data, although higher by roughly a factor of 3.
There is one data point that may be used to investigate whether spin spiral configurations allow some insight into long-ranged interactions, which is the data point to $5\sqrt{3}a/2$ in Figure 2.3. The data points corresponding to smaller supercells are described very well by the full model, as supercells of these sizes are part of the reference configurations for the model. The changes in the energy difference plotted in 2.3 which correspond to larger supercells can not be described in our model, as it does not contain interactions of a length greater than $3.84a$ which means the model predicts no change. Within the full model, the change in the energy difference observed when going from $4\sqrt{3}a/2$ to $5\sqrt{3}a/2$ in 2.3 is described by two two-atomic basis functions, one with $l = (1, 1)$ and one with $l = (2, 2)$, whose interaction coefficients are determined by spin spiral configurations, because these basis functions are larger than any supercell within the reference configurations.
These two two-atomic basis functions fail to predict the energy per atom of the configuration to the point $5\sqrt{3}a/2$ in Figure 2.3 by about $+0.4$ mRy (corresponding to $+2$ mRy on the scale in Figure 2.3). This is actually worse than not including these basis functions into the model would be. This suggests that spin spirals do not actually allow to predict long-ranged configuration changes, which is not surprising given that a model obtained solely from spin spiral configurations has been shown to not correctly describe supercells in general (see preceding result).

Efficient model: Calculating the values of all 283 basis functions requires about 330 milliseconds per atom of a supercell configuration (the time required scales linearly with the number of atoms in the supercell), using an AMD Opteron 2218 CPU. As the accuracy achievable with the available reference configurations is limited anyway due to the limited size of the supercells (see Section 2.2.2), it seems reasonable at first to stop the forward reduction procedure detailed in Section 1.8.6

when a nRSS comparable to the inherent errors due to these limits is reached. According to the data displayed in Figure 2.3, the effects of the limited range on the energy per atom are for the configurations discussed in Section 2.2.2 and for the spatial cutoff $r_{\max} = 3.84a$ chosen when selecting the considered basis functions is of the order of 0.18 mRy. This value was obtained by taking the change between data points $4\sqrt{3}a/2$ and $5\sqrt{3}a/2$ in Figure 2.3 and dividing by 4.
Achieving a comparable nRSS requires only 15 basis functions, which results in a speed of the model of about 2 milliseconds per atom on the same CPU as above, and a $\sqrt{\text{nRSS}}$ of 0.182 mRy. However, analyzing the $\sqrt{\text{nRSS}}$ separately for the smaller supercells, the larger supercells and spin spirals shows that the latter two types of configurations are described significantly worse than the smaller supercells. Therefore, the author concluded that this model might be somewhat too simple, and some more basis functions should be considered. This is somewhat supported by the observation that this model predicts an energy per atom of the configuration corresponding to point $5\sqrt{3}a/2$ in Figure 2.3 which is wrong by +0.7 mRy.

Balanced model: To give an example for a model that strikes a balance between accuracy and speed, a model with 45 basis functions was constructed by the forward reduction procedure. The speed of this model is about 8 milliseconds per atom of a supercell, and the model fits the reference configurations with a $\sqrt{\text{nRSS}} \approx 0.096$ mRy. This model predicts an energy per atom of the configuration corresponding to point $5\sqrt{3}a/2$ in Figure 2.3 which differs from the DFT value by +0.2 mRy, which is significantly more accurate than the models described above. This may be just a fortunate coincidence, and is still quite inaccurate.

Heisenberg model: Quite a lot of work has been done on "extended Heisenberg models", e.g., publications [23, 22, 24], which simply use the cosine of the angle between the magnetic moment directions of two atoms at different positions $\mathbf{r}_i, \mathbf{r}_j$ as basis functions. In the SCE, these basis functions correspond to the two-atomic interactions with $\mathbf{l}_{\alpha i} = (1, 1)$ [15]. One may select a cluster archetype for these types of basis functions that contains the origin, and the $\mathbf{l}_{\alpha i}$-tuple is fixed, which means all Heisenberg basis functions can be characterized by the "neighbour", or "shell" relative to the atom at the origin, e.g., a basis function to the nearest neighbour or first shell contains the atom at the origin and the atom at position $\mathbf{r}_j = 0.5\sqrt{3}a[111]$, a basis function to the next-nearest neighbour or second shell contains the atom at the origin and the atom at position $\mathbf{r}_j = a[100]$, etc.
These models are quite often obtained exclusively from spin spiral configurations. Doing so for our spin spiral configurations, and using up to the 25th neighbour gives a $\sqrt{\text{nRSS}}$ of 0.45 mRy, and predicts the energies of the supercell configurations with a $\sqrt{\text{RSS}}$ of 0.8 mRy per predicted configuration.

Predicting energies: To test the predictive power of the SCE, the energies per atom of configurations to the 16-atomic supercell symmetry were removed from the set of reference energies and used to test the predictive power of the various models. These energies per atom have an average deviation of 5.14 mRy from their mean. Using the full model largely fails to predict these energies, with a $\sqrt{\text{nPRSS}}$ of about 1 mRy. This is largely an effect of overfitting, because using only the basis functions from the balanced model one can predict these energies with an accuracy of 0.3 mRy, and using the basis functions of the efficient model surprisingly improves that to 0.2 mRy. The Heisenberg model - if the other supercell configurations are also used to obtain the interaction coefficients - does predict these energies with a $\sqrt{\text{nPSS}}$ of 0.4 mRy per configuration.

This does suggest that the more complicated models are rather susceptible to overfitting.

There is also a quite good argument to be made for obtaining the interaction coefficients from the energies of configurations of larger supercells and predicting the energies of smaller supercells: As it was shown in Section 1.5.4, energies to supercells where backfolding occurs can be described with arbitrary precision by the basis functions to cluster archetypes that fit within the supercell, but have effective interaction coefficients $J^{\text{SC}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$, which contains the

Comparison with [31] : In the letter [31] on the topic of the SCE, a different methodology was employed for the forward reduction described in 1.8.6, where not the most efficient basis function was selected but simply the basis function that decreases the nRSS the most. This leads to a slightly better fit of the reference data, at the cost of significantly higher computational cost. E.g., selecting the first 15 basis functions leads to a $\sqrt{\text{nRSS}}$ of 0.144 mRy, compared to the 0.182 mRy for the efficient model defined above. The computation of these basis functions takes, however, about 7.5 ms per atom of a supercell instead of just 2 ms for the basis functions forming the efficient model. The prediction for the energy per atom of the configuration corresponding to point 5 $\sqrt{3}a/2$ in Figure 2.3 is wrong by +0.6 mRy, which represents no significant improvement. Predicting the energies of the configurations to the 16-atomic supercell symmetry with these basis functions gives a $\sqrt{\text{nPSS}}$ of about 0.2 mRy, which is also no significant improvement.

These are general trends for all models obtained during the forward reduction procedure - most of the time the nRSS is about 20 % lower if the basis functions which give the best nRSS is selected compared to the case where basis functions which give the best nRSS per computation time required are selected. The predictive power of a given model, measured by predicting the 16-atomic supercells, is usually comparable with a small advantage for the most efficient approach.

Actual values for the interaction coefficients: These can be found in the Addendum, Section A.8.

Altogether, it can be said that the SCE is able to very accurately represent the reference data. The accuracy achievable is of the order of magnitude which was expected from the convergency calculations. Somewhat troubling is that more complicated models seem to be rather susceptible to overfitting, and any actual application of the SCE needs to very carefully control for this.
Selecting the basis functions for a given model by choosing the most efficient basis function seems justfied compared to the approach taken in [31], as the predictive power at a common model size is comparable while the computational speed is vastly better.
All in all, the SCE seems to be good enough for predicting the energy of general configurations for bcc iron with an accuracy of about 0.3 mRy, as estimated with the balanced model for the configurations to 16-atomic supercell symmetry. This is one order of magnitude better than a simple nearest-neighbour Heisenberg model (not discussed in above), and better by about a factor of two than the "traditional" extended Heisenberg models.
To further improve upon this would require to include configurations to larger supercell symmetries within the reference calculations. This is currently infeasible due to the limitations of Density Functional Theory implementations, and would have very likely no relevance for actual physics as the current exchange-correlation functionals are not accurate enough.

3. Results For fcc Iron

Fcc iron (also called γ-iron) is a very complicated magnetic material, with a currently unknown, but noncollinear ground state [38], competing local minima for different absolute values of magnetic moments ("high-spin" and "low-spin" states, [30, 1]), very strong dependence of magnetism on the lattice constant[1], etc. One of the goals of the original work on the SCE [1] was to use the SCE to determine the configuration of the directions of the magnetic moments in the ground state of fcc iron, by constructing a SCE for fcc iron from reference energies calculated with the MASW program and then use the SCE model for an exhaustive ground state search.

The work presented in [1] had the following shortcomings: First, the effects of the combination of space group symmetry and symmetry under global rotations of the magnetic moments were overlooked. It was effectively presumed that the different parameters $\mathbf{k}$ of the SCE basis functions characterize linear independent basis functions, because the space group symmetry operations do not act on $\mathbf{k}$. As was shown in Section 1.4, this is not the whole picture, and one has to calculate the eigenvectors $\mathbf{u}$. Because this was not taken into account in [1], the basis set used there contained linear dependencies, and while these were detected by the singular value decomposition and great effort was expended to suppress these dependencies, the results do seem somewhat questionable in hindsight. Furthermore, the investigation of how reliable limits ($n_{\max}$, $l_{\max}$, $r_{\max}$) for the parameters of the SCE basis functions that need to be considered can be obtained from DFT data (see Section 1.7) was carried out by the present author, and the corresponding results were not taken into account in [1] when the basis functions to be investigated were selected.

The resulting SCE also showed limited predictive power, and naturally it is hoped that the refined methods developed by the present author improve upon this.

3.1. Density Functional Theory (DFT)

An excellent overview and discussion of many DFT investigations of fcc iron is given in [1], and shall not be repeated here. The main points are that using or not using the different approximations mentioned in subsection 1.8.3 (ASA, AMA, etc.) leads to qualitatively differing results, and different methods that use none of these approximations (such as the MASW program) can still show large quantitative variations depending on

[1] All calculations presented here use a lattice constant of 6.7 r_{Bohr}.

the exact implementation of DFT. Therefore, while the SCE model might accurately reproduce the results of the DFT code used to obtain the reference data these results may not quantitatively represent the physical material.
Because getting reliable results in such complicated cases requires intimate knowledge of both the DFT code used and a very good overview over the literature to spot errors, the author has abstained from doing any calculations for fcc iron by himself (with one very simple exception), as he possesses neither of these requirements. All reference energies used later in this work have been calculated by R. Singer, and all information about the properties of these reference energies as detailed in the following is taken from [1].

3.1.1. Accuracy of the DFT calculations

R. Singer estimates the numerical accuracy regarding the set of **k**-points to about 30 μRy ([1], p.131). This estimate has been obtained by investigating the energy difference for specific configurations to one supercell symmetry for different sets $\{\mathbf{k}_i\}$. No analysis of the spin spiral effects detailed in Section 2.1.1 has been done, as R. Singer did not know about this and the present author did not attempt the necessary DFT calculations. Therefore, the actual numerical error may be somewhat higher.

3.1.2. Density Functionals

Naturally, the same problems as described in Section 2.1.2 for bcc iron still remain. However, for fcc iron different exchange-correlation functionals can lead to qualitatively different results regarding, e.g., the high-spin and low-spin minima. Again, an excellent overview and extensive discussion of the different results for fcc iron when using different exchange-correlation functionals can be found in [1]. After all considerations, the functional by Perdew, Burke and Ernzerhof [37] was used by R. Singer, which is a GGA functional.

3.2. SCE Parameters

As the investigation on the general convergence limits of the parameters ($n_{\max}$,$l_{\max}$,$r_{\max}$) of the necessary SCE basis functions was carried out by the present author, [1] does not contain the necessary calculations. Therefore, the present author employed the MASW code to test for the l convergence (as in subsection 2.2.1), by using the MASW program and files of sample calculations provided by R. Singer and changing only the necessary parameters.
Calculations for $r_{\max}$ as in subsection 2.2.2 were not attempted, as these would have required setting up new supercells, and starting calculations of the MASW program from scratch. As calculations for fcc iron are extremely sensitive to small errors of the user,

e.g., when choosing the starting estimate for the spin density matrix, the present author does not trust himself to obtain reliable results.

3.2.1. l Convergence

R. Singer provided the present author with the files for one calculation of a configuration to a cubic supercell symmetry with translation vectors[2] $\mathbf{T}_1 = a[100]$, $\mathbf{T}_2 = a[010]$, $\mathbf{T}_3 = a[001]$, which contains four atoms. These atoms are located at the origin of the supercell, and at $\mathbf{r}_2 = \sqrt{0.5}a[110]$, $\mathbf{r}_3 = \sqrt{0.5}a[101]$, and $\mathbf{r}_4 = \sqrt{0.5}a[011]$. The angles (in degrees, to be consistent with [1]) of the magnetic moment directions are $\vartheta_1 = 67.5$ degrees, $\varphi_1 = 70.0$, $\vartheta_2 = 165.0$, $\varphi_2 = 90.0$, $\vartheta_3 = 97.5$, $\varphi_3 = 10.0$, $\vartheta_4 = 142.5$, $\varphi_4 = 220.0$, and the absolute values $|\mathbf{m}_i|$ of the magnetic moments, as calculated from the self-consistent spin density matrix, are $|\mathbf{m}_1| = 1.845\mu_\mathrm{B}$, $|\mathbf{m}_2| = 1.342\mu_\mathrm{B}$, $|\mathbf{m}_3| = 1.591\mu_\mathrm{B}$, and $|\mathbf{m}_4| = 1.866\mu_\mathrm{B}$.
To obtain configurations that may be described by using only two-atomic SCE basis functions, the angles $\varphi_i \forall i$ and ϑ_1 were set to zero, and $\vartheta_2, \ldots, \vartheta_4$ were coherently rotated from 0 to 180 degrees in steps of 3.6 degrees. All these calculations were started from the spin density matrix of the calculation by R. Singer described above. The only actions the author took were to replace the original angles with the desired ones, and let the MASW program run until a new self-consistent spin density matrix was achieved.
Because the positions $\mathbf{r}_2$,$\mathbf{r}_3$ and $\mathbf{r}_4$ are equivalent for all new configurations selected by the present author, the magnetic moments calculated from the new self-consistent spin density matrix should have the same values at all these positions, which was always the case. Because the starting point of the spin density matrix represents very different absolute values of the magnetic moments, the author is confident that the results represent the global minimum. It may be that the comparatively high starting absolute value of the magnetic moment for the atom at position $\mathbf{r}_1$ somewhat preselects high spin states, however in the results this value varies smoothly between 1.14 μ_B for the ferromagnetic configuration and 2.4 μ_B for the configuration with $\vartheta_2, \ldots = 180$ degrees.
The resulting reference energies are fitted with the basis functions to $\alpha_1 = (\mathbf{r}_1, \mathbf{r}_2)$ and $\mathbf{l}_{\alpha i} =(1,1),(2,2), \ldots, (6,6)$. The results are collected in the following table:

[2] Miller indices of the cubic supercell are employed to denote directions in the crystal. Each $[ijk]$ denotes a normalized vector in the corresponding direction.

$\mathbf{l}_\alpha$	$j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ [mRy]	$\sqrt{\text{nRSS}}[\mu\text{Ry}]$
(1,1)	-10.7540	177
(2,2)	-0.8213	104
(3,3)	0.7692	34
(4,4)	0.1245	26
(5,5)	-0.1606	12
(6,6)	-0.0456	10

3.2.2. Discussion

When comparing the l-convergence results to the corresponding Table 2.3 for bcc iron, it is important to know that the typical differences in the energy per atom for bcc are much higher, roughly by a factor of eight. Therefore the relative magnitude of comparable errors is much lower for bcc iron than for fcc iron. It is therefore very unfortunate that the numerical errors of the DFT calculations are significantly higher for fcc iron as well. As it is, the threshold of the expected numerical error is already reached for $\mathbf{l}_{\alpha i} = (3, 3)$, and a corresponding $n_{\max} l_{\max} = 6$.
Nothing is known about $r_{\max}$, which is unfortunate.

3.3. General SCE

3.3.1. Reference Configurations

The reference configurations for the SCE for fcc iron are again a mix of configurations to spin spiral and varying supercell symmetries. The spin spiral configurations were calculated for opening angles $\vartheta_0 = 2\pi/12, 3\pi/12, \ldots, \pi/2$ for each element of a set of 119 wave vectors chosen throughout the Brillouin zone. Furthermore, 161 randomly chosen configurations for a supercell symmetry with $\mathbf{T}_1 = a[100]$, $\mathbf{T}_2 = a[010]$, and $\mathbf{T}_3 = a[001]$ were calculated by R. Singer, with 4 atoms in the supercell. Also, 50 randomly chosen configurations to a supercell symmetry with $\mathbf{T}_1 = 2a[100]$, $\mathbf{T}_2 = 2a[010]$, and $\mathbf{T}_3 = a[001]$ were calculated by R. Singer, where the supercell contains 16 atoms. Finally, 29 randomly chosen configurations to a supercell symmetry with $\mathbf{T}_1 = 2a[100]$, $\mathbf{T}_2 = 2a[010]$, and $\mathbf{T}_3 = 2a[001]$ were calculated, where the supercell contains 32 atoms.
These supercells do not use one common set $\{\mathbf{k}_i\}$ for the Brillouin zone integration during the DFT calculations, and while the effect of the different sets $\{\mathbf{k}_i\}$ on energy differences between configurations was negligible if these configurations were calculated for 4-atomic or 16-atomic supercells, the energy differences changed quite a bit when calculated for 32-atomic supercells. This resulted in a disagreement of up to 0.2 mRy for the energy difference per atom if the same two configurations were calculated once with

a 4-atomic supercell symmetry and once with a 32-atomic supercell symmetry (see [1], p. 131). As there are also very few configurations for the 32-atomic supercell symmetry, the author did not include the energies to these configurations in the reference energies (or training set) for the SCE and just used these energies as a testing set.
The average variation of the reference energies per atom of all configurations relative to the mean is 1.289 mRy, and the median of the energy per atom is 0.4 mRy higher than the mean. The largest difference in the energy per atom between any two configurations is 4.490 mRy.

3.3.2. Considered Basis Functions

The main part of the set of considered basis functions is constructed from all cluster archetypes that fit into a supercell with translation vectors $\mathbf{T}_1 = 2a[100]$, $\mathbf{T}_2 = a[010]$, and $\mathbf{T}_3 = a[001]$ and contain up to $n_{\max} = 4$ atoms. All $\mathbf{l}_{\alpha i}$-tuples to $l_{\text{tot}} \leq 6$ are considered. There are 93 such basis functions after all symmetries are taken into account, consisting of 9 basis functions to two-atomic cluster archetypes, 22 basis functions to three-atomic cluster archetypes and 62 basis functions to four-atomic cluster archetypes. Just as for bcc iron, the two-atomic basis functions up to $|\mathbf{r}| \leq 3.24$ with $l \leq 2$ were added to tentatively increase the spatial coverage of the basis functions.
Altogether, 137 basis functions are considered.

3.3.3. Results and Discussion for fcc Iron

Signs for insufficient number of reference energies: As mentioned in Section 1.8.5 when discussing cross validation, the SCE program code computes an automated nPSS (anPSS) for the current training set of energies by randomly dividing the training set into smaller parts and cross-validating within these. For bcc iron, the nRSS and anPSS do not differ significantly and are always the same to within 1.5%. Contrary to this, the quantities differ by significant amounts for fcc iron. Differences up to 20-30 %, in extreme cases 300% arise. This will be discussed further for specific models.

Full model: Fitting all 137 considered basis functions against all reference configurations, except the 32-atomic reference configurations, yields a $\sqrt{\text{nRSS}}$ of 67.3 μRy. This is of the same order of magnitude as the accuracy expected from the l-convergence data, although higher by roughly a factor of 2, just as for bcc iron. However, the prediction of the reference energies to the 32-atomic configurations fails very badly, with errors of the order of several Ry. This is due to the fact that there is an insufficient number of configurations to the larger 8-atomic supercell - there are only 50 such configurations, however there are more than 60 basis functions which are larger than the 4-atomic supercell. This means the fit matrix A

does not have full rank because there is an insufficient number of configurations, and there are correspondingly vanishing singular values. One would expect about as many vanishing singular values as the number of basis functions larger than the 4-atomic supercell minus the number of configurations to the 8-atomic supercell symmetry, however, the spin spiral configurations are of some help, and only two vanishing singular values are actually found.
Still, even if as an ad-hoc countermeasure the 10 basis functions to four-atomic cluster archetypes and highest l_{tot} are removed - which leads to no more vanishing singular values - the nRSS and anPSS differ by a factor of 3, which suggests that using all basis functions results in severe overfitting. The energies of configurations to the 32-atomic supercell symmetry are predicted with a $\sqrt{\mathrm{nPSS}}$ of 0.23 mRy. This represents no actually usable prediction, as the energies of configurations to the 32-atomic supercell symmetry have an average deviation of only 0.298 mRy from their mean.

Balanced model: To avoid overfitting, and because there is a insufficient number of reference configurations, less basis functions must be used. The model should ideally describe the 32-atomic configurations with reasonable nPSS, and the nRSS and the "automated prediction error" anPSS of the training set should be roughly the same. Therefore, the forward reduction procedure is applied, and the three quantities (nPSS, nRSS and anPSS) are plotted for each step in this procedure (Figure 3.1).
There are three interesting points in this figure: The point 0 represents the fitting of just the constant basis function, and the corresponding $\sqrt{\mathrm{nRSS}}$ is proportional to the average deviation of the energies from their mean. Therefore, this point sets the reference scale for all further points, as a model should be significantly better than fitting a straight line.
The 10th basis function added represents the first four-atomic interaction, and this leads to a sharp drop of all three quantities, and therefore a significant improvement of the model - at comparatively high computational cost. This point also represents the model that gives the best prediction for the 32-atomic supercell configurations. Afterwards, some overfitting sets in, visible from the increasing divergence of nRSS and anPSS, and the nPSS of the 32-atomic supercell configurations becomes much worse. However, the prediction is significantly bettered by the 40th interaction, which is a three-atomic interaction. 40 also happens to be the number of basis functions in the SCE model that was ultimately used in [1]. Therefore the present author selected the model with 40 basis functions, as one of the goals of this work was to assess the improvements made and thus the present author wanted a model of comparable size to the model used in [1].
This model has a $\sqrt{\mathrm{nRSS}}$ of 0.056 mRy and a $\sqrt{\mathrm{nPSS}}$ of 0.062 mRy, and predicts the 32-atomic supercells to within 0.104 mRy. This is still a rather large error, as

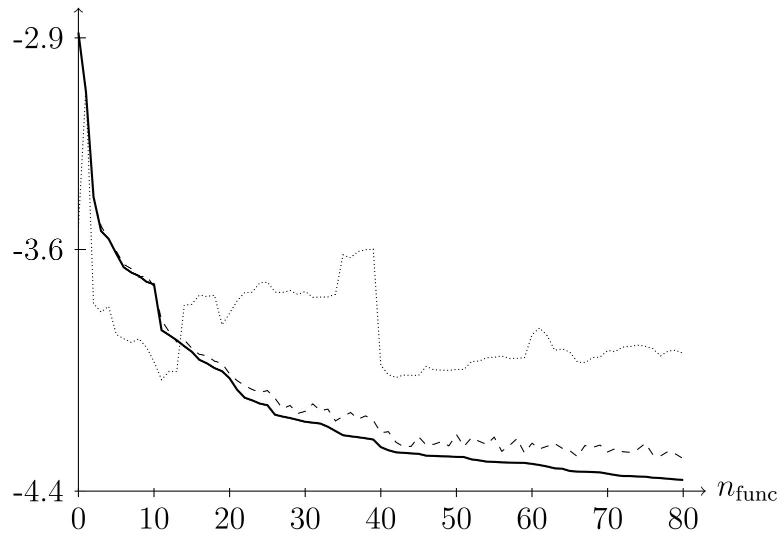

Figure 3.1.: Properties of the fit to the reference energies when the forward reduction procedure is employed. x-axis: current number n_{func} of basis functions in the model, y-axis: $\log(\sqrt{\text{nRSS}})$ (continuous) of the training set, $\log(\sqrt{\text{anPSS}})$ (dashed) obtained from subdividing the training set, $\log(\sqrt{\text{nPSS}})$ (dotted) for the testing set (energies of configurations to the 32-atomic supercell symmetry). All y values in mRy.

the average deviation is only 0.298 mRy (see "Full model").

Heisenberg model: For the same reasons as for bcc iron, an extended Heisenberg model was obtained from the spin spiral reference configurations, using Heisenberg basis functions up to the 25th neighbour. This gives a $\sqrt{\text{nRSS}}$ of 0.120 mRy, and predicts the energies of the supercell configurations with a $\sqrt{\text{RSS}}$ of 0.58 mRy per predicted configuration. This is a rather large error, as the supercell reference energies have an average deviation of only 1.477 mRy from the mean energy per atom. Limiting the predictions only to the 32-atomic supercells gives a $\sqrt{\text{RSS}}$ of 0.140 mRy per predicted configuration.

The predictive power of the models for fcc iron was only tested against the configurations to 32-atomic supercell symmetry, whose DFT energies contain large systematic differences to configurations to other supercell symmetries due to the different choices of the set $\{\mathbf{k}_i\}$ for the Brillouin zone integration during the DFT calculations by R. Singer. The author therefore decided to redo a calculation by R. Singer that tests the predictive power of the SCE for one specific ground state candidate, called $K_4^{(1)}$ in [1]. This configuration can be described within the same cubic, four-atomic supercell as was used in Section 3.2.1, with the angles $\vartheta_1 = 90.0$ degrees, $\varphi_1 = 0.0$, $\vartheta_2 = 90.0$, $\varphi_2 = 0.0$, $\vartheta_3 = 270.0$, $\varphi_3 = 0.0$, $\vartheta_4 = 90.0$, $\varphi_4 = 0.0$ (see [1], p. 145). This configuration has the lowest energy of all configurations investigated in [1], contrary to experimental evidence which suggests a ground state configuration that has a spin spiral symmetry. Therefore, it was investigated in [1] whether a modulation of the four-atomic supercell with a spin spiral symmetry leads to a reduction in energy[3]. Such configurations are different from the reference configurations, because the reference configurations contain only spin spiral modulations of the primitive cell, or configurations of the four-atomic cell without spin spiral modulation.

The investigation in [1] was done with a SCE consisting of 40 basis functions, which was compared to data obtained from the MASW program. The present author obtained the corresponding data from a plot on p.148 of [1], and compared the balanced model and the Heisenberg model obtained above against the SCE from [1] and MASW data. This comparison is shown in Figure 3.2. Somwhat satisfyingly, the improvements made seem to pay off, as the new model describes the DFT data significantly better.

However, it should not be forgotten that regarding general configurations the new model is probably still very much insufficient due to the small size of the supercells for the reference configurations, as the experience from the bcc SCE and the only rudimentary ability of the model to predict energies of the configurations to 32-atomic supercell symmetry suggests.

Altogether, one can achieve a similar accuracy for fcc-iron as for bcc iron. This is of little actual use because the typical differences in the energy per atom are significantly lower for fcc iron, leading to huge relative errors - to know a typical energy per atom of 5.14 mRy with an uncertanity of about 0.2 mRy is quite good (16-atomic supercells, efficient model, bcc iron), whereas to know an energy per atom of 0.3 mRy to within 0.1 mRy (32-atomic supercells, balanced model, fcc iron) is bad. The model presented here can only be expected to yield reliable results for configurations that have comparative symmetries as the reference configurations.

Solving this problem is as infeasible as it is for bcc iron, for the same reasons.

[3]For details on such configurations, see [1]. In the literature, these configurations can be found under the names “multi-k configurations” and “multi-k structures”.

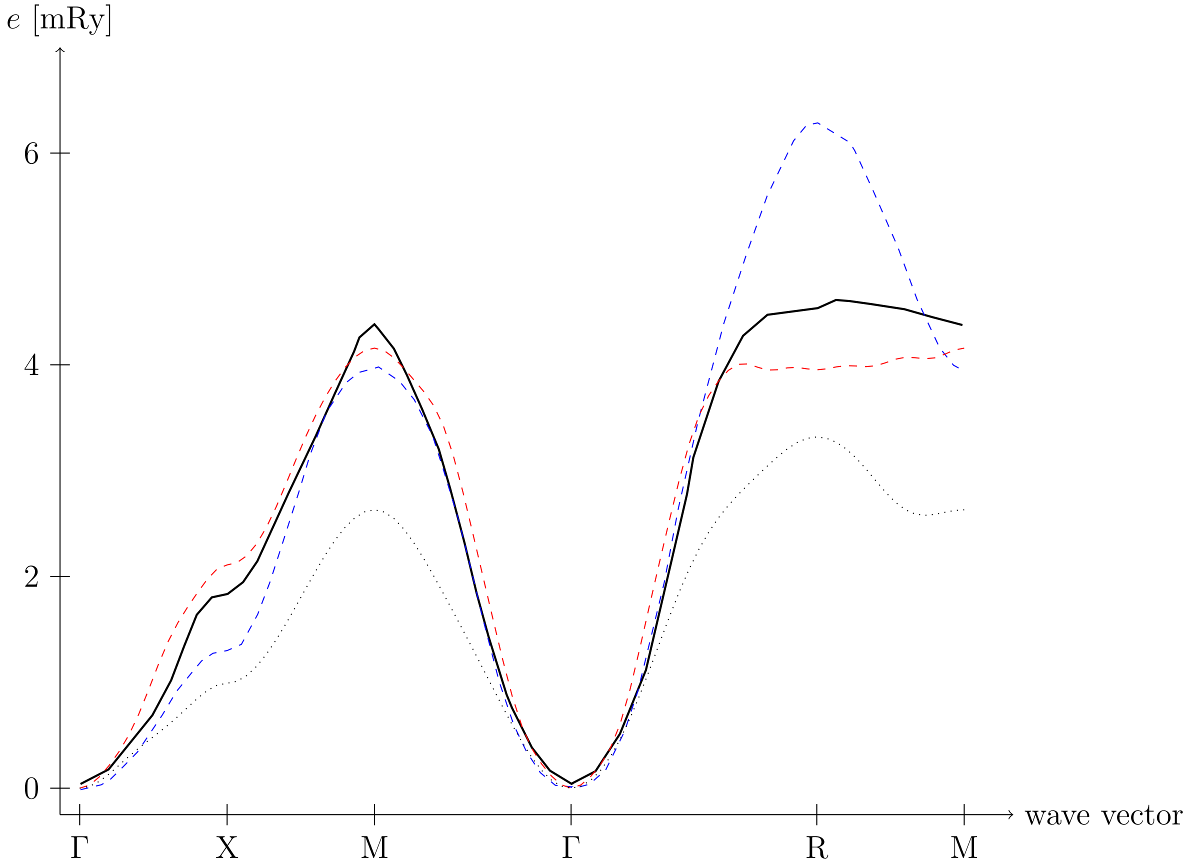

Figure 3.2.: Spin-spiral modulation of the $K_4^{(1)}$ configuration of the smallest cubic supercell of fcc iron [1]. Wave vectors along high-symmetry lines of the cubic Brillouin zone, which are described, e.g., in [2, 1]. Continuous line: DFT data (MASW program) Dashed lines: SCE of [1] (blue), balanced model of this work (red) Dotted line: Heisenberg model obtained from spin spirals

4. Failed work on other Crystal Structures, Elements and Geometries

4.1. Thin Films

One of the goals of the present work was to investigate the applicability of the SCE to other geometries, e.g., thin films. These systems have enjoyed quite a lot of theoretical interest recently and are accessible to magnetic imaging technologies such as XMCD microscopy, AFM or STM measurements, which enables more detailed experimental investigations. However, it became apparent that such an investigation is probably impossible with current DFT methods. In these methods, a thin film on a substrate is usually simulated by a so-called "slab" which consists of a sheet of atoms extending infinitely into two dimensions, and is bounded by vacuum in the third dimension on both sides. To simulate a thin film of a specific element on a substrate, the boundary layer of the slab consists of the atoms forming the film, and between the two layers enough atoms of the element forming the substrate are added to effectively separate the two layers. Naturally, how many atoms are needed can not be determined a priori. If n atoms separate the two layers, naturally the DFT calculations have to consider $n + 2$ atoms for the simplest magnetic configurations within the magnetic layer (spin spirals or ferromagnetic). If a two-dimensional supercell configuration with N atoms in the two-dimensional supercell for the magnetic layer is to be considered, the corresponding DFT supercell contains $N(n+2)$ atoms. Evidently, the necessary computation time increases quite fast as the necessary computation time still increases with the third power of the total number of atoms. Nevertheless, calculations for small two-dimensional supercells containing up to four atoms were estimated to be feasible.

Such calculations were attempted using the FLEUR program for iron on a tungsten substrate, with the surface normal along the (110) direction. However, despite significant efforts by the present author (e.g., implementing different mixing techniques), noncollinear DFT calculations did not converge, i.e, the mixing algorithms were unable to find a stable minimum. Collinear calculations yielded the expected results, as did noncollinear test calculations for simple bulk systems. Several reviews of the relevant parts of the FLEUR code by the author yielded no signs of programming errors or other

possible problems. It was therefore concluded that current mixing algorithms are probably insufficient, and the thin film calculations were abandoned.
Also, if the effects of the backfolding on the interaction coefficients in thin films are comparable to the effects in bulk bcc iron (see Section 2.2.2), which were investigated later than the attempts at thin film calculations, the validity of results that could be obtained from small two-dimensional supercells for the magnetic film if the technical difficulties were resolved seem questionable.

4.2. Other Elements and Geometries

After the failure of the investigations of thin films the author turned to investigations of other bulk materials, whose noncollinear magnetic configurations had been proven far easier to investigate by DFT. The present author had already conducted an investigation of fcc Ni during the course of his diploma thesis [39], which failed because the local magnetic moments become unstable for large cantings of the magnetic moment directions, i.e., the magnetic moments of the corresponding constrained ground state in the context of DFT are zero at some atomic positions $\mathbf{r}_i$. If one considers the cubic unit cell for Ni and chooses configurations where only the magnetic moment at the origin is canted respective to the magnetic moments of the other three atoms, the magnetic moment of the central atom vanishes when an angle of roughly 110 degrees is reached. All calculations where a larger angle is given as input to the program yield the same constrained ground state with a vanishing moment at the central position.
Clearly, a vanishing magnetic moment has no particular direction. One may disregard this complication and just presume that the energy function is simply constant with regards to changes of particular magnetic moments with an absolute value of zero. This, however, leads to significant parts of the total configuration space where the energy function is constant. While the corresponding DFT data itself can be quite well described by the SCE approach, the resulting models have shown almost no predictive power for general configurations.
Unfortunately, the problem occurs for all magnetic crystal structures of Co (fcc and hcp) as well. The only remaining magnetic elements are rare-earth elements. There, however, the principles of magnetism are vastly different as the dominant part of the magnetic moment of an atom is due to largely localized electrons in the 4f states, and only a small part is due to the itinerant 5d electrons. This leads to many open questions within the SCE, e.g., whether the two parts of the magnetic moments need to be treated as different, independent variables, which is currently not implemented in the LMTO code where the two parts are presumed to be either parallel or antiparallel, and to various other difficulties regarding the DFT calculations (see, e.g., [40, 41]). It was decided not to tackle these problems within the scope of the present work.

5. Summary

5.1. English Summary

The Spin Cluster Expansion (SCE) is a method to expand the magnetic energy in terms of a complete set of basis functions if an adiabatic approximation is made which identifies the directions of the local atomic magnetic moments as the relevant variables. The coefficients of this expansion are obtained by a least mean square fit against reference energies obtained by ab-initio Density Functional Theory (DFT) calculations. This process was shown in [1] to work for fcc iron with an accuracy of the fit that is of the same order of magnitude as the expected errors of the reference energies. The present work advances the theoretical understanding of the SCE and used this understanding to obtain improved numerical results.

On the theoretical side, an error of [1] when considering space group symmetry is resolved, partly in cooperation with R.Singer (Section 1.4). Symmetries of the configuration of the magnetic moments are studied for spin-spiral and supercell configurations (Section 1.5) and it is shown that the extended Heisenberg model commonly used for describing spin spirals [22, 23, 24] is equivalent to the SCE when the configurations of the magnetic moments are limited to spin spirals with one constant opening angle, or to spin spirals with very small opening angles. Methods are devised to determine the accuracy achievable with a specific finite subset of the infinite set of SCE basis functions by extracting information from comparatively simple DFT calculations (Section 1.7).

On the practical side the theoretical developments are tested extensively for bcc iron. The numerical results for fcc iron are shown to be improved with respect to the predictive power of the model. Several attempts are made to apply the SCE to other magnetic elements and to thin films. However, it turns out that in all other transition metals the local magnetic moments are not stable for all possible configurations of their directions, which makes the use of these directions as the relevant variables impossible. The application of the SCE to thin film geometries fails due to technical difficulties with the necessary DFT calculations.

5.2. Deutsche Zusammenfassung

Die Spin Cluster Entwicklung ist eine Methode, die Energiefunktion eines magnetischen Materials bezüglich eines eigens entwickelten vollständigen Basissatzes zu enwickeln.

Hierbei wird im Rahmen einer adiabatischen Näherung angenommen, dass die Richtungen der lokalen magnetischen Momente die wesentlichen Variablen zur Beschreibung der magnetischen Energie darstellen. Die Entwicklungskoeffizienten werden gewonnen, indem eine Anpassung an die Energien eines Satzes von Referenzkonfigurationen nach der Methode der kleinsten Fehlerquadrate durchgeführt wird. Die Energien der Referenzkonfigurationen werden hierbei durch ab-initio Dichtefunktionaltheorie (DFT) berechnet. Dass die Methode im Prinzip funktioniert und Genauigkeiten des Fits bis hin zur Genauigkeit der DFT-Referenzenergien erlaubt, wurde bereits in der Dissertation [1] von R.Singer für fcc Eisen gezeigt. Die theoretischen und praktischen Fortschritte, die im Rahmen dieser Arbeit erzielt wurden, werden im Folgenden zusammengefasst.
Ein Fehler der Arbeit [1] im Rahmen der Behandlung der Raumgruppensymmetrie eines Kristalls oder Moleküls wurde in Zusammenarbeit mit R.Singer korrigiert. Die Auswirkungen von Symmetrien der Konfiguration der magnetischen Momente wurden anhand von Spinspiralen und anderen sich periodisch wiederholenden Konfigurationen ("Superzellen") untersucht. Hierbei konnte gezeigt werden, dass das oft verwendete erweiterte Heisenberg Modell, in dem die Wechselwirkung in Form einer Kosinus nicht nur auf benachbarte klassische magnetische Momente beschränkt wird, nicht nur einen Teil der SCE darstellt wie in [15] von R.Singer gezeigt, sondern in bestimmten Fällen der exakten SCE entspricht. Diese Fälle sind, dass entweder nur Spinspiralen mit einem gemeinsamen, konstanten Öffnungswinkel behandelt werden oder Spinspiralen mit verschiedenen, aber sehr kleinen Öffnungswinkeln behandelt werden. Hierdurch kann die gute Übereinstimmung von derartigen Modellen mit den Energien von Spinspiralkonfigurationen [22, 23, 24] erklärt werden. Weiterhin wurden Methoden gefunden, wie für den im Prinzip unendlichen Satz von Basisfunktionen in der SCE wohl definierte Abschneidekriterien gefunden werden können und die resultierende Genauigkeit anhand von einfachen DFT-Rechnungen abgeschätzt werden kann.
Die theoretischen Fortschritte wurden anhand von Rechnungen für bcc Eisen überprüft. Weiterhin konnte die Vorhersagekraft der Methode für fcc Eisen deutlich verbessert werden. Die Anwendung der SCE auf andere magnetische Elemente (Co, Ni) scheiterte an dem Phänomen, dass die lokalen magnetischen Momente in diesen Metallen nicht stabil sind, es existieren viele Richtungskonfigurationen für die das magnetische Moment an einem oder mehreren Atomen verschindet. Dann sind die Richtungen der betroffenen magnetischen Momente nicht mehr wohl definiert. Die Anwendung der SCE auf andere Geometrien, im Besonderen dünne Eisenfilme, scheiterte aufgrund von technischen Problemen mit den DFT-Rechnungen.

A. Addendum

A.1. Properties of the Clebsch-Gordan coefficients

The Clebsch-Gordan coefficients are sometimes also called Wigner or Vector-Coupling coefficients. They are written as

$$\mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_2} = \mathrm{C}^{L_3}_{L_1, L_2} = (l_1 m_1 l_2 m_2 | l_1 l_2 l_3 m_3) \tag{A.1}$$

where the first notation is the standard used here, the second is a shorthand used in some lengthy formulas and the last notation is used in parts of the literature, e.g., in [16].
Relationship to the Wigner 3j-symbols often used in angular momentum theory:

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} = \frac{(-)^{l_3 - m_3}}{\sqrt{2l_3+1}} \mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_2} \tag{A.2}$$

The coefficients are nonzero if and only if:

$$l_1 + l_2 \geq l_3 \geq |l_1 - l_2| \tag{A.3}$$

$$m_3 = m_1 + m_2 \tag{A.4}$$

Four "orthogonality conditions":

$$\sum_{m_1 m_2} \mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_2} \mathrm{C}^{l_3' m_3'}_{l_1 m_1, l_2 m_2} = \delta_{l_3, l_3'} \delta_{m_3, m_3'} \delta(l_1 l_2 l_3), \tag{A.5}$$

$$\sum_{l_3 m_3} \mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_2} \mathrm{C}^{l_3 m_3}_{l_1 m_1', l_2 m_2'} = \delta_{m_1, m_1'} \delta_{m_2, m_2'}, \tag{A.6}$$

$$\sum_{m_1} \mathrm{C}^{l_3' m_3}_{l_1 m_1, l_2 m_3 - m_1} \mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_3 - m_1} = \delta_{l_3, l_3'} \delta(l_1 l_2 l_3), \tag{A.7}$$

$$\sum_{l_3} \mathrm{C}^{l_3 m_3}_{l_1 m_1', l_2 m_3 - m_1'} \mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_3 - m_1} = \delta_{m_1, m_1'}. \tag{A.8}$$

where $\delta(l_1 l_2 l_3)$ indicates the triangle condition (A.3).
Symmetries with regard to permutations of the indices:

$$\mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_2} = (-)^{l_1+l_2+l_3} \mathrm{C}^{l_3 m_3}_{l_2 m_2, l_1 m_1}. \tag{A.9}$$

$$\mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_2} = (-)^{l_1+l_2+l_3} \mathrm{C}^{l_3 -m_3}_{l_1 -m_1, l_2 -m_2}. \tag{A.10}$$

$$\begin{aligned} \mathrm{C}^{l_3 m_3}_{l_1 m_1, l_2 m_2} &= (-)^{l_1+m_1} \sqrt{\frac{2l_3+1}{2l_2+1}} \mathrm{C}^{l_2 m_2}_{l_3 m_3, l_1 -m_1} \\ &= (-)^{l_2+m_2} \sqrt{\frac{2l_3+1}{2l_1+1}} \mathrm{C}^{l_1 m_1}_{l_2 -m_2, l_3 m_3}. \end{aligned} \tag{A.11}$$

Decomposition of products[1]:

$$\begin{aligned} \mathrm{C}^{L_3}_{L_1,L_2} \mathrm{C}^{L_5}_{L_3,L_4} = & \sum_{l_6} \underbrace{\left(\sum_{m_1,m_2} \mathrm{C}^{L_3}_{L_1,L_2} \mathrm{C}^{L_5}_{L_3,L_4} \mathrm{C}^{L_6}_{L_2,L_4} \mathrm{C}^{L_5}_{L_1,L_6} \right)}_{:=C^{l_1,\dots,l_6}} \\ & \times \mathrm{C}^{L_6}_{L_2,L_4} \mathrm{C}^{L_5}_{L_1,L_6}, \end{aligned} \tag{A.12}$$

where $C^{l_1,\dots,l_6}$ (the parentheses) depends only on the values of the $l_1, \dots, l_6$. This quantity is related to the 6j-symbol common in angular momentum analysis.

special cases

In case of $m_1 = m_2 = 0$:

$$l_1 + l_2 + l_3 = \text{even}, \tag{A.13}$$

Final angular momentum of zero:

$$\mathrm{C}^{00}_{l_1 m_1, l_2 m_2} = \frac{(-)^{l_1-m_1}}{\sqrt{2l_1+1}} \delta_{l_1,l_2} \delta_{m_1,-m_2}. \tag{A.14}$$

A.2. Generalized Wigner Coefficients

The generalized Wigner coefficients

$$\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \tag{A.15}$$

are nonzero if and only if

$$\sum_i m_i = 0, \text{ and } 2j_i - \sum_{i'} j_{i'} \leq 0 \forall i \tag{A.16}$$

[1] See Eqs. (6.2.6) and (6.1.5) in [16]

Symmetry with regards to inversion in $\mathbf{m}$:

$$\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} = (-1)^{\sum l_i} \begin{pmatrix} l_1 & l_2 & \dots & l_{n_\alpha} \\ -m_1 & -m_2 & \dots & -m_{n_\alpha} \end{pmatrix}^{\mathbf{k}}. \tag{A.17}$$

Orthogonality:

$$\sum_{\mathbf{m}} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \begin{pmatrix} l_1 & \dots & l_{n_\alpha} \\ m_1 & \dots & m_{n_\alpha} \end{pmatrix}^{\mathbf{k}'} = \delta_{\mathbf{k}',\mathbf{k}}. \tag{A.18}$$

The case of $n_\alpha = 3$ is equal to the 3j-symbol,

$$\begin{aligned} \left.\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}}\right|_{n=3} &= \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \\ &= \frac{(-)^{l_3-m_3}}{\sqrt{2l_3+1}} \mathrm{C}^{l_3-m_3}_{l_1 m_1, l_2 m_2} \end{aligned} \tag{A.19}$$

The cases $n_\alpha = 2$ and $n_\alpha = 1$:

$$\begin{aligned} \left.\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}}\right|_{n_\alpha=2} &= \begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix} \\ &= \frac{(-)^{l_1-m_1}}{\sqrt{2l_1+1}} \delta_{l_1,l_2} \delta_{m_1,-m_2}. \\ \left.\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}}\right|_{n=1} &= \begin{pmatrix} l_1 \\ m_1 \end{pmatrix} = \delta_{l_1,0} \delta_{m_1,0}. \end{aligned} \tag{A.20}$$

A.3. Spherical harmonics

Definition:

$$Y^{lm}(\vartheta, \varphi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} \mathrm{e}^{\imath m \varphi} P_l^m(\cos(\vartheta)), \tag{A.21}$$

with the associated Legendre polynomials

$$P_l^m(\cos(\vartheta)) = \frac{(1-x^2)^{\frac{m}{2}}}{2^l l!} \frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}} (x^2-1)^l. \tag{A.22}$$

Complex conjugation:

$$\left(Y^{lm}(\vartheta, \varphi)\right)^* = (-)^m Y^{l-m}(\vartheta, \varphi). \tag{A.23}$$

Rotations around the z axis:

$$Y^{lm}(\vartheta, \varphi + \delta) = \mathrm{e}^{m \imath \delta} Y^{lm}(\vartheta, \varphi). \tag{A.24}$$

Orthogonality:

$$\int_S \mathrm{d}\Omega \left(Y^{l_1 m_1}(\vartheta,\varphi)\right)^* Y^{l_2 m}(\vartheta,\varphi) = \delta_{l_1,l_2}\delta_{m_1,m_2}. \tag{A.25}$$

Products of spherical harmonics with the same arguments[2]:

$$\begin{aligned} Y^{l_1 m_1}(\vartheta,\varphi)\cdot Y^{l_2 m_2}(\vartheta,\varphi) = \sum_l \left[\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}\right]^{\frac{1}{2}} \\ \times \mathrm{C}^{l\mu}_{l_1 m_1, l_2 m_2} Y^{l\mu}(\vartheta,\varphi)\mathrm{C}^{l0}_{l_1 0, l_2 0} \end{aligned} \tag{A.26}$$

with $\mu = m_1 + m_2$.

A.4. Proof of the Orbit-Stabilizer Theorem

Applying all $|\mathcal{G}|$ group elements to $\alpha_1, \mathbf{l}_{\alpha 1}$ will generate n_{diff} different "pictures" $\alpha_i, \mathbf{l}_{\alpha i}$, $i = 1, \ldots, n_{\text{diff}}$, which form the orbit $\mathcal{O}_{\alpha_1 \mathbf{l}_{\alpha 1}}$. However, one picture may be generated by several elements of $\mathcal{G}$, and therefore $\mathcal{O}_{\alpha_1 \mathbf{l}_{\alpha 1}}$ may contain less than $|\mathcal{G}|$ elements.
Assuming the stabilizer of $\alpha_1, \mathbf{l}_{\alpha 1}$ has $|\mathcal{G}_{\alpha_1 \mathbf{l}_{\alpha 1}}|$ elements $g_{\alpha_1 \mathbf{l}_{\alpha 1}}$, evidently $\alpha_1, \mathbf{l}_{\alpha 1}$ is generated $|\mathcal{G}_{\alpha_1 \mathbf{l}_{\alpha 1}}|$ times when all group elements are applied to $\alpha_1, \mathbf{l}_{\alpha 1}$.
If $\alpha_2, \mathbf{l}_{\alpha 2}$ is generated by $\tilde{g}$, one can generate $\alpha_2, \mathbf{l}_{\alpha 2}$ by at least $|\mathcal{G}_{\alpha_1 \mathbf{l}_{\alpha 1}}|$ group operations, as every operation $\tilde{g} g_{\alpha \mathbf{l}_\alpha} \forall g_{\alpha \mathbf{l}_\alpha}$ will map $\alpha_1, \mathbf{l}_{\alpha 1}$ to $\alpha_2, \mathbf{l}_{\alpha 2}$. Therefore, $\alpha_2, \mathbf{l}_{\alpha 2}$ is generated at least $|\mathcal{G}_{\alpha_1 \mathbf{l}_{\alpha 1}}|$ times when all group elements are applied to $\alpha_1, \mathbf{l}_{\alpha 1}$.
Now presume there is some other group element g_2 that maps $\alpha_1 \mathbf{l}_{\alpha 1}$ on $\alpha_2 \mathbf{l}_{\alpha 2}$, but can not be constructed in the same way as the elements discussed in the previous paragraph, $g_2 \neq \tilde{g} \circ g_{\alpha_1 \mathbf{l}_{\alpha 1}}$. However, $\tilde{g}^{-1} \circ g_2$ is evidently an element of $G_{\alpha_1 \mathbf{l}_{\alpha 1}}$. Trivially, $g_2 = \tilde{g} \circ \tilde{g}^{-1} \circ g_2$, but as the product of the latter two group operations is an element of the stabilizer, this is in direct contradiction to the presumption that g_2 may not be written as a product of $\tilde{g}$ and an element of the stabilizer. Therefore, there is no such g_2, and no other element than those constructed by $\tilde{g} g_{\alpha \mathbf{l}_\alpha} \forall g_{\alpha \mathbf{l}_\alpha}$ may map $\alpha_1, \mathbf{l}_{\alpha 1}$ to $\alpha_2, \mathbf{l}_{\alpha 2}$.
Therefore, each different picture is generated exactly $|\mathcal{G}_{\alpha_1 \mathbf{l}_{\alpha 1}}|$ times when all $|\mathcal{G}|$ group elements are applied to $\alpha_1, \mathbf{l}_{\alpha 1}$. This means there are only $n_{\text{diff}} = |\mathcal{G}| \,/\, |\mathcal{G}_{\alpha_1 \mathbf{l}_{\alpha 1}}|$ different pictures.

A.5. Details for Configurational Symmetry

A.5.1. Spin Spiral Symmetry

This section proves Eqn. (1.77). Presume,e.g., a basis function with $\alpha = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{n_\alpha})$, $\mathbf{l} = (l_1, l_2, \ldots, l_{n_\alpha})$, and a spin-spiral configuration. From the behaviour of the spherical

[2]see Eq. (4.6.5) in [16]

harmonics under rotation around the z-axis follows:

$$
\begin{aligned}
\mathcal{Y}_{\alpha\mathbf{l}_\alpha\mathbf{k}} &= \sum_{\mathbf{m}} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} Y^{l_1 m_1}(\vartheta_0, \varphi_0 + \mathbf{q}\mathbf{r}_1) \\
&\times Y^{l_2 m_2}(\vartheta_0, \varphi_0 + \mathbf{q}\mathbf{r}_2) \dots Y^{l_{n_\alpha} m_{n_\alpha}}(\vartheta_0, \varphi_0 + \mathbf{q}\mathbf{r}_{n_\alpha}) \\
&\underbrace{=}_{(A.24)} \sum_{\mathbf{m}} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \prod_{i=1}^{n_\alpha} \mathrm{e}^{m_i \imath \mathbf{q}\cdot\mathbf{r}_i} Y^{l_i m_i}(\vartheta_0, \varphi_0) \\
&= \sum_{\mathbf{m}} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \mathrm{e}^{\imath \sum_{i=1}^{n_\alpha} m_i \mathbf{r}_i \cdot \mathbf{q}} \prod_{i=1}^{n_\alpha} Y^{l_i m_i}(\vartheta_0, \varphi_0) \\
&\underbrace{\overset{*}{=}}_{(A.17)} \sum_{\mathbf{m}^+} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \left(\mathrm{e}^{\imath \mathbf{q} \sum_{i=1}^{n_\alpha} m_i \mathbf{r}_i} + \mathrm{e}^{-\imath \mathbf{q} \sum_{i=1}^{n_\alpha} m_i \mathbf{r}_i} (-1)^{\sum_{i=1}^{n_\alpha} m_i} \right) \\
&\times \prod_{i=1}^{n_\alpha} Y^{l_i m_i}(\vartheta_0, 0) \\
&= \sum_{\mathbf{m}^+} 2\cos(\mathbf{q}\cdot\mathbf{T}_{\mathbf{m}^+}) \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \prod_{i=1}^{n_\alpha} Y^{l_i m_i}(\vartheta_0, 0).
\end{aligned}
\tag{A.27}
$$

The sum over m_+ introduced in the fourth step, marked with an asterisk, indicates a decomposition of the set of all possible $\mathbf{m}$ into two parts, $\{\mathbf{m}\}_+$ and $\{\mathbf{m}\}_-$ such that to each $\mathbf{m}_+$ there is exactly one $\mathbf{m}_- = -\mathbf{m}_+$. Such a decomposition is always possible, except for the $\mathbf{m}$ where $m_i = 0 \forall i$. This $\mathbf{m}$ is defined to be part of $\{\mathbf{m}\}_+$.
Furthermore, it was presumed that φ_0=0, which can be presumed with no loss of generality, as the basis functions are invariant under global rotations of the $\{\mathbf{e}_i\}$. If $\varphi_0 = 0$, $Y^{lm}(\vartheta_0, 0) = (-1)^m Y^{l-m}(\vartheta_0, 0)$ holds. Furthermore, it was used that $\sum_i l_i =$ even, which means the generalized Wigner coefficients are invariant under inversion of $\mathbf{m}$.
The last step used that $\sum_{i=1}^{n_\alpha} m_i = 0$ as otherwise the generalized Wigner coefficient is zero, and introduced $\mathbf{T}_{\mathbf{m}^+} = \sum_{i=1}^{n_\alpha} m_i \mathbf{r}_i$ for a given $\mathbf{m}_+$. The product

$$
\prod_{i=1}^{n_\alpha} Y^{l_i m_i}(\vartheta_0, 0), \text{ with } \sum_{i=1}^{n_\alpha} l_i = \text{even}, \ \sum_{i=1}^{n_\alpha} m_i = 0 \tag{A.28}
$$

can be decomposed by applying (A.26) $n_\alpha - 1$ times to give

$$
\begin{aligned}
\prod_{i=1}^{n_\alpha} Y^{l_i m_i}(\vartheta_0, 0) &= \sum_{\tilde{l}=l_0,\text{even}}^{l_f} Y^{\tilde{l}0}(\vartheta_0, 0) N_{\mathbf{l},\mathbf{m}} \\
&= \sum_{\tilde{l}=l_0,\text{even}}^{l_f} P_{\tilde{l}}(\cos(\vartheta_0)) N_{\mathbf{l},\mathbf{m}}
\end{aligned}
\tag{A.29}
$$

where the sum runs only over even $\tilde{l}$, $l_f = \sum_{i=1}^{n_\alpha} l_i$ and $N_{\mathbf{l},\mathbf{m}}$ is some real number. $P_{\tilde{l}}(\cos(\vartheta_0))$ is the Legendre polynomial to $\tilde{l}$, which is equal to the associated Legendre polynomial $P_{\tilde{l}}^0(\cos(\vartheta_0))$ and to $Y^{\tilde{l}0}(\vartheta_0, 0)$.
The proof that m=0 for the final spherical harmonic rests on the repeated application of (A.26) to the product. The first application to $Y^{l_1 m_1}(\vartheta_0, 0) \cdot Y^{l_2 m_2}(\mathbf{e}^0_{\vartheta_0})$ gives a superposition of spherical harmonics $Y^{\tilde{l}_1 \mu_1}(\vartheta_0, 0)$ with $\mu_1 = m_1 + m_2$. Apply (A.26) again, to each product of the form $Y^{\tilde{l}_1 \mu_1}(\vartheta_0, 0) \cdot Y^{l_3 m_3}(\vartheta_0, 0)$, to give a further superposition of spherical harmonics with $\mu_2 = \mu_1 + m_3 = m_1 + m_2 + m_3$, and so on. In the final step this results in $\mu_{n_\alpha - 1} = \mu_{n_\alpha - 2} + m_{n_\alpha} = \sum_{i=1}^{n_\alpha} m_i = 0$.
The proof that each $\tilde{l}_{n_\alpha - 1} := \tilde{l}$ = even for the spherical harmonics in the final superposition is also derived along these lines:
The result of (A.26) for $Y^{l_1 m_1}(\vartheta_0, 0) \cdot Y^{l_2 m_2}(\vartheta_0, 0)$ contains $\mathrm{C}^{\tilde{l}_1 0}_{l_1 0, l_2 0}$, which requires $l_1 + l_2 + \tilde{l}_1$ = even as per (A.13). This holds for each repeated application of (A.26), and therefore

$$\begin{aligned} l_1 + l_2 + \tilde{l}_1 &= \text{even}, \\ \tilde{l}_1 + l_3 + \tilde{l}_2 &= \text{even}, \\ &\vdots \\ \tilde{l}_{n_\alpha - 2} + l_{n_\alpha} + \tilde{l}_{n_\alpha - 1} &= \text{even}. \end{aligned} \tag{A.30}$$

Taking the sum over all of these equations yields

$$\begin{aligned} &\underbrace{l_1 + l_2 + \ldots + l_{n_\alpha}}_{\text{even}} \\ &+ \underbrace{2 \cdot (\tilde{l}_1 + \tilde{l}_2 + \ldots + \tilde{l}_{n_\alpha - 2})}_{\text{even}} \\ &+ \tilde{l}_{n_\alpha - 1} = \text{even}, \end{aligned} \tag{A.31}$$

and therefore,

$$\tilde{l}_{n_\alpha - 1} = \text{even}. \tag{A.32}$$

It might be possible to calculate the coefficients $N_{\mathbf{l},\mathbf{m}}$ and the lower bound l_0 algebraically, but this was not deemed necessary.
Inserting Eqn. (A.29) into Eqn. (A.27) gives Eqn. (1.77).

A.5.2. Backfolding of Basis Functions

Within a supercell configuration backfolding of the SCE basis functions larger than the supercell occurs. This means duplicate arguments in the spherical harmonics arise. The calculations here are again carried out for the rotationally invariant basis functions, before considering space group symmetry. This means a prototypical basis functions is:

$$\mathcal{Y}_{\alpha \mathbf{l}_\alpha \mathbf{k}} = \sum_{\mathbf{m}} \begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} \times Y^{l_1 m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2 m_2}(\mathbf{e}_2^{\mathbf{r}_2}) \dots Y^{l_{n_\alpha} m_{n_\alpha}}(\mathbf{e}_{n_\alpha}^{\mathbf{r}_{n_\alpha}}) \tag{A.33}$$

To make the notation slightly easier, the presentation here limits itself to these basis functions where the **l**-tuple is already in descending order, which means

$$\begin{pmatrix} j_1 & \dots & j_{n_\alpha} \\ n_1 & \dots & n_{n_\alpha} \end{pmatrix}^{\mathbf{k}} = \begin{pmatrix} l_1 & \dots & l_{n_\alpha} \\ m_1 & \dots & m_{n_\alpha} \end{pmatrix}^{\mathbf{k}}. \tag{A.34}$$

The generalization to other orders should be evident from the results presented here.
A rather important "trick" in the following calculations is that due to the constraint $\sum m_i = 0$ for the generalized Wigner coefficients the summation over $\mathbf{m}$ has only $n_\alpha - 1$ independent indices, which may be chosen arbitrarily out of the m_i. One may also chose to sum over some of the intermediate μ_i, as those are fixed by the condition $\mu_{i-1} + m_{i+1} = \mu_i$ of each individual Clebsch-Gordan coefficient $\mathrm{C}^{k_i \mu_i}_{k_{i-1}\mu_{i-1}, l_{i+1} m_{i+1}}$, and μ_1 is fixed to $m_1 + m_2$. So instead of summing over m_{i+1} one may just as well sum over μ_i.
The basis function subject to backfolding can be reduced to a superposition of the basis functions within the cell by applying (A.26) to the product of the spherical harmonics which have the same argument. Formulas for $n_\alpha \leq 4$ will be given below.
To derive these formulas, it is convenient to first prove that the combination of Eqs. (A.7) and (A.26) results in the formula

$$\sum_{m_1} \mathrm{C}^{k\mu}_{l_1 m_1, l_2 m_2} Y^{l_1 m_1}(\vartheta, \varphi) Y^{l_2 m_2}(\vartheta, \varphi) = Q^k_{l_1, l_2} Y^{k\mu}(\vartheta, \varphi) \tag{A.35}$$

with

$$Q^k_{l_1, l_2} = \left[\frac{(2l_1 + 1)(2l_2 + 1)}{4\pi(2k + 1)} \right]^{\frac{1}{2}} \mathrm{C}^{k0}_{l_1 0, l_2 0} \tag{A.36}$$

Proof:

$$\begin{aligned}
&\sum_{m_1} \mathrm{C}^{k\mu}_{l_1 m_1, l_2 m_2} Y^{l_1 m_1}(\vartheta,\varphi) Y^{l_2 m_2}(\vartheta,\varphi) \\
&\overset{(A.26)}{=} \sum_{m_1} \mathrm{C}^{k\mu}_{l_1 m_1, l_2 m_2} \sum_{l=|l_1-l_2|}^{l_1+l_2} \underbrace{\left[\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}\right]^{\frac{1}{2}} \mathrm{C}^{l0}_{l_1 0, l_2 0}}_{:=Q^l_{l_1,l_2}} \\
&\times \mathrm{C}^{l\mu}_{l_1 m_1, l_2 m_2} Y^{l\mu}(\vartheta,\varphi) \\
&= \sum_l Q^l_{l_1,l_2} \sum_{m_1} \mathrm{C}^{l\mu}_{l_1 m_1, l_2 m_2} \mathrm{C}^{k\mu}_{l_1 m_1, l_2 m_2} Y^{l\mu}(\vartheta,\varphi) \\
&= \sum_l Q^l_{l_1,l_2} Y^{l\mu}(\vartheta,\varphi) \sum_{m_1} \mathrm{C}^{l\mu}_{l_1 m_1, l_2 \mu - m_1} \mathrm{C}^{k\mu}_{l_1 m_1, l_2 \mu - m_1} \\
&\overset{(A.7)}{=} \sum_l Q^l_{l_1,l_2} Y^{l\mu}(\vartheta,\varphi) \delta_{k,l} = Q^k_{l_1,l_2} Y^{k\mu}(\vartheta,\varphi).
\end{aligned} \tag{A.37}$$

The case $n_\alpha = 2$:

An example for the basis functions is:

$$\begin{aligned}
&\sum_{m_1} \mathrm{C}^{00}_{l_1 m_1, l_2 m_2} Y^{l_1 m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2 m_2}(\mathbf{e}_1^{\mathbf{r}_2 = \mathbf{r}_1 + \mathbf{T}_z}) \\
&\overset{(A.35)}{=} Q^0_{l_1,l_1} \underbrace{Y^{00}(\mathbf{e}_1^{\mathbf{r}_1})}_{(4\pi)^{-1/2}} \\
&\overset{(A.36)}{=} \left[\frac{(2l_1+1)(2l_1+1)}{4\pi}\right]^{\frac{1}{2}} \mathrm{C}^{00}_{l_1 0, l_1 0} \frac{1}{\sqrt{4\pi}} \\
&\overset{(A.14)}{=} \frac{2l_1+1}{4\pi} \frac{(-)^{l_1}}{\sqrt{2l_1+1}} = \frac{(-)^{l_1}}{4\pi}\sqrt{2l_1+1}.
\end{aligned} \tag{A.38}$$

The case $n_\alpha = 3$:

An example for the basis functions is:

$$\begin{aligned}
&\sum_{\mathbf{m}} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \\
&\times Y^{l_1 m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2 m_2}(\mathbf{e}_1^{\mathbf{r}_2 = \mathbf{r}_1 + \mathbf{T}Z}) Y^{l_3 m_3}(\mathbf{e}_3^{\mathbf{r}_3}),
\end{aligned} \tag{A.39}$$

Presuming that only the arguments of two spherical harmonics are equal, there are three possibilities: Either the arguments of the spherical harmonics to l_1, l_2 are equal, or those of the spherical harmonics with l_1, l_3 or l_2, l_3. However, the 3j-symbol with $\sum_i l_i =$ even is invariant under permutations of the l_i[16]. Therefore, it is sufficient to look at the case where the arguments to the spherical harmonics with l_1, l_2 are equal. Then:

$$\begin{aligned}
&\sum_{m_1,m_2} \frac{(-)^{l_3-m_3}}{\sqrt{2l_3+1}} \mathrm{C}^{l_3-m_3}_{l_1m_1,l_2m_2} \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_1^{\mathbf{r}_2=\mathbf{r}_1+\mathbf{T}z}) Y^{l_3m_3}(\mathbf{e}_2^{\mathbf{r}_3}) \\
=&\sum_{m_3} \frac{(-)^{l_3-m_3}}{\sqrt{2l_3+1}} Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) \sum_{m_1} \mathrm{C}^{l_3-m_3}_{l_1m_1,l_2m_2} \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_1^{\mathbf{r}_2=\mathbf{r}_1+\mathbf{T}z}) \\
\overset{(A.35)}{=}&\sum_{m_3} Q^{l_3}_{l_1,l_2} \frac{(-)^{l_3-m_3}}{\sqrt{2l_3+1}} Y^{l_3-m_3}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_3m_3}(\mathbf{e}_2^{\mathbf{r}_2}) \\
\overset{(A.14)}{=}&Q^{l_3}_{l_1,l_2} \sum_{m_3} \mathrm{C}^{00}_{l_3m_3,l_3m_3} Y^{l_3-m_3}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_3m_3}(\mathbf{e}_2^{\mathbf{r}_2})
\end{aligned} \tag{A.40}$$

where it was used in the second step that the two summation indices may be chosen at will out of $(m_1, m_2.m_3)$.
The result is proportional to a two-atomic basis function with the well-defined, sharp **l**-tuple (l_3, l_3).

The case $n_\alpha = 4$:

The simplest example is:

$$\begin{aligned}
&\sum_{m_1,m_2,m_3} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{k\mu}_{L_1,L_2} \mathrm{C}^{l_4-m_4}_{k\mu,L_3} \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_1^{\mathbf{r}_2=\mathbf{r}_1+\mathbf{T}z}) Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}) \\
&= \sum_{m_3,m_4} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{l_4-m_4}_{k\mu,L_3} Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}) \\
&\times \sum_{m_1} \mathrm{C}^{k\mu}_{L_1,L_2} Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_1^{\mathbf{r}_2=\mathbf{r}_1+\mathbf{T}z}) \\
&\overset{(A.35)}{=} \sum_{m_3,m_4} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{l_4-m_4}_{k\mu,L_3} Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}) \\
&\times Q^k_{l_1,l_2} Y^{k\mu}(\mathbf{e}_1^{\mathbf{r}_1}) \\
&= Q^k_{l_1,l_2} \sum_{\mu,m_3} \begin{pmatrix} k & l_3 & l_4 \\ \mu & m_3 & m_4 \end{pmatrix} \\
&\times Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}) Y^{k\mu}(\mathbf{e}_1^{\mathbf{r}_1})
\end{aligned} \tag{A.41}$$

The basis function is therefore projected solely on the three-atomic basis function with the l-tuple (k, l_3, l_4)
The pair l_3, l_4 is only slightly more difficult:

$$\begin{aligned}
&\sum_{m_1,m_2,m_3} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{K}_{L_1,L_2} \mathrm{C}^{l_4-m_4}_{K,L_3} \times \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4=\mathbf{r}_3+\mathbf{T}_Z}) \\
&\overset{(A.11)}{=} \sum_{m_1,m_2,m_3} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{K}_{L_1,L_2} \mathrm{C}^{k-\mu}_{L_3,L_4} \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4=\mathbf{r}_3+\mathbf{T}_Z}) \\
&= \sum_{m_1,m_2} \mathrm{C}^{K}_{L_1,L_2} Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_2^{\mathbf{r}_2}) \\
&\times \sum_{m_3} \frac{(-)^{k-\mu}}{\sqrt{2k+1}} \mathrm{C}^{K}_{L_3,L_4} Y^{l_3m_3}(\mathbf{e}_3^{\mathbf{r}_3}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4=\mathbf{r}_3+\mathbf{T}_Z}) \\
&= \sum_{m_1,m_2} \frac{(-)^{k-\mu}}{\sqrt{2k+1}} \mathrm{C}^{K}_{L_1,L_2} \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{k\mu}(\mathbf{e}_3^{\mathbf{r}_3})
\end{aligned} \tag{A.42}$$

which means that the basis function is projected solely on the three-atomic basis function with the l-tuple (l_1, l_2, k).
For the remaining pair note that by interchanging l_1, l_2 or l_3, l_4 the basis function changes at most its sign, therefore it is sufficient to look at the pair l_2, l_3. The solution resorts

to Eq. (A.12):

$$
\begin{aligned}
&\sum_{m_1,m_2,m_3} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{K}_{L_1,L_2} \mathrm{C}^{l_4-m_4}_{K,L_3} \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{l_3m_3}(\mathbf{e}_2^{\mathbf{r}_3=\mathbf{r}_2+\mathbf{T}_Z}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}) \\
&= \sum_{m_1,m_2,m_3} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \sum_{\tilde{l}} C^{l_1,\ldots,\tilde{l}} \mathrm{C}^{\tilde{L}}_{L_2,L_3} \mathrm{C}^{l_4-m_4}_{L_1,\tilde{L}} \times \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_2m_2}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{l_3m_3}(\mathbf{e}_2^{\mathbf{r}_3=\mathbf{r}_2+\mathbf{T}_Z}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}) \\
&= \sum_{\tilde{l}} C^{l_1,\ldots,\tilde{l}} \sum_{m_1,m_4} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{l_4-m_4}_{L_1,\tilde{L}} Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}) \\
&\times \sum_{m_3} \mathrm{C}^{\tilde{L}}_{L_2,L_3} Y^{l_2m_2}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{l_3m_3}(\mathbf{e}_2^{\mathbf{r}_3=\mathbf{r}_2+\mathbf{T}_Z}) \\
&= \sum_{\tilde{l}} C^{l_1,\ldots,\tilde{l}} C^{\tilde{l}} \sum_{m_1,\tilde{\mu}} \frac{(-)^{l_4-m_4}}{\sqrt{2l_4+1}} \mathrm{C}^{l_4-m_4}_{L_1,\tilde{L}} \\
&\times Y^{l_1m_1}(\mathbf{e}_1^{\mathbf{r}_1}) Y^{\tilde{l}\tilde{\mu}}(\mathbf{e}_2^{\mathbf{r}_2}) Y^{l_4m_4}(\mathbf{e}_4^{\mathbf{r}_4}).
\end{aligned}
\tag{A.43}
$$

which means that the basis function is now a superposition of *all* three-atomic basis functions with l-tuples $(l_1, \tilde{l}, l_4)$.

A.6. Expansion of Vector Quantities

The transformation properties of $\mathcal{Y}^{11}_{\alpha\mathbf{lk}}, \mathcal{Y}^{10}_{\alpha\mathbf{lk}}, \mathcal{Y}^{1-1}_{\alpha\mathbf{lk}}$ with respect to a global, homogeneous rotation of the $\{\mathbf{e}_i\}$, defined by the Euler angles α, β, γ, are given by

$$
\begin{aligned}
D^{(1)}_{m'm}(\alpha,\beta,\gamma) &= e^{\imath m'\gamma} d^{(1)}_{m'm}(\beta) e^{\imath m\alpha}, \\
d^{(1)}_{m'm}(\beta) &= \begin{pmatrix} \frac{1}{2}(1+\cos(\beta) & \frac{1}{\sqrt{2}}\sin(\beta) & \frac{1}{2}(1-\cos(\beta)) \\ \frac{1}{\sqrt{2}}\sin(\beta) & \cos(\beta) & \frac{1}{\sqrt{2}}\sin(\beta) \\ \frac{1}{2}(1-\cos(\beta) & -\frac{1}{\sqrt{2}}\sin(\beta) & \frac{1}{2}(1+\cos(\beta)) \end{pmatrix}.
\end{aligned}
\tag{A.44}
$$

From comparison of these transformation properties follows that any vector $\mathbf{v}$ with components v_x, v_y, v_z in some global cartesian coordinate system can be expanded as

$$v_+ = v_x + v_y = \sum_{\alpha}\sum_{\mathbf{l}}\sum_{\mathbf{k}} J_+^{\alpha\mathbf{lk}} \mathcal{Y}^{11}_{\alpha\mathbf{lk}}. \tag{A.45}$$

$$v_0 = v_z = \sum_{\alpha}\sum_{\mathbf{l}}\sum_{\mathbf{k}} J_0^{\alpha\mathbf{lk}} \mathcal{Y}^{10}_{\alpha\mathbf{lk}}. \tag{A.46}$$

$$v_- = v_x - v_y = \sum_{\alpha}\sum_{\mathbf{l}}\sum_{\mathbf{k}} J_{-1}^{\alpha\mathbf{lk}} \mathcal{Y}^{1-1}_{\alpha\mathbf{lk}}. \tag{A.47}$$

A.7. Energy Scales

The following paragraphs aim to give a rough overview over the energy scales related to magnetism within the considered materials. This should provide the reader with some context for the energies involved.
The thermodynamic energy scale is set forth by the Curie temperature multiplied by the Boltzmann constant. Furthermore, a measure for the energy differences caused by changes in the directions of the magnetic moments is given by the difference between the ferromagnetic and antiferromagnetic state, or rather the ferromagnetic and some strongly canted state for space groups which correspond to frustrated antiferromagnets. Also, the energy response to small deviations of the directions from of the ground state configuration is provided. For changes in the absolute values of the magnetic moments, a reasonable energy scale is set by the difference between the energies of the completely nonmagnetic state, where each magnetic moment vanishes, and the ground state. One can also investigate the energy response to small longitudinal deviations from the ground state.
The Tables A.1 and A.2 show the energy equivalent of the Curie temperature and information obtained by the following calculations:

- The magnetic moment per atom m_0 in the ferromagnetic state.
- The difference between the energies per atom for the ferromagnetic state e_{FM} and the nonmagnetic state e_{NM}, where all magnetic moments $\mathbf{m}$ are zero.
- The difference between the energies per atom for the ferromagnetic state e_{FM} and a strongly canted state e_{can}. For the bcc crystal structure this is simply the antiferromagnetic configuration, where the angle between neighbouring magnetic moments is π. There is no such configuration for the fcc and hcp crystal structure. For the fcc structure a cubic supercell with four atoms was chosen where the magnetic moment of the atom at the origin is canted by π relative to the magnetic moments on the faces of the cube, which are all collinear. For the hcp structure a supercell with two atoms was chosen, and the two magnetic moments are canted by π relative to each other.
- The response of the energy per atom to a small deviation from the ferromagnetic state was calculated. For this, the same supercell configuration as for the strongly canted state was used. For ferromagnets and small deviations, this response is expected to be proportional to δ^2 where δ is some appropriate measure of the deviation. As two types of deviations, either in the magnetic moment direction or in the absolute value of the magnetic moment, for ferromagnetic alignment, are considered separately, the corresponding proportionality factor is marked as a_{long} or a_{trans} respectively.

Material	$m_0[\mu_B]$	$k_B T_C$ [43, 44]	$e_{\mathrm{FM}} - e_{\mathrm{cant}}$	$e_{\mathrm{FM}} - e_{NM}$
Fe (bcc)	2.26	6.61	-31.2	-37.4
Fe (fcc)	1.14	0.28	4.3	—
Co (fcc)	1.67	8.79	-14.1 (*)	-14.1
Co (hcp)	1.64	—	-16.7	-17.1
Ni (fcc)	0.65	3.97	-3.7 (*)	-3.7

Table A.1.: Energy range in mRy of large transversal and longitudinal fluctuations of the magnetic moments for some magnetic metals.

Material	a_{long}	a_{trans}
Fe (bcc)	139	43.8
Fe (fcc)	—	—
Co (fcc)	29.6	44.8
Co (hcp)	57.0	51.3
Ni (fcc)	3.55	8.69

Table A.2.: Energy range in mRy of small transversal and longitudinal fluctuations of the magnetic moments for some magnetic metals.

The deviation δ is measured as $[\delta] = 10^{-3} m_0$ for changes in the absolute value of the magnetic moment and as $[\delta] = 10^{-3}\pi$ for changes in the directions of the magnetic moments. The proportionality factor was acquired by a quadratic fit to deviations up to $10[\delta]$, i.e., $10^{-2} m_0$ for the absolute values and $10^{-2}\pi$ for the directions.

For changes in the directions of the magnetic moments, the absolute values of the magnetic moments were not constrained to the ground state magnetic moment but instead calculated self-consistently. For the mathematical formulation of the constraints, see Section 1.8.2, details of the implementation and a comparison with other DFT codes are available in Ref. [42].
For materials where the nonmagnetic state is lower in energy than the respective strongly canted state, the procedure naturally results in the nonmagnetic state. These cases are marked with an asterisk (*). The temperature given for fcc Fe represents the Neel temperature as given in the literature. As the ground state of fcc Fe is still unknown, it is not possible to calculate small deviations from the ground state.
A last note on hcp Co: Here the respective energies of the nonmagnetic and the strongly canted state are very close to each other. While no canted configuration where all absolute values of the magnetic moments are zero has been found by the author, there are sizeable parts of the configuration space of the directions where the absolute value of

at least one magentic moment is zero. One such class of configurations is given by one strongly canted moment embedded in a supercell with three other moments that are parallel to each other. In this case the absolute value of the canted magnetic moment vanishes for cantings larger than about $5/6\pi$.
The numbers in Tables A.1 and A.2 do not show different energy scales for longitudinal and transversal fluctuations, and thus provide no support for the adiabatic approximation behind the SCE. The arguments cited in Section 1.2 do use the term "intraatomic exchange", which, to the knowledge of the present author, has never been defined explicitly anywhere. The present author presumes that the term is meant to indicate the energy changes due to local longitudinal fluctuations as calculated here. Therefore, longitudinal fluctuations should be included in thermodynamic investigations, especially for Ni and Co where they are associated with actually lower energy changes than transversal fluctuations.

A.8. Values of the Interaction Coefficients

A.8.1. bcc Iron

The following table lists the definitions of the 15 basis functions that most efficiently decrease the nRSSr. The most efficient basis function is at the top.

α_i	$\mathbf{l}_{\alpha i}$	$j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ [mRy]	$\lvert\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}\rvert$	$\sqrt{\text{nRSS}}$	N_{eff}
$(O, 0.5\sqrt{3}a[111])$	(1,1)	9.02982	12	0.839	1
$(O, \sqrt{3}a[111])$	(1,1)	-0.75898	12	0.674	6
$(O, 0.5\sqrt{3}a[111])$	(2,2)	0.20079	12	0.433	17
$(O, 1.5\sqrt{3}a[111])$	(1,1)	0.45167	12	0.384	8
$(O, a[100])$	(1,1)	0.28286	16	0.366	11
$(O, 2a[100])$	(1,1)	-0.53596	16	0.324	3
$(O, \sqrt{2}a[110])$	(1,1)	0.10105	8	0.310	68
$(O, 0.5\sqrt{3}a[111])$	(3,3)	-0.18078	12	0.284	122
$(O, a[100])$	(2,2)	-0.30559	16	0.260	116
$(O, \sqrt{2}a[110])$	(2,2)	0.17994	8	0.249	22
$(O, 2\sqrt{2}a[110])$	(1,1)	0.25809	8	0.241	–
$(O, 0.5\sqrt{3}a[111], 0.5\sqrt{11}a[311])$	(2,1,1)	0.40554	2	0.218	13
$(O, a[100])$	(3,3)	0.13151	16	0.208	–
$(O, 0.5\sqrt{59}a[553])$	(1,1)	0.24128	4	0.204	46
$(O, 2\sqrt{2}a[110], \sqrt{2}a[110])$	(2,1,1)	0.30842	2	0.182	12

The first column lists the position vectors of the atoms that form the cluster archetype, where O stands for the origin of the coordinate system. The second column gives the

$\mathbf{l}_{\alpha i}$-tuple, the third column the interaction coefficient in mRy when all these 15 basis functions are used to fit the reference energies, the fourth column the number of space group elements in the respective stabilizer, and the fifth column the $\sqrt{\text{nRSS}}$ that corresponds to the model consisting of all basis functions that appear in the current and preceding rows of the table. The $\sqrt{\text{nRSS}}$ resulting from fitting just the constant basis function is 8.425 mRy.
When selecting basis functions only by the largest decrease in the nRSS as in [31], the resulting model differs quite a bit. The last column denotes when the corresponding basis function is added to a model where the basis functions are selected only by the nRSS as in Ref. [31] instead of selecting by the most efficient basis function (see Section 1.8.6). E.g., the basis function corresponding to the second row would be added as the sixth basis function when selecting purely by the nRSS.
The decrease in $\sqrt{\text{nrSS}}$ associated with a basis function also depends quite significantly on the other terms which form the rest of the current model under consideration. This happens because there are quite often two or more basis functions which represent largely the same behaviour within the reference data. This happens if, e.g., one basis function is mostly backfolded upon one specific other basis function for small supercells. Then the reference energies for small supercells can be described by either of those basis functions correctly, and after one of these basis functions is selected the other basis function becomes irrelevant.

A.8.2. Fcc Iron

This table lists the 15 most efficient basis functions for fcc iron.

α_i	$\mathbf{l}_{\alpha i}$	$j_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}$ [mRy]	$\lvert\mathcal{G}_{\alpha_i \mathbf{l}_{\alpha i}}\rvert$	$\sqrt{\text{nRSS}}$
$(O, 0.5\sqrt{2}a[011])$	(1,1)	-1.1727	8	0.833
$(O, a[100])$	(1,1)	1.2643	16	0.375
$(O, 0.5\sqrt{6}a[211])$	(1,1)	-0.3700	4	0.289
$(O, 0.5\sqrt{10}a[310])$	(1,1)	-0.3456	4	0.272
$(O, 0.5\sqrt{18}a[330])$	(1,1)	0.2615	8	0.243
$(O, a[100])$	(2,2)	-0.1164	16	0.219
$(O, 0.5\sqrt{38}a[532])$	(1,1)	-0.1371	2	0.211
$(O, \sqrt{3}a[111])$	(1,1)	-0.1302	12	0.205
$(O, 0.5\sqrt{2}a[011])$	(3,3)	0.1777	8	0.196
$(O, \sqrt{5}a[210])$	(1,1)	0.0337	4	0.192
$(O, 0.5\sqrt{2}[101], 0.5\sqrt{6}[211], 0.5\sqrt{10}[310])$	(1,1,1,1)	0.25809	2	0.136
$(O, \sqrt{6}a[211])$	(1,1)	-0.1056	4	0.131
$(O, 0.5\sqrt{26}a[431])$	(1,1)	0.2283	2	0.126
$(O, \sqrt{8}a[220])$	(1,1)	-0.0795	8	0.120
$(O, 0.5\sqrt{14}a[321])$	(1,1)	-0.1883	2	0.115

The eigenvector $\mathbf{u}$ for the four-atomic cluster, which has been separated by horizontal lines for easier identification, is simply 1(0), i.e., the $\mathbf{k}$-tuple (0) is already a eigenvector of the projection operator to the identical representation of the space group. The $\sqrt{\text{nRSS}}$ that results from fitting the constant basis function is 1.311 mRy.

A.9. Extending the SCE basis set

While the general formalism of the SCE as detailed in Chapter 1 is certainly powerful, the self-imposed restriction of the variables to only the directions of magnetic moments of atoms of one element is somewhat drastic. Most industrially used metals are binary or ternary alloys, and most physical phenomena involve not only changes of the magnetic moment directions but changes in other variables as well, e.g., changes of the lattice constant due to heating or external pressure.

The straightforward generalization of the approach detailed in Chapter 1 is to simply use a product of the SCE basis functions and basis functions from a complete basis set in the configuration space of the other variables as basis functions for the energy. E.g., for a magnetic alloy one would use a product of the set of the Chebyshev polynomials, which

are commonly used for expanding the formation energy of alloys, and the SCE basis functions. Such an approach greatly increases the set of possible basis functions and the number of reference configurations which need to be investigated, thus requiring many more costly DFT calculations, which makes extensive investigations currently infeasible. Also, the mathematical details of this approach have not been investigated in the course of this work. To do so, one would have to carefully retrace the steps taken for the SCE basis functions and reconsider symmetry under rotations in the space of magnetic moment directions, space group symmetry, etc. with the new basis functions.
Presented below are nevertheless two simple, exploratory calculations, one for variable lattice constant and one for treating the absolute value of the magnetic moments as an independent variable. These test calculations show similar accuracy relative to the respective typical energy changes as the regular SCE, although the absolute accuracy is significantly worse.

A.9.1. Variable lattice constant

Presume that there is no structural phase transition, that is the atoms retain the same space group symmetry, and the lattice constant a varies only on some interval $\mathcal{I} = [a_{\mathrm{min}}, a_{\mathrm{max}}]$. The complete configuration space of the lattice constant and the magnetic moments is then a product of $\mathcal{I} \otimes (S^2)^N$, where N is the number of atoms.
The procedure suggested in the introduction to this section then requires a complete basis set on $\mathcal{I}$. In the context of this work it is convenient to use the Legendre polynomials, as they are just special cases of the spherical harmonics. The Legendre polynomials are denoted P_i, and are orthogonal (although not orthonormal) and complete on the interval $[-1, 1]$. One point of view on this approach is that the dependence of the SCE interaction coefficients on the lattice constant is interpolated with Legendre polynomials. However, this formulation suggests that first a SCE for each lattice constant is obtained, and then, in a second step, the resulting interaction coefficients are fitted with Legendre polynomials. This is not the approach taken here, instead one large fit was done for all data at once.
The investigation was done for bcc iron and a was varied from $0.95a_0 \leq a_0 \leq 1.05a_0$. This interval can be mapped to $[-1, 1]$ by $f : a \mapsto f(a) = (a/a_0 - 1)/0.05$. A basis set for the energy per atom is then given by the products

$$\tilde{\gamma}^{\mathcal{G}}_{i\alpha\mathbf{lu}} = P_i(f(a)) \cdot \gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\{\mathbf{e}_i\}). \tag{A.48}$$

A cubic supercell with a volume of a^3 containing two atoms was used. All possible configurations are then described by two variables, the current lattice constant a and ϑ, the angle between the directions of the magnetic moments of the two atoms. Legendre polynomials with $0 \leq i \leq 5$ were used, and all SCE basis functions to $l_{\mathrm{tot}} \leq 10$. This results in 36 possible basis functions, whereby the basis function to $i = 0, l = 0$ is a

constant.
The energy landscape was probed by DFT at 21 equidistant points in a and ϑ each, which gives a total of 441 reference configurations. Fitting only the constant basis function to these configurations gives a $\sqrt{\text{nRSS}}$ of 25.9 mRy, which represents the typical energy difference within the reference energies. Using all 36 basis functions one obtains a $\sqrt{\text{nRSS}}$ of 0.18 mRy. A $\sqrt{\text{nRSS}}$ of 1 mRy can be achieved by using just eight basis functions. Considering the higher energy scales involved, these results are comparable to those achievable with the SCE when only the magnetic moment directions are considered as variables.

A.9.2. A Basis Set for Magnetic Moments

If the adiabatic approximation for the directions of the magnetic moments is not justifiable, one for the actual magnetic moments might be. Then, not only the directions of the magnetic moments are independent variables, but also the absolute values M of the magnetic moments. Naturally, M has a lower bound of zero, corresponding to vanishing magnetic moment, and one can reasonably presume that there exists also a higher boundary M_{max}, which can be at most the number of (valence) electrons per atom times the Bohr magneton, or the boundary may be defined indirectly by some reasonable energy limit. Therefore, it is presumed that the configuration space of the absolute values is again an interval $I = [0, M_{\text{max}}]$.The derivation of a basis set follows the same outline as that for the original SCE basis set, and the symmetrization with regards to global rotations of the $\{\mathbf{e}_i\}$ can be carried out in the same way as for the original SCE basis functions, as the corresponding symmetry operations do not act on the absolute values of the magnetic moments.

A quick test calculation

To test the merit of the approach detailed above, again a simple test calculation for the same cubic, two-atomic supercell of bcc iron was done. The magnetic configuration is then described by three variables: The two absolute values $|\mathbf{m}_1|, |\mathbf{m}_1|$ and the angle ϑ between the two magnetic moments. The absolute values were chosen out of $I = [0, 3.23572\mu_B]$. The energy relative to the global ground state was within $0 \leq E \leq 124$ mRy, whereby the highest energy was expectedly found in the state corresponding to $|\mathbf{m}_1| = |\mathbf{m}_2| = 3.23572\mu_B, \vartheta = \pi$.
Again, it is convenient to use a product of the SCE basis functions for this case and the Legendre polynomials P_i, whereby the interval of the possible values for the absolute value of the magnetic moment is mapped to [-1,1] by $h : M \mapsto h(M) = (M/M_{\text{max}} - 1)/0.05$. At $M = 0$ the energy (and therefore, the basis functions) should no longer

change if the angles are changed. This is not true for all basis functions within the set considered here (e.g., the product of the constant Legendre polynomial to order 0 and a SCE basis function has nonzero derivatives in ϑ, φ at $|\mathbf{m}| = 0$), and the author did not derive a solution for this problem. It may, however, be argued that the above basis set is complete, and therefore a parametrization that reproduces the necessary behaviour of the energy function may also be constructed from the above basis set, possibly with a loss of efficiency compared to an adapted basis set. For a discussion of a similar problem see Refs. [45, 46]. Within what limits this argument holds true in our case would be an object of a more thorough investigation.

Nevertheless, using the above basis set should allow to estimate the accuracy that can be reached with a hypothetical correctly adapted basis set. The two atoms within the supercell are symmetric due to space group symmetry, and the corresponding basis functions may possibly be written as

$$\begin{aligned} &(\mathcal{P}_{i_1}(h(M_1))\mathcal{P}_{i_2}(h(M_2)) + \mathcal{P}_{i_1}(h(M_2))\mathcal{P}_{i_2}(h(M_1))) \\ &\times \gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\vartheta) \end{aligned} \tag{A.49}$$

with the SCE pair basis functions $\gamma^{\mathcal{G}}_{\alpha_i \mathbf{l}_{\alpha i} \mathbf{u}}(\vartheta)$. This is not the result of a rigid derivation (which would follow the same general outline as the derivation for the original SCE basis functions by using the projection operator), but just an educated guess.

Using the possible orders $i_1 \in [0, 5]$, $i_2 \in [0, i_1]$ for the Legendre polynomials and $l_{\mathrm{tot}} \leq 10$ for the SCE basis functions results in 125 possible basis functions. The energy landscape was probed by DFT for 1056 configurations, and fitting just the constant function gives a $\sqrt{\mathrm{nRSS}}$ of 19.1 mRy, setting the energy scale. Using all basis functions the reference energies can be fitted with a $\sqrt{\mathrm{nRSS}}$ of about 0.15mRy. A $\sqrt{\mathrm{nRSS}}$ of 1mRy can be achieved with 14 basis functions.

Index

Bibliography

[1] Singer, R. *Modellierung von Spinwechselwirkungen auf atomarer Skala und Energetik konkurrierender Spinkonfigurationen in fcc-Eisen.* PhD thesis, Max-Planck-Institut für Metallforschung, (2008).

[2] Grotheer, O. *Ab-initio Berechnung der Spinwellenspektren von Eisen, Kobalt und Nickel.* PhD thesis, Max-Planck-Institut für Metallforschung, (2001).

[3] Drautz, R. and Fähnle, M. *Phys. Rev. B* **69**, 104404 (2004).

[4] E.Avron, J. and Elgart, A. *Communications in Mathematical Physics* **203**, 445–463 (1999).

[5] Comparat, D. *Phys. Rev. A* **80**, 012106 (2009).

[6] Korenman, V., Murray, J., and Prange, R. *Phys. Rev. B* **16**, 4032–4047 (1977).

[7] Korenman, V., Murray, J., and Prange, R. *Phys. Rev. B* **16**, 4048–4057 (1977).

[8] Korenman, V., Murray, J., and Prange, R. *Phys. Rev. B* **16**, 4058–4062 (1977).

[9] Antropov, V., Katsnelson, M., Harmon, B., van Schilfgaarde, M., and Kusnezov, D. *Phys. Rev. B* **54**, 1019–1035 (1996).

[10] Gyorffy, B., Pindon, A., Staunton, J., Stocks, G., and Winter, H. *J.Phys.F:Met.Phys.* **15**, 1337–1386 (1985).

[11] J.Hubbard. *Proc.Phys.Soc.,London,Sect. A* **276**, 238–257 (1963).

[12] Uhl, M. and Kübler, J. *Phys. Rev. Lett.* **77**, 334–337 (1996).

[13] Sandratskii, L. M. and Mavropoulos, P. *Phys. Rev. B* **83**, 174408 (2011).

[14] Ruban, A. V., Khmelevskyi, S., Mohn, P., and Johansson, B. *Phys. Rev. B* **75**, 054402 (2007).

[15] Singer, R. and Fähnle, M. *Journal of Mathematical Physics* **47**, 113503 (2006).

[16] Edmonds, A. *Angular Momentum in Quantum Mechanics.* Princeton University Press, (1996).

[17] Cromwell, J. *Group Theory in Physics.* Acad. Press, (1984).

[18] Streitwolf, H. *Gruppentheorie in der Festkörperphysik.* Akademische Verlagsgesellschaft Geest & Portig K.G., (1967).

[19] S.Sternberg. *Group theory and physics*, chapter 2.5 Action on function spaces. Cambridge University Press (1994).

[20] S.Sternberg. *Group theory and physics*, chapter 1.3 The action of a group on a set. Cambridge University Press (1994).

[21] Sandratskii, L. M. *Advances in Physics* **47**, 91–160 (1998).

[22] Pajda, M., Kudrnovsky, J., I.Turek, Drchal, V., and Bruno, P. *Phys. Rev. B* **64**, 174402 (2001).

[23] Turek, I., J.Kudrnovsky, Drchal, V., and Bruno, P. *Philosophical Magazine* **86**(12), 1713–1752 (2006).

[24] Kübler, J. *Theory of Itinerant Electron Magnetism.* Oxford Science Publications, (2000).

[25] Levy, M. *Proc. Natl.Acad.Sci. USA* **76**, 6062–6065 (1979).

[26] Perdew, J. and S.Kurth. volume 620 of *Lecture Notes in Physics*, chapter Density Functionals for Non-relativistic Coulomb systems in the New Century. (2003).

[27] Srivastava, G. P. *J. Phys. A* **17**, L317–L321 (1984).

[28] Dederichs, P., Blügel, S., Zeller, R., and Akai, H. *Phys. Rev. Lett.* **53**, 2512–2515 (1984).

[29] R.Singer, Fähnle, M., and Bihlmayer, G. *Phys. Rev. B* **71**, 214435 (2005).

[30] Sjöstedt, E. and Nordström, L. *Phys. Rev. B* **66**, 014447 (2002).

[31] Singer, R., Dietermann, F., and Fähnle, M. *Phys. Rev. Lett.* **107**, 017204 (2011).

[32] Diaz-Ortiz, A., Dosch, H., and Drautz, R. *J.Phys: Condens. Matter* **19**, 406206 (2007).

[33] Diaz-Ortiz, A. and Dosch, H. *Phys. Rev. B* **76**, 012202 (2007).

[34] J.Moreno and J.M.Soler. *Phys. Rev. B* **45**, 13891–13898 (1992).

[35] von Barth, U. and Hedin, L. *Journal of Physics C* **5**(13), 1629–1642 Jul (1972).

[36] Perdew, J. and Wang, Y. *Phys. Rev. B* **33**, 8800 (1986).

[37] Perdew, J., Burke, K., and M.Ernzerhof. *Phys. Rev. Lett.* **77**, 3865 (1996).

[38] T.Naono and Y.Tsunoda. *J.Phys.:Condens.Matter* **16**, 7723 (2004).

[39] Dietermann, F. Master's thesis, University of Stuttgart, (2007).

[40] Kurz, P., Bihlmayer, G., and Blügel, S. *J.Phys.:Condens.Matter* **14**, 6353 (2002).

[41] K.Hummler and M.Fähnle. *Phys. Rev. B* **53**, 3272 (1996).

[42] Köberle, I. Master's thesis, University of Stuttgart, (2007).

[43] Stearns, M. *3d,4d and 5d Elements, Alloys and Compounds*, volume 19a of *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, chapter 1.1.2.4 Spontaneous magnetization, magnetic moments and high-field susceptibility. Springer (1997).

[44] H.Drulis and M.Drulis. *Rare Earth Metals, Elements, Hydrides and Mutual Alloys*, volume 19d1 of *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, chapter 2.1.3.7 Gadolinium. Springer (1991).

[45] Eisen, H., Heinrichs, W., and Witsch, K. *J. Comput. Phys.* **96**, 241–257 (1991).

[46] Livermore, P. W., Jones, C. A., and Worland, S. J. *J. Comput. Phys.* **227**, 1209–1224 (2007).

Danksagung

Folgenden Menschen möchte ich für ihre Unterstützung beim Gelingen dieser Arbeit danken:

- Prof. Dr. Manfred Fähnle, dessen Geist, Ohr und Herz stets offen sind.
- Dr. Reinhard Singer, der hervorragende Vorarbeit leistete und mich mit großer Geduld in das Thema einführte.
- Dr. Lothar Schimmele, mit dem fachliche und politische Diskussionen auf einem soliden Fundament ausgetragen werden konnten.
- Dr. Sergej Subkow, für seinen unerschütterlichem Optimismus und trockenen Humor.