Diplomarbeit

Ab–Initio Calculations of Phonon Contributions on the Phase Stability of Metamagnetic FeRh for Heat–Assisted Magnetic Recording (HAMR)

Institut für Festkörperphysik (IFP)
Technische Universität Wien

Unter der Anleitung von:
Priv.Doz. Dipl.–Ing. Dr.techn. Dieter Suess

Florian Hofer, BSc
Alliogasse 21/ 3/ 8
1150 Wien

Februar, 2018

(Dieter Suess) (Florian Hofer)

Für Fabian.

Contents

Abstract

This thesis presents calculations of temperature driven lattice vibrations and their impact on the metamagnetic phase transition of ordered FeRh in the cubic $B2$ structure. Computational data has been derived from first–principles on basis of Density Functional Theory (DFT) using the Vienna *ab-initio* Simulation Package (VASP). The properties of the lattice vibrations have been found by phonon calculations within a supercell approach and in the limits of a harmonic approximation, which are implemented in the *python*–based programme `phono.py`. The antiferromagnetic ground state (AFM II) of bulk FeRh has been compared to the competing ferromagnetic state (FM) with a very accurately relaxed primitive cell. Furthermore, two different GGA functionals (PBE and RPBE) have been applied and show significant differences in the Fe magnetic moments and the atomic volume concerning the AFM II state. For PBE, where the Fe magnetic moments and the cell volume are smaller compared to RPBE, the phonon band structure shows imaginary frequencies pointing to a lattice instability in the AFM phase. Calculations using RPBE, however, reveal stable modes at all high symmetry points of the reciprocal space and give a clear hint that the magnetic properties and the cell volume of FeRh can play a decisive role in order to stabilise the AFM II state. Thermodynamic properties of the vibrational modes, as the internal energy E, free energy F, entropy S and heat capacity C_V have been evaluated as well. Comparing the free energy of the two magnetic phases demonstrates that the phase transition cannot be explained by only taking lattice vibrations into account. The experimental observation of a large entropy difference at the expected transition temperature $T_\mathrm{M} \sim 350\,\mathrm{K}$, however, can be confirmed with $15.3\,\mathrm{J/(K\,kg)}$ [$26.5\,\mathrm{J/(K\,kg)}$] for PBE [RPBE]. Prior to a detailed presentation of the calculated results a comprehensive summary on the theory and thermodynamics of phase transitions, the basic principles of DFT, as well as the description and concepts of phonon calculations is given.

Kurzfassung

Diese Arbeit zeigt die Berechnungen von temperaturabhängigen Gitterschwingungen und deren Auswirkungen auf den metamagnetischen Phasenübergang von FeRh in der kubischen $B2$ Struktur. Alle Daten wurden aus sogenannten *first–principles* auf der Grundlage der Dichte Funktional Theorie (DFT) und unter Verwendung des Vienna *ab-initio* Simulation Package (VASP) gewonnen. Die Eigenschaften der Gittervibrationen wurden durch Berechnung der Phononen im Festkörper mithilfe von Simulationen von Superzellen und im Rahmen der harmonischen Näherung analysiert. Letzteres ist im Programm `phono.py`, das auf der Programmiersprache *python* basiert, implementiert. Der anti–ferromagnetische Grundzustand (AFM II) von FeRh wurde mit dem energetisch konkurrierenden ferromagnetischen Zustand (FM) mit einer sehr genau relaxierten primitiven Zelle verglichen. Darüber hinaus wurden zwei unterschiedliche GGA Funktionale (PBE und RPBE) angewandt, die im AFM II Zustand signifikante Unterschiede in den lokalen magnetischen Momenten der Fe Ionen und dem Atomvolumen zeigen. Bei der Anwendung des PBE Funktionals, bei dem die magnetischen Momente und das Zellvolumen im Vergleich zu RPBE kleiner sind, sind in der Bandstruktur der Phononen imaginäre Frequenzen zu finden, die auf eine Instabilität des Gitters in der AFM Phase hinweisen. Berechnungen mit RPBE zeigen jedoch stets stabile Moden an allen Symmetriepunkten des reziroken Raums und geben einen entscheidenden Hinweis auf die Möglichkeit den AFM II–Zustand von FeRh mithilfe seiner magnetischen Eigenschaften und des Zellvolumens zu stabilisieren. Die thermodynamischen Eigenschaften der Phononen, wie die innere Energie E, freie Energie F, Entropie S und Wärmekapazität C_V wurden ebenso berechnet. Vergleicht man die freie Energie der beiden magnetischen Phasen, so zeigt sich, dass der Phasenübergang nicht allein durch die Berücksichtigung der Gittervibrationen erklärt werden kann. Die experimentelle Beobachtung eines großen Unterschiedes in der Entropie zwischen beiden Phasen bei der erwarteten Übergangstemperatur $T_\mathrm{M} \sim 350\,\mathrm{K}$ kann jedoch mit den Werten $15.3\,\mathrm{J/(K\,kg)}$ [$26.5\,\mathrm{J/(K\,kg)}$] für PBE [RPBE] bestätigt werden. Bevor die berechneten Ergebnisse und Methoden im Detail gezeigt werden, findet man im ersten Teil dieser Arbeit des Weiteren eine weitreichende Zusammenfassung zur Theorie und Thermodynamik von Phasenübergängen, den Prinzipien zu DFT, sowie den Konzepten zur Beschreibung und Berechnung von Phononen.

Chapter 1

Phase Transitions

The fundamental laws of nature, i.e. Maxwell and many–body Schrödinger equations, ideally show significant features of symmetry. They are both *invariant* under translation of space and time, rotations and reflections (Euclidian plane isometry). The states in real condensed matter, however, do not reveal the full symmetry of these laws. A solid for example is only invariant to discrete translations, as well as rotations of a point group. This is the reason why chemical substances (matter) can be found in different *phases* which differ in their symmetry, hence show different (thermal, mechanical, electric, magnetic, ...) properties.
A *phase* is a homogeneous region of a macroscopic system with a characteristic arrangement (crystallographic, electric, magnetic, ...) derived from a given chemical composition and dependent on a set of outer thermodynamic variables (e.g. temperature T, pressure p, magn. field $\vec{\boldsymbol{H}}$, elec. field $\vec{\boldsymbol{E}}$, ...). In other words, it describes one of the many different possible forms of a state of matter in which the macroscopic physical properties of the substance are uniform and set by outer thermodynamic conditions. A "state" in this sense is represented by a macroscopic observable called *order parameter* that for each phase changes to entirely different values, for example, density $\rho(\vec{\boldsymbol{r}})$ (solid, liquid, gas), magnetisation $\vec{\boldsymbol{M}}$ (ferromagnet, anti-ferromagnet, paramagnet), and so on. When changing these outer conditions to certain values (critical limit) a transformation from one phase to another will occur, which is called a *phase transition*. Such transitions always follow the principle of a minimised thermodynamic potential which describes the system accordingly, e.g. Helmholtz free energy for magnetic systems $F = F(T, \vec{\boldsymbol{M}}, N) \to \min \to \mathrm{d}F = 0$.

One of the first to present a qualitative description of phase transitions in real gases was J.D. van der Waals in the publication of his doctoral thesis (1873). P. Weiss developed a model for the phase transition of a ferromagnet (1907) even without the techniques of quantum mechanics. A phenomenological approach of continuous (*second order*) phase transitions, see Sec.1.2.2, has been derived by L.D. Landau (1937), which also provided as a basis of describing superconductivity without its microscopic mechanism. All of those theories are accounted to the *classical theories* of phase transition. The first non-trivial model of a ferromagnet regarding microscopic interactions is named after the German physicist E. Ising (1925: Ising model). The Hamiltonian is based on localised magnetic moments at each lattice site where the strength of the interaction (overlap of the individual wavefunctions) between two sites is described by an *exchange integral* J_{ij} and can usually be restricted to the nearest–neighbours. The spin operator is represented by a classical one–dimensional vector and has only two anti-parallel states, $S_i = \pm 1$, which symbolises *spin up* $(+1, \uparrow)$ and *spin down* $(-1, \downarrow)$. Conventionally, the spin vector is therefore chosen along the z-axis.

$$\vec{\boldsymbol{S}}_i = S_i \hat{\boldsymbol{e}}_{\boldsymbol{z}}$$

Note that the spin vector is indeed in a *classical* representation, because commutation relations between spins – one of the most important features of quantum physics – are not considered in

this model. Ising's task was to prove whether a spontaneous ordering of the spins as it manifests in a ferromagnet could be described via the microscopic interaction $J_{(ij)}$. In his dissertation he was able to solve only the one–dimensional case and showed that a phase transition to the ordered state could only occur at $T = 0$. Hence, long range order is destroyed at any finite temperature which is a result opposed to the findings of P. Weiss. It was later when R. Peierls (1936) could prove the existence of phase transitions for higher dimensions $D \geq 2$ and L. Onsager (1948) presented an analytic expression of the free energy and the magnetisation for $D = 2$,

$$M(T) = \begin{cases} \left[1 - \sinh^{-4}\left(\frac{2J}{k_\mathrm{B}}\right)\right]^{\beta} & , T < T_\mathrm{C} \\ 0 & , T \geq T_\mathrm{C} \end{cases} , \tag{1.1}$$

where k_B is the Boltzmann constant and $\beta = \frac{1}{8}$ is the *critical exponent* of the order parameter. The Curie temperature can be derived as

$$T_\mathrm{C} = \frac{2}{\ln(1+\sqrt{2})} \frac{J}{k_\mathrm{B}} \sim 2.269185 \frac{J}{k_\mathrm{B}} . \tag{1.2}$$

For the three–dimensional ($D = 3$) Ising model there is still no analytic solution so far, but numerical calculations from Monte–Carlo simulations already offer convincing results.

1.1 Thermodynamics of Phase Transitions

This section is not supposed to claim a complete derivation of the equations appearing in the notion of thermodynamics (Td). It should rather serve as an overview and a reminder of the most important expressions and the terminology needed to describe phase transitions in a macroscopic sense. Starting with a quite general definition the emphasis later on will be laid on the description of magnetic systems.

The *first law of* Td can generally be written as

$$\mathrm{d}U = \delta Q + \delta W , \tag{1.3}$$

where $\mathrm{d}U$ is the differential quantity of the *internal energy*, δQ is the difference of exchanged *heat*, δW is the *work* done on ($\delta W > 0$) or by ($\delta W < 0$) the system with respect to the outer environment. If the system is *open* to its environment, i.e. particles are exchanged, there is a special form of energy usually accounted to δW named particle– or *chemical energy* δW_C which is related to the *chemical potential* μ specifying the potential energy that is needed to add *one* particle to the system. Hence, for α different *types of particles*, where the system contains $N_k\,(k = 1, \ldots, \alpha)$ particles of each type,

$$\begin{aligned} \delta W \quad &\longrightarrow \quad \delta W + \delta W_\mathrm{C} \\ \delta W_\mathrm{C} &= \sum_{k=1}^{\alpha} \mu_k \, \mathrm{d}N_k . \end{aligned} \tag{1.4}$$

The internal energy U is, and even more so, has to be a unique function of the independent *state variables* in phase space, which means that for a circular process

$$\oint \mathrm{d}U = 0\,. \tag{1.5}$$

The internal energy is therefore a *state function* and $\mathrm{d}U$ is a *total differential*, which means that for n state variables $\{x_1, x_2, \ldots, x_n\}$

$$\mathrm{d}U = \sum_j \left(\frac{\partial U}{\partial x_j}\right)_{x_{k\neq j}} \mathrm{d}x_j = \sum_j F_j(x_1, \ldots, x_n)\,\mathrm{d}x_j$$

$$\Longrightarrow \quad \left(\frac{\partial F_j}{\partial x_i}\right)_{x_{m\neq i}} \equiv \quad \frac{\partial^2 U}{\partial x_i\,\partial x_j} = \frac{\partial^2 U}{\partial x_j\,\partial x_i} \quad \equiv \left(\frac{\partial F_i}{\partial x_j}\right)_{x_{m\neq j}}, \forall\, i, j \leq n\,. \tag{1.6}$$

The latter expression in (1.6) is known as *integrability condition* (Schwarz's theorem). Fulfilling these conditions guarantees that energy is conserved and the system is well–defined. Otherwise, one would have a free source of gaining energy taking one path from initial state $A \to B$ and a different one back from final state $B \to A$ (*perpetuum mobile*).

1.1.1 Work — Expression in generalised coordinates

A general expression for the (differential, quasistatic) work δW can be denoted as

$$\delta W = \sum_j F_j\,\mathrm{d}x_j\,, \tag{1.7}$$

which closely resembles similar terms found in the fields of analytical mechanics or electrodynamics. The variables in the set $\mathbf{x} = \{x_1, \ldots, x_j, \ldots, x_n\}$ are *generalised coordinates* and in relation to their conjugated *generalised force* components $\mathbf{F} = \{F_1, \ldots, F_j, \ldots, F_n\}$. Generalised coordinates do not necessarily need to have a dimension of length, nor are generalised forces required to be seen in a strict mechanical sense. The only requisite of a generalised system is for the product $(F_j \cdot q_j)$ to always lead to a *dimension of energy*.

1.1.2 Heat — Entropy as a state function

The first law of TD (1.3) states that the temperature of a system can also be changed without any work done in a sense as it was noted above. The notion of heat is introduced as a form of energy (exchange) and closely related to the definition of *entropy* S. In TD this relation is purely based on a macroscopic phenomenological point of view. A better way to derive a smooth definition for this quantity and to understand its meaning can be done in statistical mechanics within the theory by L. Boltzmann and his famous formula $S = k_\mathrm{B} \ln \Omega_N(E, V)$. In the kinetic gas theory heat is described as the kinetic energy of the individual gas molecules in a chaotic motion. From this form of energy a common non-formal description of entropy is followed as a *measure of disorder* in a many–body system, where the latter actually refers to the number of different microscopic states a system can be in.

The *Clausius theorem* (or *Clausius inequality*) states the important result

$$\oint \frac{\delta Q}{T} \leq 0\,. \tag{1.8}$$

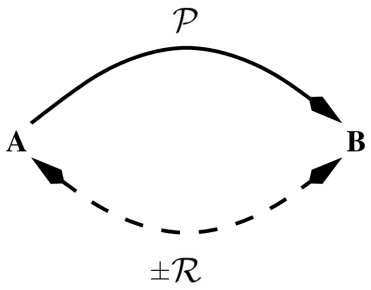


Fig. 1.1: Virtual substitute process $\mathcal{R}$ of an irreversible process $\mathcal{P}$ between an initial state A and a final state B.

It is derived from a system that undergoes a cyclic process and exchanges heat with $N \to \infty$ reservoirs (*thermodynamic limit*) which themselves are coupled with one another by *Carnot engines*. For every set of initial and final states there is a path between them that is *reversible*. A reversible cyclic process just changes the flow into opposite direction which mathematically manifests in a change of sign in (1.8). Since both directions deliver a correct statement in order to avoid a contradiction only the *equality* sign holds for reversible processes,

$$\oint \frac{\delta Q_{\text{rev}}}{T} = 0\,. \tag{1.9}$$

The integral in the equation above from an arbitrary state A to another state $A \to B$ is therefore *path independent* and the quantity of entropy is defined as

$$\begin{aligned} S(B) &= S_0(A) + \int_A^B \frac{\delta Q_{\text{rev}}}{T} \\ \Longrightarrow\ \mathrm{d}S &= \frac{\delta Q_{\text{rev}}}{T}\,, \end{aligned} \tag{1.10}$$

where $\mathrm{d}S$ is an exact (total) differential and therefore a *state function*, similar to $\mathrm{d}U$ in (1.6). Note that S_0 is solely an *additive constant* and a *calculation* of the entropy is only possible if a reversible path between A and B is defined. This is also the case when $A \to B$ is an *irreversible process*. In this case one needs to find a *virtual substitute process* between the same two (initial and final) states which is reversible and delivers the same change in entropy. In the following an irreversible path $\mathcal{P}$ and the according virtual reversible path $\mathcal{R}$ are defined, as it is figuratively shown in Fig.1.1 . Both paths start from a mutual initial state A and end in the same final state B, so that they can also be combined to a cyclic process via

$$A \overset{\mathcal{P}}{\rightsquigarrow} B \underset{-\mathcal{R}}{\leftsquigarrow} A\,.$$

Starting from the Clausius inequality (1.8) one derives

$$\begin{aligned} \int_{A\,(\mathcal{P})}^{B} \frac{\delta Q}{T} \quad + \int_{B\,(-\mathcal{R})}^{A} \frac{\delta Q}{T} \quad &\leq 0 \\ \int_{A\,(\mathcal{P})}^{B} \frac{\delta Q}{T} \quad \leq -\int_{B\,(-\mathcal{R})}^{A} \frac{\delta Q}{T} \quad &= \int_{A\,(\mathcal{R})}^{B} \frac{\delta Q}{T}\,. \end{aligned} \tag{1.11}$$

The expression on the right side is equivalent to the integral in (1.10), hence

$$\begin{aligned} \int_{A\,(\mathcal{R})}^{B} \frac{\delta Q}{T} \quad &\equiv \int_{A}^{B} \frac{\delta Q_{\text{rev}}}{T} \\ \implies \mathrm{d}S \geq \frac{\delta Q}{T}\,. \end{aligned} \tag{1.12}$$

The latter is known as the mathematical expression of the *second law of* Td and a generalisation of the previously derived formula seen in (1.10). Comparison of these two shows again that equality can only be true for reversible processes. This is an important result which states that a process of an isolated system ($T\mathrm{d}S \geq 0$) always seeks the equilibrium of the system where the entropy is at its maximum. In other words, when heat is absorbed (1.12) denotes a change in entropy which for an irreversible process is larger than necessary. For such a process the entropy increases after a complete cycle which after (1.3) results in a loss of internal energy.

Note that δW as well as δQ are *not* total differentials (which is symbolised by the letter *delta* "δ"), so that they do not fulfil the integrability condition in (1.6). However, it is always possible to define an *integrating factor* $\lambda(x_1, \ldots, x_n) \neq 0$, which is a function of state variables that is multiplied to an inexact differential δA to derive a new integrable function f by

$$\begin{aligned} \mathrm{d}f = \lambda(x_1, \ldots, x_n)\,\delta A = \sum_j \lambda\,\tilde{F}_j(x_1, \ldots, x_n)\,\mathrm{d}x_j \\ \implies \quad \left(\frac{\partial(\lambda \cdot \tilde{F}_j)}{\partial x_i}\right)_{x_{m\neq i}} = \left(\frac{\partial(\lambda \cdot \tilde{F}_i)}{\partial x_j}\right)_{x_{m\neq j}}, \forall\, i, j \leq n\,. \end{aligned} \tag{1.13}$$

The definition of $\lambda(x_1, \ldots, x_n)$ though is not unique. In fact, there is an infinite number of possibilities available. Hence, δW and δQ are not unique in phase space and cannot be declared as state functions. A plausible explanation is that per definition *work* and *heat* are a transfer or a flow of energy into or out of a system. This means that they represent a process where an energy (state) variable is changed from an initial to a final value under certain conditions. However, the conditions in order to reach this explicit and unique final state do not necessarily need to be always the same, making it impossible to define a system solely based on those functions.
Opposed to the latter it can be seen from (1.9) and the definition of entropy (1.10) that $\mathrm{d}S$ is indeed an exact (total) differential. The transformation of δQ is thereby achieved by an integrating factor $\lambda(T) = \frac{1}{T}$, which is only dependent on temperature alone and is therefore unique (except for a constant scaling factor).

1.1.3 Free Energies

The results from the previous sections show that the entropy of a system can also be a state function. Therefore, we can generally combine the first and the second law, (1.3) and (1.12),

respectively, in order to derive the fundamental relation of TD,

$$T\,\mathrm{d}S \geq \mathrm{d}U - \delta W - \delta W_{\mathrm{C}}\,. \tag{1.14}$$

Internal Energy (adiabatic) An explicit general expression of the latter can be given for open systems, using (1.4) and (1.7),

$$\mathrm{d}U \leq T\,\mathrm{d}S + \sum_j F_j\,\mathrm{d}x_j + \sum_{k=1}^{\alpha} \mu_k\,\mathrm{d}N_k\,, \tag{1.15}$$

which yields a formula for the differential of the *internal energy*. It can be seen that $\mathrm{d}U$ is defined by the *independent* state variables $(S, \mathbf{x}, \mathbf{N})$. In systems where the inner energy can be described as a function

$$U = U(S, \mathbf{x}, \mathbf{N})\,,$$

the used state variables can be freely varied (or kept constant) in experiments. Even more so, by (1.15) their conjugated dependent variables can be derived directly by partial differentiation of the inner energy with the according entity keeping the others constant (see also Apx.A). Therefore, a state equation similar to $\mathrm{d}U$ is called a *thermodynamic potential*, in analogy to a potential in classical mechanics, and the independent state variables are called *natural variables* of the system. If the system is closed from the environment ($\mathrm{d}N_k = 0$) (1.15) states that for an *adiabatic process* ($\delta Q = 0 \iff \mathrm{d}S = 0$) $\mathrm{d}U$ measures the work the system is able to perform. The *internal energy* is thus the *free energy for an adiabatic process*.

Helmholtz Free Energy (isothermal) The natural variables of the internal energy are not always the best choice to describe a specific system. For example, the entropy is a system parameter that is experimentally quite hard to control. Hence, there is a need to define further thermodynamic potentials which natural variables are more suitable to the experiment. The transformation from one set of (natural) variables to another uses the fact that all entities of energy in TD come in pairs of conjugated parameters. A differential form of those pairs can be easily calculated by $\mathrm{d}(x\,y) = x\,\mathrm{d}y + y\,\mathrm{d}x$, from which one can interchange natural variables solely by addition or subtraction. This is called a *Legendre transformation*. In order to swap entropy S by its conjugate variable temperature T one defines the *Helmholtz free energy* F as

$$F = U - T\,S \tag{1.16}$$

$$\implies \quad \mathrm{d}F = \mathrm{d}U - \mathrm{d}(T\,S) = \underbrace{(\delta Q - T\,\mathrm{d}S)}_{\substack{=0\,,\text{ reversible}\\ <0\,,\text{ irreversible}}} - S\,\mathrm{d}T + \sum_j F_j\,\mathrm{d}x_j + \sum_{k=1}^{\alpha} \mu_k\,\mathrm{d}N_k$$

$$\mathrm{d}F \leq -S\,\mathrm{d}T + \sum_j F_j\,\mathrm{d}x_j + \sum_{k=1}^{\alpha} \mu_k\,\mathrm{d}N_k\,. \tag{1.17}$$

If the process is reversible the first term vanishes with the relation (1.10) and temperature T has indeed been transformed into a natural variable. As a consequence one can state from (1.17) that for reversible *isothermal* ($\mathrm{d}T = 0$) processes the Helmholtz free energy equals the amount of work a system can use or gain.

$$F = F(T, \mathbf{x}, \mathbf{N})$$

For an irreversible isothermal process (1.12) shows that $\mathrm{d}F < (\delta W + \delta W_\mathrm{C})$ which follows that $\mathrm{d}F$ gives the *maximal* amount of work for the system to exchange, which only can be used if the process is reversible. Since temperature is now a control parameter and can be kept constant it is required for the entropy to change during a process, i.e. there will always be an exchange of heat with the outer environment which is unavailable to be used by the system.

Gibbs Free Energy (isothermal – iso"forces") For systems where it is possible to control the temperature T and the forces F_j that define the work of the system, for example pressure p, the according Legendre transformation

$$G = H - T\,S \tag{1.18}$$

$$\implies \quad \mathrm{d}G = \mathrm{d}U - \sum_j \mathrm{d}(F_j\,x_j) - \mathrm{d}(T\,S)$$

$$\mathrm{d}G \leq -S\,\mathrm{d}T - \sum_j x_j\,\mathrm{d}F_j + \sum_{k=1}^{\alpha} \mu_k\,\mathrm{d}N_k \tag{1.19}$$

defines the *Gibbs free energy* G. The introduced new variable $H = U - \sum F_j\,x_j$ is the *enthalpy* of the system. G is thus sometimes also referred to as *free enthalpy*. In order to maintain a constant temperature $\mathrm{d}T = 0$ (isothermal) and constant generalised forces $\mathrm{d}F_j = 0$ (in terms of pressure $\mathrm{d}p = 0$ *isobaric*) heat and work have to constantly be exchanged with the outer environment and can therefore not be used by the system.

$$G = G(T, \mathbf{F}, \mathbf{N})$$

All that is left is the previously in (1.4) defined chemical energy for open systems. This is an exact result, which will be shown in Sec.1.1.6. In summary, it can be stated that for reversible *isothermal* processes with *constant generalised forces*, e.g. pressure, the Gibbs free energy gives the amount of *chemical work* available to open systems.

1.1.4 Homogeneity

In TD the scalability of a system is treated by *extensive* variables and reveals important results which shall be presented in Sec.1.1.6.

Tab. 1.1: Examples for combinations of conjugated intensive and extensive variables.

	intensive	extensive
thermal	T	S
work	F_i	x_i
expansion	p	V
magnetic	$\vec{\boldsymbol{H}}$	$\vec{\boldsymbol{M}}$
chemical	μ	N

Extensive & Intensive Variables Thermodynamic variables can behave differently when the homogeneous phase of the system they determine is scaled by a certain factor λ. *Extensive* variables multiply with λ as well, where *intensive* variables maintain their value everywhere within the system. Free energies are *always* extensive. Energy expressions derived by two conjugated thermodynamic variables thus are always a combination of an extensive and intensive variable. Tab.1.1 shows examples of conjugated variables. The natural variables of the internal energy are $U = U(S, \mathbf{x}, \mathbf{N})$, and therefore changes in the internal energy are only related to changes in extensive variables. As shown in Sec.1.1.3 the extensive variables can be exchanged with their conjugated intensive variables by the according Legendre transformations defining the thermodynamic potentials. If the system is *homogeneous* and all surface effects are neglected (*thermodynamic limes*) the following *homogeneity relations* for the thermodynamic potentials read

$$U(\lambda S, \lambda \mathbf{x}, \lambda \mathbf{N}) = \lambda U(S, \mathbf{x}, \mathbf{N}) , \tag{1.20a}$$

$$F(T, \lambda \mathbf{x}, \lambda \mathbf{N}) = \lambda F(T, \mathbf{x}, \mathbf{N}) , \tag{1.20b}$$

$$G(T, \mathbf{F}, \lambda \mathbf{N}) = \lambda G(T, \mathbf{F}, \mathbf{N}) . \tag{1.20c}$$

1.1.5 Response Functions

In experiments investigating the physics of a material is usually done by changing a natural variable or applying an external field and measure the response of the system. This can be traced over a certain range of an outer variable, so that the behaviour of the system is fitted into a *response function*. Unfortunately, the characteristics of these functions are rarely linear. Linearity can generally only be granted for a defined (small) range of the outer entity Y causing the response Z. For changes within that range it is then possible to write

$$\mathrm{d}Z = \left(\frac{\partial Z}{\partial Y}\right)_{X \neq Y} \mathrm{d}Y , \tag{1.21}$$

where the expression in parenthesis is the according response function and the statement at the bottom of the parenthesis means that *any other* quantity $X \neq Y$, which Z is also dependent on, is kept *constant*.

Heat Capacities & Specific Heat Heating a system exhibits an influence — a response — on its temperature. The reference of how much heat δQ is needed to achieve a temperature change $\mathrm{d}T$ is given by the *heat capacity* C_X of the system

$$C_X \doteq \left(\frac{\delta Q}{\mathrm{d}T}\right)_X . \tag{1.22}$$

Again, X stands for all the natural variables different from temperature that are kept constant. Since C_X is dependent on the total mass M it is an *extensive* variable (for a description of

extensive and intensive variables see Sec.1.1.4). It is practical to define the *specific heat* c_x by dividing (1.22) by the total mass

$$c_x \doteq \left(\frac{\delta Q}{M\,\mathrm{d}T}\right)_x , \tag{1.23}$$

which now is an *intensive* variable.

Consider a closed system where the internal energy is only dependent on temperature T and the *set* of generalised coordinates $\mathbf{x} = \{x_1, \ldots, x_j, \ldots, x_n\}$ that within terms of linearity

$$\begin{gathered} U = U(T, \mathbf{x}) \\ \Longrightarrow \quad \mathrm{d}U = \left(\frac{\partial U}{\partial T}\right)_{\mathbf{x}} \mathrm{d}T + \sum_j \left(\frac{\partial U}{\partial x_j}\right)_{T,\, x_{i\neq j}} \mathrm{d}x_j \,. \end{gathered} \tag{1.24}$$

Using (1.7) in the first law (1.3) the heat differential can be written as

$$\delta Q = \mathrm{d}U - \sum_j F_j \,\mathrm{d}x_j \,. \tag{1.25}$$

Entering (1.24) in the latter yields a *differential form* of the *first law of* TD,

$$\delta Q = \left(\frac{\partial U}{\partial T}\right)_{\mathbf{x}} \mathrm{d}T + \sum_j \left[\left(\frac{\partial U}{\partial x_j}\right)_{T,\, x_{i\neq j}} - F_j\right] \mathrm{d}x_j \,. \tag{1.26}$$

For $\mathbf{x} = \{\text{const}\}$ it is relatively simple to derive an expression for $C_{\mathbf{x}}$ from (1.26) using its definition (1.22). For the derivation of $C_{\mathbf{F}}$, with $\mathbf{F} = \{\text{const}\}$, a transformation to $x_j = x_j(T, F_1, \ldots, F_n)$ is necessary

$$\begin{gathered} F_j = F_j(T, \mathbf{x}) \quad \longrightarrow \quad x_j = x_j(T, \mathbf{F}) \\ \Longrightarrow \quad \mathrm{d}x_j = \left(\frac{\partial x_j}{\partial T}\right)_{\mathbf{F}} \mathrm{d}T + \sum_k \left(\frac{\partial x_j}{\partial F_k}\right)_{T,\, F_{i\neq k}} \mathrm{d}F_k \,, \end{gathered} \tag{1.27}$$

which can be entered into (1.26). The heat capacities for both cases result in

$$\mathbf{x} = \{\text{const}\} \longrightarrow \quad C_{\mathbf{x}} = \left(\frac{\partial U}{\partial T}\right)_{\mathbf{x}} , \tag{1.28a}$$

$$\mathbf{F} = \{\text{const}\} \longrightarrow \quad C_{\mathbf{F}} = \left(\frac{\partial U}{\partial T}\right)_{\mathbf{x}} + \sum_j \left[\left(\frac{\partial U}{\partial x_j}\right)_{T,\, x_{i\neq j}} - F_j\right] \left(\frac{\partial x_j}{\partial T}\right)_{\mathbf{F}} , \tag{1.28b}$$

so that by subtraction a relation between the two expressions can be followed,

$$C_{\mathbf{F}} - C_{\mathbf{x}} = \sum_j \left[\left(\frac{\partial U}{\partial x_j}\right)_{T,\, x_{i\neq j}} - F_j\right] \left(\frac{\partial x_j}{\partial T}\right)_{\mathbf{F}} . \tag{1.29}$$

Bulk Modulus & Compressibility These two entities are related to volume and pressure changes. Measurements are either conducted adiabatically ($\mathrm{d}S = 0$) or isothermally ($\mathrm{d}T = 0$). The adiabatic (isothermal) *bulk modulus* K is defined as

$$K_{S\,(T)} = -V \left(\frac{\partial p}{\partial V} \right)_{S\,(T)} = {\kappa_{S\,(T)}}^{-1} \,, \tag{1.30}$$

and the inverse is called *compressibility* κ. Mechanical stability of the system always requires (without proof) that

$$\kappa_{S\,(T)} \geq 0 \,. \tag{1.31}$$

Thermal Expansion The response of matter to modify its taken volume when temperature is changed is described by a *thermal expansion coefficient* α at constant pressure

$$\alpha_p = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p \,. \tag{1.32}$$

(Magnetic) Susceptibility & Expansion The response function for *magnetic systems* is usually seen as an analogy to the previously mentioned compressibility, where this time a relation between the *exterior magnetic field* $H = |\vec{\boldsymbol{H}}|$ (source) and the *magnetisation* $M = |\vec{\boldsymbol{M}}|$ (effect) is given by

$$\chi_{S\,(T)} = \left(\frac{\partial M}{\partial H} \right)_{S\,(T)} \,, \tag{1.33}$$

and also with the possibility for χ to become *negative* (diamagnetism).
For completeness there is also a magnetic analogy to thermal expansion without any specific name

$$\beta_H = \left(\frac{\partial M}{\partial T} \right)_H \,. \tag{1.34}$$

1.1.6 Gibbs – Duhem Relation

The homogeneity relations (1.20) presented in Sec.1.1.4 hold for *any* scaling factor λ. As an example the expression for the internal energy is used to differentiate both sides of (1.20a) with respect to λ, letting $(\lambda \to 1)$

$$\begin{aligned} U(S, \mathbf{x}, \mathbf{N}) &= \lim_{\lambda \to 1} \left\{ \frac{\partial U}{\partial(\lambda\, S)} \frac{\partial(\lambda S)}{\partial \lambda} + \sum_j \frac{\partial U}{\partial(\lambda\, x_j)} \frac{\partial(\lambda x_j)}{\partial \lambda} + \sum_{k=1}^{\alpha} \frac{\partial U}{\partial(\lambda\, N_k)} \frac{\partial(\lambda\, N_k)}{\partial \lambda} \right\} \\ \implies \quad U &= \left(\frac{\partial U}{\partial S} \right) S + \sum_j \left(\frac{\partial U}{\partial x_j} \right) x_j + \sum_{k=1}^{\alpha} \left(\frac{\partial U}{\partial N_k} \right) N_k \,. \end{aligned} \tag{1.35}$$

Expressions for $F(T, \mathbf{x}, \mathbf{N})$ and $G(T, \mathbf{F}, \mathbf{N})$ can be derived in a similar way using (1.20b) and (1.20c), respectively. Appropriately entering the partial derivatives of (A.1) brings the result

$$U = T\,S + \sum_j F_j\,x_j + \sum_{k=1}^{\alpha} \mu_k\,N_k\,, \tag{1.36a}$$

$$F = \sum_j F_j\,x_j + \sum_{k=1}^{\alpha} \mu_k\,N_k\,, \tag{1.36b}$$

$$G = \sum_{k=1}^{\alpha} \mu_k\,N_k\,. \tag{1.36c}$$

There are three conclusions which can be followed from the equations in (1.36):

1. It can now be directly seen by subtraction that

$$F = U - T\,S \quad \text{, and} \quad G = (U - \textstyle\sum F_j\,x_j) - T\,S = H - T\,S\,.$$

2. As it was stated before in Sec.1.1.3, (1.36c) shows that the free enthalpy indeed can be interpreted as the *total chemical energy* available to the system.

3. From the (total) differential of the internal energy $\mathrm{d}U$, which was previously derived in (1.15), it follows from (1.36a) that

$$\begin{aligned} \mathrm{d}U &= \mathrm{d}(T\,S) + \sum_j \mathrm{d}(F_j\,x_j) + \sum_{k=1}^{\alpha} \mathrm{d}(\mu_k\,N_k) \overset{(1.15)}{=} \\ &= T\,\mathrm{d}S + \sum_j F_j\,\mathrm{d}x_j + \sum_{k=1}^{\alpha} \mu_k\,\mathrm{d}N_k \\ \Longrightarrow \quad & S\,\mathrm{d}T + \sum_j x_j\,\mathrm{d}F_j + \sum_{k=1}^{\alpha} N_k\,\mathrm{d}\mu_k = 0\,. \end{aligned} \tag{1.37}$$

   The latter is called *Gibbs–Duhem relation*, implying that it is *impossible* to define a thermodynamic potential with only independent *intensive* variables.

1.1.7 Equilibrium Conditions

When two or more subsystems are brought into contact with each other there are certain boundary conditions that have to be met in order to describe the equilibrium state for the system in total. They all rely on the fundamental relation (1.14) and only differ by the several possible ways the system interacts with its outer surroundings. Therefore, it is clear that there are also special cases which can be conveniently described using thermodynamic potentials and shall be presented here. Whether the boundaries separating the subsystems are interpreted as walls or just different phases of a substance will not change the outcome of the conditions for equilibrium (see Sec.1.2.3).

Isolated System An isolated system is not able to exchange heat with the surroundings and is characterised by

$$\begin{array}{ccccc} U = \text{const} & , & x_j = \text{const} & , & N_k = \text{const} \\ \Updownarrow & & \Updownarrow & & \Updownarrow \\ \mathrm{d}U = 0 & , & \mathrm{d}x_j = 0 & , & \mathrm{d}N_k = 0 \quad . \end{array} \tag{1.38a}$$

This implies with (1.15) that

$$\boxed{\mathrm{d}S \geq 0 \implies \mathrm{d}S = 0 \text{ (in equilibrium)}} \, . \tag{1.38b}$$

As long as irreversible processes are still possible in this kind of system the entropy can only *increase*. Transition to equilibrium is therefore irreversible and characterised by the entropy reaching its *maximum*.

Closed System in Heat Bath There are two different configurations for closed systems that can exchange heat to keep a constant temperature.

The first one is set up so that there is *no change of work* as in (1.7),

$$\begin{array}{ccccc} T = \text{const} & , & x_j = \text{const} & , & N_k = \text{const} \\ \Updownarrow & & \Updownarrow & & \Updownarrow \\ \mathrm{d}T = 0 & , & \mathrm{d}x_j = 0 & , & \mathrm{d}N_k = 0 \quad , \end{array} \tag{1.39a}$$

which yields with (1.17)

$$\boxed{\mathrm{d}F \leq 0 \implies \mathrm{d}F = 0 \text{ (in equilibrium)}} \, . \tag{1.39b}$$

All irreversible processes that are possible under conditions (1.39a) result in a *decreasing* Helmholtz free energy. In equilibrium F is at its *minimum*.

The second possible isolated system is configured in a way that the *forces are constant*,

$$\begin{array}{ccccc} T = \text{const} & , & F_j = \text{const} & , & N_k = \text{const} \\ \Updownarrow & & \Updownarrow & & \Updownarrow \\ \mathrm{d}T = 0 & , & \mathrm{d}F_j = 0 & , & \mathrm{d}N_k = 0 \quad . \end{array} \tag{1.40a}$$

For irreversible processes under these conditions the Gibbs free energy (free enthalpy),

$$\boxed{\mathrm{d}G \leq 0 \implies \mathrm{d}G = 0 \text{ (in equilibrium)}} \, , \tag{1.40b}$$

is *decreasing* at all times and *minimal* at equilibrium.

1.2 Classifications

A commonly used nomenclature to classify phase transitions was introduced by P. Ehrenfest (1933). His definition is based on the non-analytical behaviour of the free energy (or the potential that describes the system) as a function of one of the thermodynamic control parameters at such a transition. Since those quantities can be expressed as the n^{th} derivative of a free energy, a change in phase can accordingly be categorised as to be of n^{th} *order*.

1.2.1 Ehrenfest Classification

As it was introduced above, when there is a transition from one phase to the other its *order* is defined by the lowest derivative of the free energy which becomes discontinuous at that state. Considering the Helmholtz and Gibbs free energy, F and G respectively, the notion of a *first–order* phase transition manifests when the first derivatives

$$F = U - T \cdot S \quad \longrightarrow \quad S = -\left(\frac{\partial F}{\partial T}\right)_V , \quad p = -\left(\frac{\partial F}{\partial V}\right)_T , \tag{1.41a}$$

$$G = H - T \cdot S \quad \longrightarrow \quad S = -\left(\frac{\partial G}{\partial T}\right)_p , \quad V = \left(\frac{\partial G}{\partial p}\right)_T , \tag{1.41b}$$

show a discontinuity at the critical point. The only assumed generalised coordinate here is the *volume* ($\mathrm{d}x_j \to \mathrm{d}V$) and its conjugated force, which is the *pressure* ($-p$), so that the derivative with respect to volume (∂V) in (1.41a) leads to pressure and vice versa (except for the sign) in the case of Gibbs energy in (1.41b). A finite jump in the entropy is associated with a *latent heat* $\Delta Q = T_{12}(S_2 - S_1)$ keeping the transition temperature T_{12} constant and is typical for a first–order transition.

For *second–order* transitions the second derivatives reveal non-analytic behaviour at the critical points. The first derivatives, however, are smooth functions over their entire space, which implies there is no latent heat in such transitions since $S_1 = S_2 = S$. The terms for $F = F(T, V, \mathbf{N})$ yield

$$\left(\frac{\partial^2 F}{\partial T^2}\right)_V = \frac{\partial}{\partial T}\left(\frac{\partial F}{\partial T}\right)_V = -\left(\frac{\partial S}{\partial T}\right)_V = -\frac{C_V}{T} \quad \leq 0 , \tag{1.42a}$$

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_T = \frac{\partial}{\partial V}\left(\frac{\partial F}{\partial V}\right)_T = -\left(\frac{\partial p}{\partial V}\right)_T = +\frac{1}{V \cdot \kappa_T} \geq 0 , \tag{1.42b}$$

where the expressions in (1.41a) have been used.
For the Gibbs free energy $G = G(T, p, \mathbf{N})$ one derives

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_p = \frac{\partial}{\partial T}\left(\frac{\partial G}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial T}\right)_p \quad = -\frac{C_p}{T} \leq 0 , \tag{1.43a}$$

$$\left(\frac{\partial^2 G}{\partial p^2}\right)_T = \frac{\partial}{\partial p}\left(\frac{\partial G}{\partial p}\right)_T = \left(\frac{\partial V}{\partial p}\right)_T \quad = -V \cdot \kappa_T \leq 0 , \tag{1.43b}$$

by applying (1.41b).
As it is shown in Apx.A the second derivatives lead to response functions, such as the heat capacity at constant volume and pressure, respectively,

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V , \tag{1.44a}$$

$$C_p = \left(\frac{\partial H}{\partial T}\right)_p = T\left(\frac{\partial S}{\partial T}\right)_p . \tag{1.44b}$$

(1.42b) and (1.43b) both use the expression (1.30) for the compressibility κ_T at constant temperature. Note that in a *magnetic system* the magnetic susceptibility χ_T (1.33) usually replaces κ_T,

$$M = -\left(\frac{\partial F}{\partial H}\right)_T \quad \longrightarrow \quad \frac{\partial^2 F}{\partial H^2} = -\left(\frac{\partial M}{\partial H}\right)_T = -\chi_T \,, \tag{1.45a}$$

$$H = \quad \left(\frac{\partial G}{\partial M}\right)_T \quad \longrightarrow \quad \frac{\partial^2 G}{\partial M^2} = \quad \left(\frac{\partial H}{\partial M}\right)_T = {\chi_T}^{-1} \,, \tag{1.45b}$$

with the absolute value of the exterior *magnetic field* $H = |\vec{\boldsymbol{H}}|$ and the *magnetisation* $M = |\vec{\boldsymbol{M}}|$. Due to reasons of stability C_X and κ_X always need to have a positive value, so that the second derivatives also show that the Helmholtz free energy (Gibbs free energy) is a *concave* (*concave*) function of T, and *convex* (*concave*) as a function of V. The susceptibility χ is usually positive, however, can also take negative values as it is the case in diamagnetic materials.

The change of matter from solid to liquid holds as a good example of a *first–order* transition, since the density ρ (or better the difference of the densities of the two phases $[\rho_{\mathrm{ph1}} - \rho_{\mathrm{ph2}}]$) appears to be the order parameter due to its vast change at the critical point and is solely the inverse of the volume $\rho = V^{-1}$ in (1.41b).

Even though after the scheme by Ehrenfest phase transitions of higher order ($n > 2$) can exist per definition, the difference of coexisting phases becomes physically less significant up to the point where the definition of a phase is not sensible any more. Basically, only first and second–order transitions are of practical interest, although the number of examples for second–order transitions in the sense of Ehrenfest are quite limited.

1.2.2 Beyond Ehrenfest Classification – Modern Modifications

The objections to the applicability of the Ehrenfest classification are all due to experimental observations. For example, the existence of *metastable phases*, which typically exist at first–order transitions can be observed. This implies that for each phase the according thermodynamic potential is an individual smooth function which can be continued into another phase. However, the discontinuity in the first derivative of the total energy function is indeed apparent, because at the transition temperature, where the individual free energies intersect, only the phase with the smaller free energy is stable leaving a "kink" in the overall curve. In many systems with second–order transitions measurements show that the critical thermodynamic entities are rather *singularities* than finite discontinuations. Note that the notion of individual free energy functions for each phase that intersect are not applicable for higher–order transitions, since first derivative has to be a smooth function. Thus, there is no real analogy between first– and higher–order transitions, which contradicts the concept of Ehrenfest. Furthermore, as stated before in Sec.1.2.1, to classify phase transitions up to arbitrarily high order is not practical.

Therefore, the classification of phase transitions is now split into only *two* categories. Based on the behaviour of the *first derivative* of the free energy as a function of its (energy–)conjugated natural variable it is now common to use the definition of a *discontinuous* and *continuous* phase transition.

A discontinuous phase transition is the same as the previously discussed first–order phase transition after Ehrenfest. Due to the sudden change in entropy there is a latent heat $\Delta Q = T_{12}\Delta S$

serving as a qualitative observable for the discontinuity, but does *not* hold as a *matter constant*, i.e. the transition temperature T_{12}, where two phases are in coexistence, is adjusted by other variables of the system $T_{12} = T_{12}(x_j, Fj)$. Also for most systems it is observed that ΔS is decreasing for rising T_{12} until one reaches a *critical temperature* T_c, where the phase boundary (discontinuity) disappears and the first derivatives are smooth again

$$\lim_{T_{12}\to T_\mathrm{c}} \Delta S = 0\,.$$

When there are singularities found in the second derivatives of the free energy at T_c one thus denotes it as a continuous phase transition. Continuous transitions often show anomalous phenomena appearing around a critical point of the system, due to the phases becoming indistinguishable. This anomalous behaviour is treated and explained by the *theory of critical phenomena*.

Order Parameter Beside the discontinuity in the according derivative there is another tool to identify a phase transition. The notion of an *order parameter* is a macroscopic variable that can be properly defined in each phase, but signals the breakdown of an ordered state by an abrupt change of its own value. The best known example of an order parameter is the magnetisation of a ferromagnet which for $(T < T_\mathrm{C})$ has a finite value due to the spontaneous order of the spins without an exterior field, and for $(T \geq T_\mathrm{C})$ indicates a disordered state by abruptly cancelling to zero. Other examples are the *gap* Δ of a superconductor or the difference of the densities, $\Delta\rho = \rho_\mathrm{f} - \rho_\mathrm{g}$, in a fluid – gas phase.

1.2.3 Coexistence Conditions for Closed Systems

In Sec.1.1.7 equilibrium conditions have been derived for bringing two (or more) subsystems into contact under certain conditions. These subsystems shall now be identified by the individually distinguishable *phase* $(\nu = 1, \ldots, \pi)$ the substance is in. Each phase ν consists of α different *type of particles*, where $N_\nu^{(k)}$ $(k = 1, \ldots, \alpha)$ is the *number of particles* of type k in phase ν. Hence, the total number N in a closed system reads

$$N = \sum_{\nu=1}^{\pi} N_\nu = \sum_{\nu=1}^{\pi}\left(\sum_{k=1}^{\alpha} N_\nu^{(k)}\right) \equiv \sum_{k=1}^{\alpha}\sum_{\nu=1}^{\pi} N_\nu^{(k)} = \sum_{k=1}^{\alpha} N_k$$
$$\Longrightarrow \quad N_\nu = \sum_{k=1}^{\alpha} N_\nu^{(k)} \quad , \quad N_k = \sum_{\nu=1}^{\pi} N_\nu^{(k)}\,,$$

with N_ν and N_k being the *total number of particles* in phase ν and of type k, respectively. Note that in a closed system $\mathrm{d}N_\nu \neq 0$ is possible whereas N_k cannot change. Any other *extensive* natural variable is also additive,

$$x_j = \sum_{\nu=1}^{\pi} x_{j\nu}\,,$$

and since this is also true for the free energies one can write

$$F(T, \mathbf{x}, \mathbf{N}) = \sum_{\nu=1}^{\pi} F_\nu(T, \mathbf{x}_\nu, \mathbf{N}_\nu) \quad , \quad G(T, \mathbf{F}, \mathbf{N}) = \sum_{\nu=1}^{\pi} G_\nu(T, \mathbf{F}, \mathbf{N}_\nu) \tag{1.46}$$

From (1.39a) we derive the following *boundary conditions*

$$\mathrm{d}x_j = \sum_{\nu=1}^{\pi} \mathrm{d}x_{j\nu} = 0\,, \tag{1.47a}$$

$$\mathrm{d}N_k = \sum_{\nu=1}^{\pi} \mathrm{d}N_\nu^{(k)} = 0\,, \tag{1.47b}$$

where the latter ($\mathrm{d}N_k = 0$) is also valid for (1.40a).

In order to find the state where all the phases are in equilibrium under the boundary conditions (1.47) based on (1.39b) one needs to minimise the Helmholtz free energy by

$$\mathrm{d}F = \sum_{\nu=1}^{\pi} \left\{ \sum_j \left(\frac{\partial F_\nu}{\partial x_{j\nu}} \right)_{\mathbf{N}_\nu} \mathrm{d}x_{j\nu} + \sum_{k=1}^{\alpha} \left(\frac{\partial F_\nu}{\partial N_\nu^{(k)}} \right)_{\mathbf{x}_\nu,\, N_\nu^{(l\neq k)}} \mathrm{d}N_\nu^{(k)} \right\} = 0\,. \tag{1.48}$$

If all $\mathbf{x}_\nu$, $\mathbf{N}_\nu$ would be independent from one another, the equilibrium state could be easily calculated by letting the coefficients in (1.48) vanish. Unfortunately, due to the boundary conditions this is not the case and one has to use the mathematical optimisation method of *Lagrange Multipliers*, where without changing anything one can state

$$(1.47) \implies \begin{aligned} &\lambda_j \cdot \sum_{\nu=1}^{\pi} \mathrm{d}x_{j\nu} = 0\,, && (1.49a)\\ &\lambda_k \cdot \sum_{\nu=1}^{\pi} \mathrm{d}N_\nu^{(k)} = 0 && (1.49b) \end{aligned}$$

$\{\lambda_j\}$ and $\{\lambda_k\}$ are real numbers not defined yet and called *Lagrange parameter*. Using (1.49) in (1.48) reads

$$\sum_{\nu=1}^{\pi} \left\{ \sum_j \left[\left(\frac{\partial F_\nu}{\partial x_{j\nu}} \right)_{\mathbf{N}_\nu} - \lambda_j \right] \mathrm{d}x_{j\nu} + \sum_{k=1}^{\alpha} \left[\left(\frac{\partial F_\nu}{\partial N_\nu^{(k)}} \right)_{\mathbf{x}_\nu,\, N_\nu^{(l\neq k)}} - \lambda_k \right] \mathrm{d}N_\nu^{(k)} \right\} = 0\,. \tag{1.50}$$

Since there is only *one* boundary condition per natural parameter it is possible to split them into $(\pi - 1)$ *independent* variables and 1 variable *dependent* on all the others. Hence, the Lagrange parameters can be arbitrarily set to

$$F_{j1} = \left(\frac{\partial F_1}{\partial x_{j1}} \right)_{\mathbf{N}_1} = \lambda_j \quad , \quad \mu_1^{(k)} = \left(\frac{\partial F_1}{\partial N_1^{(k)}} \right)_{\mathbf{x}_1,\, N_1^{(l\neq k)}} = \lambda_k\,,$$

so that for $(\nu = 1)$ all expressions in squared brackets in (1.50) are zero and only independent variables in the terms of the sums are left. However, this means due to independence that all terms in square brackets have to be zero, and one can follow the final results

$$F_{j\nu} = F_j, \quad \forall \nu\,, \tag{1.51a}$$

$$\begin{array}{ccccccc} \mu_1^{(1)} & = \cdots & = \mu_\nu^{(1)} & = \cdots & & = \mu_\pi^{(1)} \\ \vdots & & \vdots & & \vdots & & \\ \mu_1^{(\alpha)} & = \cdots & = \mu_\nu^{(\alpha)} & = \cdots & & = \mu_\pi^{(\alpha)} \end{array}\,, \tag{1.51b}$$

which implies that for a system in equilibrium an *intensive* variable not natural to the system has the *same value* for *all* phases ν. This can also be shown via statistical mechanics by optimising the *inverse of the partition sum*, which gives the most probable configuration (state) a system can be found in equilibrium.

1.2.4 Gibbs' Phase Rule

Dependent on the free energy describing the system the chemical potentials in (1.51b) show the following dependencies,

$$F(T,\mathbf{x},\mathbf{N}) \longrightarrow \mu_\nu^{(k)}\,(T,\xi,\mathbf{c}) \quad , \quad G(T,\mathbf{F},\mathbf{N}) \longrightarrow \mu_\nu^{(k)}\,(T,\mathbf{F},\mathbf{c})\,, \tag{1.52}$$

where $\mu_\nu^{(k)}$ as an intensive quantity can only be dependent on intensive variables. Hence, it is necessary to use the intensive *concentrations*

$$c_\nu^{(k)} \doteq \frac{N_\nu^{(k)}}{N_\nu} \implies \sum_{k=1}^{\alpha} c_\nu^{(k)} = c_\nu = 1\,,$$

and *molar generalised coordinates* defined as

$$\xi_{j\nu} \doteq \frac{x_{j\nu}}{N_\nu}\,, \quad j = 1,\ldots,n$$

rather than the extensive number of particles $N_\nu^{(k)}$ and generalised coordinates $x_{j\nu}$, respectively.

The chemical potential is therefore dependent on $(1+n+\alpha\pi)$ (*intensive!*) variables which are not independent from each other due to the $[\alpha\cdot(\pi-1)]$ equations from (1.51b) and the π side conditions of concentrations ($c_\nu = 1$). In total one therefore derives f *degrees of freedom* for a non-reactive multi-component heterogeneous system in thermal equilibrium,

$$\begin{aligned} f &= (1+n+\alpha\cdot\pi) - \alpha\cdot(\pi-1) - \pi \\ \implies &\quad \boxed{f = (n+1) + \alpha - \pi} \quad . \\ &\quad (f \geq 0) \end{aligned} \tag{1.53}$$

The latter is known as *Gibbs' phase rule* and shows that the number of possible phases has an upper limit

$$\pi \leq (n+1) + \alpha\,,$$

since the number of degrees of freedom can never be negative.

1.2.5 Clausius–Clapeyron Equation

After deriving general conditions for the coexistence of π phases we shall now assume two arbitrary phases ($\nu = 1,2$) separated by a *phase line* in some parameter plane which defines the thermodynamic system with only one component ($\alpha = 1$). The two variables (T,X) that generate this plane shall be identified as temperature T and a general entity X. A point on the phase

line is defined after (1.51b) where the chemical potential of both phases are equal. Considering two points on the line that are separated by a small difference $(\Delta T, \Delta X)$ brings

$$\begin{aligned} \mu_1(T, X) &= \mu_2(T, X) \\ \mu_1(T + \Delta T, X + \Delta X) &= \mu_2(T + \Delta T, X + \Delta X) \end{aligned}$$

$$\implies \quad \mu_1(T + \Delta T, X + \Delta X) - \mu_1(T, X) = \mu_2(T + \Delta T, X + \Delta X) - \mu_2(T, X)\,,$$

where the latter has been derived by subtracting the second from the first equation. Each side is expanded into a Taylor series of *linear* order in $(\Delta T, \Delta X)$ and then solved for the ratio of the two differences in the limes $(\Delta T, \Delta X) \to 0$,

$$\left(\frac{\partial \mu_1}{\partial T}\right)_X \Delta T + \left(\frac{\partial \mu_1}{\partial X}\right)_T \Delta X \approx \left(\frac{\partial \mu_2}{\partial T}\right)_X \Delta T + \left(\frac{\partial \mu_2}{\partial X}\right)_T \Delta X$$

$$\implies \quad \frac{\mathrm{d}X}{\mathrm{d}T} = \lim_{\Delta \to 0} \frac{\Delta X}{\Delta T} = \frac{\left(\frac{\partial \mu_1}{\partial T}\right)_X - \left(\frac{\partial \mu_2}{\partial T}\right)_X}{\left(\frac{\partial \mu_2}{\partial X}\right)_T - \left(\frac{\partial \mu_1}{\partial X}\right)_T} \quad . \tag{1.54}$$

(1.54) relates the slope at a certain point of a phase line where two phases are in thermodynamic equilibrium to the thermodynamic properties of those phases. This is the general form of the *Clausius–Clapeyron equation*. The evaluation of the partial differentials shall be shown by the following example.

Magnetic System For such a system the Gibbs free energy can be written as $G_\nu = G_\nu(T, \vec{\boldsymbol{H}}, N_\nu)$. Note that usually $N_1 \neq N_2$, so that by (1.36c)

$$\mu_\nu(T, \vec{\boldsymbol{H}}) = \frac{1}{N_\nu} G_\nu(T, \vec{\boldsymbol{H}}, N_\nu)\,, \tag{1.55}$$

which by using (1.19) and (1.40b) yields

$$\mathrm{d}\mu_\nu = -s_\nu \mathrm{d}T - \vec{\boldsymbol{m}}_\nu \mathrm{d}\vec{\boldsymbol{H}}\,. \tag{1.56}$$

Note that $s_\nu = \frac{S_\nu}{N_\nu}$ and $\vec{\boldsymbol{m}}_\nu = \frac{\vec{\boldsymbol{M}}_\nu}{N_\nu}$ are divided by the according number of particles and thus are the intensive entropy and magnetisation, respectively, per particle in a given phase. It is explicitly assumed that $s_1 \neq s_2$, as well as $m_1 \neq m_2$ with $m_\nu = |\vec{\boldsymbol{m}}_\nu|$. This implies that (1.54) can only be valid for *first–order* phase transitions.
Finally, the *Clausius–Clapeyron equation for magnetic systems* reads

$$\frac{\mathrm{d}H}{\mathrm{d}T} = \frac{s_2 - s_1}{(m_1 - m_2)}\,. \tag{1.57}$$

1.3 Mean–Field Models

For the studies of phase transitions the task is to derive models that are appropriate to the system and solve them by means of statistical (quantum-)mechanics. Unfortunately, there are

only a few that have been solved exactly. In the case of an N-spin Ising model, for example, the number of states equals to 2^N (due to the two possible spin configurations per state). If N is increased just by one order of magnitude, the number of states scales significantly by several orders making it impossible to calculate the partition function in a usual manner. Hence, with the exception of some ingenious solutions one is generally restricted to approximation methods which offer a possibility to understand the essential features of the physical phenomena under investigation. One of the most common and widely used is the *mean–field theory*, where basically a many–body interaction is averaged and represented by a mean–field that interacts only with a single variable reducing the number of degrees of freedom drastically. It turns out that this description is equivalent to ignoring fluctuations from the average value of the observable.
The starting point for a mean–field approximation of a (microscopic) Hamiltonian $\mathcal{H}$ is the *Bogoliubov inequality*

$$\begin{gathered} F \leq \Phi \,, \\ \text{where } \Phi = F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle \ . \end{gathered} \tag{1.58}$$

F is the *true* free energy of the system, $\mathcal{H}_0$ a *trial* Hamiltonian with the corresponding free energy F_0. The average $\langle \ldots \rangle$ is taken in the ensemble defined by $\mathcal{H}_0$. If the trial Hamiltonian is dependent on a parameter H_0 the mean–field free energy is then derived by minimising Φ with respect to H_0,

$$F_{\mathrm{MF}} = \min_{H_0} \Phi \,. \tag{1.59}$$

The latter is analogous to the variational principle in quantum mechanics, where (1.58) guarantees that the mean–field free energy cannot be smaller than the true free energy. For $\mathcal{H}_0$ usually a free Hamiltonian (no interaction between individual particles) is considered.

In magnetic systems the *Heisenberg Hamiltonian* describes the interaction between neighbouring spins and it shall now serve as a basis for the following derivations of this section. In general the latter reads

$$\mathcal{H} = -\sum_{i<j} J_{ij} \vec{\boldsymbol{S}}_i \cdot \vec{\boldsymbol{S}}_j - g_{\mathrm{J}}\, \mu_{\mathrm{B}}\, \vec{\boldsymbol{H}} \sum_i \vec{\boldsymbol{S}}_i \,. \tag{1.60}$$

The Heisenberg model treats the spins as microscopic quantum mechanical observables. The first term stands for the cooperative behaviour of pairwise interaction between two spins on different lattice sites i and $j \neq i$. The coupling constants J_{ij} are the respective *exchange integrals*, coming from the quantum mechanical *exchange interaction* of indistinguishable particles. They are *positive* for a *ferromagnetic* interaction and *negative* in case of an *anti–ferromagnetic* one. If only *nearest–neighbour* sites are considered — denoted as indices in angle brackets — the first term in (1.60) reads

$$-J \sum_{\langle k,\delta_k \rangle} \vec{\boldsymbol{S}}_k \cdot \vec{\boldsymbol{S}}_{k+\delta_k} \,,$$

where δ_k is the counter that symbolically points to all the closest neighbouring spins of site k. The second term in (1.60) with the Landé factor (g_{J}) and the Bohr magneton ($\mu_{\mathrm{B}} = \frac{e\hbar}{2m}$) is the *Zeeman interaction* of the spins with an external field $|\vec{\boldsymbol{H}}| > 0$. For $T = 0\,\mathrm{K}$ all spins are aligned to that field and form the ground state of the system.

1.3.1 Ising Model

The Ising model restricts the spin vector operator to only one dimension, most conveniently to the z-axis

$$\vec{\boldsymbol{S}}_i = S_i \hat{\boldsymbol{e}}_{\boldsymbol{z}} \,,$$

and defines *two* antiparallel states it can occupy as *spin up* ($S_i = +1$) and *spin down* ($S_i = -1$). The resulting Hamiltonian using (1.60) reads

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i \, S_j - g_{\mathrm{J}} \, \mu_{\mathrm{B}} \, H \sum_i S_i \,, \tag{1.61}$$

where the external field $\vec{\boldsymbol{H}} = (0, 0, H)$ is pointing in z-direction.

It is desired that the interaction between the spins is averaged in a *mean–field* H_0. This can be achieved by replacing the $\aleph$ nearest–neighbour spins in (1.61) by an average value $S_j \to \langle S \rangle$, so that one derives a non-interacting (trial) Hamiltonian in the form of

$$\mathcal{H}_0 = -\left(H_0 + g_{\mathrm{J}} \, \mu_{\mathrm{B}} \, H\right) \sum_i S_i \,, \tag{1.62}$$

$$\text{where } H_0 = J \, \aleph \, \langle S \rangle \,.$$

The formal derivation of the expression for H_0 can be reviewed in Apx.B . It will be shown in the following that this is equivalent to neglecting the fluctuations (deviations) δS from the thermal average $\langle S \rangle$ of the spins.
The Ising spin variables are separated into

$$S = \langle S \rangle + \delta S$$

$$\implies \quad S_i \, S_j = S_i \, \langle S_j \rangle + \langle S_i \rangle S_j - \langle S_i \rangle \, \langle S_j \rangle + \left[S_i - \langle S_i \rangle\right] \left[S_j - \langle S_j \rangle\right] \,. \tag{1.63}$$

The latter expression is entered into the first term of (1.61) ignoring the second–order terms of δS, which yields

$$\mathcal{H} = -J \, \langle S \rangle \overbrace{\sum_{\langle i,j \rangle} (S_i + S_j)}^{\equiv \aleph \sum_i S_i} + J \, \langle S \rangle^2 \underbrace{\sum_{\langle i,j \rangle} 1}_{= \aleph \frac{N}{2}} - J \overbrace{\sum_{\langle i,j \rangle} \left[S_i - \langle S_i \rangle\right] \left[S_j - \langle S_j \rangle\right]}^{\to 0 \text{ (neglected)}} - g_{\mathrm{J}} \, \mu_{\mathrm{B}} \, H \sum_i S_i$$

$$\implies \mathcal{H} \approx \mathcal{H}_0 + \frac{M_0 \, H_0}{2} \,, \tag{1.64}$$

which is the same expression as in (1.62), except for a constant magnetic energy term with the average magnetisation $M_0 = N \, \langle S \rangle$. Note that $\langle S_i \rangle = \langle S_j \rangle = \langle S \rangle$ has been assumed due to translation invariance (space translation symmetry). The average results from the ensemble

defined by $\mathcal{H}_0$ from (1.62), which can be denoted as

$$\langle S\rangle = \langle S_i\rangle = \frac{\sum\limits_{S_i=\pm1} S_i\, \mathrm{e}^{\beta\,(H_0+\dots)\,S_i}}{\sum\limits_{S_i=\pm1} \mathrm{e}^{\beta\,(H_0+\dots)\,S_i}} = \frac{\sinh\left(\beta\,(H_0+\dots)\right)}{\cosh\left(\beta\,(H_0+\dots)\right)}$$

$$\Longrightarrow \qquad \langle S\rangle = \tanh\Big(\beta\,(J\,\aleph\,\langle S\rangle + \dots)\Big)\,. \tag{1.65}$$

Since $\beta = (k_\mathrm{B}\,T)^{-1}$ is the *inverse temperature* the latter yields an equation that can be solved *self–consistently* in order to determine the average magnetisation $\langle S\rangle$ dependent on the external parameters of temperature T and magnetic field H. As indicated before it is not possible to solve (1.65) explicitly, but the qualitative behaviour of the latter can be understood by plotting the expression on the left and the right side of the equation individually as a function of $\langle S\rangle$ (see Fig.1.2). The intersections of the curves then correspond to the values for $\langle S\rangle$ that solve (1.65). If for simplicity $H = 0$, it can be seen that a solution different from zero is only possible when the slope of $\tanh(\beta\,J\,\aleph\,\langle S\rangle)$ at the origin is greater than 1, i.e. $J\,\aleph\,\beta > 1$. The according value indeed minimises the mean–field free energy, which means that the system shows a stable *ferromagnetic* phase with a finite *spontaneous magnetisation*. With rising temperature the slope at the origin becomes smaller until $J\,\aleph\,\beta < 1$ and the two curves do not intersect anymore (except for the origin), representing the *paramagnetic* phase. It is therefore obvious that where the transition between those two phases occurs a *critical temperature* T_c can be defined as

$$k_\mathrm{B}\,T_\mathrm{c} = J\,\aleph\,. \tag{1.66}$$

As an estimation for the temperature dependence of the mean–field magnetisation only small $\langle S\rangle$ and τ, where

$$\tau \doteq \frac{T - T_\mathrm{c}}{T_\mathrm{c}} \tag{1.67}$$

is the *reduced temperature*, are assumed. The temperature is thus close to T_c so that the hyperbolic tangent can be expanded to

$$\tanh(x) = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \dots\,,\ \text{with}\ x = \frac{1}{1+\tau}\,\langle S\rangle\,.$$

Using the latter it is possible to approximate (1.65) to

$$\begin{aligned}\langle S\rangle &\approx \frac{1}{1+\tau}\,\langle S\rangle - \frac{1}{3}\frac{1}{(1+\tau)^3}\,\langle S\rangle^3 + \mathcal{O}(\,\langle S\rangle^5,\,\tau^5)\\ &\approx (1-\tau)\,\langle S\rangle - \frac{1}{3}\,\langle S\rangle^3 + \mathcal{O}(\,\langle S\rangle\tau^2,\ \langle S\rangle^3\tau)\,,\end{aligned} \tag{1.68}$$

$$\text{where}\qquad \frac{1}{1+\tau} = 1 - \tau + \tau^2 - \dots\,.$$

Neglecting all terms of order τ^2 and higher the temperature dependence can be estimated to

$$\langle S\rangle \sim (-\tau)^{\frac{1}{2}}\,. \tag{1.69}$$

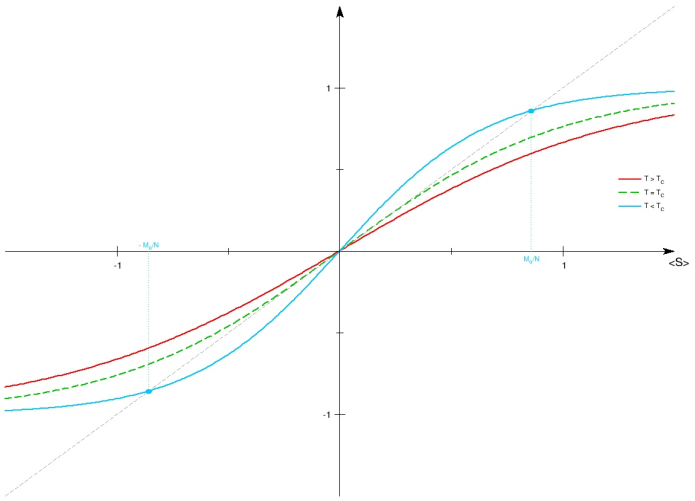


Fig. 1.2: Depicted solution for $\langle S \rangle$ derived from the according self–consistent mean–field approach (1.65). The straight dashed line *(grey)* only intersects with the hyperbolic tangent if the argument $J \aleph \beta > 1$ *(blue)*. For $J \aleph \beta < 1$ *(red)* there is no other possible solution than zero. One can therefore define a critical temperature T_c according to (1.66) where the transition from ferro– to paramagnetic phase occurs *(green)*.

It is worth noting that the critical temperature only depends on the number of nearest neighbours $\aleph$ and other important details of the lattice structure, such as *dimensionality*, are completely neglected. This results, for example, in an incorrect prediction of a phase transition at finite temperature for the one–dimensional Ising model.

Ising Model — Antiferromagnetic Ordering Models within the mean–field theory generally assume a parallel spin alignment of two neighbouring sites, so that all spins have the tendency to point in the same direction and ferromagnetic (FM) interaction can be described. In an antiferromagnetic (AFM) material, however, the stable configuration of a pair of neighbouring spins is antiparallel. The Hamiltonian (1.61) is still valid, but since now $S_i\, S_j = -1$ the exchange integral has to be *negative* ($J < 0$) in order to minimise the Hamiltonian $\mathcal{H}$.
The simplest way to treat AFM interaction is to split the system into two sublattices, A and B, where within each subsystem all spins are parallel to one another, but point in the exact opposite direction of the sites described by the other subsystem. Many lattices, such as the simple square (cubic) lattice in two (three) dimensions, respectively, allow such a description where the neighbouring spins of one site of sublattice A (B) always belong to the other sublattice B (A). The result from (1.65) for the self–consistent calculation of the average magnetisation in the FM case has to remain valid if one considers the change of sign of the exchange by substituting $J \to -J$. Furthermore, all neighbouring sites of a spin belong to the other sublattice, so that

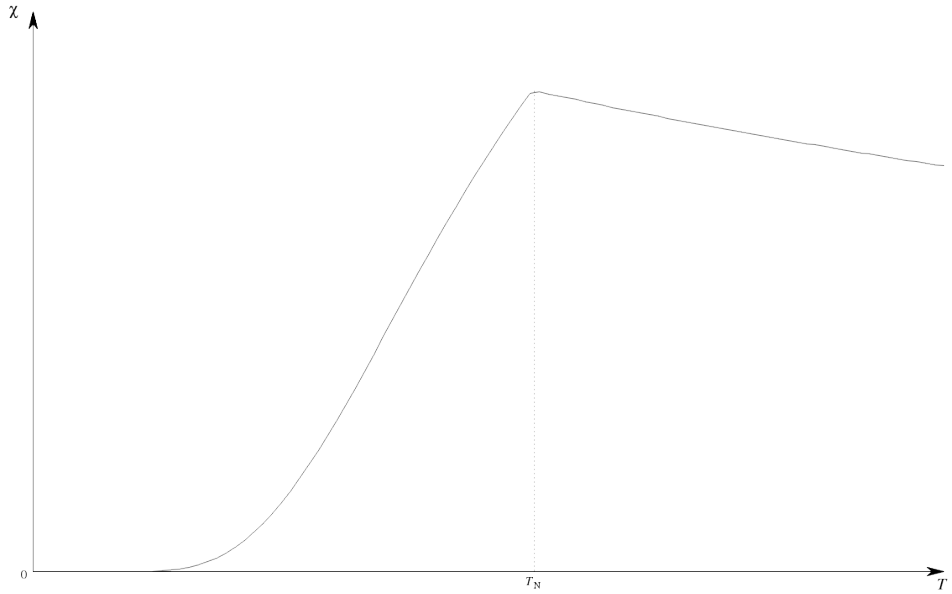


Fig. 1.3: Temperature dependence of the AFM susceptibility χ_{AFM}. The critical temperature T_N can be identified by a cusp in the curve's continuity.

the self–consistent equation (1.65) is therefore split into two coupled equations,

$$\langle S_\mathsf{A}\rangle = \tanh\Big(\beta\left(-J\,\aleph\,\langle S_\mathsf{B}\rangle + \dots\right)\Big)\,, \tag{1.70a}$$

$$\langle S_\mathsf{B}\rangle = \tanh\Big(\beta\left(-J\,\aleph\,\langle S_\mathsf{A}\rangle + \dots\right)\Big)\,. \tag{1.70b}$$

$\langle S_\mathsf{A}\rangle$ and $\langle S_\mathsf{B}\rangle$ are the averaged values of the spins from sublattice A and B, respectively. It is quite obvious that a replacement of $\left\{\langle S_\mathsf{A}\rangle = -\langle S_\mathsf{B}\rangle\right\} \to \langle S\rangle$ once more yields the FM solution again, which means that — except for the opposite orientation of the spins — sublattice A and B show exactly the same properties as it is expected for AFM. Thus, it follows that there is also a *critical point* in case of an AFM configuration, $T_\text{N} = J\,\aleph$, called *Néel temperature*, where spontaneous *staggered* magnetisation occurs.
The behaviour of the magnetic susceptibility can be analysed by using its definition (1.33) in the zero–field limit ($H \to 0$) and that the magnetisation reads $M = N\,\langle S\rangle$,

$$\chi_\mathsf{A} \sim \lim_{H\to 0}\frac{\partial\,\langle S_\mathsf{A}\rangle}{\partial H} = \frac{\beta\left(-J\,\aleph\,\chi_\mathsf{B}+1\right)}{\cosh^2\left(\beta\,J\,\aleph\,\langle S_\mathsf{B}\rangle\right)}\,, \tag{1.71a}$$

$$\chi_\mathsf{B} \sim \lim_{H\to 0}\frac{\partial\,\langle S_\mathsf{B}\rangle}{\partial H} = \frac{\beta\left(-J\,\aleph\,\chi_\mathsf{A}+1\right)}{\cosh^2\left(\beta\,J\,\aleph\,\langle S_\mathsf{A}\rangle\right)}\,. \tag{1.71b}$$

These equations are again a coupled set which is satisfied by the sublattice susceptibilities per spin $\chi_{\mathsf{A},\mathsf{B}}$. Note that the symmetry relation of the hyperbolic cosine, $\cosh(-x) = \cosh(x)$, has been applied. For AFM materials the set of equations (1.71a) has to meet the conditions $\chi_\mathsf{A} = \chi_\mathsf{B} \to \chi_{\text{AFM}}$ and $\langle S_\mathsf{A}\rangle = -\langle S_\mathsf{B}\rangle \to \langle S\rangle$ in order to derive a self–consistent equation for the

AFM susceptibility per spin,

$$\chi_{\text{AFM}} = \frac{\beta\,(-J\,\aleph\,\chi_{\text{AFM}} + 1)}{\cosh^2\left(\beta\,J\,\aleph\,\langle S\rangle\right)} = \beta\left(-J\,\aleph\,\chi_{\text{AFM}} + 1\right)\left(1 - \langle S\rangle^2\right), \tag{1.72}$$

where relation (1.65) can be entered into the latter by expressing the hyperbolic tangent

$$\frac{1}{\cosh^2(x)} = \frac{\cosh^2(x) - \sinh^2(x)}{\cosh^2(x)} = 1 - \tanh^2(x)\,.$$

Since temperature is implicitly taken into account by the inverse β it is possible to analyse the behaviour of χ_{AFM} with varying temperature. Considering (1.72) for $T \geq T_{\text{N}}$ there is no spontaneous magnetisation and thus $\langle S\rangle = 0$, which yields the simple expression

$$\chi_{\text{AFM}}(T) = \frac{1}{T + T_{\text{N}}} \qquad , T \geq T_{\text{N}}\,. \tag{1.73}$$

It can already be seen that the susceptibility does not diverge at $T = T_{\text{N}}$, as it would be the case for a FM material at critical temperature T_{c}. The reason is that a uniform field due to the opposite orientation of the spins below T_{N} cannot effectively cause a macroscopic response around the critical temperature. In case of $T < T_{\text{N}}$ the susceptibility reads

$$\chi_{\text{AFM}}(T) = \frac{1 - \langle S\rangle^2}{T + T_{\text{N}}\left(1 - \langle S\rangle^2\right)} \qquad , T < T_{\text{N}}\,, \tag{1.74}$$

and shows a rapid decrease to zero as $T \to 0$. Both curves are plotted in Fig.1.3, where it can be seen that instead of divergence the AFM susceptibility shows a cusp in the continuity of the curve when reaching the Néel temperature.

1.3.2 Landau Theory

Landau theory of phase transitions is a phenomenological approach of the free energy F of the system based on very simple assumptions. It does not consider microscopic entities of a statistical model, such as Ising spins for example. The free energy is only dependent on one variable known as *order parameter*. As long as the symmetry of the problem is preserved the latter can be arbitrarily chosen.

In magnetic systems the order parameter is usually the magnetisation M. For a *ferromagnet* without an external field the free energy can then be expanded as a power series in M, where only *even* powers are taken into account,

$$F(M) = F_0 + a_2\,M^2 + a_4\,M^4 + \ldots\,. \tag{1.75}$$

Only those terms are compatible with the symmetry of a magnetic system, since *invariance* has to be guaranteed when the sign of the magnetisation is reversed. The coefficients $a_i = a_i(T)$ are dependent on temperature, but constant as functions of M. For a qualitative analysis it is sufficient to stop the expansion after the fourth order term (M^4). As a matter of fact, subsequent higher terms cannot alter the critical behaviour of the system. It is required that the coefficient a_4 always has a *positive* value, since otherwise the free energy would decrease infinitely with increasing M. Additionally, for a *positive* coefficient $a_2 > 0$ the free energy is

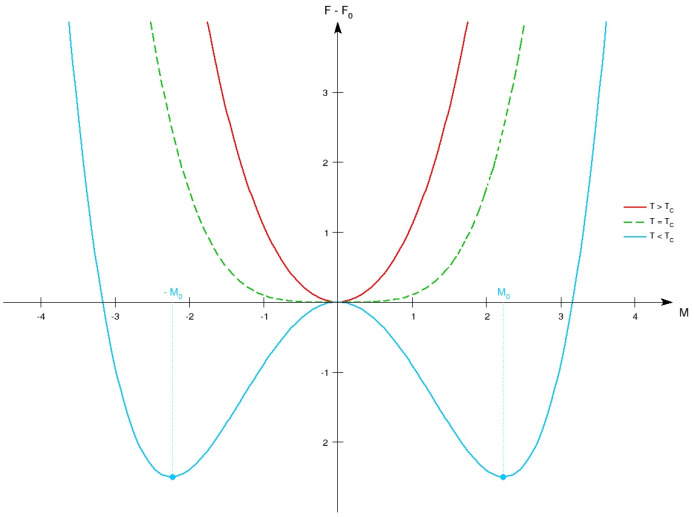


Fig. 1.4: Plot of the Landau model (1.75) up to M^4. The paramagnetic phase *(red)* is determined by a positive coefficient $a_2 > 0$. If $a_2 < 0$ *(blue)*, the curve shows two minima at $\pm M_0$, where only one of the values is actually observed in a real ferromagnetic phase (spont. symmetry breaking). The phase transition occurs when $a_2 = 0$ *(green)*.

minimal only at the origin ($M = 0$), thus the system is in a *paramagnetic phase*. If due to temperature influence the coefficient decreases and becomes *negative*, $a_2 < 0$, two minima are formed at finite values $\pm M_0$ of the order parameter which emerge away from origin into positive and negative direction, respectively, as a_2 decreases, indicating a *ferromagnetic phase*. Both cases are depicted in Fig.1.4. The transition from one phase to the other occurs when $a_2 = 0$, which corresponds to a critical temperature T_c. One can thus parametrise the coefficient by

$$a_2 = a\,\tau\,, \tag{1.76}$$

where τ is again the reduced temperature from (1.67) and a is a plain scaling constant. With the variation of τ it is now possible to describe a continuous phase transition of the magnetisation.

The equilibrium magnetisation M_0 can be calculated by the first derivative of the free energy

$$\frac{\mathrm{d}F}{\mathrm{d}M} = 2a_2\,M + 4a_4\,M^3\,, \tag{1.77}$$

where the latter is zero and for $\tau < 0$ yields

$$\begin{aligned} M_0 &= \sqrt{-\frac{a_2}{2\,a_4}} = \sqrt{-\frac{a}{2\,a_4}\,\tau} \\ \Longrightarrow \qquad M_0 &\sim (-\tau)^{\frac{1}{2}}\,, \end{aligned} \tag{1.78}$$

which is the same result that has been previously derived with (1.69) in the case of the Ising model. Differentiating (1.77) a second time,

$$\frac{\mathrm{d}^2F}{\mathrm{d}M^2} = 2a_2 + 12a_4\,M^2\,, \tag{1.79}$$

furthermore yields the susceptibility $\chi = \left(\frac{\mathrm{d}^2 F}{\mathrm{d}M^2}\right)^{-1}$ for $\tau < 0$,

$$\chi = \frac{1}{2a_2 + 12a_4 \left(-\frac{a_2}{2a_4}\right)} = -\frac{1}{4a\,\tau}$$
$$\implies \quad \chi \sim (-\tau)^{-1} \,. \tag{1.80}$$

It is thus shown that these two theories are equivalent in the sense of approximating a mean–field interaction (although details of the dimensionality and number of components is completely lost, which is generally not appropriate). Using (1.78) the expression for the free energy (1.75) can be rewritten in terms of the reduced temperature for $\tau < 0$,

$$F = F_0 - \frac{a^2}{4a_4}\,\tau^2 \,. \tag{1.81}$$

Note that the symmetry inherent in (1.75) consequently establishes two stable states $\pm M_0$ that coexist in the ferromagnetic order. However, it is evident that in a physical system the actual state only shows one of the two minima truly realised. It follows that the realised equilibrium state does not show the above mentioned symmetry, which is called *spontaneous symmetry breaking*. Only a part of the phase space is reached by the system; the tendency to one of the equilibriums is developed by a small external field or an initial condition in time evolution of the system. This is a common phenomenon seen in mean–field approximations.

1.3.3 The Limits of Mean–Field Approximations

The application of approximative methods also always demands the question of its limitations. Usability and validity are the two decisive factors one generally has to consider. The limit for mean–field approximations clearly lies within the disregard of fluctuations around the average of the physical quantities. In order to give an estimation for the latter and thus find a rule where the mean–field theory in D dimensions can reliably be applied, one integrates all fluctuations of magnetisation up to the length scale of correlation length ξ,

$$\sigma^2 = \int_0^\xi \mathrm{d}^D r\, \Gamma(r) \,, \tag{1.82}$$

with $r = |i - j|$ as the difference to the specific lattice sites, and $\Gamma(r)$ is the *correlation function* of the spins, where

$$\Gamma(i,j) = \Big\langle \big(S_i - \langle S_i \rangle\big)\big(S_j - \langle S_j \rangle\big) \Big\rangle$$
$$\implies \quad \Gamma(r) = \langle S_r S_0 \rangle - \langle S_r \rangle \langle S_0 \rangle = \langle S_r S_0 \rangle - \langle S \rangle^2 \,.$$

For the last term *translation invariance* and *isotropy* are assumed. The correlation function $\Gamma(r)$ remains at an almost constant value for $r < \xi$ and exponentially decreases beyond. Fluctuations at a length scale larger than ξ are therefore uncorrelated and a change of the upper limit beyond ξ will not effectively change the value of the integral in (1.82). It is shown in Apx.C that the integral with extension to infinity is proportional to the magnetic susceptibility χ. The latter is

known as *susceptibility sum rule*, which is derived from linear response theory, and is a special case of the *fluctuation–dissipation theorem*. As a consequence we can therefore write

$$\chi = \beta \int_0^\infty \mathrm{d}^D r \, \Gamma(r) \approx \beta \, \sigma^2 \, . \tag{1.83}$$

The mean–field approximation is consistent where the fluctuations σ^2 are extensively smaller than the square of the magnetisation integrated up to the correlation length, which gives a self–consistent condition called *Ginzburg criterion*,

$$\int_0^\xi \mathrm{d}^D r \; \langle S_r \rangle \, \langle S_0 \rangle \sim M^2 \, \xi^D \quad \gg k_\mathrm{B} T \, \chi \, . \tag{1.84}$$

Near criticality the criterion can be rewritten using the according critical exponents,

$$\begin{aligned} \tau^{2\beta} \, \tau^{-\nu D} \; &\gg \; k_\mathrm{B} T \, \tau^{-\gamma} \, , \\ \implies \quad \nu D - 2\beta \; &> \; \gamma \, , \end{aligned} \tag{1.85}$$

which implies when the mean–field values (see Tab.1.2) are entered that

$$D > 4 \, .$$

Hence, it is shown that mean–field theory is consistent for $D > 4$, where the critical exponents are granted to take the mean–field values and mean–field theory forms a universality class. The threshold $D_\mathrm{uc} = 4$ is the *upper critical dimension*. The argument above only concerns the asymptotic behaviour of thermodynamic functions near criticality. Although the exponents are correct beyond D_uc, critical temperatures T_c and other non-universal quantities are generally not. It is quite plausible that a theory, where the interaction is based on the average of all other spins, should approximate to better results when the system is viewed in a wider region. Beside a higher dimensionality this could be achieved by an increase of interacting neighbours or a greater range of interactions. Generally, the mean–field approximation is a better theory outside a certain critical region with size R^D, where it is expected that fluctuations are small and even almost neglectable. The size of the region is material dependent (non-universal) and can be estimated again by the Ginzburg criterion (1.84); considering that $M^2 \propto |\tau|$, $\xi \propto R|\tau|^{-1/2}$, and using the result (1.80) from Landau theory,

$$R^D |\tau|^{1-\frac{D}{2}} \; \gg \; \frac{1}{4a \, |\tau|} \quad \implies \quad R^D \; \gg \; C \, |\tau|^{-\frac{(4-D)}{2}} \, , \tag{1.86}$$

where C is a constant of order one. The latter equation shows that fluctuations become more relevant as the dimensionality D is lower.

Albeit the reliability of the mean–field theory is quite limited, there are many cases where its application, for example as series expansion from four dimensions, brings fruitful qualitative predictions, as it often delivers (in some cases) the only feasible approach to complicated issues. It can then be seen as the groundwork for more sophisticated calculations. Finally, the most peculiar feature is that mean–field theory — as opposed to many other numerical approaches — actually improves with increasing dimensionality, which highlights the importance of symmetry of the order parameter.

1.3.4 Universality

The theory of continuous phase transition reveals interesting effects when approaching the *critical parameters*, for example T_c, where the transition occurs. Especially in magnetic systems there are some entities that diverge (e.g. susceptibility χ_T) and others that vanish (e.g. magnetisation M as order parameter). There are many (physically) different systems that can show the exact same behaviour near criticality, which leads to the notion of *universality*. If a system shows universal features, the order parameter will become less dependent on the details of the system the closer the *system parameter* approaches its critical value. Mathematically, the ansatz

$$A = A_0 \left|T - T_c\right|^{\pm\lambda} \sim A_0 \left|\tau\right|^{\pm\lambda} \tag{1.87}$$

describes this fact, where $\lambda > 0$ is the *critical exponent* which takes typical values for the according quantity A to parametrise its near-critical behaviour when approaching T_c. As it has been shown previously in Sec.1.3.1 and Sec.1.3.2 the critical exponent for the order parameter (magnetisation) is $\beta = \frac{1}{2}$. All further critical exponents for other magnetic quantities can be derived in a similar manner or as a consequence of β taking this value. They are shown in Tab.1.2 . It might seem surprising that two different models derive the same critical exponents under the class of mean–field approximations; however, this is not a coincidence. The fact that very different systems can have the same critical exponents classifies them as universal entities. Opposed to these the values of the critical parameters $(T_c, \ldots)$ are dependent on the Hamiltonian of the system in a complex manner and are therefore *non-universal*. The phenomena of universality is based on quantities which only depend on a few fundamental global parameters (e.g. dimension), but not on the knowledge of any dynamical details of the system. Moreover, systems with the same set of critical exponents can be unified to so-called *universality classes* where it is important to note that these systems do not necessarily need to have the same physical basis, but all commonly show the same critical behaviour. A phase transition is therefore sufficiently characterised by finding and relying on the proper universality class. The mean–field approximation is a good example for such a class, as the individual critical exponents always take the same value.

Tab. 1.2: The critical exponents of Mean–Field approximations.

Crit. Exponent	Mean–Field
α	0
β	$\frac{1}{2}$
γ	1
δ	3
ν	$\frac{1}{2}$

Renormalisation Group and Critical Phenomena In a wider sense universality can be explained within the theoretical frame of *renormalisation group*. A system with n_f degrees of freedom can be equivalently described by a reduced system, where $n_f' < n_f$ (*scale invariance*). The concept of renormalisation is to observe the change of physical quantities as the length scale is successively increased by applying mathematical operations embodied within this group. These

are known as *coarse graining* and *(re)scaling*, where parts of the microscopic degrees of freedom are traced out and the spatial scale is again normalised after being changed by a *scaling factor* b, respectively. Repeating these operations draws attention to the long-length behaviour of a macroscopic system, but neglects short-range contributions. Since for phase transitions the correlation length ξ goes to infinity fluctuations near the critical point are thereby systematically taken into account. On the other hand short-ranged interaction is gradually cancelled out and the critical behaviour of the system is understood as the *asymptotic behaviour* of infinitely many iterations of the renormalisation process. Under such a transformation, where a set of variables $\{S\}$ is traced out and a set $\{S'\}$ remains the partition function (and hence the free energy) has to be invariant,

$$\mathcal{Z}_{n_\mathrm{f}}(\mathcal{H}) = \sum_{\{S'\}} \sum_{\{S\}} \mathrm{e}^{-\mathcal{H}} = \sum_{\{S'\}} \mathrm{e}^{-\mathcal{H}'} \equiv \mathcal{Z}_{n_\mathrm{f}'}(\mathcal{H}') \,, \tag{1.88}$$

where the Hamiltonians are derived by a generally *non-linear transformation* $R_b(\cdot)$ that reduces the number of degrees of freedom accordingly,

$$\begin{aligned} & \mathcal{H}' = R_b(\mathcal{H}) \\ \Longrightarrow \quad & {n_\mathrm{f}}' = b^{-D}\, n_\mathrm{f} \,, \end{aligned} \tag{1.89}$$

where D stands for the dimensionality. Coarse graining is not a unique operation. Choosing one procedure leads to a particular renormalisation group scheme and since for $b > 1$ information is irretrievably lost as degrees of freedom are traced out, it is impossible to define an inverse transformation ${R_b}^{-1}$. There is, however, still an *identity map* ($b = 1$) and two successive mappings are equivalent to a single map

$$\mathcal{H}'' = R_{b_2}\, R_{b_1}(\mathcal{H}) \equiv R_{b_1\, b_2}(\mathcal{H}) \,, \tag{1.90}$$

so that mathematically the set of transformations $\{R_b\}$ forms what is called a *semi–group*.

If the system is at the critical point there are fluctuations of all length scales, which is the reason it should stay unchanged after many steps of renormalisation. If the number of steps goes to infinity the Hamiltonian approaches a *fixed point* $\mathcal{H}^*$ defined by

$$\mathcal{H}^* = \lim_{n\to\infty} {R_b}^n(\mathcal{H}_\mathrm{c}) \,. \tag{1.91}$$

It should be clear that the Hamiltonian at the critical point $\mathcal{H}_\mathrm{c}$ is not the same as the one at the fixed point $\mathcal{H}^*$, but rather approaches the latter asymptotically. Furthermore, a fixed point can be identified by its invariance

$$\mathcal{H}^* = R_b(\mathcal{H}^*) \,. \tag{1.92}$$

The Hamiltonian can be written in a *parameter space* $\{u_1, \ldots, u_n\}$ and introduced as the sum of products between parameter u_j and operator $\hat{A}_j$,

$$\mathcal{H} = \sum_j u_j\, \hat{A}_j \,. \tag{1.93}$$

The operators are well-defined, so that analogously to (1.89) one can use a similar map for the parameter space

$$u' = R_b(u) \,. \tag{1.94}$$

It is not explicitly stated that obviously lengths are again reduced by a scale factor b for each transformation. As this is also valid for the correlation length ξ it now can be seen that in the limit of infinitely repeated renormalisation group operations the asymptotic behaviour of the parameters $\{u_j\}$ determines the critical exponents that describe non-analyticities of the physical quantities at criticality. A *fixed point* u^* in parameter space is again identified by invariance

$$u^* = R_b(u^*)\,, \tag{1.95}$$

and can be associated with the according fixed point Hamiltonian $\mathcal{H}^*$ from (1.91). The correlation length at the fixed point reads

$$\xi\left[R_b(u^*)\right] = \xi\left[u^*\right] = b^{-1}\xi\left[u^*\right]\,, \tag{1.96}$$

where $\xi\left[u^*\right]$ can either be zero (away from criticality $\to 0$) or infinity (stays at criticality $\to \infty$). The latter manifests a *critical fixed point* where the former is a *trivial fixed point* and expresses that there is no characteristic length scale at such point. As the interest lies in the singular behaviour near criticality it is instructive to study the properties of the system with parameter values slightly away from the fixed point. The parameters before and after renormalisation can thus be written as a small deviation from it,

$$u = u^* + \delta u\,, \tag{1.97a}$$
$$u' = u^* + \delta u'\,. \tag{1.97b}$$

The transformation appearing in the recursion relation (1.94) is in general non-linear, but analytic and can be linearised by expansion around the fixed point and neglecting terms of second–order and beyond,

$$u' = R_b(u^* + \delta u) \approx \overbrace{R_b(u^*)}^{=u^*\,(1.95)} + \underbrace{\frac{\partial R_b}{\partial u}(u^*)}_{=L_b(u^*)}\,\delta u\,. \tag{1.98}$$

$L_b(u^*)$ is a not necessarily symmetric real matrix which can be diagonalised with real eigenvalues. Comparing the latter with (1.97b) one derives the *linearised recursion relation*, which is only valid close to the fixed point

$$\delta u' = L_b(u^*)\,\delta u\,, \tag{1.99}$$
$$\text{where}\quad \left[L_b(u^*)\right]_{ij} = \frac{\partial u_i'}{\partial u_j}(u^*)\,.$$

Since the semi–group property of (1.90) must also be valid for $L_{b_2}\,L_{b_1} \equiv L_{b_1\,b_2}$, where L_b generally is a function of scaling factor b, the *eigenvalues* $\lambda_i(b)$ of the transformations as a consequence of linearity must also read

$$\begin{aligned} \lambda_i(b_1\,b_2) &= \lambda_i(b_1)\,\lambda_i(b_2) \\ \Longrightarrow \qquad \lambda_i(b) &= b^{v_i}\,, \end{aligned} \tag{1.100}$$

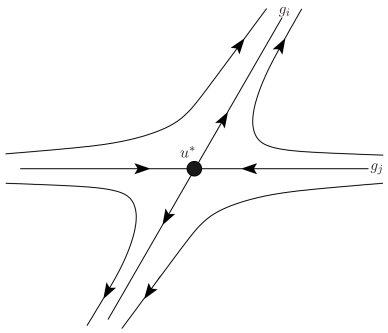


Fig. 1.5: Schematic parameter flow of the scaling fields g_i and g_j, where one of the according exponents is *positive* ($\upsilon_i > 0$) and the other is *negative* ($\upsilon_j < 0$). The fixed point $u^*(g_i = g_j = 0)$ can only be reached when there are no contributions of g_i.

which can only be satisfied if the eigenvalues are powers of b. The deviations δu, $\delta u'$ are then expanded by the set of eigenvectors $\{\phi_i\}$ of L_b, so that the equations in (1.97) read

$$u = u^* + \sum_i g_i \, \phi_i \,, \tag{1.101a}$$

$$u' = u^* + \sum_i g_i{}' \, \phi_i \,. \tag{1.101b}$$

The variables g_i are called *scaling fields* and are important quantities to characterise the parameter space $\mathbf{u}$. Using (1.99) implies that g_i and g_i' are related by

$$\begin{aligned} \sum_i g_i{}' \, \phi_i = L_b \sum_i g_i \, \phi_i = \sum_i g_i \, L_b \phi_i = \\ = \sum_i g_i \, \lambda_i(b) \, \phi_i = \sum_i g_i \, b^{\upsilon_i} \, \phi_i \\ \implies \quad g_i{}' = b^{\upsilon_i} \, g_i \end{aligned} \tag{1.102}$$

The renormalisation group analysis has therefore been reduced to gaining information about the fixed points u^*, the exponents $\{\upsilon_i\}$ of the eigenvalues of L_b and the scaling fields $\{g_i\}$.

Critical phenomena can be characterised by the eigenvalues and eigenvectors of the linear transformation L_b. Especially the concept of universality can be extracted from them, where the sign of the eigenvalue's exponents takes a decisive part. It can be seen from (1.102) that if υ_i is *positive*, the scaling field g_i is *amplified* ($b > 1$) after each renormalisation group transformation. The parameter u' would therefore move further away from the fixed point with every additional step. For *negative* υ_i, on the other hand, the parameter would converge towards u^*. Fig.1.5 shows the parameter flow schematically for a positive and a negative exponent. In order for the system parameters to be attracted to a fixed point it is therefore necessary that all scaling fields with a positive exponent are gauged to zero,

$$\forall g_i \, : \, \{\upsilon_i > 0 \implies g_i = 0\} \therefore \{u^* = \lim_{n\to\infty} {R_b}^n(u)\} \,. \tag{1.103}$$

Since this is a decisive adjustment to the behaviour of the system a scaling field with positive exponent is called a *relevant variable* ($\upsilon > 0$). The rapid diminishment of scaling fields with negative exponents makes them an *irrelevant variable* ($\upsilon < 0$), as they have no effect on the critical properties. For $\upsilon = 0$ variables are called *marginal*, which are associated with logarithmic corrections to scaling. The described features of a variable are relative to the according fixed point. A variable may be relevant to one particular fixed point, but irrelevant (or marginal) to another.
In a ferromagnetic system only two parameters, temperature and external magnetic field, need to be adjusted to critical values in order to observe critical phenomena. Slight deviation from these values will lead the system away from criticality, which shows that the two quantities τ (reduced temperature) and h (reduced magnetic field) are related — even proportional near the fixed point due to analyticity — to the relevant variables g_τ and g_h, respectively. Variables represented by other scaling fields with negative exponents do not affect the essential features of critical phenomena. Hence, there is usually no influence on the critical exponents, except for the positive exponents of υ_τ and υ_h. Since the parameters irrelevant to the description of criticality can be quite diverse, the observed critical exponents can cover a wide range of different systems described by these parameters. This leads a way to clarify the notion of universality from the view of renormalisation group theory by showing that the only traceable difference of systems with the same critical phenomena are the irrelevant parameters, while the relevant ones are commonly shared. Therefore, a *universality class* specifies amongst other fundamental entities the shared set of relevant variables.

The critical exponents can be linked to the positive exponents υ_i of the eigenvalues of the linear transformation L_b which are related to the relevant scale variables g_i. The main concept lies behind the analysis of the free energy per degree of freedom $f(\mathcal{H})$ under a renormalisation group transformation (remember that the free energy is invariant per se). With the definition $f \doteq \frac{F}{n_\mathrm{f}}$ and the result from (1.89) one can define a renormalisation group transformation as

$$f(g_1, g_2, g_3, \ldots) = b^{-D}\, f(g_1', g_2', g_3', \ldots) + w(g_1, g_2, g_3, \ldots)\,, \tag{1.104}$$

where for completeness the regular term w has been added, which corresponds to the logarithm of a coefficient product in the partition function and should be regarded when the exact value of the free energy is demanded. This multiplicative factor, however, does not have any influence on the fixed point or the critical exponents and *is thus omitted*. In the ferromagnetic case g_τ and g_h become proportional to τ and h close to the critical point, respectively. One can therefore use (1.104) with (1.102), so that after n steps of renormalisation,

$$f(\tau, h) = b^{-nD}\, f(b^{n\upsilon_\tau}\,\tau, b^{n\upsilon_h}\,h)\,. \tag{1.105}$$

For $\tau > 0$ the number n is chosen so that the first argument on the right goes to unity,

$$b^{n\upsilon_\tau}\,\tau = 1 \iff n = -\frac{1}{\upsilon_\tau}\,\frac{\ln\tau}{\ln b}\,.$$

For $\tau < 0$ the sign is just reversed to -1. Replacing b^n the expression above is equivalent to a *scale transformation*,

$$b^n = \tau^{-\frac{1}{\upsilon_\tau}} \tag{1.106}$$

$$\implies \qquad f(\tau, h) = \tau^{\frac{D}{\upsilon_\tau}}\, f\left(1, h\,\tau^{-\frac{\upsilon_h}{\upsilon_\tau}}\right) \equiv \tau^{\frac{D}{\upsilon_\tau}}\,\Phi\left(h\,\tau^{-\frac{\upsilon_h}{\upsilon_\tau}}\right)\,. \tag{1.107}$$

The latter is known as *scaling law* and has effectively transformed a function of two independent variables into a single–variable function, which is called *scaling function* Φ. The scaling law implicitly assumes that the free energy f is a *generalised homogeneous function*. The repetition of the renormalisation group transformation shifts the reduced temperature τ away from the critical region and the usual critical condition $|\tau| \ll 1$ has to be replaced by

$$|\tau| \ll 1 \quad \longrightarrow \quad |\tau| = b^{-n\upsilon_\tau} \, .$$

It is possible to establish special relations between the critical exponents and the exponents of the eigenvalues $\upsilon_\tau, \upsilon_h$ by using the scaling law. Not explicitly shown, they result in the following expressions,

$$\alpha = 2 - \frac{D}{\upsilon_\tau}, \quad \beta = \frac{D - \upsilon_h}{\upsilon_t}, \quad \gamma = \frac{2\upsilon_h - D}{\upsilon_\tau}, \quad \delta = \frac{\upsilon_h}{D - \upsilon_h} \, . \tag{1.108}$$

The critical exponents are not independent from each other. Eliminating υ_τ and υ_h from (1.108) yields the *scaling relations*

$$\gamma = \beta \, (\delta - 1) \, , \tag{1.109a}$$

$$\alpha + 2\beta + \gamma = 2 \, , \tag{1.109b}$$

which are satisfied by the exponents of mean–field theory.

Chapter 2

Density Functional Theory

The Density Functional Theory (DFT) is a quantum mechanical modelling method used to calculate the electronic structure of many–body systems. The idea is to work with a simple quantity — in this case the electron density ϱ — instead of solving Hamiltonians through complicated many–body wave functions. The basic principles are presented on a very pragmatic level by the Thomas–Fermi model (free electron gas). The fundamental steps, however, have been made by the two theorems of Pierre Hohenberg and Walter Kohn. The first Hohenberg–Kohn theorem shows that the energy of the system can be expressed as a functional $E[\varrho(\vec{\boldsymbol{r}})]$ of the spatially dependent electron density $\varrho(\vec{\boldsymbol{r}})$, and the second theorem states that this functional is found to be minimal at the unique ground state density $\varrho(\vec{\boldsymbol{r}}) \to \varrho_0(\vec{\boldsymbol{r}})$. DFT is among the most versatile methods available in computational physics. The computational effort is relatively low compared to the traditional methods based on the complex many–body wave functions, such as the Hartree–Fock theory.

2.1 Many–Body Systems: Schrödinger Representation

The dynamics of a quantum system is described by the Hamiltonian $\hat{\mathcal{H}}$. If $|\Psi\rangle$ is a quantum state wave function of the system, its time–evolution is given by

$$i\hbar\partial_t\,|\Psi\rangle \;=\hat{\mathcal{H}}\,|\Psi\rangle\;, \tag{2.1}$$

$$\text{where } \partial_t \doteq \tfrac{\partial}{\partial t} \text{ and } \Psi = \Psi(\vec{\boldsymbol{r}}_1,\ldots,\vec{\boldsymbol{r}}_N,t)\,.$$

Stationary states with discrete energy eigenvalues are obtained by the time independent Schrödinger equation using the separation ansatz $\Psi = \psi(\vec{\boldsymbol{r}}_1,\ldots,\vec{\boldsymbol{r}}_N)\cdot \mathrm{e}^{-i\frac{E}{\hbar}t}$

$$\hat{\mathcal{H}}\,|\psi\rangle \;=\; |\psi\rangle\,E, \quad \langle\psi|\psi\rangle \;=1\,, \tag{2.2}$$

or alternatively derived from the variational problem

$$E[\psi] = \frac{\langle\psi|\hat{\mathcal{H}}|\psi\rangle}{\langle\psi|\psi\rangle}\,. \tag{2.3}$$

In the latter case the ground state energy $E_0 \le E[\psi]$ is identified as the minimum of the energy functional,

$$E_0 = \min_{\psi} E[\psi]\,, \tag{2.4}$$

which can be calculated using the variational principle under the constraints $\delta E[\psi] = 0$ and $\langle\psi|\psi\rangle = 1$,

$$\delta\Big\{\,\langle\psi|\hat{\mathcal{H}}|\psi\rangle \;-\lambda\big(\,\langle\psi|\psi\rangle\;-1\big)\Big\} = 0\,. \tag{2.5a}$$

λ is the *Lagrange parameter*, which is equivalent to the energy eigenvalue E,

$$\lambda \equiv E \Longrightarrow \delta\Big\{ \langle\psi|\hat{\mathcal{H}}|\psi\rangle - E\big(\langle\psi|\psi\rangle - 1\big)\Big\} = 0\,. \tag{2.5b}$$

The equivalence is justified due to the fact that the variation with respect to $\langle\psi| = \psi^*$ leads again to the Schrödinger equation (2.2).

The system to be considered consists of N identical particles (in most cases these particles are electrons) moving in an external field and interacting with each other by pair forces. The number of particles is conserved; furthermore there are no excitations, nor a temperature dependence. The Hamiltonian for this case is built up by the kinetic energy operator $\hat{T}$, the external potential operator $\hat{U}$ of the interaction of the particles with the field of the nuclei, and the two–particle interaction operator $\hat{W}$,

$$\hat{\mathcal{H}} = \hat{T} \oplus \hat{U} \oplus \hat{W}\,. \tag{2.6}$$

The spatial representation uses the eigenstates of the coordinates $\vec{\boldsymbol{r}}$ (and possibly the spin projection s with respect to a given quantisation axis z) of the particles as basis vectors in the Hilbert space of one–particle quantum states:

$$\begin{aligned} \hat{\boldsymbol{r}}\,|\boldsymbol{r}\rangle &= |\mathbf{r}\rangle\,\vec{\boldsymbol{r}} \\ \hat{\sigma}_{\mathrm{z}}\,|s\rangle &= |s\rangle\,s \end{aligned}$$

$$x \doteq (\vec{\boldsymbol{r}}, s)\,, \quad \int \mathrm{d}x \doteq \sum_s \int \mathrm{d}^3 r\,.$$

If the Hamiltonian in (2.6) is now acting on spatially dependent (stationary) wave functions, it then can be explicitly given as

$$\begin{aligned} \hat{\mathcal{H}} &= \quad \hat{T} \quad + \quad \hat{U} \quad + \quad \hat{W} \\ \hat{\mathcal{H}} &= -\frac{\hbar^2}{2m}\sum_{i=1}^{N} \vec{\boldsymbol{\nabla}}_i^{\,2} + \sum_{i=1}^{N} u(x_i) + e^2 \sum_{i<j}^{N} \frac{1}{r_{ij}}\,, \end{aligned} \tag{2.7}$$

$$\text{where } r_{ij} = |\vec{\boldsymbol{r}}_i - \vec{\boldsymbol{r}}_j| \quad \text{and} \quad u(x_i) = -e^2 \sum_{\alpha} \frac{Z_\alpha}{r_{i\alpha}}\,.$$

$u(x_i)$ is the potential of the external field for a particle with position and spin x_i. It should be noted that this field can also be spin dependent. Generally, this dependency would be treated by four spatial functions, where two spinor indices (ss') of a (spatially local) $S = \frac{1}{2}$ operator define these functions by $u_{ss'}(\vec{\boldsymbol{r}})$. This operator is only used in density functional treatments of ground states with *non-collinear* spin structure. In most cases, however, the spin dependence is visualised as an external magnetic field $B(\vec{\boldsymbol{r}})$, which is restricted to the z–direction and only acting on the spin while its effect on the orbital motion is neglected. The result is a simple potential energy contribution of $-2sB(\vec{\boldsymbol{r}})$ and allows a treatment similar to the spin–independent case.

For a better overview *Natural Units* will be used within this chapter, which gives (2.7) the form

$$\boxed{\hbar = m = e = 1}$$

$$\hat{\mathcal{H}} = -\frac{1}{2}\sum_{i=1}^{N} \vec{\boldsymbol{\nabla}}_i^{\,2} + \sum_{i=1}^{N} u(x_i) + \sum_{i<j}^{N} \frac{1}{r_{ij}} \,. \tag{2.8}$$

2.2 Thomas–Fermi Theory

In the modern progress of DFT one tries to find a functional expression for the ground state energy $E[\rho]$ via the ground state density $\rho(\vec{\boldsymbol{r}})$ and then to base a variational principle with respect to the density on this functional relation. The *Thomas–Fermi theory* is the earliest and least complex version of these approximating theories, but learning from its concept is still of importance.
Thomas and Fermi independently considered that there is an important contribution to the total energy due to the interaction between the electrons themselves. However, they could only derive a classical Coulomb term. The quantum mechanical phenomena they were not aware of at that time was the exchange energy of the electrons (fermions), and thus they neglected the correlation term. All potential terms could therefore be expressed through the particle density $\rho(\vec{\boldsymbol{r}})$. The only contribution not readily related to $\rho(\vec{\boldsymbol{r}})$ was the kinetic energy.

It is the assumption that the electrons surrounding an atomic nucleus behave like a *homogeneous interaction–free fermion gas* ($S = \frac{1}{2}$). In the atomic ground state they will fill out a spherical momentum space volume V_p up to the Fermi momentum p_F,

$$V_p = \frac{4}{3}\pi\, {p_\mathrm{F}}^3 \,. \tag{2.9}$$

Considering the spin of the electrons and the Pauli exclusion principle one electronic state occupies a phase space volume of $\frac{h^3}{2}$, which can easily be shown by calculating the phase space volume of a non-spin state in a three dimensional cube with length L. The periodic boundary conditions yield a $\Delta k = \frac{2\pi}{L}$. Therefore, the phase space volume for one state is

$$(L \cdot \Delta p)^3 = (L \cdot \hbar\, \Delta k)^3 = h^3 \,. \tag{2.10}$$

The total phase space volume of a small volume element $\mathrm{d}V$ can be calculated by

$$V_\mathrm{phase} = V_p \cdot \mathrm{d}V = \frac{4}{3}\pi\, {p_\mathrm{F}}^3 \cdot \mathrm{d}V \,. \tag{2.11}$$

Using (2.10) and (2.11) the number $\mathrm{d}N$ of occupied states is

$$\begin{aligned} \mathrm{d}N = \frac{2}{h^3} \cdot V_\mathrm{phase} = \frac{8\pi}{3}\,\frac{1}{h^3}\,{p_\mathrm{F}}^3\,\mathrm{d}V = \\ = \frac{1}{3\pi^2}\left(\frac{p_\mathrm{F}}{\hbar}\right)^3 \mathrm{d}V \,. \end{aligned}$$

With $\rho \doteq \frac{\mathrm{d}N}{\mathrm{d}V}$ and $p_\mathrm{F} = \hbar\, k_\mathrm{F}$ the expression for the electron density of the Thomas–Fermi model reads,

$$\boxed{\rho = \frac{1}{3\pi^2}\left(\frac{p_\mathrm{F}}{\hbar}\right)^3 = \frac{1}{3\pi^2}\,{k_\mathrm{F}}^3} \quad . \tag{2.12}$$

The *average kinetic energy per particle* $\langle t \rangle$ now can be calculated (classically) as a function of the electron density,

$$\langle t \rangle = \frac{\int_0^{p_\mathrm{F}} \mathrm{d}^3p \, \frac{p^2}{2m}}{\int_0^{p_\mathrm{F}} \mathrm{d}^3p} = \frac{4\pi}{2mV_p} \int_0^{p_\mathrm{F}} \mathrm{d}p \, p^4 = \frac{1}{m} \, \frac{3}{10} \, {p_\mathrm{F}}^2 \, . \tag{2.13}$$

Using (2.12) yields

$$\langle t \rangle \longrightarrow t(\rho) = \frac{\hbar^2}{m} \, \frac{3}{10} \, \left(3\pi^2\right)^{\frac{2}{3}} \, \rho^{\frac{2}{3}} \, . \tag{2.14a}$$

Transforming to Natural Units results in

$$\boxed{t(\rho) = C_\mathrm{F} \, \rho^{\frac{2}{3}}} \, , \tag{2.14b}$$

$$\text{where } C_\mathrm{F} = \frac{3}{10} \, \left(3\pi^2\right)^{\frac{2}{3}} \approx 2.8712 \, .$$

The *kinetic energy per unit volume* $\mathrm{d}V$ is $\rho \cdot t(\rho)$, so that

$$\mathrm{d}T = C_\mathrm{F} \, \rho^{\frac{5}{3}} \mathrm{d}V \, . \tag{2.15}$$

Under the assumption that the electron density varies slowly enough in space for a *real* atomic configuration, then

$$T\left[\rho(\vec{\boldsymbol{r}})\right] \approx \int_V \mathrm{d}V \, \rho(\vec{\boldsymbol{r}}) \, t(\vec{\boldsymbol{r}}) = C_\mathrm{F} \int_V \mathrm{d}V \, \rho^{\frac{5}{3}}(\vec{\boldsymbol{r}}) \tag{2.16}$$

can be conducive to an acceptable approximation for the kinetic energy functional $T[\rho]$ of the electron density.
For the potential terms of (2.1) the transfer from a discrete number of particles to a (continuous) particle density has to be accomplished by

$$N = \sum_{i=1}^{N} \rho(\vec{\boldsymbol{r}}_i) \longrightarrow \int_V \mathrm{d}^3r \, \rho(\vec{\boldsymbol{r}}) \, . \tag{2.17}$$

Without proof the *potential functionals* are

$$U\left[\, \rho(\vec{\boldsymbol{r}}); u(\vec{\boldsymbol{r}}) \,\right] = \int_V \mathrm{d}^3r \, u(\vec{\boldsymbol{r}}) \, \rho(\vec{\boldsymbol{r}}) \, , \tag{2.18}$$

$$W\left[\, \rho(\vec{\boldsymbol{r}}) \,\right] = \frac{1}{2} \int_V \mathrm{d}^3r \, u_\mathrm{H}(\vec{\boldsymbol{r}}) \, \rho(\vec{\boldsymbol{r}}) \doteq \frac{1}{2} V_\mathrm{H}[\rho(\vec{\boldsymbol{r}})] \, , \tag{2.19}$$

$$\text{where } u(\vec{\boldsymbol{r}}) = \sum_\alpha \frac{Z_\alpha}{|\vec{\boldsymbol{r}} - \vec{\boldsymbol{R}}_\alpha|} \, , \quad u_\mathrm{H}(\vec{\boldsymbol{r}}) = \int_V \mathrm{d}^3r' \, \frac{\rho(\vec{\boldsymbol{r}}')}{|\vec{\boldsymbol{r}}' - \vec{\boldsymbol{r}}|} \quad .$$

Using (2.16), (2.18) and (2.19) the Thomas–Fermi functional for the total energy is

$$E_\mathrm{TF}\left[\, \rho(\vec{\boldsymbol{r}}); u(\vec{\boldsymbol{r}}) \,\right] = C_\mathrm{F} \int_V \mathrm{d}^3r \, \rho^{\frac{5}{3}}(\vec{\boldsymbol{r}}) + \int_V \mathrm{d}^3r \, u(\vec{\boldsymbol{r}}) \, \rho(\vec{\boldsymbol{r}}) + \frac{1}{2} \int_V \mathrm{d}^3r \, u_\mathrm{H}(\vec{\boldsymbol{r}}) \, \rho(\vec{\boldsymbol{r}}) \, . \tag{2.20}$$

It is a functional explicitly given by the density $\rho(\vec{\boldsymbol{r}})$ and the external potential $u(\vec{\boldsymbol{r}})$. The factor $\frac{1}{2}$ in the last term prevents double counting. It is the so called *Hartree–term* V_{H} with the *Hartree potential* $u_{\mathrm{H}}(\vec{\boldsymbol{r}})$, which describes the classical Coulomb interaction of the electrons.
For a Coulomb system (2.20) becomes

$$\begin{aligned} E_{\mathrm{TF}}[\,\rho(\vec{\boldsymbol{r}});\vec{\boldsymbol{R}}_\alpha,\vec{\boldsymbol{R}}_\beta\,] =& E[\,\rho(\vec{\boldsymbol{r}})\,]+V_{\mathrm{nucl}}[\,\vec{\boldsymbol{R}}_\alpha,\vec{\boldsymbol{R}}_\beta\,]= \\ &= C_{\mathrm{F}}\int_V \mathrm{d}^3r\;\rho^{\frac{5}{3}}(\vec{\boldsymbol{r}}) - \sum_\alpha \int_V \mathrm{d}^3r\;\frac{Z_\alpha\,\rho(\vec{\boldsymbol{r}})}{|\vec{\boldsymbol{r}}-\vec{\boldsymbol{R}}_\alpha|}+ \\ &+\frac{1}{2}\iint_V \mathrm{d}^3r\,\mathrm{d}^3r'\;\frac{\rho(\vec{\boldsymbol{r}}')\,\rho(\vec{\boldsymbol{r}})}{|\vec{\boldsymbol{r}}'-\vec{\boldsymbol{r}}|}+\frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha\,Z_\beta}{|\vec{\boldsymbol{R}}_\alpha-\vec{\boldsymbol{R}}_\beta|}\,. \end{aligned} \tag{2.21}$$

The last term in the equation above is the Coulomb interaction energy of the nuclei. It is independent from the electron density and can therefore be separated in terms of a *Born–Oppenheimer approximation*.
In order to find the ground state, (2.20) will be used as a variational expression, where the density $\rho(\vec{\boldsymbol{r}})$ will be varied under the constraint that the number of particles N from (2.17) is conserved. It is the usual practice that one assumes a potential $u(\vec{\boldsymbol{r}})$ to be negative and to approach zero at infinity, which is a proper description in most of the cases

$$u(\vec{\boldsymbol{r}})\leq 0\,,\qquad \lim_{|\vec{\boldsymbol{r}}|\to\infty} u(\vec{\boldsymbol{r}})=0\,.$$

If a particle is added, the potential can either bind it $\left(E_{\mathrm{TF}}^{N+1}[u]<E_{\mathrm{TF}}^{N}[u]\right)$, or the particle disappears at infinity in an $E\to 0$ state. For the latter there would be no minimal $\rho(\vec{\boldsymbol{r}})$ for $(N+1)$ particles. But there is always a maximum number of particles $N_{\max}$ that the potential can bind, and hence the particle density minimises the functional $E_{\mathrm{TF}}[\rho;u]$ for *all* N, because the excess particles are disappearing at infinity. Precisely, the condition for variation has to be changed to

$$\int_V \mathrm{d}^3r\;\rho(\vec{\boldsymbol{r}})\leq N_{\max}\,. \tag{2.22}$$

In other words this expression states that a minimum is taken for some ρ, where

$$\int_V \mathrm{d}^3r\;\rho(\vec{\boldsymbol{r}})=N\,,\quad \forall N\in\mathbb{N}\,:\,N<N_{\max}\,,$$

and that a minimum is also taken for $N\geq N_{\max}$, but this time with

$$\int_V \mathrm{d}^3r\;\rho(\vec{\boldsymbol{r}})=N_{\max}\,,\quad \forall N\in\mathbb{N}\,:\,N\geq N_{\max}\,.$$

The variational principle with respect to ρ yields

$$\delta\left\{E_{\mathrm{TF}}[\,\rho(\vec{\boldsymbol{r}})\,]-\mu\left[\int_V \mathrm{d}^3r\;\rho(\vec{\boldsymbol{r}})-N\right]\right\}=0\,, \tag{2.23}$$

where μ is the *Lagrange parameter*, which can be identified as the chemical potential. The use of functional derivatives will be necessary in order to be able to determine the kinetic and potential

terms of (2.20). Per definition a *test function* $\varphi(\vec{r})$ has to be introduced,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon} F\left[\, f(\vec{r}) + \varepsilon\varphi(\vec{r}) \,\right]_{\varepsilon=0} &\doteq \lim_{\varepsilon\to 0}\left[\frac{F[f+\varepsilon\varphi] - F[f]}{\varepsilon}\right] = \\ &= \int \mathrm{d}^3 r\, \frac{\delta F}{\delta f(\vec{r})} \cdot \varphi(\vec{r})\end{aligned} \tag{2.24}$$

As an example the derivative of (2.16) can be determined by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon} T\left[\, \rho(\vec{r}) + \varepsilon\varphi(\vec{r}) \,\right]_{\varepsilon=0} &= \lim_{\varepsilon\to 0} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} C_\mathrm{F} \int \mathrm{d}^3 r\, \left[\, \rho(\vec{r}) + \varepsilon\varphi(\vec{r}) \,\right]^{\frac{5}{3}} = \\ &= \lim_{\varepsilon\to 0} \frac{5}{3} C_\mathrm{F} \int \mathrm{d}^3 r\, \left[\, \rho(\vec{r}) + \varepsilon\varphi(\vec{r}) \,\right]^{\frac{2}{3}} \cdot \varphi(\vec{r}) = \\ &= \frac{5}{3} C_\mathrm{F} \int_V \mathrm{d}^3 r\, \rho^{\frac{2}{3}}(\vec{r}) \cdot \varphi(\vec{r})\end{aligned}$$

$$\Longrightarrow \qquad \boxed{\frac{\delta T}{\delta\rho(\vec{r})} = \frac{5}{3} C_\mathrm{F}\, \rho^{\frac{2}{3}}(\vec{r})} \quad . \tag{2.25}$$

The derivatives of the potential terms can be calculated in an analogous way, but they are only dependent on the external potential $u(\vec{r})$ and the Hartree–potential $u_\mathrm{H}(\vec{r})$.
Using (2.25) in (2.23) results in

$$\mu = \frac{\delta E_\mathrm{TF}}{\delta\rho(\vec{r})} = \frac{5}{3} C_\mathrm{F}\, \rho^{\frac{2}{3}}(\vec{r}) + u(\vec{r}) + u_\mathrm{H}(\vec{r}) \,, \tag{2.26}$$

however, in order to ensure that the minimal density is found there has to be the restriction

$$\frac{5}{3} C_\mathrm{F}\, \rho^{\frac{2}{3}}(\vec{r}) = \max\left\{\mu - u(\vec{r}) - u_\mathrm{H}(\vec{r}) \,,\, 0\right\} \,, \tag{2.27a}$$

with the definiton

$$\left[\, \mu - u(\vec{r}) - u_\mathrm{H}(\vec{r}) \,\right]_+ \doteq \max\left\{\mu - u(\vec{r}) - u_\mathrm{H}(\vec{r}) \,,\, 0\right\}$$

$$\Longrightarrow \qquad \rho(\vec{r}) = \frac{2\sqrt{2}}{3\pi^2}\left[\, \mu - u(\vec{r}) - u_\mathrm{H}(\vec{r}) \right]_+^{\frac{3}{2}} . \tag{2.27b}$$

Equation (2.27a) is called the *Thomas–Fermi equation* which yields a relation between the electron density and the interacting potentials. The density $\rho(\vec{r})$ remains non-zero where $u(\vec{r}) + u_\mathrm{H}(\vec{r}) < \mu$. In (2.27b) the numeric expression of C_F has been entered. The Thomas–Fermi equation can be brought into a differential form by using the Poisson equation. The Hartree potential brings the relation

$$-\Delta u_\mathrm{H}(\vec{r}) = 4\pi\rho(\vec{r}) \,, \tag{2.28}$$

which together with (2.27b) yields

$$-\Delta u_\mathrm{H}(\vec{r}) = \frac{8\sqrt{2}}{3\pi}\left[\, \mu - u(\vec{r}) - u_\mathrm{H}(\vec{r}) \right]_+^{\frac{3}{2}} . \tag{2.29}$$

The latter equation shows the advantage of the Thomas–Fermi approximation; instead of solving the N particle Schrödinger equation (eigenvalue problem in $3N$ coordinates), approximative solutions of the ground state energy and the ground state density are now obtained by solving a three dimensional integral equation. In the following there will be a qualitative analysis of the Thomas–Fermi atom and a short discussion of the limits within this model.

2.2.1 The Thomas–Fermi Atom

Neutral Atom ($N = Z$) For a neutral atom one has,

$$N = Z\,, \qquad u(\vec{\boldsymbol{r}}) = -\frac{Z}{r} \quad \Longrightarrow \quad \Delta\, u(\vec{\boldsymbol{r}}) = 4\pi Z\, \delta(\vec{\boldsymbol{r}})\,. \tag{2.30}$$

In this case, the Thomas–Fermi equation (2.27) is solved by a spherical $\rho(r)$. For a region outside of a given radius r_0 we can assume

$$\rho(r) = 0\,, \quad r > r_0\,. \tag{2.31}$$

Thus, the potentials are retrieved by applying the laws of electrostatics,

$$u_{\text{eff}} \doteq u + u_{\text{H}} = 0\,, \quad r > r_0\,, \tag{2.32}$$

where the *effective potential* u_{eff} has been introduced. Entering these results into (2.27) yields

$$\mu = 0\,, \quad r > r_0\,. \tag{2.33}$$

Positively Charged Ion ($N < Z$) In this case $u_{\text{eff}} < 0$ for all $r > r_0$ and therefore,

$$\mu = u_{\text{eff}}(r_0) < 0\,. \tag{2.34}$$

Negatively Charged Ion ($N > Z$) Since electrostatics now deliver a positive effective potential $u_{\text{eff}} > 0, r > r_0$, where $\rho(r) = 0, r > r_0$, it is impossible to derive a finite r_0 according to (2.27). Furthermore, both the effective potential and the density approach to zero for $r \to \infty$, hence

$$\mu = 0\,, \quad \forall r \tag{2.35}$$

is the only possible solution.

For $N \geq Z$ (neutral atoms and negatively charged ions) the result was $\mu = 0$, so that by using (2.29) and (2.30) the Thomas–Fermi equation takes the form

$$\begin{aligned} -\Delta\, u_{\text{H}}(\vec{\boldsymbol{r}}) - \Delta\, u(\vec{\boldsymbol{r}}) &= \frac{8\sqrt{2}}{3\pi}\Big(- u_{\text{H}}(\vec{\boldsymbol{r}}) - u(\vec{\boldsymbol{r}})\Big)^{\frac{3}{2}} - 4\pi Z\, \delta(\vec{\boldsymbol{r}}) \\ -\Delta\, u_{\text{eff}}(\vec{\boldsymbol{r}}) &= \frac{8\sqrt{2}}{3\pi}\Big(- u_{\text{eff}}(\vec{\boldsymbol{r}})\Big)^{\frac{3}{2}} - 4\pi Z\, \delta(\vec{\boldsymbol{r}})\,. \end{aligned} \tag{2.36}$$

Assuming that the potentials are only dependent on r, and not on the angles (ϑ, ϕ), the Laplace operator becomes

$$\Delta \bullet = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \bullet}{\partial r}\right) + \mathcal{O}(\vartheta, \phi) = \frac{1}{r}\frac{\partial^2}{\partial r^2}\Big(r\ \bullet\Big) + \mathcal{O}(\vartheta, \phi)\,. \tag{2.37}$$

Hence, for $N \geq Z$ and $r \neq 0$ (2.36) reads

$$-\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\Big(r\,u_{\mathrm{eff}}(r)\Big) = \frac{8\sqrt{2}}{3\pi}\Big(-u_{\mathrm{eff}}(r)\Big)^{\frac{3}{2}} - 4\pi Z\,\delta(r)\,. \tag{2.38}$$

The delta function leads to two boundary conditions,

$$\begin{aligned} &(1) \qquad r \to 0 \;:\; r\,u_{\mathrm{eff}} \to -Z\,, \\ &(2) \qquad r \to \infty \;:\; u_{\mathrm{eff}} \to 0\,, \end{aligned}$$

and determines a unique solution for u_{eff} in (2.38). Condition (1) can be derived directly by integrating both sides of (2.36) over the boundary of a sphere with arbitrary r. For $r \to 0$ the delta function and the integration $\int_\Omega \sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\phi = 4\pi$ deliver the result for (1). For $r \to \infty$ the particle density $\rho(r)$ goes to zero in order to describe the physical situation. The density can be obtained by entering u_{eff} into the Poisson equation. These considerations justify the second boundary condition (2). For large r the solution of (2.38) is $u_{\mathrm{eff}} \sim r^{-4}$.

In a different approach (2.38) can be rewritten using Gauss' theorem. If $N(r)$ denotes the number of electrons inside the radius r, then this theorem states

$$r^2\frac{\mathrm{d}}{\mathrm{d}r}u_{\mathrm{eff}}(r) = Z - N(r) \quad \Longrightarrow \quad -\Delta\,u_{\mathrm{eff}}(r) = \frac{1}{r^2}\frac{\partial}{\partial r}\big(N(r) - Z\big)\,, \tag{2.39}$$

where the Laplace operator in spherical coordinates (2.37) has been used again. The only possible solution for (2.38) in this case is $Z - N = Z - N(\infty) = 0$. In other words, under the assumption $N \geq Z$ only $N = Z$ can be found as an adequate solution. The summary of these three cases implies that within the Thomas–Fermi Theory

- there is a *unique solution* for every neutral atom ($N = Z$),
- there is *always a finite* r_0 for positively charged ions ($N < Z$), and
- there is *no solution* for negatively charged ions ($N > Z$).

Using the ansatz

$$u_{\mathrm{eff}}(r) = -\frac{Z}{r}\chi(\alpha r)\,,$$

the Thomas–Fermi equation for the neutral atom can be transferred into a *universal equation*

$$\frac{\mathrm{d}^2\chi(x)}{\mathrm{d}x^2} = \frac{1}{x^{\frac{1}{2}}}\chi(x)^{\frac{3}{2}}\,, \quad x = \alpha r\,. \tag{2.40}$$

It is $\chi(0) = 1$ and $\chi(\infty) = 0$. The universal equation (2.40) can only be solved numerically. Without proof, the *Thomas–Fermi electron density* $\rho(r)$ can be derived as

$$\boxed{\rho(r) = \frac{32}{9\pi^3}\left[\frac{\chi(\alpha r)}{\alpha r}\right]^{\frac{3}{2}} Z^2} \quad , \tag{2.41}$$

$$\begin{aligned} r \to 0 \;&:\; \rho(r) \to \sim r^{-3/2} \\ r \to \infty \;&:\; \rho(r) \to \sim r^{-6} \end{aligned}$$

and therefore the *Thomas–Fermi energy* of the atom becomes

$$E_{\mathrm{TF}}[Z] \approx -0.7687\, Z^{\frac{7}{3}} \,. \tag{2.42}$$

This value is approximately 55 % lower than the one expected for hydrogen, and about 15 % too low for heavy elements ($Z \approx 100$), compared to the exact quantum mechanical result $E_n = \frac{Z^2}{n^2} E_{\mathrm{R}}$. In this case (natural units) the *Rydberg constant* $E_{\mathrm{R}} = -0.5$. The monotonous function $\chi(\alpha r)$ leads to a monotonous charge density. There is no indication of a *shell structure of the atom* in the Thomas–Fermi model and the charge density diverges in a wrong asymptotic behaviour (since there is no exponential falloff). This "failure" is corrected by the theorems of Hohenberg and Kohn.

2.3 Hohenberg–Kohn Theory

The theory of Thomas and Fermi was a rather simply figured aim to derive a first description of the structure of heavier atoms when the new apparatus of quantum physics was applied. The first results were quite promising. It was about 30 years later when *Teller* stated with his analytical proof that there is no possibility of a chemical binding in the Thomas–Fermi theory (without the Weizsäcker term). Even though this was meant to be a destructive verdict to their quite pragmatic approach, it was maybe due to Teller himself that many scientists kept an interest in that topic. Teller's theorem had to be confirmed, and as a consequence many important results have been derived from these activities.

The first method to deliver promising results was Slater's Xα method and considered as an approach from quantum chemistry and its closely related Hartree–Fock theory. In his attempt to interpret the role of the exchange interaction in metals he assumed that it is possible to approximate the latter by a potential averaged over the occupied states of a homogeneous electron gas. The *Slater exchange potential* then only depends on the local electron density pointed at by the vector $\vec{\boldsymbol{r}}$, which in full terms reads

$$u_{\mathrm{xcS}}\big[\rho\big] = -3e^2 \left(\frac{3}{8\pi}\right)^{\frac{1}{3}} \rho^{\frac{1}{3}}(\vec{\boldsymbol{r}}) \,. \tag{2.43}$$

The main parameter, however, is the correction factor α, which is multiplied by the Slater exchange term u_{xcS} in order to meet the respective Hartree–Fock value of the total energy of an isolated atom. It appears that again a rather pragmatic step turned out to become most significant for the success of this method.

The biggest impact on the understanding of the electronic structure of solids has been made with the development of the *Density Functional Theory* (DFT). Formally based on the two theorems of Walter Kohn and Pierre C. Hohenberg, as well as a decisive measure Kohn elaborated in cooperation with Lu J. Sham, this formalism instantly offered a much broader theoretical basis and hence created many ways to generalise the Thomas–Fermi and the Xα approach. The first theorem basically states that the ground state energy can be uniquely determined by a universal functional of the electron density $\rho(\vec{\boldsymbol{r}})$. The second theorem is related to the variation of the total energy with respect to the electron density, which leads to an effective one–electron Schrödinger equation. Several applications of this theory are derived by making reasonable assumptions and compromises; some of which will be presented in this chapter. The gap between the rigorous execution of the theory and its pragmatic execution will probably never be closed.

2.3.1 The Theorems of Hohenberg and Kohn

The following derivations assume a fixed particle number N and Hamiltonians as shown in Sec.2.1 [see for example (2.7)]. It is important to emphasise that minimising the energy functional as done in (2.4) leads to a ground state energy E_0 as a lower limit ($E_0 \leq E[\psi]$). A given Hamiltonian with a certain external potential $U[N, u(\vec{r})]$ — $\hat{T}$ and $\hat{W}$ only depend on the particle number N — yields a ground state wave function ψ_0 and thereafter a ground state density ρ_0 can be determined only if a *non–degenerated* ground state is assumed,

$$N, u(\vec{r}) \to \hat{\mathcal{H}} \to \psi_0 \to \rho_0 \quad \Longrightarrow \quad E = E[N, u(\vec{r})] \,.$$

Therefore, the energy functional depends on N and $u(\vec{r})$. Clearly, this is not desired within the DFT. If it can be shown that both the number of particles and the external potential can be uniquely expressed by the particle density $\rho(\vec{r})$, the energy functional would only depend on one variable and a $3N$ dimensional problem could be reduced to a single 3 dimensional one. This is where the first theorem of Hohenberg and Kohn can serve as an answer [1].

Theorem I

The determination of the particle number N is trivial. By the nature of the particle density,

$$N = \int_V \mathrm{d}^3 r \, \rho(\vec{r}) \,,$$

is the obvious result. The second problem of showing that an external potential $u(\vec{r})$ can be traced back by the according density in a unique fashion is proved by *reductio ad absurdum*. In other words, a proposition is shown to be true if the opposite of this proposition will always lead to a contradiction.

One assumes that for two different external potentials $u_1(\vec{r})$ and $u_2(\vec{r})$ ($u_1 \neq u_2$) the Hamiltonians $\hat{\mathcal{H}}[u_i]$ yield two different ground state energies $E_0[u_i]$ with the according wave functions $\psi_0[u_i] \doteq \psi_i$. Both states, however, derive the *same* ground state density $\rho_0(\vec{r})$. Hard potential barriers are excluded and therefore ψ_i is *non-zero* where the potentials are different, so that ψ_2 is not an eigenstate of the Schrödinger equation in a potential u_1 and vice versa. Hence, the general variation principle leads to an inequality in both cases,

$$\begin{aligned} E_0[u_1] < \langle\psi_2|\,\hat{\mathcal{H}}[u_1]\,|\psi_2\rangle &= \langle\psi_2|\,\hat{\mathcal{H}}[u_2]\,|\psi_2\rangle + \langle\psi_2|\left(\hat{\mathcal{H}}[u_1] - \hat{\mathcal{H}}[u_2]\right)|\psi_2\rangle = \\ &= E_0[u_2] + \int \mathrm{d}^3 r \, \rho_2(\vec{r}) \left(u_1(\vec{r}) - u_2(\vec{r})\right) \,, \end{aligned} \tag{2.44a}$$

$$\begin{aligned} E_0[u_2] < \langle\psi_1|\,\hat{\mathcal{H}}[u_2]\,|\psi_1\rangle &= \langle\psi_1|\,\hat{\mathcal{H}}[u_1]\,|\psi_1\rangle - \langle\psi_1|\left(\hat{\mathcal{H}}[u_1] - \hat{\mathcal{H}}[u_2]\right)|\psi_1\rangle = \\ &= E_0[u_1] - \int \mathrm{d}^3 r \, \rho_1(\vec{r}) \left(u_1(\vec{r}) - u_2(\vec{r})\right) \,. \end{aligned} \tag{2.44b}$$

After building the sum over both equations and using the initial assumption that $\rho_0[u_1] = \rho_0[u_2] = \rho_0$, one derives the presumed *contradiction*,

$$E_0[u_1] + E_0[u_2] < E_0[u_2] + E_0[u_1] \,. \tag{2.45}$$

The ground state energy E_0 of a system with N particles and an external potential $u(\vec{r})$ can be determined by a unique functional of the particle density $\rho(\vec{r})$, so that $E_0 = E[\rho_0(\vec{r})]$.

Theorem II

As stated before, the kinetic energy T and the electron interaction term W do not depend on $u(\vec{r})$. These term can thus be unified to one general expression in the form of a functional known as the universal *Hohenberg–Kohn functional*. It is defined as

$$F_{\mathrm{HK}}[\rho_0] \doteq \langle\psi_0|\,\hat{T} + \hat{W}\,|\psi_0\rangle \ . \tag{2.46a}$$

Equation (2.46a) only holds for non–degenerated ground states. In that case the mapping between ρ_0 and $u(\vec{r})$ is in equal relation (1:1) and, consequently, this is also true for the mapping between ρ_0 and the ground state wave function $\psi_0[u(\vec{r})]$. However, as degeneracy of the ground state is quite common one has to find a more general expression. Therefore, the previous assumption of non–degeneracy is from now on neglected and

$$\begin{aligned} F_{\mathrm{HK}}[\rho] &\doteq E_u[\rho] - \int \mathrm{d}^3r\ \rho(\vec{r})\, u(\vec{r}) \quad , \\ \text{where} \quad E_u[\rho] &= \int \mathrm{d}^3r\ \rho(\vec{r})\, \hat{\mathcal{H}}[u] \quad , \end{aligned} \tag{2.46b}$$

will be used as a basic definition of the Hohenberg–Kohn functional.

As a starting point of the theorem consider

$$E_{\tilde{u}}[\rho] = F_{\mathrm{HK}}[\rho] + \int \mathrm{d}^3r\ \rho(\vec{r})\, \tilde{u}(\vec{r}) \ ,$$

as a functional of the two *independent* variables ρ and an *arbitrary* $\tilde{u} \neq u$, where it is assumed that ρ is a *ground state density* for a system with potential u. The following yields

$$\begin{aligned} E_{\tilde{u}}[\rho] &= F_{\mathrm{HK}}[\rho] + \int \mathrm{d}^3r\ \rho(\vec{r})\, u(\vec{r}) + \int \mathrm{d}^3r\ \rho(\vec{r}) \Big(\tilde{u}(\vec{r}) - u(\vec{r})\Big) = \\ &= E_0[\rho] + \int \mathrm{d}^3r\ \rho(\vec{r}) \Big(\tilde{u} - u\Big) = \\ &= \langle\psi_0|\,\hat{\mathcal{H}}[u]\,|\psi_0\rangle + \int \mathrm{d}^3r\ \rho(\vec{r}) \Big(\tilde{u} - u\Big) = \\ &= \langle\psi_0|\,\hat{\mathcal{H}}[\tilde{u}]\,|\psi_0\rangle \\ \implies \quad & E_{\tilde{u}}[\rho] \geq E_0[\rho] = E_u[\rho_0] \quad . \end{aligned} \tag{2.47}$$

The inequality derives from the general variational principle of quantum mechanics and follows since ρ is the ground state density ρ_0 for a system in an external potential u. Hence, the total energy can be minimised as a functional $E_0 = \min E_u[\rho] = E_u[\rho_0]$, however, (2.47) only

holds under the condition that

$$\delta\left\{ E_u[\rho(\vec{\boldsymbol{r}})] - \mu\left(\int_V \mathrm{d}^3 r\, \rho(\vec{\boldsymbol{r}}) - N\right)\right\} = 0$$

$$\Longrightarrow \qquad \mu = \frac{\delta E_u[\rho]}{\delta\rho(\vec{\boldsymbol{r}})} = \frac{\delta F_{\mathrm{HK}}[\rho]}{\delta\rho(\vec{\boldsymbol{r}})} + u(\vec{\boldsymbol{r}}) \quad . \tag{2.48}$$

The latter is the Hohenberg–Kohn variational principle. μ serves as the Lagrange parameter and can be identified as the *chemical potential*. The problem of the Hohenberg–Kohn theory becomes obvious now. The density, and thus the ground state energy, could be easily derived by (2.48) for any known potential $u(\vec{\boldsymbol{r}})$ if only the functional $F_{\mathrm{HK}}[\rho]$ would be explicitly known. Unfortunately, this is not the case and many attempts to solve (2.48) directly have remained unsuccessful. However, the second theorem generally guarantees the *existence* of the functional.

2.3.2 Kohn–Sham Equations

Sec.2.3.1 provides a solid proof for the existence of an energy functional depending on the density matrix that is rigorously based on quantum theory. It does not, however, offer any explicit values. Some of the most important contributions to modern DFT have been made by adaptation and amendment of the many–fermion theory. The interpretation of the density as a basis set of *electron orbitals*, as well as using the occurring exchange energy term as a possibility for several approximations, has made DFT a powerful tool to tackle the problems of modern computational solid state physics.

The basic principles of this chapter are tightly related to the Hartree–Fock theory (see Apx.D). In Sec.2.2 it was shown that the weakest part of the Thomas–Fermi theory was the treatment of the kinetic energy functional $T[\rho]$. Kohn and Sham fixed this problem by mapping a *non–interacting* system to a fully interacting one; they were able to treat the whole interaction itself in an *effective potential*. The advantage of the approach from a non–interacting system is that all orbitals ϕ_i $(i = 1, \ldots, N)$ are derived by a single–particle Hamiltonian. Hence, they are independent of one another and the general wave function is a single (Slater-)determinant. This method is less elaborate to implement into algorithms than the Hartree–Fock approximation, for example, and has led to the great success of programs based on DFT. The ingenious part of the Kohn–Sham equations is the treatment of the exchange– and correlation potential of the electron–electron interaction [2].

First, one considers an *interaction free* N particle system ($\hat{W} = 0$). The Hohenberg–Kohn functional (2.46) is identical to the kinetic energy functional when it is considered for the ground state density of the system,

$$F_{\mathrm{HK}}[\rho_0] \equiv \qquad T[\rho_0] = E_u^0[\rho_0] - \int \mathrm{d}^3 r\, \rho_0(\vec{\boldsymbol{r}})\, u(\vec{\boldsymbol{r}})\,. \tag{2.49}$$

Still the explicit form of this functional is not known but, as it was stated before, the second Hohenberg–Kohn theorem grants its existence. For a system of N non–interacting fermions there are always single particle states $\phi_i(\vec{\boldsymbol{r}})$ (*orbitals*) which deliver an *exact solution* of the ground state when they are composed in form of a *single* (Slater-)determinant. From the general expression

of the quantum mechanical probability density an ansatz for the ground state density of the system can be derived,

$$\rho(\vec{\boldsymbol{r}}) = \lim_{\tilde{N}\to\infty} \sum_{i=1}^{\tilde{N}} n_i\, \phi_i^*(\vec{\boldsymbol{r}})\phi_i(\vec{\boldsymbol{r}}) \quad \longrightarrow \quad \sum_{i=1}^{N} \phi_i^*(\vec{\boldsymbol{r}})\phi_i(\vec{\boldsymbol{r}})\,, \tag{2.50}$$

$$0 \le n_i \le 1 \quad \text{(occupation numbers)} \quad n_i = 1$$

$$\text{only if } \Psi_0 = \frac{1}{\sqrt{N!}} \det \|\phi_i(\vec{\boldsymbol{r}}_j)\|\,, \quad \langle\phi_i|\phi_j\rangle = \delta_{ij}\,.$$

The orbital states are eigenstates of the kinetic energy operator $\hat{T} \propto \hat{p}^2$. Under the same conditions as before, the kinetic energy of the system is the sum over the single eigenvalues,

$$T[\rho] = \langle\Psi_0|\,\hat{T}\,|\Psi_0\rangle = -\frac{1}{2}\sum_{i=1}^{N} \langle\phi_i|\vec{\boldsymbol{\nabla}}|\phi_i\rangle\,. \tag{2.51a}$$

Likewise, the density ρ in the integral of the external potential U can also be replaced by,

$$U[\rho] = \int \mathrm{d}^3r\, \rho(\vec{\boldsymbol{r}})\, u(\vec{\boldsymbol{r}}) = \sum_{i=1}^{N} \langle\phi_i|u|\phi_i\rangle \tag{2.51b}$$

These two expressions can be entered into (2.49) to form an ansatz for the HK variational principle as in (2.48), in order to calculate the total energy $E_u[\rho_0] = E_0$ of the ground state for a set of determinant states $\{\phi_i\}$ that minimise the energy. The side conditions now are provided by the normalisation of the orbitals. Note that the Hermitian Lagrange multipliers ε_{ij} have already been diagonalised by a unitary transformation and $\varepsilon_i \doteq \varepsilon_{ii}$,

$$\delta\left\{\sum_{i=1}^{N}\left(-\frac{1}{2}\langle\phi_i|\vec{\boldsymbol{\nabla}}^2|\phi_i\rangle + \langle\phi_i|u|\phi_i\rangle - \varepsilon_i\big(\langle\phi_i|\phi_i\rangle - 1\big)\right)\right\} = 0\,. \tag{2.52}$$

It is sufficient to only vary over the imaginary part $\langle\phi_i| = \phi_i^*(\vec{\boldsymbol{r}})$, and furthermore, each term i of the sum can be treated separately. This yields a *single particle* Schrödinger equation for the N orbitals ϕ_i, which are the *exact lowest eigenstates of a non-interacting system*,

$$\hat{h}^0\,\phi_i = \left(-\frac{1}{2}\vec{\boldsymbol{\nabla}}^2 + u(\vec{\boldsymbol{r}})\right)\phi_i = \varepsilon_i\,\phi_i \qquad (i = 1,\ldots,N)\,. \tag{2.53}$$

Taking both equations of (2.51) into account, one can easily derive an expression for the (ground state) energy functional by the following actions,

$$\sum_i \int \mathrm{d}^3r \quad \phi_i^*(\vec{\boldsymbol{r}}) \cdot (2.53)$$

$$\Longrightarrow \qquad E_u[\rho_0] = T[\rho_0] + U[\rho_0] = \sum_{i=1}^{N} \varepsilon_i \qquad (\varepsilon_1 \le \varepsilon_2 \le \ldots)\,. \tag{2.54}$$

For the *interacting system* ($\hat{W} \neq 0$) a rearrangement of the Hohenberg–Kohn functional is necessary,

$$F_{\mathrm{HK}}[\rho] = T[\rho] + W[\rho] = T[\rho] + \frac{1}{2}V_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho]\,, \tag{2.55}$$

where V_{H} is the Hartree energy as in (2.19) which describes the classical part of the interaction. The last expression E_{xc} is the *exchange- and correlation energy* which contains *any other* energy contribution caused by quantum effects of the interaction of indistinguishable particles. Composing the terms in (2.55) differently reveals

$$E_{\mathrm{xc}}[\rho] = \Big(W - \frac{1}{2}V_{\mathrm{H}}\Big)[\rho] = -\frac{1}{2}V_{\mathrm{F}}[\rho]\,, \tag{2.56}$$

which can be interpreted as the non-classical Fock (exchange) energy. If (2.55) is used in (2.48), the Lagrange parameter of the Hohenberg–Kohn variational principle reads

$$\begin{aligned}\mu &= \frac{\delta T[\rho]}{\delta\rho} + \frac{1}{2}\,\frac{\delta V_{\mathrm{H}}[\rho]}{\delta\rho} + \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho} + u(\vec{\boldsymbol{r}}) \\ &= \frac{\delta T[\rho]}{\delta\rho} + \underbrace{u_{\mathrm{H}}(\vec{\boldsymbol{r}}) + u_{\mathrm{xc}}(\vec{\boldsymbol{r}}) + u(\vec{\boldsymbol{r}})}_{=u_{\mathrm{eff}}(\vec{\boldsymbol{r}})}\,.\end{aligned} \tag{2.57}$$

The latter expression contains the *exchange- and correlation potential* $u_{\mathrm{xc}}(\vec{\boldsymbol{r}})$, which is defined as

$$u_{\mathrm{xc}}(\vec{\boldsymbol{r}}) \doteq \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho}\,. \tag{2.58}$$

Note that the factor $\frac{1}{2}$ of the Hartree term vanishes, due to the derivative of the biquadratic density expression when the Hartree term is explicitly written out. The problem is that the functional for the kinetic energy $T[\rho]$ is unknown. One thus again falls back to the variation of a set of (determinant) single states which, similar to (2.52), reads

$$\delta\Bigg\{\sum_{i=1}^{N}\Big(-\frac{1}{2}\,\langle\phi_i|\vec{\boldsymbol{\nabla}}^2|\phi_i\rangle \;+\; \langle\phi_i|u_{\mathrm{eff}}|\phi_i\rangle \;-\varepsilon_i\big(\,\langle\phi_i|\phi_i\rangle \;-1\big)\Big)\Bigg\} = 0\,, \tag{2.59}$$

with the decisive difference that now it is applied to a system of non-interacting electrons moving in a (given) *effective potential* u_{eff}. The corresponding single particle Schrödinger equations,

$$\hat{h}\,\phi_i = \Big(-\frac{1}{2}\vec{\boldsymbol{\nabla}}^2 + u_{\mathrm{eff}}(\vec{\boldsymbol{r}})\Big)\,\phi_i = \varepsilon_i\,\phi_i \qquad (i = 1,\ldots,N)\;, \tag{2.60}$$

are called *Kohn–Sham equations*, which allow to compose the ground state particle density using the optimised *Kohn–Sham orbitals*,

$$\rho_0(\vec{\boldsymbol{r}}) = \sum_i |\phi_i(\vec{\boldsymbol{r}})|^2\,. \tag{2.61}$$

It should be noted that the individual eigenfunctions ϕ_i and according eigenvalues ε_i in (2.60) as a consequence of the mapping to a non-interacting system do not have any direct physical meaning. Without any approximation of the exchange term and further considerations an explicit

expression for total energy of the ground state cannot be derived. However, it is still sensible to calculate the sum over the eigenvalues analogous to (2.54), which gives

$$\sum_i \int \mathrm{d}^3 r \quad \phi_i^*(\vec{\boldsymbol{r}}) \cdot (2.60)$$

$$T[\rho_0] + U[\rho_0] + V_\mathrm{H}[\rho_0] + \int \mathrm{d}^3 r\, \rho_0(\vec{\boldsymbol{r}})\, u_\mathrm{xc}(\vec{\boldsymbol{r}}) = \sum_{i=1}^{N} \varepsilon_i$$

$$\Longrightarrow \qquad E_u[\rho_0] = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} V_\mathrm{H}[\rho_0] + E_\mathrm{xc}[\rho_0] - \int \mathrm{d}^3 r\, \rho_0(\vec{\boldsymbol{r}})\, u_\mathrm{xc}(\vec{\boldsymbol{r}})\,, \tag{2.62}$$

where the total energy expression of (2.46) together with (2.55) has been compared to the sum of eigenvalues. The equation in (2.60) describes a non-linear problem because the effective potential u_eff depends on the density ρ by (2.58). Therefore, the Kohn–Sham algorithm is a *self–consistent* method, similar to the Hartree–Fock method. On the contrary, however, it still offers an *exact* solution because the many–body interactions, as well as exchange and correlation effects, are fully integrated into the Hamiltonian. Unfortunately, the missing link to an exact DFT calculation remains unrevealed due to the lacking knowledge of the precise functional E_xc and the potential u_xc, where the complexity of many–electron systems is entirely hidden (the only exception is the free electron gas). The introduction of an effective potential u_eff and the thereby arising simple form of the Kohn–Sham equations are an indisputable advantage in the computational implementation which led to the known success of DFT. The major difference in quality of nowadays available programs is to what extent the functional E_xc and its derivative have been approximated in the source code.

2.3.3 Approximations of the Exchange and Correlation Functional

As previously mentioned in Sec.2.3.2, the advantage of DFT is that its treatment is rigorously based on many–body quantum theory. It should be clear now that the Thomas–Fermi theory (see Sec.2.2) is a crude but *explicit* approximation of the Kohn–Sham equations. The key to an explicit formulation of the latter is to model the exchange and correlation energy according to the specific problem and probe those models to the phenomenology by comparison. Handling the formalism of Kohn and Sham in the spirit of the overly pragmatic Thomas–Fermi approach leads to the *Local Density Approximation*, which can be considered as the starting point of all other approximative variants of DFT.

Local Density Approximation (LDA) This case is based upon the assumption of a *slow and weak spatial variaton* of the density in an inhomogeneous assembly, which can be locally approximated by the well–known model system of a *homogeneous electron gas*. Let $\varepsilon_\mathrm{xc} = \varepsilon - \varepsilon_\mathrm{H}$ be the exchange and correlation energy per particle of a homogeneous electron gas, then $\rho(\vec{\boldsymbol{r}})\, \varepsilon_\mathrm{xc}[\rho]$ is the energy per unit volume and we derive

$$E_\mathrm{xc} \approx \int \mathrm{d}^3 r\, \rho(\vec{\boldsymbol{r}})\, \varepsilon_\mathrm{xc}[\rho(\vec{\boldsymbol{r}})] \doteq E_\mathrm{xc}^{\mathrm{LDA}}[\rho]\,. \tag{2.63}$$

The Kohn–Sham potential can be calculated by its definition (2.58) and the functional derivative now reads,

$$u_{\mathrm{xc}} \approx \frac{\delta}{\delta\rho(\vec{\boldsymbol{r}})}\Big(\rho(\vec{\boldsymbol{r}})\,\varepsilon_{\mathrm{xc}}[\rho]\Big) = \varepsilon_{\mathrm{xc}}[\rho] + \rho(\vec{\boldsymbol{r}})\,\frac{\delta\varepsilon_{\mathrm{xc}}}{\delta\rho(\vec{\boldsymbol{r}})}[\rho] \doteq u_{\mathrm{xc}}^{\mathrm{LDA}}\,. \tag{2.64}$$

The main task within the LDA is to find a proper expression for $\varepsilon_{\mathrm{xc}}$ and set the parameters that are according to the described system. As one of many examples an approximation formula by Gunnarsson and Lundqvist is shown [3, 4]. They executed long numerical calculations focusing on the properties of every individual electron in a homogeneous electron gas and, subsequently, fitted their numerical data to a simple analytic function, which reads

$$\varepsilon_{\mathrm{xc}}[\rho] = \alpha\,\frac{1}{r_{\mathrm{S}}} - \frac{1}{2}\beta\,G\big(\tfrac{r_{\mathrm{S}}}{11.4}\big)\,, \tag{2.65}$$

$$\text{where } \alpha = -\tfrac{1}{2}\,\tfrac{3}{2\pi\alpha'} = -0.4581\,,\ \beta = 0.0666\,,\ \text{and}$$
$$G(x) = \Big[\Big(1+x^3\Big)\ln\Big(1+\tfrac{1}{x}\Big) - x^2 + \tfrac{x}{2} - \tfrac{1}{3}\Big]\,.$$

$\alpha' = \big(\frac{4}{9\pi}\big)^{\frac{1}{3}}$ is a numerical constant, and r_{S} is the Wigner–Seitz radius of the homogeneous electron gas. Using (2.64) and the fact that

$$\rho(r_{\mathrm{S}}) = \frac{1}{V_{\mathrm{S}}} = \frac{1}{\frac{4\pi}{3}r_{\mathrm{S}}^3} \longrightarrow \frac{\partial}{\partial\rho} = \frac{\partial r_{\mathrm{S}}}{\partial\rho}\frac{\partial}{\partial r_{\mathrm{S}}} = -\frac{4\pi r_{\mathrm{S}}^4}{9}\frac{\partial}{\partial r_{\mathrm{S}}}$$

$$\Longrightarrow \qquad \rho(r_{\mathrm{S}})\frac{\partial}{\partial\rho} = -\frac{1}{3}\,r_{\mathrm{S}}\,\frac{\partial}{\partial r_{\mathrm{S}}}\,, \tag{2.66}$$

the Kohn–Sham potential reads

$$u_{\mathrm{xc}}(r_{\mathrm{S}}) = \frac{4}{3}\alpha\,\frac{1}{r_{\mathrm{S}}} - \frac{1}{2}\beta\,\ln\Big(1+\tfrac{11.4}{r_{\mathrm{S}}}\Big)\,. \tag{2.67}$$

Calculations based on LDA reveal that the approximation can be used as a powerful tool. Sometimes this is also true in systems where the assumption of a slowly varying density is not fulfilled. Further derivatives of this kind are known as the LSDA, a generalisation that also considers the electron–spin, and LDA+U, which offers a freely adjustable interaction parameter, the *Hubbard* U to simulate an excitation gap in the band structure due to strong electron–electron correlation (e.g. *Mott insulators* cannot be described by pure mean–field theory). Naturally, with all approximations there are also configurations where LDA fails and should be avoided, namely systems with a small number of electrons ($N < 4$) that are spatially well separated from each other. Especially in the case of the hydrogen atom the Hartree–Fock theory delivers an effective potential which is equal to the external potential u. Hence, u_{H} and u_{xc} have to cancel each other out, however, this is not obtained by the approach of a homogeneous electron gas. The so–called *self interaction* is obviously incorrect and yields in an error of approximately 5 % for the hydrogen energy.

Generalized Gradient Approximations (GGA) A clear improvement to LDA, but nevertheless still a local approach, is implemented within the GGA [5–7]. This method also implements

the *derivative (gradient) of the density* and delivers a somewhat semi-local information. The exchange and correlation energy is again determined by

$$E_{\mathrm{xc}} \approx \int \mathrm{d}^3\vec{\boldsymbol{r}}\, \rho(\vec{\boldsymbol{r}})\, \varepsilon_{\mathrm{xc}}\big[\rho,\, |\vec{\boldsymbol{\nabla}}\rho|\big] \doteq E_{\mathrm{xc}}^{\mathrm{GGA}}\big[\rho,\, \vec{\boldsymbol{\nabla}}\rho\big] \,. \tag{2.68}$$

Functionals that also take the *second derivative* (Laplacian) of the electron density into account, and further include a kinetic energy density τ, are known as meta–GGA functionals,

$$\varepsilon_{\mathrm{xc}} \quad \longrightarrow \quad \varepsilon_{\mathrm{xc}}\big[\rho,\, |\vec{\boldsymbol{\nabla}}\rho|,\, \Delta\rho;\, \tau\big] \,.$$

There is no standard functional for GGA. The most commonly used are the PBE functional by Perdew, Burke and Ernzerhof for GGA [8–10], and (R)TPSS for meta–GGA [11, 12].

Hybrid (Hartree–Fock) Functionals The treatment of many–body systems is tightly associated with the Hartree–Fock approximation. In Apx.D a comprehensive description is offered concerning this matter. The derived Hartree–Fock equations (D.7) reveal an (exact) exchange term (Fock–term), which lowers the energy of the system. The latter is used in a mixture with the exchange and correlation terms of the previously discussed approximations in order to establish so-called *Hybrid Functionals* [13]. As an example the PBE0 functional is shown, which has been independently developed by Ernzerhof and Scuseria [14], as well as Adamo and Barone [15],

$$E_{\mathrm{xc}}^{\mathrm{PBE0}} = a\, V_{\mathrm{F}} + (1-a)\, E_{\mathrm{x}}^{\mathrm{GGA}} + E_{\mathrm{c}}^{\mathrm{GGA}} \,. \tag{2.69}$$

The PBE0 functional derived its name from being free of any fitted empirical parameter. Note that the correlation term is only considered by GGA. The exchange terms, however, are being combined by a *mixing parameter* $a = [0,1] \in \mathbb{R}$ (Standard for PBE0 is $a = 0.75$). Using this functional without any further considerations would require an immense computational effort because of the slow decrease of the Coulomb potential (r^{-1}) in the Fock–term. A more powerful hybrid functional has been introduced by Heyd, Scuseria and Ernzerhof (HSE) [16], with the assumption of a *screened Coulomb potential* for the exchange term(s). Furthermore, the exchange functionals are separated into *long ranging* (LR) and *short ranging* (SR) parts with a remarkably larger computational efficiency [17],

$$E_{\mathrm{xc}}^{\mathrm{HSE}} = a\, V_{\mathrm{F}}^{\,(\mathrm{SR})}(\omega) + (1-a)\, E_{\mathrm{x}}^{\mathrm{GGA}\ (\mathrm{SR})}(\omega) + E_{\mathrm{x}}^{\mathrm{GGA}\ (\mathrm{LR})}(\omega) + E_{\mathrm{c}}^{\mathrm{GGA}} \,. \tag{2.70}$$

The parameter ω describes the weight of short–ranged interaction acting on the particles, where a is again the mixing parameter (Standard values for the commonly used HSE06 are $a = 0.25$ and $\omega = 0.2$). For $\omega = 0$ the hybrid functional degenerates to PBE0. The use of these functionals for materials with strong Coulomb screening significantly improves the band gap values compared to theory.

2.4 Vienna Ab-Initio Simulation Package (VASP)

The **V**ienna **A**b-Initio **S**imulation **P**ackage (VASP) is a computer programme for atomic scale materials modelling, i.e. electronic structure calculations and quantum mechanical molecular dynamics (MD) from first principles [18].

VASP computes an approximate solution of the many–body Schrödinger equation, either within DFT solving the Kohn–Sham equations or within the Hartree–Fock (HF) approximation solving the Roothaan equations. Hybrid functionals, like the HSE06, are also implemented [19].

In VASP, major quantities like the one-electron orbitals, the electronic charge density and the local potential are expressed in a plane–wave basis set [20, 21]. The interactions between the electrons and ions are described using norm–conserving or ultrasoft pseudopotentials (PP) [22–24], lately implemented by a special type named projector–augmented–wave (PAW) method [25, 26].

2.4.1 Methods

The KS equations build the bridge from a complicated many–body wave function to a set of orthogonal single state orbitals, which of course drastically changes the computational demand of necessary grid points,

$$\begin{array}{ccc} \Psi(\vec{r}_1, \ldots, \vec{r}_N) & \longrightarrow & \left\{\phi_1(\vec{r}), \ldots, \phi_N(\vec{r})\right\} . \\ (\#\text{ of grid points})^N & & N \times (\#\text{ of grid points}) \end{array}$$

However, for the sake of efficiency there is the need for periodic boundary conditions, which are implemented by the use of lattice periodic Bloch functions

$$\phi(\vec{k}s\,;\,\vec{r}+\vec{R}) = \phi(\vec{k}s\,;\,\vec{r})\;\mathrm{e}^{i\vec{k}\vec{R}}\,,$$

$$\text{where}\quad \phi(\vec{k}s\,;\,\vec{r}) = \upsilon(\vec{k}s\,;\,\vec{r})\;\mathrm{e}^{i\vec{k}\vec{r}}\quad\text{and}\quad \upsilon(\vec{k}s\,;\,\vec{r}+\vec{R}) = \upsilon(\vec{k}s\,;\,\vec{r})\,.$$

All states are labelled by a Bloch vector $\vec{k}$, which usually lies within the first Brillouin zone of the reciprocal space lattice, and the *band index* s. Computing the charge density is achieved by integrating over the first Brillouin zone,

$$\rho(\vec{r}) = \frac{1}{V_{\mathrm{BZ}}}\sum_s \int_{\mathrm{BZ}} \mathrm{d}\vec{k}\, f(\vec{k}s)\,|\phi(\vec{k}s\,;\,\vec{r}|^2 \sim \sum_{\vec{k},s} w(\vec{k})\, f(\vec{k}s)\,|\phi(\vec{k}s\,;\,\vec{r}|^2\,\Delta\vec{k}\,, \tag{2.71}$$

where the integral has been replaced by a weighted sum under the assumption that the orbitals at Bloch vector $\vec{k}$ are close to one another and thus almost identical. As a consequence one derives a discrete set of $\vec{k}$ points that sample the reciprocal space for a number of bands that is of the same order as the number of electrons per unit cell (coarse grain sampling). The number of points can be arbitrarily set by the user (KPOINTS file) and it is recommended to choose a mesh that is centred by the Γ symmetry point of the crystal. This number can be further reduced by applying the according symmetry relations of the crystal in order to derive the *irreducible* number of $\vec{k}$ points (IBZ) with a proper weighting.

Plane Wave Basis Set In solid state physics it is common to express the solutions of the KS equations (2.60) and the derived properties of a crystal using *plane waves*. Hence, the cell periodic part of the Bloch functions is described in a basis set of plane waves. Then the treatment of these functions reduces to Fourier analysis. One thus expands

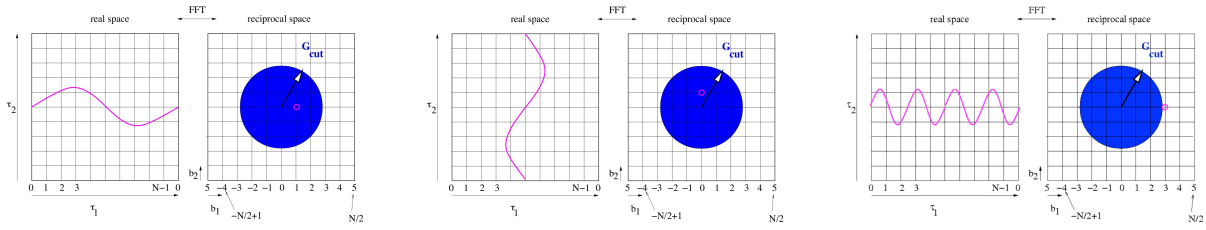


Fig. 2.1: Simplified depiction of a plane wave FFT with energy cutoff radius. (a) Plane wave in τ_1 direction. (b) Plane wave in τ_2 direction. (c) Plane wave with a kinetic energy larger than the energy cutoff. All pictures taken from [18, Documentation:VASP Workshop Lectures].

$$\upsilon(\vec{\boldsymbol{k}}s\,;\,\vec{\boldsymbol{r}}) = \frac{1}{\sqrt{V_{\mathrm{BZ}}}} \sum_{\vec{\boldsymbol{G}}} C(\vec{\boldsymbol{G}},\,\vec{\boldsymbol{k}}s)\; \mathrm{e}^{i\vec{\boldsymbol{G}}\vec{\boldsymbol{r}}} \quad \Longleftrightarrow \quad \phi(\vec{\boldsymbol{k}}s\,;\,\vec{\boldsymbol{r}}) = \frac{1}{\sqrt{V_{\mathrm{BZ}}}} \sum_{\vec{\boldsymbol{G}}} C(\vec{\boldsymbol{G}},\,\vec{\boldsymbol{k}}s)\; \mathrm{e}^{i(\vec{\boldsymbol{G}}+\vec{\boldsymbol{k}})\vec{\boldsymbol{r}}}\,, \tag{2.72}$$

but all other cell periodic properties can be expanded as well, for example

$$\rho(\vec{\boldsymbol{r}}) = \sum_{\vec{\boldsymbol{G}}} \rho_{\vec{\boldsymbol{G}}}\; \mathrm{e}^{i\vec{\boldsymbol{G}}\vec{\boldsymbol{r}}}\;. \tag{2.73}$$

In principle there is an infinite amount of plane waves, so that for computational purposes the basis set must be restricted to a *cutoff energy* E_{cut} including all plane waves for which

$$\frac{\hbar^2}{2}\,|\vec{\boldsymbol{G}}+\vec{\boldsymbol{k}}|^2 < E_{\mathrm{cut}}\;.$$

This property is the maximal kinetic energy associated with a plane wave of wave vector $\left(\vec{\boldsymbol{G}}+\vec{\boldsymbol{k}}\right)$. It is again chosen by the user (ENCUT tag in INCAR file) and needs to be carefully tested for every system. It turns out that there are operations where the evaluation becomes more convenient in real space and others that find an easier application in reciprocal space. The main advantage of plane waves becomes obvious as there is a computationally convenient mapping of those two spaces by *fast Fourier transformation* (FFT) that scales with $t \sim N \ln N$,

$$C(\vec{\boldsymbol{r}},\,\vec{\boldsymbol{k}}s) = \sum_{\vec{\boldsymbol{G}}} C(\vec{\boldsymbol{G}},\,\vec{\boldsymbol{k}}s)\; \mathrm{e}^{i\vec{\boldsymbol{G}}\vec{\boldsymbol{r}}} \tag{2.74a}$$

$$\Updownarrow$$

$$C(\vec{\boldsymbol{G}},\,\vec{\boldsymbol{k}}s) = \frac{1}{N} \sum_{\vec{\boldsymbol{r}}} C(\vec{\boldsymbol{r}},\,\vec{\boldsymbol{k}}s)\; \mathrm{e}^{-i\vec{\boldsymbol{G}}\vec{\boldsymbol{r}}}\quad, \tag{2.74b}$$

where N is the number of plane waves within the FFT grid (unit cell). The *action* of the Hamiltonian on the orbitals can thus be efficiently evaluated. Another reason to use plane waves is that many elements exhibit a band structure similar to the free electron picture, for example those which are metallic in the s and p shell [21], and that energy expressions become practically easy to implement into a programme.

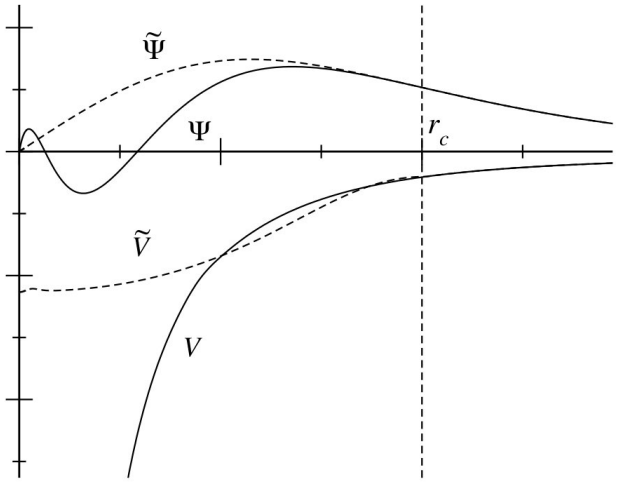


Fig. 2.2: General principle of a pseudopotential. The nodal features close to the nucleus of an orbital wave function Ψ caused by a potential V are replaced by a non-nodal orbital $\tilde{\Psi}$ with the according (pseudo)potential $\tilde{V}$. Beyond the cutoff radius r_c, where the interaction with other orbitals and thus the chemical bonding should occur, the two wave functions and potentials are (in the ideal case) identical, respectively. Taken from [18, Documentation:VASP Workshop Lectures].

Projector–Augmented–Wave Method (PAW) The number of plane waves that are needed in order to describe spatially strongly localised states, i.e. $3s$ and beyond, or rapid oscillations of the orbitals near the nucleus would be (except for the very first light elements as for example, hydrogen) prohibitively large. However, there are methods that are commonly applied in computational physics to escape this situation. The first solution is a process known as *frozen core approximation*, where the core electrons are pre-calculated and remain in the same (frozen) state creating an effective potential during the remaining calculations. The justification of this method is that in a heavy element the inner electrons are well shielded by the outer ones and do not contribute to the chemical binding. The second approach builds up on the former by introducing the idea of (norm–conserving or ultrasoft) *pseudopotentials* [22–24] that act around the atom instead of using an exact potential in order to neglect the capacity increasing, but often not necessary rapid oscillations near the nucleus. A sketch of the general idea born with pseudo–wave functions and the according potentials is described in Fig.2.2. The problem of pseudopotential theory in general is that the element is reduced to only have effective (nodeless) valence states, where orthogonality to the (frozen) core states is not guaranteed anymore. However, there is a special kind of pseudopotential theory, which still incorporates the nodal feature of the wave functions, known as the *Projector–Augmented–Wave* method by Blöchl [25] and is implemented in VASP.

The essential part of the PAW method is the representation of the true orbital ψ_n in two different bases,

$$|\psi_n\rangle = |\tilde{\psi}_n\rangle + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle \right) \langle \tilde{p}_i | \tilde{\psi}_n \rangle \ . \tag{2.75}$$

The first term $\tilde{\psi}_n$ is the pseudo–orbital as explained in Fig.2.2 and is expressed in the basis set of plane waves. The second part is built up by *local functions* around the according atomic sites

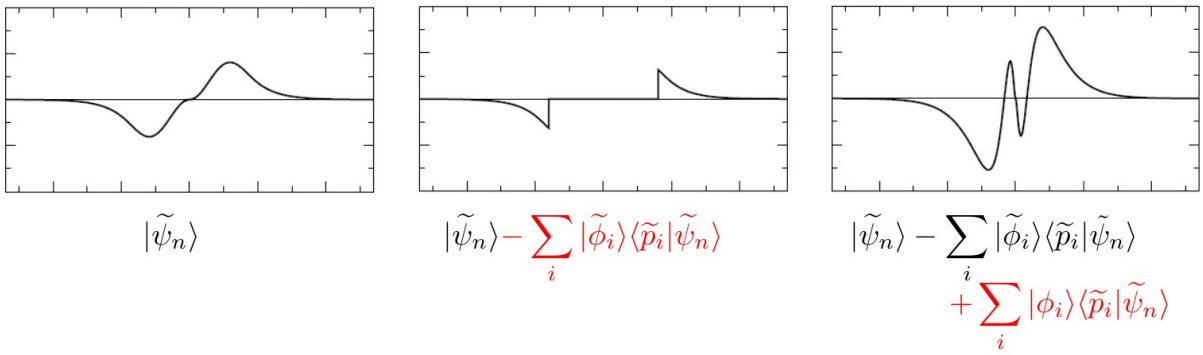


Fig. 2.3: The PAW method in three steps. The true orbital is split according to (2.75) first into a pseudo–wave function described in plane waves, which is the variational part to be determined by VASP calculations *(left)*. Within a certain radius around the atomic site all plane–wave features are cancelled out by the subtraction of pseudo partial–waves $\tilde{\phi}_i$, which are described by an additional complete basis set *(center)*, and finally nodal features associated with the true atomic wave function are locally added by the all–electron partial–waves ϕ_i within another independent basis set *(right)*. Taken from [18, Documentation:VASP Workshop Lectures].

in an additional independent basis set of radial logarithmic grids which describes local on–site spheres that conserve the nodal structure and the orthogonality to the core states without the necessity to be represented in plane waves. The employment of radial logarithmic grids incorporates the description of highly resolvable features with little computational effort. The sum in (2.75) can be interpreted as the measure of how strong these localised functions shall be mixed with the plane wave solution, which is determined by a projection $\tilde{p}_i$ onto the according plane wave solution. One distinguishes again between localised (true) all–electron *partial–waves* ϕ_i and *pseudo partial–waves* $\tilde{\phi}_i$. A schematic illustration of the mixing is shown in Fig.2.3. In VASP the (pseudo)partial–waves and the projection functions are computed in advance for almost every single element of the periodic table and stored in special files that need to be provided for every calculation (POTCAR files). The all–electron partial–waves are derived as solutions to the radial scalar relativistic non-spin polarised Schrödinger equations on a radial logarithmic grid, which are similar to the KS equations (2.60),

$$\Big(-\frac{1}{2}\vec{\boldsymbol{\nabla}}^2 + u_{\text{eff}}\Big)\,|\phi_i\rangle \;= \varepsilon_i\,|\phi_i\rangle \;. \tag{2.76}$$

The set of partial–waves can then be transformed by a pseudisation process into

$$|\phi_i\rangle \longrightarrow |\tilde{\phi}_i\rangle \quad\Longrightarrow\quad u_{\text{eff}} \longrightarrow \tilde{u}_{\text{eff}} \quad\Longrightarrow\quad \langle\tilde{p}_i|\tilde{\phi}_j\rangle = \delta_{ij}\;,$$

where the only requirement to the projection functions $\tilde{p}_i$ is to be dual to the pseudo partial–waves which obey

$$\Big(-\frac{1}{2}\vec{\boldsymbol{\nabla}}^2 + \tilde{u}_{\text{eff}} + \sum_{ij}\,|\tilde{p}_i\rangle\,D_{ij}\,\langle\tilde{p}_j|\Big)\,|\tilde{\phi}_k\rangle \;= \varepsilon_k\Big(1 + \sum_{ij}\,|\tilde{p}_i\rangle\,Q_{ij}\,\langle\tilde{p}_j|\Big)\,|\tilde{\phi}_k\rangle \;. \tag{2.77}$$

The matrix elements D_{ij} and Q_{ij} are called the *strength parameters* and *augmentation charges*, respectively,

$$D_{ij} = \langle \phi_i | \, -\frac{1}{2}\vec{\boldsymbol{\nabla}}^2 + u_{\text{eff}} \, | \phi_j \rangle \, - \, \langle \tilde{\phi}_i | \, -\frac{1}{2}\vec{\boldsymbol{\nabla}}^2 + \tilde{u}_{\text{eff}} \, | \tilde{\phi}_j \rangle \ , \tag{2.78a}$$

$$Q_{ij} = \langle \phi_i | \phi_j \rangle \, - \, \langle \tilde{\phi}_i | \tilde{\phi}_j \rangle \ . \tag{2.78b}$$

The augmentation charges account for the fact that the norm of the pseudo partial–waves is not necessarily equal to the norm of the all–electron partial–waves. The relation in (2.77) thus grants that the pseudo eigenvalue spectrum is identical to the all–electron problem. It is important to highlight that quantities within the PAW method can never be expressed on a common grid. The essential advantage of this method manifests in the rigorous separation of quantities on the regular plane wave grid and those on the radial logarithmic grid with the assurance that there will never be any cross terms between the two spaces.

Electronic Minimisation Processes One distinguishes between two methods in order to find the ground state by solving the KS equations. There is the possibility to either minimise the functional $E[\{\psi_n\}]$ directly by using a set of initial orbitals (random numbers) and optimise them by propagation along the gradient (the residual, see (2.82) below) of the energy curve (*direct minimisation*), or to start with a trial density that is diagonalised with the corresponding Hamiltonian expressed in a plane wave basis in order to derive a new set of orbitals and repeat these steps until a convergence criterion is reached (*self–consistency–cycle*),

$$\bar{\mathrm{H}} = \langle \vec{\boldsymbol{G}} | \ \mathcal{H}[\rho] \ | \vec{\boldsymbol{G}}' \rangle \ ,$$

$$\rho_0 \longrightarrow \bar{\mathrm{H}}_0 \longrightarrow \rho' \longrightarrow \rho_1 = \{\rho', \rho_0\} \longrightarrow \bar{\mathrm{H}}_1 \longrightarrow \ldots \quad .$$

In the latter case the new density ρ' from the derived orbitals is *admixed* to the old one to avoid a phenomena known as *charge sloshing*, which is especially evident in small gap systems. VASP uses a *Broyden mixer* for the new density, and afterwards the diagonalisation is executed iteratively (subspace diagonalisation) in order to calculate only the minimal amount of eigenstates needed for the problem and not all possible eigenstates of the FFT grid. There are two possible flavours implemented in the program, namely the RMM–DIIS residual or the more reliable *Blocked Davidson* algorithm. These methods yield efficiently sized diagonalisation problems of a certain subspace, where the Rayleigh–Ritz problem with the Hamiltonian $\bar{\mathrm{H}}$ of the according subspace and the overlap operator $\bar{\mathrm{S}}$ reads,

$$\sum_n H_{mn} B_{nk} = \sum_n \tilde{\varepsilon}_k \, S_{mn} B_{nk} \ , \tag{2.79}$$

$$\text{where} \quad H_{mn} = \langle \psi_m | \, \mathcal{H} \, | \psi_n \rangle \quad \text{and} \quad S_{mn} = \langle \psi_m | \, \mathcal{S} \, | \psi_n \rangle \ .$$

One derives an approximation to the lowest eigenstates of $\mathcal{H}$ within the subspace spanned by the current orbitals by

$$| \psi'_k \rangle \ = \sum_n B_{nk} \, | \psi_n \rangle \ , \tag{2.80}$$

with the approximative eigenvalues

$$\tilde{\varepsilon} = \frac{\langle \psi_m | \, \mathcal{H} \, | \psi_n \rangle}{\langle \psi_m | \, \mathcal{S} \, | \psi_n \rangle} \ . \tag{2.81}$$

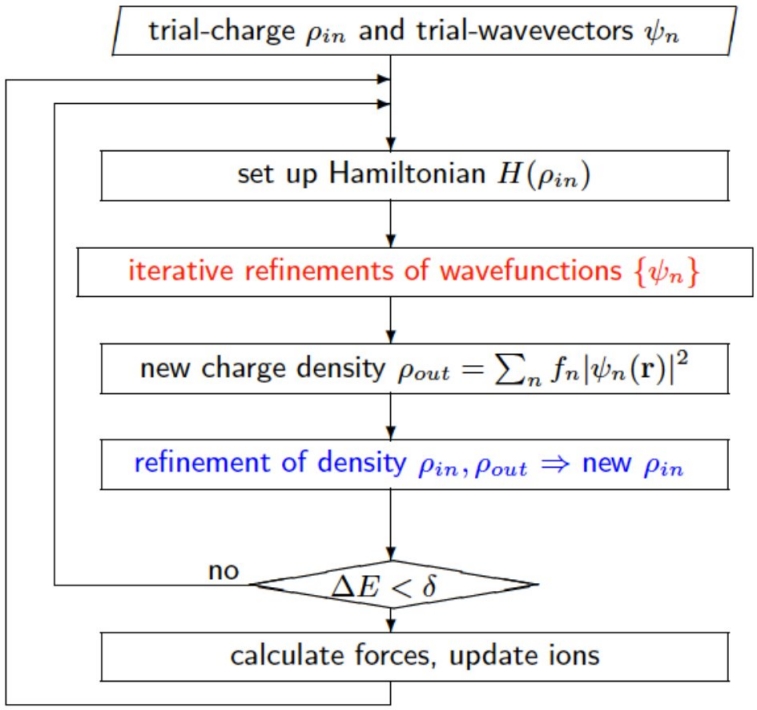


Fig. 2.4: Diagram of the steps in a self–consistency–cycle implemented in VASP. Iterative diagonalisation (optimisation of $\{\psi_n\}$) is executed by the use of RMM–DIIS or *Blocked Davidson* algorithms and density mixing that is responsible for the construction of a new input density is covered by a *Broyden mixer*. Taken from [18, Documentation:VASP Workshop Lectures].

The norm of the residual

$$|\mathrm{R}(\psi_n)\rangle = \left(\mathcal{H} - \tilde{\varepsilon}\,\mathcal{S}\right)|\psi_n\rangle \ , \tag{2.82}$$

is a measure for the error within the new eigenvector. Hence, for an exact state the residual would be zero.

Chapter 3

Phonons

Lattice vibrations are treated in the classical picture of elastic waves propagating through a (polyatomic) crystal. Orthonormality guaranties a simple treatment of the modes which jointly allow a certain perturbation spreading over the atomic positions of the crystal. The quantum mechanical analogue is the creation of *phonons*, which are quasiparticles derived from the treatment of quantised elastic waves in a harmonic potential. Quasiparticles are constituents of the principle of elementary excitations, where the complex behaviour of a many–body system (simplifying a quantum mechanical many–body problem) is investigated by perturbing the ordered ground state by an outer influence to an excited state of higher energy and afterwards measure the according response function. An elementary excitation cannot be composed out of other excitations at lower energy. If the quasiparticles are only weakly interacting with one another most of the low level excited states of a solid can be considered as an ensemble of elementary excitations. A phonon shows bosonic feature (integral spin) and is therefore seen as a collective excitation rather than a quasiparticle, although it should be emphasised that there is no strict distinction between these two expressions. However, as mentioned before, the picture of an elastic wave propagating in a certain direction of the solid is much simpler to describe than observing every perturbed atom individually.

3.1 Classical Picture of Elastic Waves

Within small displacements of the atomic equilibrium position the elastic response of the crystal is a linear function of the forces, so that the elastic energy is a quadratic function of the displacement of any two points in the crystal. Consider a linear chain of equal masses M which are connected by springs with equal force constants C. This description would be equivalent to an elastic wave propagating in certain directions of cubic crystals, for example [100] or [111]. For these directions entire planes of atoms of the same type (and equal mass M) are in phase with displacements either parallel or perpendicular to the direction of the wave vector and can thus be described by a single displacement variable u_n for plane n. The problem is therefore one dimensional (as is the chain). Assuming only nearest neighbour interactions the total force on the n^{th} atom in the chain is

$$F_n = M\,\ddot{u}_n = C\,(u_{n+1} - u_n) - C\,(u_n - u_{n-1}) = C\,(u_{n+1} - 2u_n + u_{n-1})\,. \tag{3.1}$$

The solutions for u_n shall all have the same time dependence of $\mathrm{e}^{-i\omega t}$. These *normal modes* — normal in a sense that every solution with a fixed frequency (resonance) being orthonormal, i.e. independent from the others — are well suited for the description of elementary excitations.

Hence, the equation of motion (3.1) transforms into

$$\begin{aligned} -M\omega^2 u_n &= C\,(u_{n+1} - 2u_n + u_{n-1}) \\ \Longrightarrow \qquad u_{n\pm1} &= u\ \mathrm{e}^{i(n\pm1)ka}\ . \end{aligned} \tag{3.2}$$

The latter is known as a (finite) difference equation which yields propagating wave solutions for $u_{n\pm1}$, where k is the wave vector of the elastic wave and a is the spacing between the atoms and will depend on the direction of k. Note that the time dependent factor $\mathrm{e}^{-i\omega t}$ is not added, as it is the same for all solutions and usually cancels out. The solutions are explicitly entered into (3.2) and the cancellation of $u \cdot \mathrm{e}^{inka}$ leaves

$$-M\omega^2 = C\left(\mathrm{e}^{ika} + \mathrm{e}^{-ika} - 2\right). \tag{3.3}$$

With the identity of the cosine, $\cos(x) = \frac{1}{2}(\mathrm{e}^{ix} + \mathrm{e}^{-ix})$, and the trigonometric relation $\cos(2x) = \cos^2(x) - \sin^2(x)$ one derives the dispersion relation

$$\begin{aligned} \omega^2 &= \frac{2C}{M}\Big(1 - \cos(ka)\Big) = \frac{4C}{M}\,\sin^2\left(\tfrac{ka}{2}\right) \\ \Longrightarrow \qquad \omega &= \omega_0 \left|2\,\sin\left(\tfrac{ka}{2}\right)\right|, \end{aligned} \tag{3.4}$$

where

$$\omega_0 = \sqrt{\frac{C}{M}}\,, \tag{3.5}$$

is the resonance frequency.
The group velocity of the wave packet reads

$$v_\mathrm{G} = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \omega_0\, a \left|\cos\left(\tfrac{ka}{2}\right)\right|, \tag{3.6}$$

which from the supremum at $k = 0$ decreases with increasing $k \to \frac{\pi}{a}$ until reaching zero at the boundary of the first Brillouin zone at $k = \frac{\pi}{a}$ (showing symmetric behaviour in the negative direction of $k \to -\frac{\pi}{a}$). At these boundaries the solution (3.2) for u_n represents a standing wave

$$u_n = u\ \mathrm{e}^{inka} \longrightarrow u\ \mathrm{e}^{\pm in\pi} = u\,(-1)^n\,. \tag{3.7}$$

The latter describes neighbouring atoms in opposite phases. It can be seen by (3.6) that $v_\mathrm{G} = 0$ and the wave does not move at all, but successive reflections at the boundaries build up a standing wave.
For small wave vectors, i.e. the phase $k\,a \ll 1$ (*long wavelength limit*), the sine in (3.2) can be approximated to $\sin(x) \sim x$, so that the dispersion relation yields

$$\begin{aligned} \omega &\sim \omega_0\, ka \\ \frac{\omega}{k} &= \omega_0\, a = \sqrt{\frac{C}{M}\,a^2} = c_\mathrm{S} \\ \Longrightarrow \qquad v_\mathrm{G} &= c_\mathrm{S}\,. \end{aligned} \tag{3.8}$$

In the limit of long wavelengths (sometimes referred to as *continuum limit*) the frequency is directly proportional to the wave vector, which means that the group velocity is independent of frequency and a constant, namely the speed of sound c_S (if you could see it, then you'd understand).

Different Masses M_ν The phonon dispersion exhibits more than the previously presented features when there are two or more different masses M_ν ($\nu = 1, \ldots, p$) in a coupled chain, which is equivalent to a primitive cell of a crystal containing N_u atoms of p different types. As a consequence of the coupling, each polarisation mode develops different branches named *acoustical* and *optical* due to their energy bandwidth. For a unit cell in D dimensional reciprocal space in total there are DN_u branches to the dispersion relation, which are distributed over D acoustical and $D(N_\mathrm{u}-1)$ optical ones. The number of modes available at each wave vector k is appropriate to the number of degrees of freedom (DoF) provided by every atom in the crystal. Individually, every atom has (D) DoF (for $D = 3$ it would be: x, y, z), hence, for N unit cells with each containing N_u atoms there is a total of ($DN_\mathrm{u}\,N$) DoF for the whole crystal. It follows that the number of allowed k values in one Brillouin zone has to be N per single branch. This can be easily verified by applying periodic boundary conditions to the modes of a cubic crystal of volume $V = L^D$, similar to (2.10). One k value occupies the reciprocal space volume of

$$(\Delta k)^D = \left(\frac{2\pi}{L}\right)^D = \frac{(2\pi)^D}{V}\,.$$

With a volume V_p of the primitive cell the volume of the Brillouin zone is $\frac{(2\pi)^D}{V_\mathrm{p}}$, and the number of allowed k values is thus derived by the ratio

$$\frac{\dfrac{(2\pi)^D}{V_\mathrm{p}}}{\dfrac{(2\pi)^D}{V}} = \frac{V}{V_\mathrm{p}} = N\,, \tag{3.9}$$

which is exactly the number of primitive cells N in the crystal. Therefore, the acoustical branches occupy DN modes, whereas the optical branches accommodate the other $(N_\mathrm{u}-1)\,DN$ modes.

As an example a cubic crystal containing two different type of atoms with according mass (M_1, M_2) shall be considered. One set of planes only consists of atoms with mass M_1 and another set of planes only of atoms with mass M_2, where the planes are arranged in alternating fashion (see Fig.3.1). As before, only elastic waves that propagate in a symmetry direction of the cubic crystal are described such that a single plane contains just one type of atom. Within a plane–set of one type the repeat equilibrium distance of one plane to the other is a, which is not the nearest–neighbour distance $d \leq \frac{a}{2}$. The equation of motion for each plane interacting with its nearest–neighbour and identical force constants C reads

$$M_1\,\ddot{u}_n = C\,(v_n + v_{n-1} - 2u_n)\,, \tag{3.10a}$$

$$M_2\,\ddot{v}_n = C\,(u_{n+1} + u_n - 2v_n)\,. \tag{3.10b}$$

Again the solutions shall have a form of a propagating wave similar to (3.2) with the same time dependency, but this time with different amplitudes (u, v) on alternate planes. Entering these solutions explicitly into (3.10) and cancel all $\mathrm{e}^{i(nka-\omega t)}$ leaves

$$-M_1\omega^2\,u = C\,v\left(1 + \mathrm{e}^{-ika}\right) - 2C\,u\,, \tag{3.11a}$$

$$-M_2\omega^2\,v = C\,u\left(\mathrm{e}^{+ika} + 1\right) - 2C\,v\,. \tag{3.11b}$$

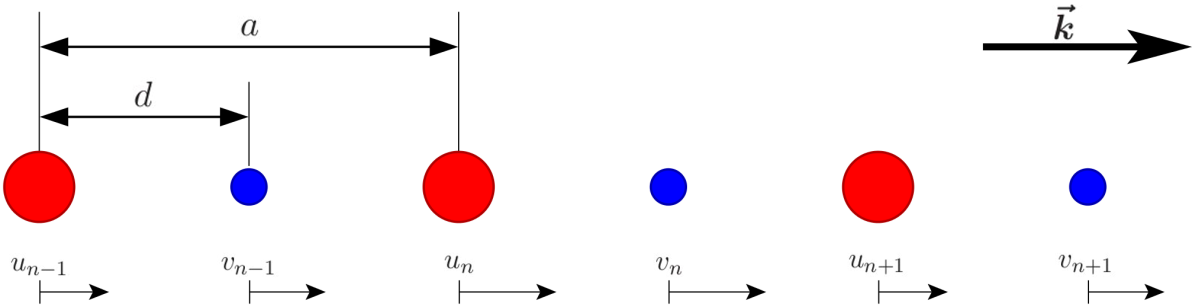


Fig. 3.1: Illustrating the derivation of (3.10). Cubic crystal containing two different type of atoms with M_1 *(red)* and M_2 *(blue)* with displacements u_i and v_i from equilibrium position, respectively. The repeat distance of one plane to another is a, which should not be confused with the nearest–neighbour distance $d \leq \frac{a}{2}$.

The latter is a homogeneous linear equation which only has a solution if the determinant of the coefficients is zero,

$$\begin{vmatrix} 2C - M_1\omega^2 & -C\left(1 + \mathrm{e}^{-ika}\right) \\ -C\left(1 + \mathrm{e}^{+ika}\right) & 2C - M_2\omega^2 \end{vmatrix} = 0$$

$$\Longrightarrow \qquad M_1 M_2\, \omega^4 - 2C\big(M_1 + M_2\big)\, \omega^2 + 2C^2\Big(1 - \cos(ka)\Big) = 0\,. \tag{3.12}$$

Due to the biquadratic form of the determinant it is quite easy to find the exact solutions for ω^2 which yield a dispersion relation that reads

$$\omega_\pm^2 = \frac{C}{\mu} \pm \frac{C}{\mu}\sqrt{1 - \frac{2\mu}{M_1 + M_2}\Big(1 - \cos(ka)\Big)}\,, \tag{3.13}$$

$$\text{where} \quad \mu = \frac{M_1 M_2}{M_1 + M_2} \quad \text{(reduced mass)}\,.$$

However, from a pragmatic point of view it will be more informative to estimate the results in the limits of small wave vectors $k \ll \frac{1}{a}$ and at the boundaries of the first Brillouin zone, $k = \pm\frac{\pi}{a}$. For small arguments the cosine can be simplified to $\cos(x) \sim 1 - \frac{x^2}{2}$, so that within the long wavelength limit ($k\,a \ll 1$) the root in (3.13) can be written as

$$\sqrt{1 - \frac{\mu}{M_1 + M_2} k^2\, a^2} \sim 1 - \frac{1}{2}\frac{\mu}{M_1 + M_2} k^2\, a^2\,,$$

where within the same limit also the root itself can be approximated by $\sqrt{1 - cx^2} \sim 1 - \frac{1}{2}cx^2$. The two branches of (3.13) are then evaluated to

$$\Longrightarrow \qquad \omega_+^2 \simeq \frac{2C}{\mu} = 2C\left(\frac{1}{M_1} + \frac{1}{M_2}\right) \qquad (\textit{optical})\,, \tag{3.14a}$$

$$\Longrightarrow \qquad \omega_-^2 \simeq \frac{1}{2}\frac{C}{M_1 + M_2} k^2\, a^2 \qquad (\textit{acoustical})\,. \tag{3.14b}$$

The former yields an expression for the *optical branch* and hardly shows any dispersion, where the latter is called the *acoustical branch* with a dispersion relation reminiscent of (3.8). The names of the branches appear to be a sensible choice when either of the expressions are substituted into the equations of motion (3.10). In the limit of $k \to 0$ the ratio of the amplitudes for the optical branch (3.14a) reads

$$\frac{u}{v} = -\frac{M_2}{M_1} \,.$$

The atoms vibrate against each other, while their centre of mass does not change. A motion of this type can be excited by the electric field of a light wave (infrared), especially when the vibrating atoms carry opposite charges. The acoustical expression (3.14b) vanishes for $k \to 0$, so that the amplitude ratio is

$$\frac{u}{v} = 1 \iff u = v \,.$$

The atoms oscillate in phase together with their centre of mass, as it is the case for long wavelength vibrations in acoustics, hence the name. It should be noted that for any arbitrary k the expressions for both amplitude ratios become more complex. The tendencies of both dispersion relation branches are shown in Fig.3.2a . The optical branch is usually found at higher frequencies, which becomes clear when deriving (3.14a) can also be achieved by simply neglecting the ω–independent term in (3.12). On the other hand, (3.14b) is found as well by ignoring the higher order term of the frequency in (3.12), which shows that the acoustical branch is dominant at lower frequencies.
At the boundaries of the first Brillouin zone, where $k = \pm\frac{\pi}{a}$, the off-diagonal elements of the coefficient determinant (3.12) are zero, so that the problem simplifies to

$$\left(2C - M_1\omega^2\right) \cdot \left(2C - M_2\omega^2\right) = 0 \,,$$

and therefore only allows two constant values as solutions

$$\omega_j{}^2 = \frac{2C}{M_j} \,, \qquad j = 1, 2 \,. \tag{3.15}$$

Which one of these roots is part of the optical or acoustical branch is dependent on the masses. If $M_1 > M_2$, then ω_1 would be of an acoustic type and ω_2 of an optical one, and vice versa. More important, however, is the fact that there are no waveform solutions for frequencies between ω_1 and ω_2. At the first Brillouin zone boundary one thus finds a *frequency gap*, which turns out to be characteristic for elastic waves in polyatomic lattices.

3.2 Quantum Mechanical Treatment – Phonons as Quasiparticles

Phonons are quantised lattice vibrations with frequency ω and thus energy $\hbar\omega$ and (quasi) momentum $\hbar k$. However, as it was shown in the beginning of Sec.3.1 stationary lattice vibrations can be interpreted as standing waves in the crystal where the group velocity is zero. What seems to be a discrepancy at first sight can be explained by the description of inelastic scattering. When photons or neutrons scatter inelastically with the crystal the resulting energy difference ΔE is mostly transformed into lattice vibrations. During the scattering process the momentum transfer creates a travelling acoustic wave which is reflected at crystal surfaces and finally relaxes into a stationary oscillating state, so that within the process the momentum $\hbar\Delta k$ is actually taken

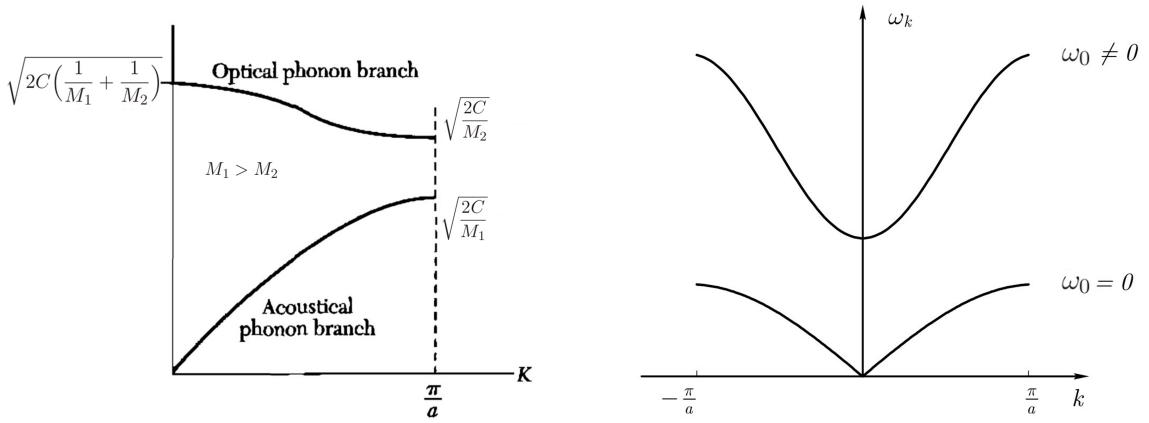


Fig. 3.2: Comparison of phonon dispersion relations (a) *classically* derived with a frequency gap at the first Brillouin zone boundary and (b) for a *quantised* elastic wave based on the model of a harmonic oscillator. (a) and (b) taken from [27] and [28], respectively.

by the whole crystal. This rather complicated context of an inelastic scattering process in a crystal can be significantly simplified by the introduction of a phonon as a carrier of energy and quasi–momentum and therefore allows a correct application of the conservation laws. Phonons are thus also referred to as *quasiparticles*, in order to emphasise that because they are driving static lattice vibrations they neither can have any mass, nor (real) momentum, but are used as a simplification in the description of the complicated energy and momentum transfer throughout the whole crystal when particles are inelastically scattered.

3.2.1 Quantised Elastic Waves in a Harmonic Potential

A simple model for phonons in a crystal is a periodic linear lattice of N particles of mass M that are connected by springs with a force constant C and length a (lattice constant). Each particle interacts in a harmonic potential and is further harmonically coupled to its nearest neighbours. The particle motion can be quantised exactly as it would be the case for a harmonic oscillator (or a set of coupled harmonic oscillators), but in order to reach consistent expressions a transformation to normal coordinates (*phonon coordinates*) that will represent a travelling wave is additionally necessary.

The Hamiltonian of a harmonic oscillator reads

$$\begin{aligned} \mathcal{H}\,|\phi_\xi\rangle &= \varepsilon_\xi\,|\phi_\xi\rangle \ , \\ \text{with} \quad \mathcal{H} &= \frac{1}{2M}\,p^2 + \frac{M{\omega_0}^2}{2}\,x^2 \quad , \end{aligned} \tag{3.16}$$

and the according energy eigenvalues with $\xi = 0, 1, 2, \ldots$ are

$$\varepsilon_\xi = \left(\xi + \frac{1}{2}\right)\hbar\omega_0 \ . \tag{3.17}$$

Let u_n be the displacement of particle n out of its equilibrium position, and p_n is its momentum. Then the Hamiltonian of the system is

$$\mathcal{H} = \sum_{n=1}^{N} \frac{1}{2M} {p_n}^2 + \sum_{n=1}^{N} \left(\frac{M\omega^2}{2} \left(u_{n+1} - u_n\right)^2 + \frac{M{\omega_0}^2}{2} {u_n}^2 \right) . \tag{3.18}$$

The latter used in an eigenvalue problem as in (3.16) is solvable for a chain in the limit of an infinitely expanded system, which means that *periodic boundary conditions* ($u_{n+N} = u_n$) can be used and the problem is thus translation invariant,

$$\begin{aligned} & x_n = n \cdot a + u_n \\ \Longrightarrow \quad & x_{n+N} = (n+N) \cdot a + u_{n+N} = n \cdot a + u_n + N \cdot a = x_n + N \cdot a = x_n \, . \end{aligned} \tag{3.19}$$

We can therefore diagonalise the Hamiltonian by a Fourier transformation from the coordinates $\{u_n, p_n\}$ to $\{U_k, P_k\}$ known as *normal* or *phonon coordinates*,

$$u_n = \frac{1}{\sqrt{N}} \sum_k U_k \, \mathrm{e}^{inka} \qquad , \qquad p_n = \frac{1}{\sqrt{N}} \sum_k P_k \, \mathrm{e}^{-inka} \tag{3.20a}$$

$$\Updownarrow \qquad\qquad\qquad \Updownarrow$$

$$U_k = \frac{1}{\sqrt{N}} \sum_n u_n \, \mathrm{e}^{-inka} \qquad , \qquad P_k = \frac{1}{\sqrt{N}} \sum_n p_n \, \mathrm{e}^{inka} \, . \tag{3.20b}$$

The N values for wave vector k are given by the periodic boundary conditions,

$$\begin{aligned} & u_{n+N} = u_n \implies \mathrm{e}^{iNka} = 1 \implies \boxed{k = \frac{2\pi\ell}{Na}} \quad , \\ & \text{where} \quad \forall\, \ell \in \mathbb{Z} : \begin{cases} |\,\ell\,| \leq \frac{N-1}{2} & (N \text{ odd}) \\ -\frac{N}{2} < \ell \leq \frac{N}{2} & (N \text{ even}) \end{cases} \, . \end{aligned} \tag{3.21}$$

The thereby confined interval for values of k defines the first Brillouin zone in reciprocal space. Using (3.19) and the quantum commutator relation $[x_m, p_n] = i\hbar\,\delta_{mn}$, one can prove that the canonical variables $\{u_n, p_n\}$ satisfy the following commutator relations,

$$[u_m, p_n] = i\hbar\,\delta_{mn} \qquad , \qquad [u_m, u_n] = 0 \qquad , \qquad [p_m, p_n] = 0 \, , \tag{3.22}$$

and thus are conjugate operators, which also justifies the opposite signs of the exponential function in the Fourier transform. One now needs to prove that $\{U_k, P_k\}$ are conjugate as well. Taking a closer look at the commutator relation with the expressions from (3.20b) and (3.22) yields

$$[U_k, P_{k'}] = \frac{1}{N} \sum_m \sum_n [u_m, p_n] \, \mathrm{e}^{-i(mk - nk')a} = \frac{i\hbar}{N} \sum_n \mathrm{e}^{-in(k-k')a} = i\hbar\,\delta(k - k') \, , \tag{3.23}$$

where the latter summation has been evaluated as

$$\sum_n \mathrm{e}^{-in(k-k')a} = \sum_n \mathrm{e}^{-i2\pi(\ell-\ell')\frac{n}{N}} = N\,\delta_{\ell\ell'} \equiv N\,\delta(k - k') \, . \tag{3.24}$$

The other relations, $[U_k, U_{k'}] = 0$ and $[P_k, P_{k'}] = 0$ follow directly from their definitions. Note that $\delta(k-k')$ is generally valid for unlimited $k = \frac{2\pi\ell}{Na}$. For an index ℓ that is restricted to the first Brillouin zone according to (3.21) the Kronecker–delta will be denoted as $\delta(k-k') \to \delta_{kk'}$. The transformation is executed by entering (3.20a) into the Hamiltonian (3.18) and make again use of the summation rule (3.24). The individual terms thus read

$$\sum_n {p_n}^2 = \frac{1}{N} \sum_n \sum_k \sum_{k'} P_k P_{k'} \, \mathrm{e}^{-in(k+k')a} = \sum_k \sum_{k'} P_k P_{k'} \, \delta_{-k,k'}$$

$$\Longrightarrow \quad \boxed{\sum_n {p_n}^2 = \sum_k P_k \, P_{-k}} \qquad \Longrightarrow \quad \boxed{\sum_n {u_n}^2 = \sum_k U_k \, U_{-k}} \quad , \tag{3.25a}$$

$$\sum_n \left(u_{n+1} - u_n\right)^2 = \sum_n \left(u_{n+1} - u_n\right)_k \left(u_{n+1} - u_n\right)_{k'} =$$

$$= \frac{1}{N} \sum_n \sum_k \sum_{k'} U_k U_{k'} \, \mathrm{e}^{in(k+k')a} \left(\mathrm{e}^{ika} - 1\right) \left(\mathrm{e}^{ik'a} - 1\right) =$$

$$= \sum_k \sum_{k'} U_k U_{k'} \left(\mathrm{e}^{i(k+k')a} - \mathrm{e}^{ika} - \mathrm{e}^{ik'a} + 1\right) \delta_{-k,k'} = \sum_k U_k U_{-k} \left(2 - \left(\mathrm{e}^{ika} + \mathrm{e}^{-ika}\right)\right)$$

$$\Longrightarrow \quad \boxed{\sum_n \left(u_{n+1} - u_n\right)^2 = 2 \sum_k U_k U_{-k} \left(1 - \cos(ka)\right)} \quad . \tag{3.25b}$$

Hence, in total the phonon Hamiltonian in normal coordinates becomes

$$\mathcal{H} = \sum_k \frac{1}{2M} \, P_k \, P_{-k} + \sum_k \frac{M {\omega_k}^2}{2} \, U_k \, U_{-k} \, , \tag{3.26}$$

with the according dispersion relation

$$\begin{aligned} {\omega_k}^2 &= 2\,\omega^2 \left(1 - \cos(ka)\right) + {\omega_0}^2 \\ \Longrightarrow \quad \omega_k &= \sqrt{\omega^2 \left(2 \sin\left(\tfrac{ka}{2}\right)\right)^2 + {\omega_0}^2} \, . \end{aligned} \tag{3.27}$$

One again distinguishes between acoustical phonons, which are derived for $\omega_0 = 0$, and optical phonons for a finite $\omega_0 \neq 0$. The Hamiltonian (3.26) in Fourier space is very similar to (3.16) — this will be proved next in Sec.3.2.2 — and thus yields uncoupled oscillators with frequency ω_k (only oscillators with wave vector $\pm k$ are still connected, so that also $\omega_k = \omega_{-k}$).

3.2.2 Second Quantisation of Phonon Coordinates

Since phonons are indistinguishable particles it is more convenient to diagonalise the Hamiltonian (3.26) in second quantisation. Analogue to the harmonic oscillator there is a possibility to

introduce a *creation* and *annihilation operator* by a transformation of the phonon coordinates,

$$a_k = \frac{1}{\sqrt{2\hbar M\omega_k}} \left(M\omega_k \, U_k + i \, P_{-k} \right) , \tag{3.28a}$$

$$a^\dagger{}_k = \frac{1}{\sqrt{2\hbar M\omega_k}} \left(M\omega_k \, U_{-k} - i \, P_k \right) , \tag{3.28b}$$

where one uses the fact that $U^\dagger{}_k = U_{-k}$ and $P^\dagger{}_{-k} = P_k$, which can be deduced from (3.20b) and the requirement that u_n and p_n need to be Hermitian operators. With the properties of (3.23) the commutation relations for the ladder operators give the well-known relations,

$$[a_k, a^\dagger{}_{k'}] = \frac{1}{2\hbar} \Big(- i \, [U_k, P_{k'}] + i \overbrace{[P_{-k'}, U_{-k}]}^{=[U_k,P_{k'}]^\dagger} \Big) = \frac{1}{2\hbar} \Big(\hbar \, \delta_{kk'} + \hbar \overbrace{\delta_{-k',-k}}^{=\delta_{k'k}=\delta_{kk'}} \Big)$$

$$\Longrightarrow \qquad [a_k, a^\dagger{}_{k'}] = \delta_{kk'} \qquad , \qquad [a_k, a_{k'}] = 0 \qquad , \qquad [a^\dagger{}_k, a^\dagger{}_{k'}] = 0 \, , \tag{3.29}$$

which already indicate the bosonic behaviour of phonons.
The inverse of this transformation yields the following relations,

$$U_k = \sqrt{\frac{\hbar}{2M\omega_k}} \left(a_k + a^\dagger{}_{-k} \right) , \tag{3.30a}$$

$$P_k = i \sqrt{\frac{\hbar M\omega_k}{2}} \left(a^\dagger{}_k - a_{-k} \right) . \tag{3.30b}$$

These expressions are entered into the Hamiltonian (3.26) in order to derive

$$\begin{aligned}
&\frac{\hbar\omega_k}{4} \Big[\big(a_{-k} - a^\dagger{}_k\big)\big(a^\dagger{}_{-k} - a_k\big) + \big(a_k + a^\dagger{}_{-k}\big)\big(a_{-k} + a^\dagger_k\big) \Big] = \\
&= \frac{\hbar\omega_k}{4} \left[\begin{array}{c} a_{-k}a^\dagger{}_{-k} + a^\dagger{}_k a_k - a_{-k}a_k - a^\dagger{}_k a^\dagger{}_{-k} + \\ + a^\dagger{}_{-k}a_{-k} + a_k a^\dagger{}_k + a_k a_{-k} + a^\dagger{}_{-k}a^\dagger{}_k \end{array} \right] = \\
&= \frac{\hbar\omega_k}{4} \Big[\big(1 + 2\, a^\dagger{}_{-k}a_{-k}\big) + \big(1 + 2\, a^\dagger{}_k a_k\big) + \underbrace{[a_k, a_{-k}]}_{=0} + \underbrace{[a^\dagger{}_{-k}, a^\dagger{}_k]}_{=0} \Big] \equiv \frac{\hbar\omega_k}{2} \big(1 + 2\, a^\dagger{}_k a_k\big)
\end{aligned}$$

$$\Longrightarrow \qquad \mathcal{H} = \sum_k \hbar\omega_k \left(a^\dagger{}_k a_k + \tfrac{1}{2} \right) . \tag{3.31}$$

Consequently, it was shown by the application of the commutator rules (3.29) that the Hamiltonian can be interpreted as a system of N uncoupled oscillators, where the sum in (3.31) is over all N wave vectors $\{k_1, \ldots, k_j, \ldots, k_N\}$ within the first Brillouin zone.

In *second quantisation* a many–body state is represented in an occupation number basis, also known as Fock state, where the basis state is denoted by a set of occupation numbers and the tensor product of eigenfunctions is implicitly symmetrised. Consider a wave function $\Psi = |n_{k_1}, \ldots, n_{k_j}, \ldots, n_{k_N}\rangle$ composed out of eigenstates of (3.31) with according symmetry to

solve the Schrödinger equation,

$$\mathcal{H}\,\Psi = E\,\Psi \quad \Longleftrightarrow \quad \mathcal{H}\,|n_k\rangle = \varepsilon_k\,|n_k\rangle\ ,$$
$$\text{with} \quad E = \sum_k \varepsilon_k\ .$$

Then it is possible to prove that $a^{(\dagger)}\,|n\rangle$ are also eigenstates of the same Hamiltonian. In this case the ladder operators commute with the Hamiltonian (3.31) the following way,

$$[\mathcal{H}, a_k] = -\hbar\omega_k\, a_k \qquad , \qquad [\mathcal{H}, {a^\dagger}_k] = +\hbar\omega_k\, {a^\dagger}_k\ , \tag{3.32}$$

which is easy to verify using (3.29) and a standard commutator rule, $[AB, C] = A[B, C]+[A, C]B$. The latter yields

$$\mathcal{H}\,a_k\,|n_k\rangle = a_k(\mathcal{H} - \hbar\omega_k)\,|n_k\rangle = (\varepsilon_k - \hbar\omega_k)\,a_k\,|n_k\rangle\ , \tag{3.33a}$$

$$\mathcal{H}\,{a^\dagger}_k\,|n_k\rangle = {a^\dagger}_k(\mathcal{H} + \hbar\omega_k)\,|n_k\rangle = (\varepsilon_k + \hbar\omega_k)\,{a^\dagger}_k\,|n_k\rangle\ . \tag{3.33b}$$

$a^{(\dagger)}\,|n\rangle$ are indeed eigenstates with the respective eigenvalues of $\varepsilon \mp \hbar\omega$ and the names of the operators now become evident: As the action of the operators onto the eigenfunctions can be repeated the creation operator *rises* the energy by an *integer multiple* of $\hbar\omega_k$, where the annihilation operator *lowers* it by the same measure. This points to a (quasi)particle nature, whence these integral excitations are called *phonons*; ${a^\dagger}_k$ *creates* a phonon with wave vector k and frequency ω_k and a_k *destroys* (annihilates) one with k and ω_k. As it was mentioned before, wave vector k is confined to the first Brillouin zone by the N values it can take, however, the appertaining occupation numbers n_k of the phonon states are unlimited natural numbers. Thus, phonons underlie *boson* statistics.
The ground state of Ψ is denoted as $|0\rangle$ (sometimes also called *vaccum state*) and shall indicate that all occupation numbers are zero. There is no state lower in energy than the ground state, so that

$$a_k\,|0\rangle = 0\ , \quad \forall k$$

immediately follows. One can therefore determine the ground state energy as,

$$\left(\mathcal{H} - \frac{\hbar\omega_k}{2}\right)|0\rangle = \sum_k \hbar\omega_k\,{a^\dagger}_k a_k\,|0\rangle = 0 = \left(\mathcal{H} - E_0\right)|0\rangle$$

$$\Longrightarrow \qquad E_0 = \sum_k \varepsilon_{0\,k} = \sum_k \frac{\hbar\omega_k}{2}\ . \tag{3.34}$$

Due to the treatment of quantum mechanics the oscillating system thus always has a *zero–point energy*, as it is the case for the quantum harmonic oscillator and opposed to the classical result. The total energy must therefore be

$$E = \sum_k \varepsilon_k = \sum_k n_k\,\hbar\omega_k + E_0 = \sum_k \hbar\omega_k\left(n_k + \tfrac{1}{2}\right), \tag{3.35}$$

which compared to (3.31) lets us identify the Hermitian *particle number operator* as

$$\hat{n}_k = {a^\dagger}_k a_k \qquad \Longrightarrow \qquad \hat{n}_k\,|n_k\rangle = n_k\,|n_k\rangle\ , \tag{3.36}$$

and has $|n_k\rangle$ as eigenstate giving the occupancy of the phonon mode k. Similar to (3.32) the number operator commutes with the ladder operators by

$$[\hat{n}_k, a_{k'}] = -a_k\,\delta_{kk'} \qquad , \qquad [\hat{n}_k, {a^\dagger}_{k'}] = {a^\dagger}_k\,\delta_{kk'}\,, \tag{3.37}$$

and one can show that $a^{(\dagger)}\,|n\rangle$ are also eigenstates of $\hat{n}$ by using the latter relations in

$$\hat{n}_k\left(a_k\,|n_k\rangle\right) = a_k(\hat{n}_k - 1)\;|n_k\rangle = (n_k - 1)\left(a_k\,|n_k\rangle\right), \tag{3.38a}$$

$$\hat{n}_k\left({a^\dagger}_k\,|n_k\rangle\right) = {a^\dagger}_k(\hat{n}_k + 1)\;|n_k\rangle = (n_k + 1)\left({a^\dagger}_k\,|n_k\rangle\right). \tag{3.38b}$$

This result is not surprising as it is related to (3.33) (the operator $\hat{n}_k$ is part of the Hamiltonian) and also yields the same findings considering the respective eigenvalues; the creation operator ${a^\dagger}_k$ *increases* the occupation number n_k for a phonon with wave vector k, where the annihilator operator a_k *reduces* it (note that only in this case $n_k > 0$). However, it can also be found that the states

$$\hat{n}_k\,|n_k - 1\rangle = (n_k - 1)\,|n_k - 1\rangle \qquad , \qquad \hat{n}_k\,|n_k + 1\rangle = (n_k + 1)\,|n_k + 1\rangle\,,$$

derive the same eigenvalues as before, respectively, so that the two according states must be proportional to one another,

$$a_k\,|n_k\rangle = C_{k-}\,|n_k - 1\rangle \qquad , \qquad {a^\dagger}_k\,|n_k\rangle = C_{k+}\,|n_k + 1\rangle\,.$$

The constants $C_{k\pm}$ can be determined by calculating the square norm,

$$\begin{aligned} |C_{k-}|^2 = \|C_{k-}\,|n_k - 1\rangle\,\|^2 = \|a_k\,|n_k\rangle\,\|^2 = \langle n_k|{a^\dagger}_k a_k|n_k\rangle = \langle n_k|\hat{n}_k|n_k\rangle = n_k\,, \\ |C_{k+}|^2 = \|C_{k+}\,|n_k + 1\rangle\,\|^2 = \|{a^\dagger}_k\,|n_k\rangle\,\|^2 = \langle n_k|a_k{a^\dagger}_k|n_k\rangle = \langle n_k|\hat{n}_k + 1|n_k\rangle = n_k + 1\,, \end{aligned} \tag{3.39}$$

which yields in total

$$a_k\,|n_k\rangle = \sqrt{n_k}\,|n_k - 1\rangle\,, \tag{3.40a}$$

$${a^\dagger}_k\,|n_k\rangle = \sqrt{n_k + 1}\,|n_k + 1\rangle\,. \tag{3.40b}$$

3.2.3 Phonon Dynamics

All of the derived Hamiltonians (3.18), (3.26), and (3.31) are in general time–independent and thus valid at any point. It is therefore best to investigate the phonon dynamics in the *Heisenberg representation* (matrix mechanics) in which only operators are evolved in time, but all state vectors are static (in contrast to the Schrödinger representation, where it is exactly the opposite). In this case an arbitrary operator X is expanded by

$$X(t) = \mathrm{e}^{+i\frac{\mathcal{H}}{\hbar}t}\;X\;\mathrm{e}^{-i\frac{\mathcal{H}}{\hbar}t} = \mathcal{T}^\dagger(t)\,X\,\mathcal{T}(t)\,,$$

where $\mathcal{T}(t)$ is the *time–evolution operator*. The Heisenberg equation, whose derivation can be found in Apx.E, provides an equation of motion for the operator X in the according representation and reads

$$\dot{X}(t) = \frac{i}{\hbar}\,[\mathcal{H}, X(t)]\,\Big(+\,\partial_t X(t)\Big)\,, \tag{3.41}$$

where it is assumed that the operator *itself* is not explicitly time–dependent; $\partial_t X \to 0$ is here stated for the sake of completeness.

If the Hamiltonian is used in form of (3.26) one can describe an equation of motion for the phonon coordinate operator U_k by using (3.41) together with the commutator (3.23) and finds,

$$\dot{U}_k = \frac{i}{\hbar}\,[\mathcal{H}, U_k] = \frac{i}{\hbar}\,\frac{1}{2M}[P_k\,P_{-k}, U_k] = \frac{i}{\hbar}\,\frac{1}{2M}\Big(P_k\,\underbrace{[P_{-k}, U_k]}_{=-i\hbar\,\delta_{-k,k}} + \underbrace{[P_k, U_k]}_{=-i\hbar}\,P_{-k}\Big) = \frac{1}{2M}\Big(2\,P_{-k}\Big)$$

$$\Longrightarrow \qquad P_{-k} = M\,\dot{U}_k\,. \tag{3.42}$$

Using the latter result one can further write,

$$\ddot{U}_k = \frac{i}{\hbar}\,[\mathcal{H}, \dot{U}_k] = \frac{i}{\hbar M}\,[\mathcal{H}, P_{-k}] = \frac{i}{\hbar}\,\frac{{\omega_k}^2}{2}[U_k\,U_{-k}, P_{-k}] =$$

$$= \frac{i}{\hbar}\,\frac{{\omega_k}^2}{2}\Big(U_k\,\underbrace{[U_{-k}, P_{-k}]}_{=i\hbar} + \underbrace{[U_k, P_{-k}]}_{=i\hbar\,\delta_{k,-k}}\,U_{-k}\Big) = -\frac{{\omega_k}^2}{2}\Big(2\,U_k\Big)$$

$$\Longrightarrow \qquad \ddot{U}_k + {\omega_k}^2\,U_k = 0\,. \tag{3.43}$$

For the quantum mechanical treatment the derived differential equation is transformed back into particle coordinates $u_n(t)$ by (3.20b),

$$\frac{1}{\sqrt{N}}\sum_n \Big(\ddot{u}_n(t)\;\mathrm{e}^{-ika_n} + {\omega_k}^2 u_n(t)\;\mathrm{e}^{-ika_n}\Big) = 0$$

$$\Longrightarrow \qquad \ddot{u}_n(t) + {\omega_k}^2 u_n(t) = 0\,, \tag{3.44}$$

which is the equivalent to the (classical) equation of motion (3.1) of a harmonic oscillator with frequency ω_k from (3.27). Here we define $a_n = n \cdot a$. The time evolution again has been achieved by the Heisenberg representation

$$u_n(t) = \mathcal{T}^\dagger(t)\,u_n(0)\,\mathcal{T}(t)\,, \tag{3.45}$$

where the particle displacement operator $u_n(0)$ can be found from its Fourier transform and written in second quantisation by entering (3.30a) into (3.20a),

$$\begin{aligned} u_n(0) = u_n &= \sum_k \sqrt{\frac{\hbar}{2N\,M\omega_k}}\;\mathrm{e}^{ika_n}\Big(a_k + {a^\dagger}_{-k}\Big) = \\ &= \sum_k \sqrt{\frac{\hbar}{2N\,M\omega_k}}\Big(a_k\;\mathrm{e}^{ika_n} + {a^\dagger}_k\;\mathrm{e}^{-ika_n}\Big)\,. \end{aligned} \tag{3.46}$$

The time dependency (3.45) of $u_n(t)$ can thus be written as

$$\begin{aligned}\mathcal{T}^\dagger(t)\, u_n\, \mathcal{T}(t) &= \sum_k \sqrt{\frac{\hbar}{2N\, M\omega_k}} \left[\mathcal{T}^\dagger(t)\, a_k\, \mathcal{T}(t)\, \mathrm{e}^{+ika_n} + \mathcal{T}^\dagger(t)\, a^\dagger{}_k\, \mathcal{T}(t)\, \mathrm{e}^{-ika_n} \right] = \\ &= \sum_k \sqrt{\frac{\hbar}{2N\, M\omega_k}} \left[\exp\left(i \sum_k \omega_k (\hat{n}_{-k} + \tfrac{1}{2})\, t \right) a_k\, \exp\left(-i \sum_k \omega_k (\hat{n}_k + \tfrac{1}{2})\, t + ika_n \right) + \right. \\ &\qquad \left. + \exp\left(i \sum_k \omega_k (\hat{n}_k + \frac{1}{2})\, t \right) a^\dagger{}_k\, \exp\left(-i \sum_k \omega_k (\hat{n}_{-k} + \frac{1}{2})\, t - ika_n \right) \right] \end{aligned}$$

$$\implies \qquad u_n(t) = \sum_k \sqrt{\frac{\hbar}{2N\, M\omega_k}} \left(a_k\, \mathrm{e}^{i(ka_n - \omega_k t)} + a^\dagger{}_k\, \mathrm{e}^{-i(ka_n - \omega_k t)} \right) \quad . \tag{3.47}$$

The structure of this expression is similar to the classical solution, except that the amplitudes of the waves are replaced by creation and annihilation operators, respectively. For $k = \frac{\pi}{a}$ we would derive an equation for a standing wave again. The fact that

$$\mathcal{T}^\dagger(t)\, a_k\, \mathcal{T}(t) = a_k\, \mathrm{e}^{-i\omega_k t} \qquad , \qquad \mathcal{T}^\dagger(t)\, a^\dagger{}_k\, \mathcal{T}(t) = a^\dagger{}_k\, \mathrm{e}^{+i\omega_k t} \, ,$$

is not shown here, but these relations are derived by expanding the exponential functions into their respective power series and building up commutator relations that are suited for comparison to (3.32) or (3.37). The expansion is then set back to the exponential function in the final result. Details of this quite elaborate calculation can be found in literature [28].

3.3 Phonons in Computation – phono.py

Phonon calculations within this thesis have been achieved by the use of a *python*–based programme called `phono.py` [29], which is implemented and maintained by Associate Professor Atsushi Togo (atztogo.github.com) from the Isao Tanaka Research Group of Kyoto University. It is based on the supercell approach and derives phononic features within the limit of a harmonic approximation. However, as a kind of pre-process precisely converged forces between the atoms need to be evaluated from first–principles calculations, for example, as it is done in VASP. `phono.py` was built to replace and extend `fropho` (fropho.sourceforge.net) and is thus also based on the *Parlinski–Li–Kawazoe method* [30], a numerical fitting approach to obtain force constants by given forces and displacements. An outstanding feature of the programme structure is the well-implemented symmetry finder with the use of `spglib` (spglib.sourceforge.net). As the source code is written in *python* within a logic structure, it is easy to understand and can as well be imported as a *python* module. Further information and a comprehensive manual of the programme can be found at atztogo.github.io/phonopy .

3.3.1 Phonon Methods

Assuming a three dimensional ($D = 3$) perfect crystal lattice with boundary conditions the expressions derived in Sec.3.2 need to be further elaborated. One defines $\vec{r}_l$ as the position vector to the Bravais lattice of unit cell l $\{l \in \mathbb{N} : l = 1, \ldots, N\}$, and within a unit cell the equilibrium

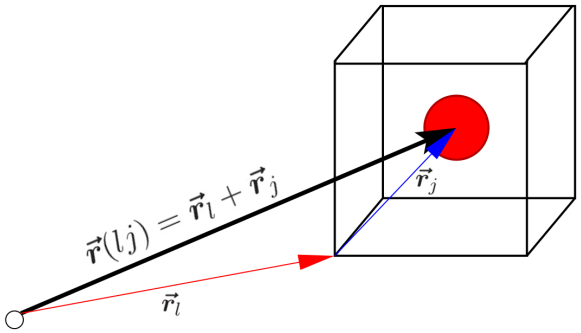


Fig. 3.3: Definition of the unit cell vector $\vec{r}_l$ and the relative equilibrium position vector $\vec{r}_j$.

position vector to basis atom j $\{j \in \mathbb{N} \ : \ j = 1, \ldots, N_u\}$ is denoted by $\vec{r}_j$ (see Fig.3.3 for reference). Combined the equilibrium position of atom j in unit cell l is thus written as

$$\vec{r}(lj) = \vec{r}_l + \vec{r}_j \ , \tag{3.48}$$

and the total number of DoF is $3N_u\, N$.

Second Order Force Constants Similar to (2.6) we can write Hamiltonian (3.18) generally as

$$\mathcal{H} = \hat{T} \oplus \hat{V} \ ,$$

where $\hat{T}$ and $\hat{V}$ are the kinetic and potential energy, respectively. The latter is represented as a function of the *momentary* atomic position vectors $\vec{x}(lj) = \vec{r}(lj) + \vec{u}(lj)$,

$$V = V\left[\vec{x}(11), \ldots, \vec{x}(lj), \ldots, \vec{x}(N\, N_u)\right] = V(\mathbf{x}) \ ,$$

so that it can be expanded with respect to the atomic displacements $\vec{u}(lj)$ around the equilibrium states $\mathbf{x} = \vec{r}(lj)$ (denoted by a subscript "0"),

$$\begin{aligned} V = V\left[\vec{r}(lj)\right] + \sum_{lj\alpha} \left(\frac{\partial V(\mathbf{x})}{\partial x_\alpha(lj)}\right)_0 u_\alpha(lj) + \\ + \frac{1}{2} \sum_{lj\alpha} \sum_{l'j'\beta} \left(\frac{\partial^2 V(\mathbf{x})}{\partial x_\alpha(lj)\, \partial x_\beta(l'j')}\right)_0 u_\alpha(lj)\, u_\beta(l'j') + \mathcal{O}(u^3) \approx \\ \approx V_0 - \sum_{lj\alpha} \Big(F_\alpha(lj)\Big)_0 u_\alpha(lj) + \frac{1}{2} \sum_{lj\alpha} \sum_{l'j'\beta} \Phi_{\alpha\beta}(lj, l'j')\, u_\alpha(lj)\, u_\beta(l'j') \quad , \end{aligned} \tag{3.49}$$

where the newly introduced indices (α, β) are for the Cartesian components ($D = 3 \ : \ \alpha, \beta = \{1, 2, 3\}$ or $\{x, y, z\}$). The first term $V_0 = 0$ can be identified as a constant contribution at the equilibrium positions (thus sometimes referred to as *binding energy* due to the form of the potential) and is for simplicity arbitrarily set to zero. The first and second derivative are respectively found to be a force due to a displacement of atom j in unit cell l within a potential $V(\mathbf{x})$ of the other atoms in the crystal,

$$F_\alpha(lj) = -\frac{\partial V(\mathbf{x})}{\partial x_\alpha(lj)} \ , \tag{3.50}$$

which vanishes in the linear term as the expansion is around equilibrium positions, and a second order *force constant*,

$$\Phi_{\alpha\beta}(lj, l'j') = \frac{\partial^2 V(\mathbf{x})}{\partial x_\alpha(lj)\, \partial x_\beta(l'j')} = -\frac{\partial F_\alpha(lj)}{\partial x_\beta(l'j')} \,. \tag{3.51}$$

The displacements of the atomic positions in a crystal are rarely larger than 3 % of the lattice constant. Therefore it is sensible to stay within the *harmonic approximation* (small displacements) and neglect terms of third order or higher, which finally yields

$$V(\mathbf{x}) \approx \frac{1}{2} \sum_{lj\alpha} \sum_{l'j'\beta} \Phi_{\alpha\beta}(lj, l'j')\, u_\alpha(lj)\, u_\beta(l'j') \,. \tag{3.52}$$

Using a potential approximated by (3.52) one derives a harmonic Hamiltonian, which can be diagonalised and delivers a solution of independent bases.

Dynamical Matrix Equivalent to the diagonalisation of the harmonic Hamiltonian is to solve the eigenvalue problem of the *dynamical matrix*. The Hamiltonian yields an equation of motion that reads

$$M_j\, \ddot{u}_\alpha(lj, t) = -\sum_{l'j'\beta} \Phi_{\alpha\beta}(lj, l'j')\, u_\beta(l'j', t) \,. \tag{3.53}$$

This is a linear force law and the ansatz,

$$u_\alpha(lj, t) = \frac{1}{\sqrt{M_j}}\, w_\alpha(lj, \vec{\boldsymbol{k}})\, \mathrm{e}^{-i\omega(\vec{\boldsymbol{k}})t} \,, \tag{3.54}$$

results in an eigenvalue equation for the $(3N_u\, N)$ normal frequencies $\omega(\vec{\boldsymbol{k}})$,

$$\sum_{l'j'\beta} \frac{1}{\sqrt{M_j\, M_{j'}}}\, \Phi_{\alpha\beta}(lj, l'j')\, w_\beta(l'j', \vec{\boldsymbol{k}}) = \omega^2(\vec{\boldsymbol{k}})\, w_\alpha(lj) \,, \tag{3.55}$$

where the direction of the oscillations is given by the eigenvectors $w_\alpha(lj, \vec{\boldsymbol{k}})$ (*polarisation*). Taking into account that due to crystal symmetry the force constants $\Phi_{\alpha\beta}(lj, l'j')$ show a translational invariance and thus are only dependent on the difference $\Delta l = l - l'$, one can develop the eigenvectors $\vec{\boldsymbol{w}}(lj)$ similar to the electronic band model by the use of the *Bloch theorem*,

$$w_\alpha(lj, \vec{\boldsymbol{k}}s) = c_\alpha(j, \vec{\boldsymbol{k}}s)\, \mathrm{e}^{i\vec{\boldsymbol{k}}\vec{\boldsymbol{r}}_l} \,,$$

where s is known as *band index* and wave vector $\vec{\boldsymbol{k}}$ is restricted only to the first Brillouin zone. (3.55) can then be simplified to

$$\sum_{j'\beta} D_{\alpha\beta}(jj', \vec{\boldsymbol{k}})\, c_\beta(j', \vec{\boldsymbol{k}}s) = \omega^2(\vec{\boldsymbol{k}}s)\, c_\alpha(j, \vec{\boldsymbol{k}}s) \,, \tag{3.56}$$

where $D_{\alpha\beta}(jj', \vec{\boldsymbol{k}})$ is the real symmetric *dynamical matrix*

$$D_{\alpha\beta}(jj', \vec{\boldsymbol{k}}) \doteq \frac{1}{\sqrt{M_j\, M_{j'}}} \sum_{l'} \Phi_{\alpha\beta}(l', jj')\, \mathrm{e}^{i\vec{\boldsymbol{k}}(\vec{\boldsymbol{r}}_{l'} - \vec{\boldsymbol{r}}_l)} \,, \tag{3.57}$$

which is reduced to a $(3N_u)$ dimensional eigenvalue problem. Consequently, the polarisation vector $\vec{\boldsymbol{w}}(\vec{\boldsymbol{k}}s)$ and frequencies $\omega^2(\vec{\boldsymbol{k}}s)$ are real as well and obtained by diagonalisation. The dynamical matrix can be determined by *ab-initio* calculations of the force constants; either directly with methods of perturbation theory, for example DF(P)T, or indirectly by finite displacement methods and the according evaluation of the resulting forces on the atoms (see also Sec.3.3.3). The atomic displacements $\vec{\boldsymbol{u}}(t)$ are given similar to (3.47),

$$u_\alpha(lj,t) = \sum_{\vec{\boldsymbol{k}},s} \sqrt{\tfrac{\hbar}{2N_u\, M_j \omega(\vec{\boldsymbol{k}}s)}} \Big(a(\vec{\boldsymbol{k}}s)\, \exp\Big[i\Big(\vec{\boldsymbol{k}}\vec{\boldsymbol{r}}(lj) - \omega(\vec{\boldsymbol{k}}s)t\Big)\Big] + \\ + a^\dagger(\vec{\boldsymbol{k}}s)\, \exp\Big[-i\Big(\vec{\boldsymbol{k}}\vec{\boldsymbol{r}}(lj) - \omega(\vec{\boldsymbol{k}}s)t\Big)\Big]\Big)\, c_\alpha(j,\vec{\boldsymbol{k}}s) \quad . \tag{3.58}$$

A frequency $\omega^2 < 0$ (*imaginary mode*) leads to a dynamically unstable crystal structure, which sometimes can relate to a phase transition or simply a relaxation of the crystal lattice. In `phono.py` imaginary frequencies are displayed by negative values in the plot of the phonon band structure.

Modified Parlinski–Li–Kawazoe Method This method is a numerical fitting approach and implemented in `phono.py` in order to derive force constants from given displacements and according forces [30]. In `phono.py` (3.51) is used as a displacement method based on finite differences and approximated by

$$\Phi_{\alpha\beta}\big(lj,l'j'\big) \sim -\frac{F_\alpha\big(lj,u_\beta(l'j')\big) - F_\alpha\big(lj\big)}{u_\beta\big(l'j'\big)}\,, \tag{3.59}$$

where the difference is taken from

$$\Delta\vec{\boldsymbol{x}}(l'j') = \big[\vec{\boldsymbol{r}}(l'j') + \vec{\boldsymbol{u}}(l'j')\big] - \vec{\boldsymbol{r}}(l'j') = \vec{\boldsymbol{u}}(l'j')\,.$$

The latter is written in matrix representation for an atomic pair $\{lj,l'j'\}$ as

$$\bar{\mathrm{F}} = -\bar{\mathrm{U}}\,\bar{\mathrm{P}}$$

$$\big(F_x\,,\,F_y\,,\,F_z\big) = -\big(u_x\,,\,u_y\,,\,u_z\big) \begin{pmatrix} \Phi_{xx} & \Phi_{xy} & \Phi_{xz} \\ \Phi_{yx} & \Phi_{yy} & \Phi_{yz} \\ \Phi_{zx} & \Phi_{zy} & \Phi_{zz} \end{pmatrix} \quad , \tag{3.60}$$

where the symmetry of the force constants ($\Phi_{\alpha\beta} = \Phi_{\beta\alpha}$) has been taken into account. The matrix equation can be expanded for a sufficient number of atomic displacements and forces, and solved using the *Moore–Penrose inverse* (pseudoinverse), so that

$$\bar{\mathrm{P}} = -\begin{pmatrix} \bar{\mathrm{U}}_1 \\ \bar{\mathrm{U}}_2 \\ \vdots \end{pmatrix}^+ \begin{pmatrix} \bar{\mathrm{F}}_1 \\ \bar{\mathrm{F}}_2 \\ \vdots \end{pmatrix} \quad . \tag{3.61}$$

The required number of sufficient displacements to solve the equations is reduced by applying site–point symmetry operations onto (3.60),

$$\hat{R}(\bar{\mathrm{F}}) = -\hat{R}(\bar{\mathrm{U}})\,\bar{\mathrm{P}}\,, \tag{3.62}$$

where $\hat{R}$ is the symmetry operation centring at $\vec{\boldsymbol{u}}(lj)$, which can be written in the matrix representation of the rotation operator $\bar{\mathrm{R}}$

$$\begin{aligned}\hat{R}(\bar{\mathrm{F}}) &\longrightarrow \bar{\mathrm{R}}\,\bar{\mathrm{F}}\big(\hat{R}^{-1}(l'j')\big)\,,\\ \hat{R}(\bar{\mathrm{U}}) &\longrightarrow \bar{\mathrm{R}}\,\bar{\mathrm{U}}\big(l'j'\big)\,.\end{aligned}$$

Group Velocity This property can be calculated by

$$\begin{aligned}v_{\mathrm{G}}(\vec{\boldsymbol{k}}s) &= \vec{\boldsymbol{\nabla}}_k\,\omega(\vec{\boldsymbol{k}}s) = \\ &= \frac{1}{2\omega(\vec{\boldsymbol{k}}s)}\,\partial_{\vec{\boldsymbol{k}}}\big(\omega(\vec{\boldsymbol{k}}s)\big)^2 = \frac{1}{2\omega(\vec{\boldsymbol{k}}s)}\,\langle\vec{\boldsymbol{c}}(\vec{\boldsymbol{k}}s)|\partial_{\vec{\boldsymbol{k}}}\,\hat{D}(\vec{\boldsymbol{k}})|\vec{\boldsymbol{c}}(\vec{\boldsymbol{k}}s)\rangle\ ,\end{aligned} \tag{3.63}$$

where the diagonalised dynamical matrix solves the eigenvalue problem (3.56) for the values of $\omega^2(\vec{\boldsymbol{k}}s)$. `phono.py` implements the calculation of the group velocity using finite differences by

$$v_{\mathrm{G}}(\vec{\boldsymbol{k}}s) \sim \frac{1}{2\omega(\vec{\boldsymbol{k}}s)}\,\langle\vec{\boldsymbol{c}}(\vec{\boldsymbol{k}}s)|\tfrac{\Delta\bar{\mathrm{D}}(\vec{\boldsymbol{k}})}{\Delta\vec{\boldsymbol{k}}}|\vec{\boldsymbol{c}}(\vec{\boldsymbol{k}}s)\rangle\ . \tag{3.64}$$

The difference $\Delta\vec{\boldsymbol{k}} = (\Delta k_x\,,\,\Delta k_y\,,\,\Delta k_z)$ is described in reciprocal space by Cartesian coordinates. The method employs *central difference*, where the distance $\pm\Delta q_\alpha\,(\alpha = x, y, z)$ is specified by the user in units of the reciprocal space.

3.3.2 Thermodynamic Properties

As it was demonstrated before in Sec.3.2.2 phonons show bosonic behaviour. Thermodynamic properties are thus derived from Bose–Einstein statistics. The partition function $\mathcal{Z}$ for a system of $\tilde{N}$ phonons reads,

$$\mathcal{Z} = \prod_{\vec{\boldsymbol{k}},s}\frac{\mathrm{e}^{-\frac{1}{2}\beta\,\hbar\omega(\vec{\boldsymbol{k}}s)}}{1-\mathrm{e}^{-\beta\,\hbar\omega(\vec{\boldsymbol{k}}s)}} = \prod_{\vec{\boldsymbol{k}}s}\Big[2\,\sinh\big(\tfrac{1}{2}\beta\,\hbar\omega(\vec{\boldsymbol{k}}s)\big)\Big]^{-1}\,, \tag{3.65}$$

with $\beta = (k_{\mathrm{B}}\,T)^{-1}$ being the inverse temperature. Its derivation can be found in Apx.F . The average occupation number of a phonon mode $\{\vec{\boldsymbol{k}}s\}$ is then according to a Bose–Einstein distribution,

$$\langle n(\vec{\boldsymbol{k}}s)\rangle = \frac{1}{\mathrm{e}^{\beta\,\hbar\omega(\vec{\boldsymbol{k}}s)}-1} \quad\Longrightarrow\quad \langle\tilde{N}\rangle = \sum_{\vec{\boldsymbol{k}},s}\langle n(\vec{\boldsymbol{k}}s)\rangle\,, \tag{3.66}$$

which yields the (harmonic) internal energy of the phonon system to be

$$U = \sum_{\vec{\boldsymbol{k}},s}\hbar\omega(\vec{\boldsymbol{k}}s)\,\Big(\,\langle n(\vec{\boldsymbol{k}}s)\rangle + \tfrac{1}{2}\Big)\,. \tag{3.67}$$

The free energy is according to statistical mechanics and (3.65),

$$F = -k_{\mathrm{B}}T\,\ln\mathcal{Z} = \frac{1}{2}\sum_{\vec{\boldsymbol{k}},s}\hbar\omega(\vec{\boldsymbol{k}}s) + k_{\mathrm{B}}T\sum_{\vec{\boldsymbol{k}},s}\ln\Big(1-\mathrm{e}^{-\beta\,\hbar\omega(\vec{\boldsymbol{k}}s)}\Big)\,, \tag{3.68}$$

which with (1.41a) gives an expression for the entropy,

$$
\begin{aligned}
S &= -\left(\frac{\partial F}{\partial T}\right)_V = \\
&= -\frac{1}{T}\sum_{\vec{k},s}\frac{\hbar\omega(\vec{k}s)}{\mathrm{e}^{\beta\,\hbar\omega(\vec{k}s)}-1} - k_\mathrm{B}\sum_{\vec{k},s}\ln\left[1-\mathrm{e}^{-\beta\,\hbar\omega(\vec{k}s)}\right] = \\
&= \frac{1}{2T}\sum_{\vec{k},s}\hbar\omega(\vec{k}s)\coth\left(\tfrac{1}{2}\beta\,\hbar\omega(\vec{k}s)\right) - k_\mathrm{B}\sum_{\vec{k},s}\ln\left[2\,\sinh\left(\tfrac{1}{2}\beta\,\hbar\omega(\vec{k}s)\right)\right] .
\end{aligned}
\tag{3.69}
$$

Finally, the (harmonic) response of the system can be described by entering (3.67) into (1.28) and derive the heat capacity at constant volume,

$$
\begin{aligned}
C_V &= \left(\frac{\mathrm{d}U}{\mathrm{d}T}\right)_V = T\cdot\left(\frac{\partial S}{\partial T}\right)_V = \\
&= k_\mathrm{B}\sum_{\vec{k},s}\left(\beta\,\hbar\omega(\vec{k}s)\right)^2\frac{\mathrm{e}^{\beta\,\hbar\omega(\vec{k}s)}}{\left(\mathrm{e}^{\beta\,\hbar\omega(\vec{k}s)}-1\right)^2} \quad .
\end{aligned}
\tag{3.70}
$$

3.3.3 Workflow in `phono.py`

The derivation of phonon properties from a material of interest is done by a combination of `phono.py` and an interface to any external calculator that features the calculation of interatomic forces (preferably from first–principles). The workflow of a calculation is shown in the diagram of Fig.3.4 . Boxes and diamonds show tasks related to `phono.py` and external calculators, respectively. Terms in circles symbolise input and intermediate output data for further analysis. The starting point is a carefully relaxed unit cell with very small uncertainties in the interacting forces. `phono.py` analyses the applicable symmetry operations of the crystal and creates a supercell of desired and necessary size (set by the DIM tag or the `--dim` option). A supercell too small in size might result in an incomplete response caused by an atomic displacement, which would cut off important contributions to the phonon dispersion, especially in the localised bands. The interaction range thus needs to be carefully tested, but in most cases as a rule of thumb a supercell should contain around 100 atoms in order to derive satisfying results. As a next step the user can choose between two ways that indirectly yield the dynamical matrix $\bar{\mathrm{D}}(\vec{k})$,

(1) from the *set of atomic forces* due to finite displacements, or
(2) from the *force constants* calculated by means of perturbation theory.

The default amplitude for the displacements is 0.01 Å, but can be arbitrarily set by the user (DISP or `--disp`) with caution; a displacement value that is too small causes numerical errors in the force constants, where a large value leaves the confinement of the harmonic approach and demands anharmonic contributions to the potential. In default mode `phono.py` applies *plus–minus* displacements onto a minimum number of atoms in the supercell under symmetry aspects for an elastic wave whose propagating direction is one of the basis vectors of the crystal. Errors of residual forces are thus often cancelled out. Forces or force constants can be calculated by VASP

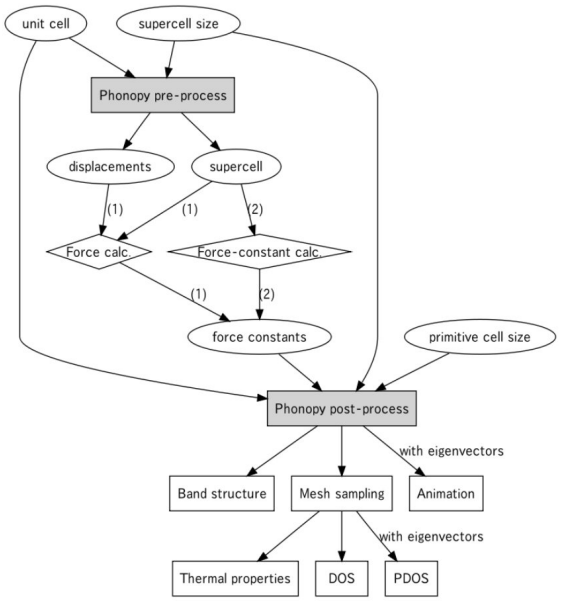
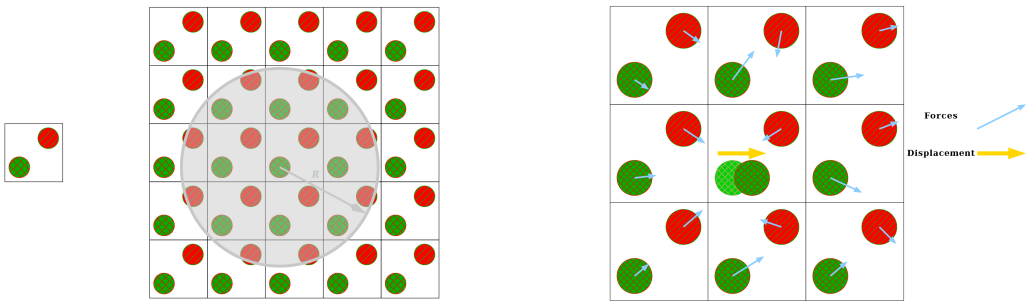


Fig. 3.4: The workflow of a phonon calculation with `phono.py` used in pre– and post–processing. Illustration taken from [29].

using the settings IBRION=−1 (no relaxation) or IBRION=8 (DFPT), respectively. `phono.py` then builds the dynamical matrix (3.57) at an arbitrary wave vector $\vec{\boldsymbol{k}}$ from the force constants of the supercell and solves (3.56) in order to obtain the according phonon frequencies $\omega(\vec{\boldsymbol{k}}s)$ and polarisation vectors $\vec{\boldsymbol{c}}(j, \vec{\boldsymbol{k}}s)$. Finally, the post–process includes the calculation and plotting of features as the *band structure*, phonon (P)DOS and *thermal properties* according to Sec.3.3.2, as well as an output of the polarisation eigenvectors and an animation–interface for the vibrational modes.



(a) (b)

Fig. 3.5: (a) Phonon interaction range R within a supercell. Next to it is a unit cell for reference. In this example a $4 \times 4 \times 4$ supercell would be necessary to describe all phonon contributions correctly. (b) Schematic picture of the finite displacement method. An atom is displaced from its equilibrium position which causes forces according to (3.50) that can be calculated by first–principles. The drawn displacement is exaggerated.

Chapter 4

Methodology

In the last couple of years bilayer films consisting of FeRh and FePt have attracted a great interest to be used as an application for *Heat–Assisted Magnetic Recording* (HAMR) media. The most prominent features exploited from these materials are in case of FeRh the metamagnetic first–order phase transition from an antiferromagnetic (AFM) to ferromagnetic (FM) phase with increasing temperature at $\sim$ 320K, and in case of FePt the reduction of the coercive field by heating.

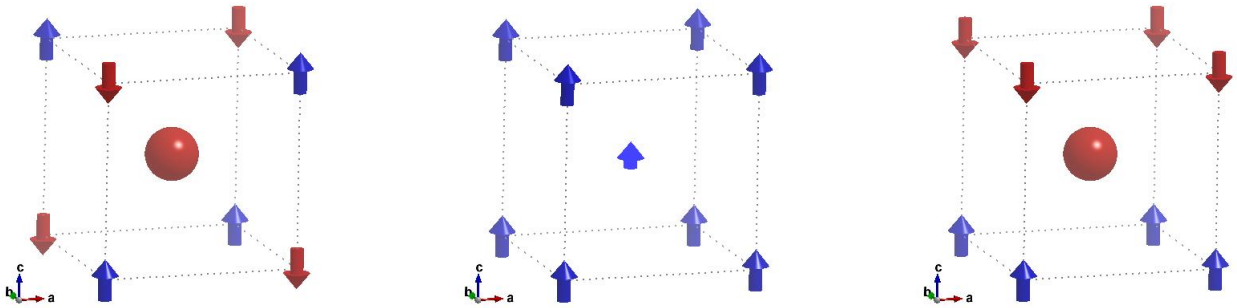
4.1 Unit Cell & Magnetic Configuration of FeRh

The unit cell of FeRh crystallises in a cubic $B2$ structure (CsCl) with space group symmetry $Pm\bar{3}m$ (#221). As it can be seen in Fig.4.1, the individual Fe and Rh ions arrange in two resembled simple cubic lattices, where the Rh lattice is shifted by half a lattice constant in [111] direction. The structure is similar to a bcc cell, however, the centered ion (Rh) is not the same as those in the corner of the cell (Fe).

In this structure there are three possible types of magnetic assembly. Lowest energy, thus the ground state, can be reached by the *AFM II* spin structure, where all Fe spins in the {111} planes are FM coupled and alternate between *spin up* and *spin down* for each individual plane (Fig.4.1a). In case of AFM coupling the Rh magnetic moments are all zero. It was shown (see e.g. Shirane *et al.* [31]) that there is another stable solution with higher energy and larger volume that forms a *FM* phase (Fig.4.1b). The magnetic moment for the individual Fe sites remains approximately constant around $3\,\mu_B$ for both phases, where the Rh moments change from zero to about $1\,\mu_B$ in the FM phase. The third, but most improbable state is *AFM I*. In this case the Fe spins are ferromagnetically coupled in the {001} lattice planes instead of a diagonal propagation, and alternate accordingly (Fig.4.1c). The transition temperature to the paramagnetic state lies at about $T_C \sim 675\,\mathrm{K}$ [32].

Models for the Metamagnetic Phase Transition Recently, FeRh has attracted attention for the use in recording media as it undergoes a temperature–induced metamagnetic phase transition between the AFM II and FM state at $T_M \sim 340$ – $350\,\mathrm{K}$, where the exact transition temperature sensitively depends on the conditions of sample preparation. This *first–order* transition was experimentally discovered already more than 70 years ago and is not only identified by a change in the magnetic ordering, but also by a large volume (about 1 %) and entropy change ($\Delta S \sim 14.0\,\mathrm{mJ/(gK)}$) between the different phases due to the electronic contribution related to spin fluctuations of the Rh atoms. The mechanism behind the transition is not well understood. The finding of an appropriate model is still in process and under debate.

One of the first explanations was given by *Kittel* suggesting an *exchange inversion model* [33], where the transition occurs due to volume–dependent exchange interaction inversion. The ex-



(a) (b) (c)

Fig. 4.1: Cubic FeRh unit cell in (a) the magnetic ground state AFM II, (b) the stable FM state, (c) the AFM I state. The arrows symbolise Fe sites and their spin direction. The Rh site is always body–centred (in red, except for the smaller spin in the FM state).

change parameters are in a linear relation with the lattice parameters and change their sign at a critical value, so that the system is driven by the change of magnetoelastic energy. This model is, however, incompatible with the observed large entropy change.

Mryasov proposed a model based on a Heisenberg Hamiltonian, but also including effective higher–order exchange interactions up to the bi-quadratic terms $(\vec{\boldsymbol{S}}_i \cdot \vec{\boldsymbol{S}}_j)^2$ [34]. Size and direction of the magnetic moment of Rh are dictated by the Weiss field of the surrounding Fe sites in a non-linear fashion (compared to FePt, for example, where it is linear [35]). The effective spin Hamiltonian therefore only contains degrees of freedom from the Fe sites, since the induced Rh moments are incorporated by an (Fe–Rh–Fe) superexchange–interaction, which generally favours AFM coupling and is still within the frame of a Heisenberg model. However, as the bi-quadratic expansion is degenerate in energy for both magnetic states the transition cannot be driven by those terms alone. *Barker* and *Chantrell* thus extended this model and emphasised the importance of a full expansion of the quadratic interactions into four–spin exchange terms $(\vec{\boldsymbol{S}}_i \cdot \vec{\boldsymbol{S}}_j)(\vec{\boldsymbol{S}}_k \cdot \vec{\boldsymbol{S}}_l)$, however, simplified the Hamiltonian to the competition between the standard bilinear and the newly introduced four–spin interactions [36]. Since fluctuations are now determined by four independent spins instead of two, the effect of the four–spin term will decrease more rapidly with temperature than the bilinear term. After a fit of the interaction parameters ($J_{\langle 001\rangle}$, $J_{\langle 011\rangle}$, $J_{\langle Q\rangle}$) from experimental data they solved the *Landau–Lifshitz–Gilbert* (LLG) equation using atomistic spin dynamics (ASD), which led to a good approximation of the transition temperature T_M.

A similar, but slightly different approach is taken by *Polesya* et al. [37]. In earlier first–principle studies by *Sandratskii and Mavropoulos* the density of states (DOS) for both magnetic phases show a strong hybridisation of the Fe and Rh states in both spin channels [38]. Especially in the case of an AFM configuration this means that instead of applying a previously suggested Stoner model, where the exchange interaction of the Fe moments at the Rh sites is treated as an external mean–field, there is a strong covalency with the spin–polarised Fe states leading to a non-zero intra–atomic spin polarisation of the Rh states (implicit spin splitting). In other words, the electronic states of the two equal spin projections of the Rh–DOS hybridise with the according

Fe sublattices (*up* and *down*) and the total moment of the Rh site averages to zero due to symmetry, but not because of a vanishing spin density (as it would be the case in a Stoner picture). In the FM case, however, the spin density is redistributed and a magnetic moment appears on the Rh site. Polesya *et al.* showed a linear dependence of the Rh magnetic moment on the FM ordering of the Fe sublattices and that the former contributes to the stabilisation of the FM state. They as well emphasise the importance of the biquadratic exchange interaction to be included in the Heisenberg Hamiltonian in order to describe the metamagnetic phase transition. However, whilst Barker and Chantrell account the extended terms to an (Fe–Rh–Fe) superexchange, the model Hamiltonian of Polesya acts solely on the Fe atoms, where the first part includes the indirect Fe–Fe exchange $J_{ij}^{\text{Fe–Fe}}$ (AFM) and the other term scales the FM exchange interaction via a response function $\chi^{\text{Fe–Rh}}$ dependent on the induced Rh moments. The parameters are obtained self–consistently from uncompensated Disordered Local Moment (uDLM) calculations and the Hamiltonian implemented in Monte Carlo (MC) simulations yields reasonable results with a transition temperature $T_{\text{M}} \sim 320\,\text{K}$.

4.2 Optimisation

In order to derive reliable results it is crucial for DFT calculations to test and optimise most of the necessary parameters, such as unit cell, reciprocal space parametrisation (k–points), energy cutoff of the plane–wave basis set, and so on. The following shall present an overview of the first rough cell optimisation (relaxation) and the testing of the total energy difference at varying k–points and energy cutoff settings.

4.2.1 Relaxations

Before the optimal sampling of the Brillouin zone (k–points) and a usable energy cutoff for the plane–wave basis set can be determined it is necessary to create unit cells where every individual ion of the cell is relaxed to its equilibrium position in the potential of the respective neighbouring sites (including electrons). For comparison, the exchange and correlation effects are treated by two slightly different functionals in the *generalised gradient approximation* (GGA) for two individual runs, namely the version of Perdew–Burke–Ernzerhof (PBE) and the revised Perdew–Burke–Ernzerhof (RPBE), respectively. The Fe (Rh) PAW potential considers $3\text{p}\,, 4\text{s}^{1}, 3\text{d}^{7}$ ($4\text{p}\,, 5\text{s}^{1}, 4\text{d}^{8}$) as valence electrons (both *pv*–potentials). Two further Fe potentials that either additionally treat 3s states as valence (*sv*–potential), or on the other hand skip the 3p states (*standard*), have been tested as well and deliver slightly different results. The cutoff–energy of the plane–wave basis was selected to a quite high value of 600 eV. It will turn out (see Sec.4.2.2) that this is about +9 % higher than the determined optimal value for E_{cut}. The Brillouin zone was also selected to be sampled with a dense $(24 \times 24 \times 24)$ Γ–centred mesh of k–points. In order to ensure accurate forces during relaxation furthermore a superfine Fast Fourier Transform (FFT) grid for the augmentation charges and a Methfessel and Paxton smearing with $\sigma \leq 0.2\,\text{eV}$ has been applied. The unit cell has been relaxed by a conjugate–gradient algorithm, as well as allowing changes of the cell structure and volume until forces between the ions were less than $1\,\text{meV/Å}$.

The unit cell shown in Fig.4.1 could only be used to simulate the FM state due to periodic boundary conditions (symmetry) implemented in VASP. It is therefore inevitable to use

larger cells in order to be able to establish all magnetic configurations and compare them with one another. The most convenient primitive cell that also has been used throughout this thesis is built up by fcc basis vectors and contains two Fu (2 FeRh). See Tab.4.1 and Fig.4.2 for reference.

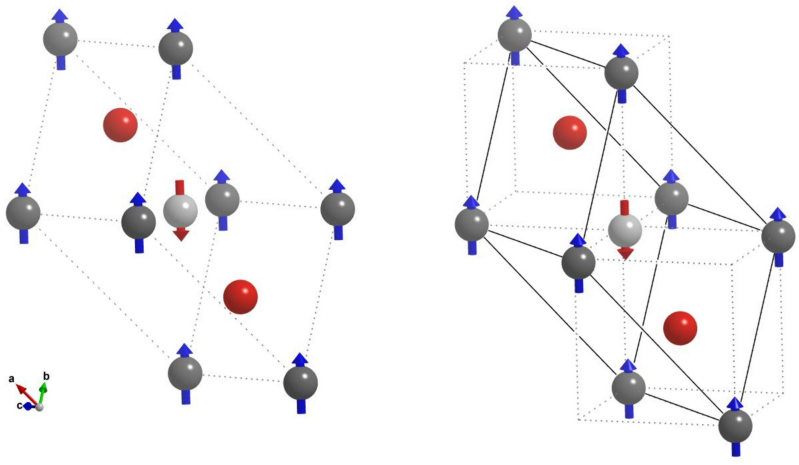
Tab. 4.1: Basis vectors of the fcc–like primitive cell and direct coordinates for the lattice sites of FeRh. $\tilde{a} = 2a$ is twice the lattice constant of the cubic unit cell a.

$$\vec{a}_1 = \tilde{a}\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad , \quad \vec{a}_2 = \tilde{a}\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \quad , \quad \vec{a}_3 = \tilde{a}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

(fcc basis)

	1	2
Fe	$\left(0\,,\,0\,,\,0\right)$	$\left(\frac{1}{2}\,,\,\frac{1}{2}\,,\,\frac{1}{2}\right)$
Rh	$\left(\frac{1}{4}\,,\,\frac{1}{4}\,,\,\frac{1}{4}\right)$	$\left(\frac{3}{4}\,,\,\frac{3}{4}\,,\,\frac{3}{4}\right)$

(direct coordinates)



(a) (b)

Fig. 4.2: (a) fcc conventional cell as the primitive structure for the AFM II ground state. (b) Comparison between a fcc *(solid)* and a cubic structure *(dotted)*. Fe sites with *spin up* in grey, with *spin down* in silver, Rh sites in red.

The second possible primitive cell containing two Fu as well has rhombohedral basis vectors and is depicted in Fig.4.3. As it can be seen from Tab.4.2 the rhombohedral basis is not as elegant to treat as it is the case with the fcc basis, due to the radical expressions of the vectors and the three different angles ($\alpha = 60°$, $\beta = 90°$, $\gamma = 120°$) between the vectors.
It should be noted that both cells are related to each other due to the cubic class of the system

which is reflected in the length of all basis vectors

$$|\vec{a}_i| = \frac{\tilde{a}}{\sqrt{2}} = \frac{2a}{\sqrt{2}} = \sqrt{2}\,a\,, \tag{4.1}$$

where $\tilde{a} = 2a$ is the lattice constant of the primitive cells and a is the one of the basic cubic unit cell. There are further possibilities for structures containing equal or more Fu, such as a simple *double stack* of the basic cell or a larger tetragonal cell as it was used for example for the first–principle calculations of Moruzzi and Marcus [39]. For comparison, an eight–fold supercell (8 FeRh) showing the equivalence of the two latter structures is shown in Fig.4.4.

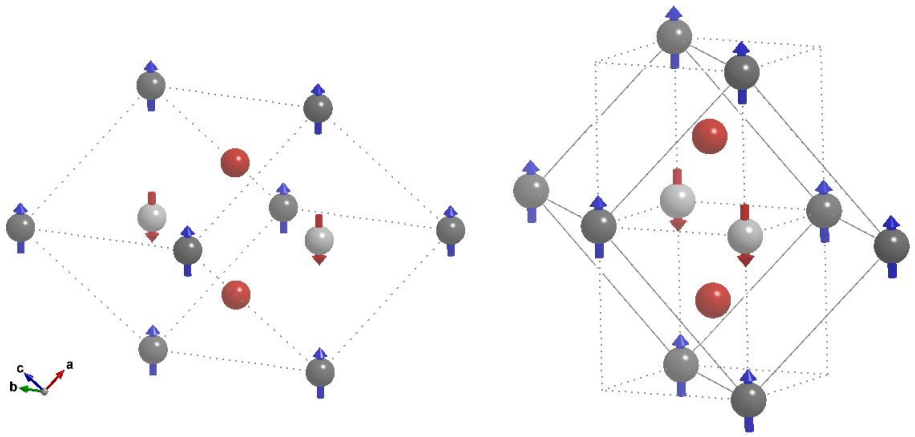
Tab. 4.2: Basis vectors of the rhombohedral primitive cell and direct coordinates for the lattice sites of FeRh. $\tilde{a} = 2a$ is twice the lattice constant of the cubic unit cell a.

$$\vec{a}_1 = \tilde{a}\left(\tfrac{1}{\sqrt{2}} \quad -\tfrac{\sqrt{3}}{\sqrt{2}} \quad 0\right) \quad , \quad \vec{a}_2 = \tilde{a}\left(\tfrac{1}{\sqrt{2}} \quad \tfrac{\sqrt{3}}{\sqrt{2}} \quad 0\right) \quad , \quad \vec{a}_3 = \tilde{a}\left(\tfrac{1}{\sqrt{2}} \quad \tfrac{1}{\sqrt{6}} \quad \tfrac{2}{\sqrt{3}}\right)$$

(rhombohedral basis)

	1	2
Fe	$\left(0\,,\,0\,,\,0\right)$	$\left(\tfrac{1}{2}\,,\,0\,,\,\tfrac{1}{2}\right)$
Rh	$\left(\tfrac{1}{4}\,,\,\tfrac{1}{2}\,,\,\tfrac{1}{4}\right)$	$\left(\tfrac{3}{4}\,,\,\tfrac{1}{2}\,,\,\tfrac{3}{4}\right)$

(direct coordinates)



(a) (b)

Fig. 4.3: (a) Primitive cell with rhombohedral basis vectors as another possibility to depict an AFM II state under periodic boundary conditions. (b) Comparison between rhombohedral *(solid)* and basic cubic structure *(dotted)*. Fe sites with *spin up* in grey, with *spin down* in silver, Rh sites in red.

The results of the relaxation runs with the settings from above are summarised in Tab.4.3. Generally, the AFM II state is individually compared to the respective FM state for the fcc–like

and rhombohedral structure under different exchange–correlation (xc) potential setups. In the case of fcc, PBE and RPBE exchange are compared with one another both with *pv* electrons for the Fe sites (as stated before the valence configuration for the Rh ions has not been changed and always kept at the recommended *pv* setting), where in the rhombohedral case the influence of the three different setups for the valence band has been investigated. As expected, both cells yield similar results, which can be seen when comparing the calculations with PBE(*pv*) exchange.
For PBE the ground–state lattice constant relaxes to $a = 2.990\,\text{Å}$ ($3.007\,\text{Å}$) for the AFM II (FM) state, which is in good agreement with experimental data and previous calculations. Note that for RPBE the cell volume is slightly larger, $a = 3.016\,\text{Å}$ ($3.033\,\text{Å}$), which together with the increased local magnetic moments will have an effect on the later investigated phonon band structure (see Sec.5.1). The energy difference of the FM to the ground state (AFM II) varies significantly with the chosen potential and valence band in a range of $\pm 20\,\text{meV}$ per FU. The smallest energy difference per FU can be found by using RPBE(*pv*), while the latter yields about 1 – 2 % larger local magnetic moments at the Fe sites, however, the induced Rh local moment in the FM state remains approximately constant compared to PBE(*pv*). PBE(*sv*) clearly underestimates the Fe moments in both phases.

Tab. 4.3: Comparison of the energy, energy difference, basic cubic lattice constant, length of lattice vector, and local magnetic moment of the Fe and Rh sites for the ground–state (AFM II) and FM state. The last column shows the used exchange and correlation potential with the according configuration of the valence band. Energy differences are given per formula unit (FU).

		E(eV)	ΔE(meV/FU)	a(Å)	$\lvert\vec{a}_i\rvert$	m_{Fe} (μ_{B})	m_{Rh} (μ_{B})	xc–pot	
FCC	AFM II	−31.289	—	2.990	4.229	±3.145	0.000	*pv*	PBE
	FM	−31.182	53.3	3.007	4.252	3.205	1.049		
	AFM II	−29.236	—	3.016	4.266	±3.205	.	.	RPBE
	FM	−29.148	44.1	3.033	4.289	3.245	1.052		
RHOMBO	AFM II	−31.289	—	2.991	4.229	±3.145	0.000	*pv*	PBE
	FM	−31.183	53.0	3.007	4.253	3.204	1.050		
	AFM II	−31.470	—	2.990	4.228	±3.022	.	*sv*	
	FM	−31.355	57.2	3.006	4.251	3.098	1.047		
	AFM II	−31.259	—	2.990	4.228	±3.123	.	std	
	FM	−31.130	64.4	3.007	4.252	3.173	1.055		

4.2.2 KPOINTS & ENCUT Optimisation

Appropriate values for the KPOINTS– and ENCUT–tag, which establishes a reciprocal space grid (with reciprocal space point volume $\{N_{b_1} \times N_{b_2} \times N_{b_3}\}$) and includes plane–waves with $E_{\text{kin}} < E_{\text{cut}}$ in the basis set, respectively, have been tested for different cell sizes and magnetic configurations of the FeRh system. The total energy convergence with rising KPOINTS/ENCUT values is shown in Fig.4.5 as energy difference $\mathrm{d}E = E_i - E_{\text{ref}}$ with respect to a reference energy of the highest KPOINTS/ENCUT value.

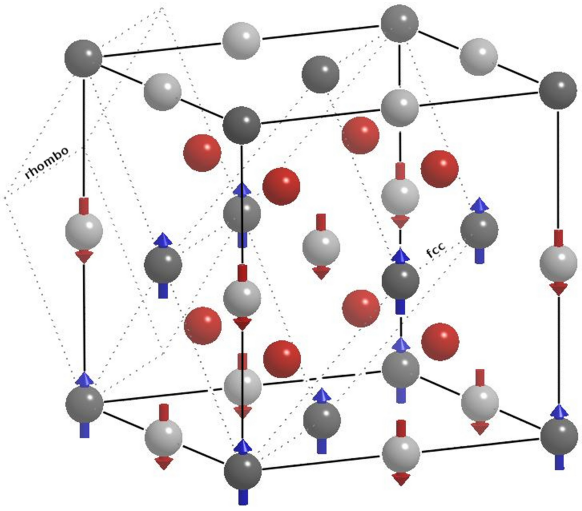
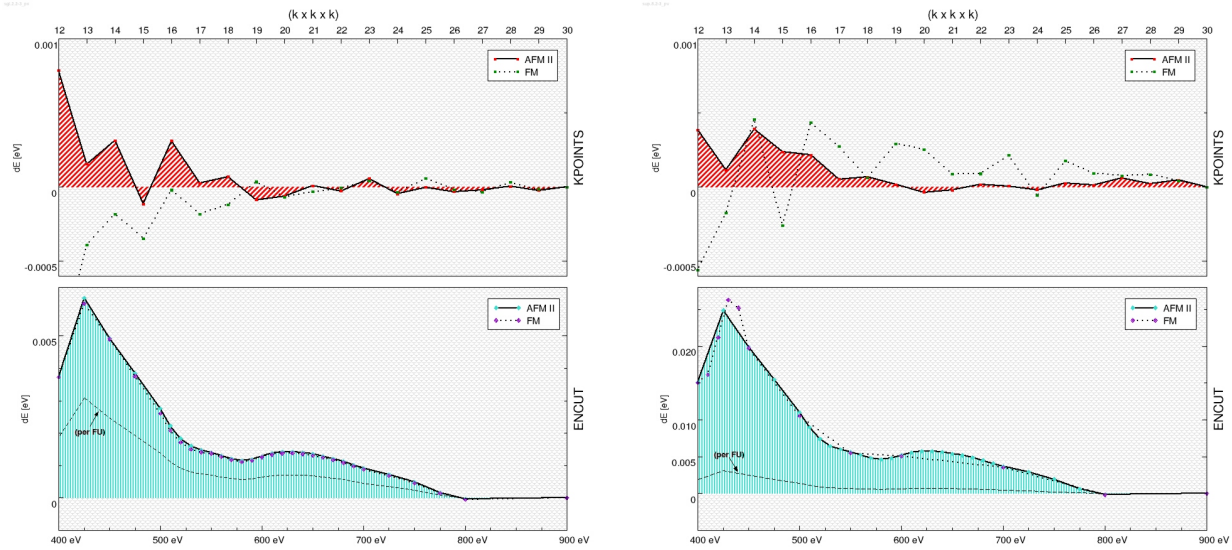


Fig. 4.4: Comparison between the fcc conventional (centred) and the rhombohedral cell (outside) in an eight–fold supercell structure (8 FeRh). Fe sites with *spin up* in grey, with *spin down* in silver, Rh sites in red.

For the calculations presented a functional that treats the exchange and correlation within the *generalised gradient approximation* in a version of Perdew–Burke–Ernzerhof (PBE GGA) has been applied, where the Fe (Rh) PAW potential considers $3p\,, 4s^{1}, 3d^{7}$ ($4p\,, 5s^{1}, 4d^{8}$) as valence electrons. Two further potentials that either additionally treat 3s (4s) states as valence, or on the other hand skip the 3p (4p) states, have been tested as well, but did not lead to satisfying results and are thus not shown here.

Two unit cells with different size have been tested. In Fig.4.5a the fcc–like cell as shown in Fig.4.2 that contains the primitive FeRh structure twice has been used, whereas in Fig.4.5b a supercell structure with 8×FeRh has been selected. In both cases two magnetic configurations (AFM II, FM) have been compared.

The increase of the number of k–points is connected with oscillations of the total energy around E_{ref}, however, from a mesh of $(12 \times 12 \times 12)$ points onwards the energy difference remains within a range of ± 1 meV for both cell types and magnetic configurations. Finally, a dense Γ–centred mesh of $(18 \times 18 \times 18)$ k–points has been used to sample the Brillouin zone for subsequent calculations (unless explicitly stated otherwise) in order to ensure a proper convergence of the total energy as it is shown in Fig.4.5.
The graphs of varying ENCUT value for both cells are qualitatively similar which can be seen by comparing the total energy curves per formula unit (FU). The latter are actually almost identical, which cannot be seen from the graphs directly as they are plotted on a different scale. Unfortunately, convergence to a non-varying total energy is quite slow and a high energy cutoff for the plane–wave energy E_{cut} is needed. Since high values for the cutoff energy would be timewise prohibitive a compromise is taken by a selecting the the latter to 550 eV (unless stated otherwise). This is already +88 % higher than the standard value (set by the ENMAX–tag) for the used Fe (PAW) potential, however, together with the above mentioned setting for KPOINTS it ensures a convergence of the total energy to less than 1 meV/FU.



(a) (b)

Fig. 4.5: Optimisation of the reciprocal space point density (KPOINTS, top) and the energy cutoff E_{cut} of the plane–wave basis set (ENCUT, bottom) on basis of the total energy convergence. Plane–waves with $E_{\text{kin}} < E_{\text{cut}}$ are included in the basis set. Energy convergence is demonstrated as the difference with respect to the energy value of the highest KPOINTS/ENCUT value. Dashed line shows the energy difference per formula unit (FU). (a) Conventional cell with 2 FeRh. (b) Supercell containing 8 FeRh.

4.2.3 Volume Variation

With the optimised number of k–points and energy cutoff a total energy curve with varying volume of the primitive cell has been calculated for the AFM II, as well as the FM state, and the result per FU is shown in Fig.4.6. The global energy minima at 26.73 Å^3 and 27.27 Å^3 for the AFM II and FM state are equivalent to an according lattice constant of 2.99 Å and 3.01 Å, respectively, thus confirm the correct cell relaxations from Sec.4.2.1. The ratio of both energetically favoured volumes indicates a magnetovolume effect of about 2 %, which slightly overestimates the experimental value of 1 %. The two curves cross at 30.23 Å^3, and beyond this point $E_{\text{FM}} < E_{\text{AFM}}$, however, the according lattice constant is already far beyond the dissolving point of the cell ($> +3\,\%$) and important contributions to the total energy such as entropy and lattice vibrations (phonons) are not considered, so that a simple picture of thermal expansion can be excluded and the metamagnetic phase transition indeed drives the lattice expansion.

4.3 Heat–Assisted Magnetic Recording

A new standard of achieving an increase of bit or areal density in magnetic hard disk drives (HDD) is the new technology of *heat–assisted magnetic recording* (HAMR). In order to reduce thermal write errors current concepts employ a combination of a hard and soft magnetic layer which are

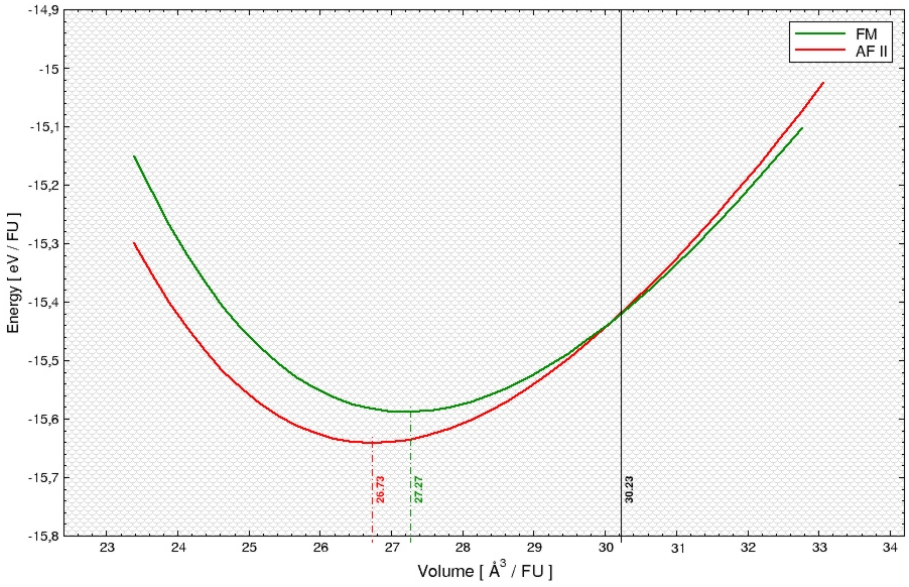


Fig. 4.6: Energy graph with varying volume of the primitive cell for the AFM II (red) and FM state (green). The minima of the curves correspond to the accordingly relaxed lattice constants. The two lines cross at a volume that is far beyond the dissolving point. Energy and volume are given per formula unit (Fu).

coupled by the formalism of exchange spring media [40–42]. The hard layer consists of FePt known for the high magnetic anisotropy and Curie temperature, where for the soft layer FeRh is chosen as it undergoes a metamagnetic transition from an Afm to a Fm state at about $T \sim 350\,\mathrm{K}$. The writing process in Hamr media is triggered by a laser that locally heats the FeRh above the transition temperature via a sophisticated lens system (plasmonic *Near–Field Transducer* (Nft) [43], focus of about 5 nm) while a head field changes the magnetic orientation. Due to the coupling to the hard layer and its large anisotropy it is possible to store this magnetic state for a long period of time, as the FeRh falls back to a state with no net magnetic moment while cooling. It is the hope that by these measures the distance between every individual bit can be lowered or, in other words, the number of bits per square inch can be significantly increased. Current Hdd products on the market reach an areal density of $\sim$ 640 GB in^{-2}, but with Hamr it is the hope to reach $\sim$ 1 TB in^{-2} and beyond, which has already been shown to be a reasonable goal to reach [44, 45]. The measured bit error rate is required to be Ber$\sim 10^{-2}$ or better, which is dependent on the signal–to–noise ratio (Snr), where the signal is determined by the remanent magnetisation and the head–disk spacing (Hds). The media noise is produced by transition position fluctuations or jitter which is dependent on the number of grains per bit and the thermal gradient while writing.

Chapter 5

Phonon Analysis

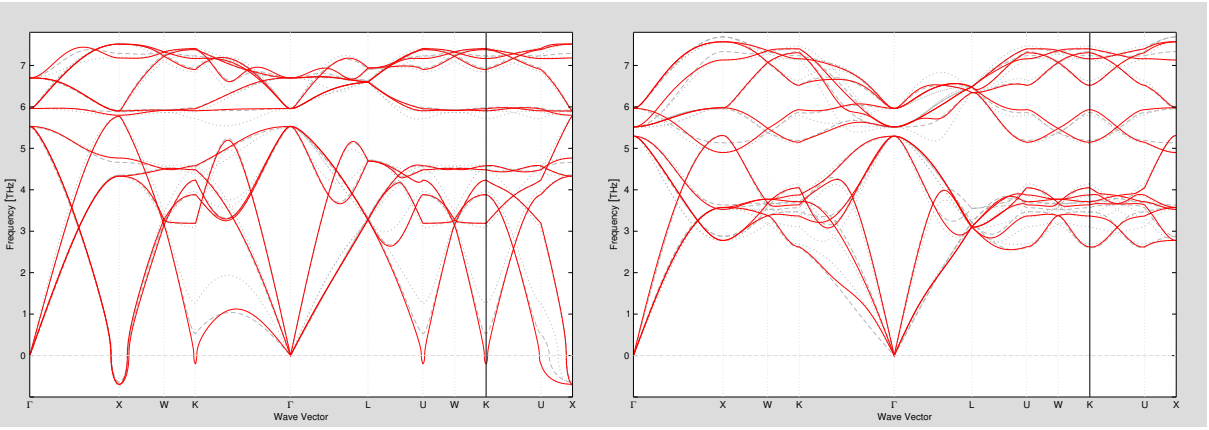
The phononic contribution of the crystal lattice to the magnetic phase transition has been investigated by calculating properties like the phonon band structure, the vibrational density of states (vDos), and the thermodynamic contributions of the phonons to the free energy F, the entropy S and the heat capacity at constant volume C_V in the two cubic magnetic phases, AFM II and FM. For convenience the primitive cell with fcc basis vectors (see Fig.4.2) again has been used. As it was mentioned before, in order to have all the stable phonon branches properly converged it is necessary to enlarge the cell size till the required accuracy of the forces is reached. It will be shown that the latter can be achieved by a supercell of size–factor $4 \times 4 \times 4$ respective to the primitive cell which then contains 256 atoms. Also, the behaviour of two different Gga potentials has been compared and delivered surprising results leading to interesting conclusions.

5.1 Phonon Band Structure

The details of the band structure for both magnetic configurations are shown in this section. The results for calculations using a PBE functional are shown in Fig.5.1. In the FM case (Fig.5.1b) a stable dispersion is found within the entire Brillouin zone throughout different cell sizes. The bands of the $4 \times 4 \times 4$ supercell (solid red lines) are compared with those of smaller cells (grey lines) regarding convergence. It can be shown that a band structure with full details can only be achieved with a considerably large super–cell of 256 atoms.

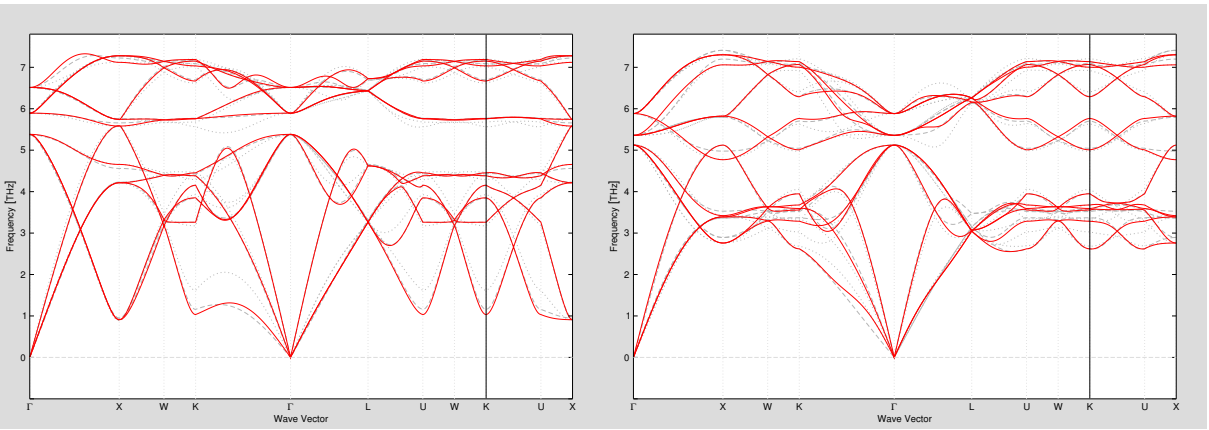
At symmetry point X the lowest band of the AFM II state shows imaginary frequencies (*negative* values in Fig.5.1a) independent of the cell size, indicating a *dynamic instability* of the crystal. With enlargement of the cell further instabilities develop at the reciprocal space points K and U, which means that the instability is not only confined to the Γ–X–W direction of the Brillouin zone, but gradually becomes stronger along X–U and X–K. Hence, a displacement of the ions according to the wave vector at X should lead to a relaxation and a lowering of the total energy. The latter points into the direction of one of the cubic axes and produces a transverse optical phonon with a periodicity of $2a_{\mathrm{AFM}}$.

In order to clarify the impact of the applied functional on the instabilities observed with PBE the calculations have been repeated using a different Gga functional. The phonon band structure employing rPBE exchange (see also Sec.4.2.1 and Tab.4.3) is shown in Fig.5.2. Compared to Fig.5.1 the stable phonon branches are qualitatively similar, however, although the optical modes at X, K, and U are considerably soft the imaginary frequencies in case of the AFM II configuration completely vanish with the use of rPBE, which reveals that the instability cannot be a universal feature. The magnetic moments from Tab.4.3 for the two functionals show a considerable discrepancy, $\pm 3.145\,\mu_{\mathrm{B}}$ (PBE) to $\pm 3.205\,\mu_{\mathrm{B}}$ (rPBE). More comprehensive investigations of the magnetic moments associated with different functionals together with related



(a) (b)

Fig. 5.1: Comparison of the phonon band structure of FeRh in the (a) AFM II, and (b) FM state calculated with a PBE functional. Bands of $2 \times 2 \times 2$ *(grey dotted)* and $3 \times 3 \times 3$ *(grey dashed)* supercells are plotted against $4 \times 4 \times 4$ *(red solid)* regarding convergence. Imaginary frequencies display as negative values.



(a) (b)

Fig. 5.2: Comparison of the phonon band structure of FeRh in the (a) AFM II, and (b) FM state calculated with an RPBE functional. Bands of $2 \times 2 \times 2$ *(grey dotted)* and $3 \times 3 \times 3$ *(grey dashed)* supercells are plotted against $4 \times 4 \times 4$ *(red solid)* regarding convergence. The imaginary frequencies from Fig.5.1a vanish due to the different treatment of interaction within the applied functional.

volume effects indeed reveal that a sole increase of the magnetic moments is sufficient for a stabilisation of the cubic AFM II phase [46, 47].

Taking a closer look at the most prominent instability at the X point (transverse optical phonon branch) reveals a wave vector that is parallel to one of the cubic axes. The direction of the ionic displacements caused by this wave vector are displayed in Fig.5.3. It can be shown

that by gradually displacing the atoms without allowing relaxations the total energy can be lowered by about 0.2 meV per Fu, and that after full cell and ionic relaxations even a new *monoclinic structure* ($P2/m$) with an energy gain of 24.3 meV per Fu can be found [46]. Unfortunately, this structure has not yet had any experimental confirmation. It is assumed that the extremely shallow energy minimum found by the displacement can easily be smeared out by kinetic fluctuations at temperatures that are already larger than 2 K. Furthermore, a small number of anti–site defects is enough to suppress AFM order down to low temperatures [48].

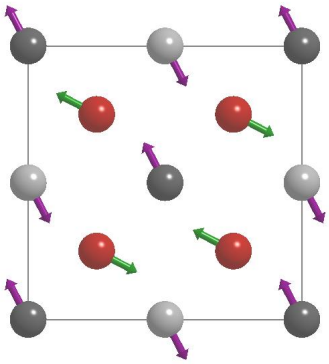


Fig. 5.3: Direction of the ionic displacements caused by the phonon wave vector along one of the cubic axes at the most prominent instability at the X point for calculations using the PBE functional.

5.2 Thermodynamic Aspects

Thermodynamic properties have been calculated within the harmonic approximation for all four configurations AFM,FM(PBE) shown in Fig.5.4a,b, and AFM,FM(RPBE) in Fig.5.4c,d. The internal energy $E(T)$, free energy $F(T)$, entropy $S(T)$, and heat capacity at constant volume $C_V(T)$ have been derived according to the expressions (3.67)–(3.70), respectively. All these thermodynamic features contributed by the phonons can be directly related to experiments and could deliver a not negligible contribution in the explanation of the metamagnetic phase transition. Note that magnetic contributions to the free energies have not been included.

Qualitatively all curves in Fig.5.4 show similar results. A detailed comparison of the energy curves (Fig.5.5) reveals that albeit phonons might contribute to the phase transition the phonon energies alone cannot display any crossing and thus no phase transition at room temperature. Investigating the entropy S of both phases in Fig.5.6 reveals a considerable difference of about $\Delta S = 2.44\,\mathrm{J/(K\,mol)} \rightarrow 15.3\,\mathrm{J/(K\,kg)}$ [$4.21\,\mathrm{J/(K\,mol)} \rightarrow 26.5\,\mathrm{J/(K\,kg)}$] for PBE [RPBE] at $T \sim 350\,\mathrm{K}$ where the transition should take place, which has also been observed experimentally [49]. From $T = 40\,\mathrm{K}$ onwards the FM state exhibits a larger heat capacity than the other phase. Above room temperature the vibrational specific heat for constant V reaches the limit that comes with the law of Dulong–Petit. Below $T = 40\,\mathrm{K}$ the lattice specific heat of the AFM phase exceeds the one from the FM phase. This has also been observed in thin film experiments [50].

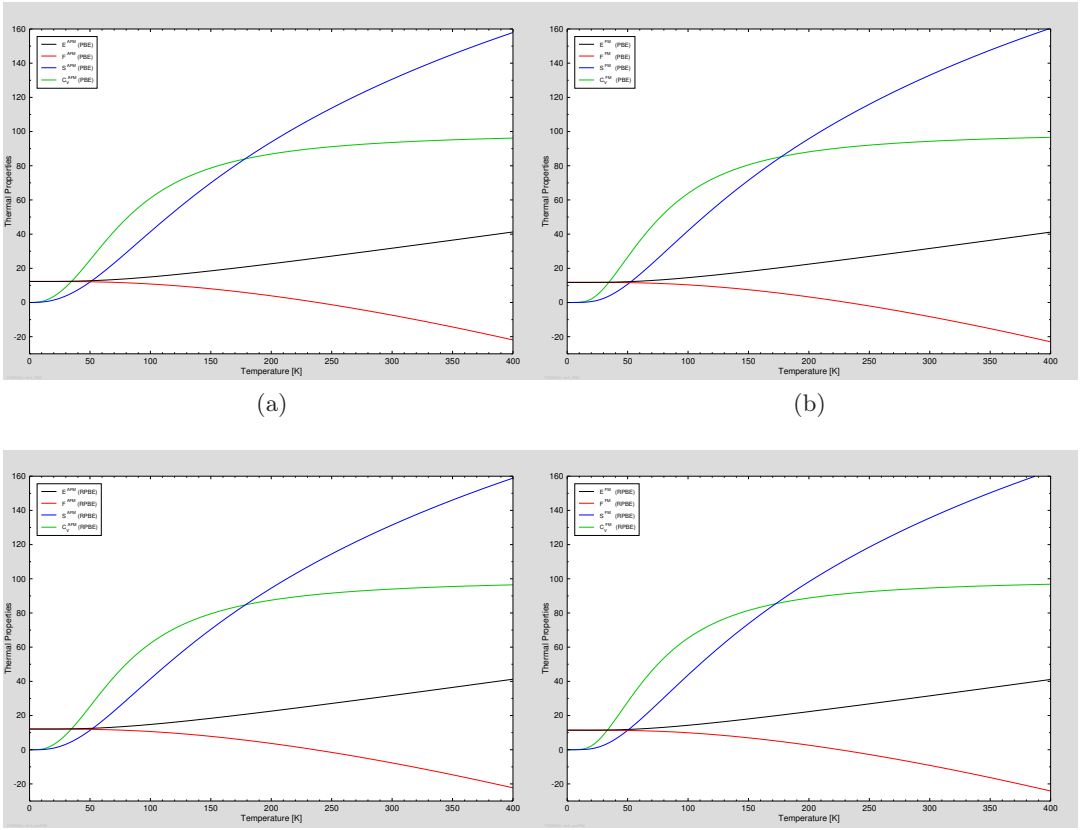
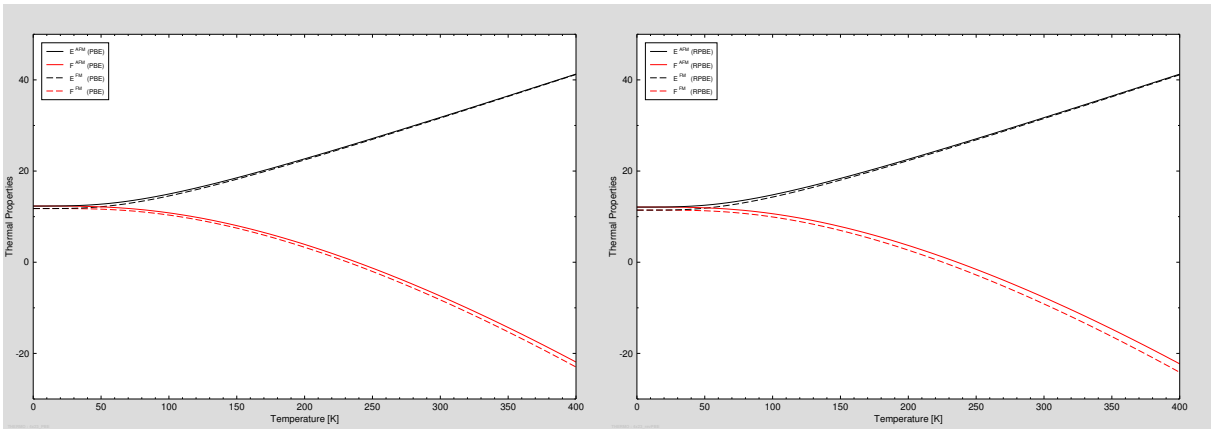


Fig. 5.4: Phonon contributions of the thermodynamic properties in FeRh. All four plots show the internal energy *(black)*, the free energy *(red)*, the entropy *(blue)*, and the heat capacity at constant volume *(green)*. Energies in kJ/mol, entropy and heat capacity in J/(K mol). (a) AFM II (PBE) (b) FM (PBE) (c) AFM II (RPBE) (d) FM (RPBE). All plots show the same qualitative behaviour.

5.3 Conclusions

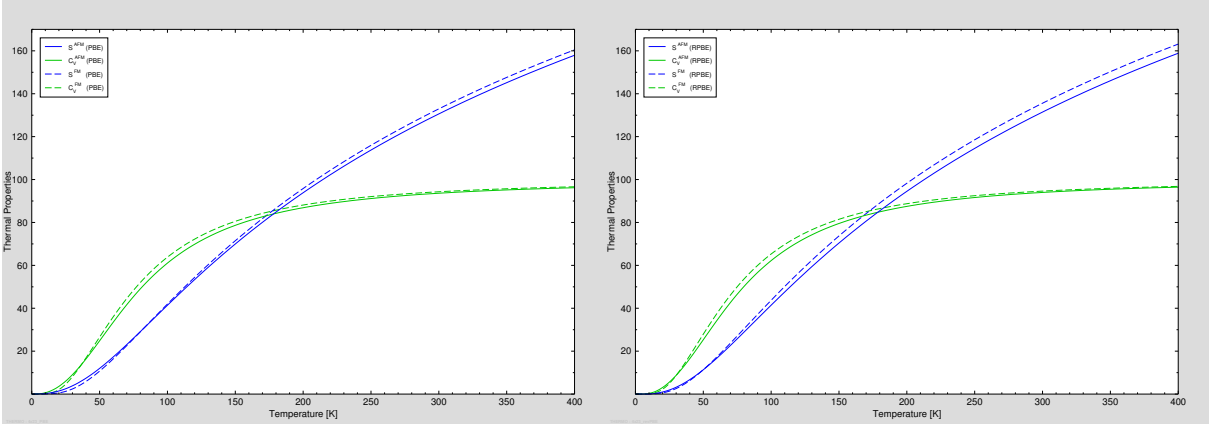
It has been shown that the AFM II phase in cubic FeRh manifests soft phonon branches which are imaginary at some high symmetry points in the reciprocal space, especially around the X point. These instabilities are sensitive to the magnitude of the stabilised total magnetic moment and related volume effects, where larger moments are apparent to stabilise the cubic structure. Further investigations [46] show that by careful relaxation of the atoms in constrained direction of the displacements a monoclinic AFM structure lower in energy can be found, however, the energy gain along the calculated transformation pathway is only of about 0.2 meV per FU, so that already small thermal fluctuations are sufficient to suppress the transition and it thus becomes almost impossible to confirm this structure experimentally. Thermodynamic properties have been calculated and the according energies deliver no direct signs of a phase transition at room temperature. However, it could yield important contributions to more dominant interactions that drive the transition. Furthermore, the large entropy difference that has been observed



(a) (b)

Fig. 5.5: Detailed comparison of the phonon contribution to the internal energy *(black)* and the free energy *(red)*, both in kJ/mol. None of the curves cross at room temperature, which means that the metamagnetic phase transition is not based on phonons.

when reaching the transition temperature can be confirmed. As a consequence, the knowledge and detailed information on the thermodynamic contributions of the lattice could play a decisive role in the many existing magnetic models and could be an important step forward in the final discovery of the origin of the metamagnetic transition in FeRh.



(a) (b)

Fig. 5.6: Detailed comparison of the phonon contribution to the entropy *(blue)* and the heat capacity at constant volume *(green)*, both in J/(K mol). At $T \sim 350\,\mathrm{K}$, where the transition should occur, the entropy of both states shows a considerable difference. Beyond $T = 40\,\mathrm{K}$ the heat capacity of the AFM II phase exceeds the one of the FM phase.

Appendix

A. Derivatives of the Free Energies

As it was presented in Sec.1.1.3 all defined free energies are thermodynamic potentials if they can be described with their natural variables. That is, the *first derivatives* always result in the according conjugated dependent state variables.

$$U = U(S, \mathbf{x}, \mathbf{N}) \quad \longrightarrow \quad T = \left(\frac{\partial U}{\partial S}\right)_{\mathbf{x}} , \quad F_j = \left(\frac{\partial U}{\partial x_j}\right)_{S,\, x_{i\neq j}} , \tag{A.1a}$$

$$F = F(T, \mathbf{x}, \mathbf{N}) = U - T \cdot S \quad \longrightarrow \quad S = -\left(\frac{\partial F}{\partial T}\right)_{\mathbf{x}} , \quad F_j = \left(\frac{\partial F}{\partial x_j}\right)_{T,\, x_{i\neq j}} , \tag{A.1b}$$

$$G = G(T, \mathbf{F}, \mathbf{N}) = H - T \cdot S \quad \longrightarrow \quad S = -\left(\frac{\partial G}{\partial T}\right)_{\mathbf{F}} , \quad x_j = -\left(\frac{\partial G}{\partial F_j}\right)_{T,\, F_{i\neq j}} . \tag{A.1c}$$

All the other natural variables have to be kept *constant*. This includes the independent variables $\mathbf{N}$, which is not explicitly stated at the derivatives for a better overview. In all three cases the derivative of N_k yields the chemical potential μ_k

$$\bullet = \bullet(X, \mathbf{N}) \quad \longrightarrow \quad \mu_k = \left(\frac{\partial \bullet}{\partial N_k}\right)_{X,\, N_l \neq N_k} \tag{A.2}$$

The *second derivatives*, using the results from (A.1), read

$$\left(\frac{\partial^2 U}{\partial S^2}\right)_{\mathbf{x}} = \frac{\partial}{\partial S}\left(\frac{\partial U}{\partial S}\right)_{\mathbf{x}} = \left(\frac{\partial T}{\partial S}\right)_{\mathbf{x}} = \frac{T}{C_{\mathbf{x}}} \geq 0 , \tag{A.3a}$$

$$\left(\frac{\partial^2 U}{\partial {x_j}^2}\right)_{S,\, x_i \neq x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial U}{\partial x_j}\right)_{S,\, x_i \neq x_j} = \left(\frac{\partial F_j}{\partial x_j}\right)_{S,\, x_i \neq x_j} , \tag{A.3b}$$

$$\left(\frac{\partial^2 F}{\partial T^2}\right)_{\mathbf{x}} = \frac{\partial}{\partial T}\left(\frac{\partial F}{\partial T}\right)_{\mathbf{x}} = -\left(\frac{\partial S}{\partial T}\right)_{\mathbf{x}} = -\frac{C_{\mathbf{x}}}{T} \leq 0 , \tag{A.4a}$$

$$\left(\frac{\partial^2 F}{\partial {x_j}^2}\right)_{T,\, x_i \neq x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial F}{\partial x_j}\right)_{T,\, x_i \neq x_j} = \left(\frac{\partial F_j}{\partial x_j}\right)_{T,\, x_i \neq x_j} , \tag{A.4b}$$

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_{\mathbf{F}} = \frac{\partial}{\partial T}\left(\frac{\partial G}{\partial T}\right)_{\mathbf{F}} = -\left(\frac{\partial S}{\partial T}\right)_{\mathbf{F}} = -\frac{C_{\mathbf{F}}}{T} \leq 0 , \tag{A.5a}$$

$$\left(\frac{\partial^2 G}{\partial {F_j}^2}\right)_{T,\, F_i \neq F_j} = \frac{\partial}{\partial F_j}\left(\frac{\partial G}{\partial F_j}\right)_{T,\, F_i \neq F_j} = -\left(\frac{\partial x_j}{\partial F_j}\right)_{T,\, F_i \neq F_j} , \tag{A.5b}$$

where from the definition (1.28)

$$C_{\mathbf{x}} = \left(\frac{\delta Q}{\mathrm{d}T}\right)_{\mathbf{x}} = \left(\frac{\mathrm{d}U}{\mathrm{d}T}\right)_{\mathbf{x}} = \left(\frac{\partial U}{\partial S}\right)_{\mathbf{x}} \left(\frac{\partial S}{\partial T}\right)_{\mathbf{x}} = T \cdot \left(\frac{\partial S}{\partial T}\right)_{\mathbf{x}} ,$$
$$C_{\mathbf{F}} = \left(\frac{\delta Q}{\mathrm{d}T}\right)_{\mathbf{F}} = \left(\frac{\mathrm{d}H}{\mathrm{d}T}\right)_{\mathbf{F}} = \left(\frac{\partial H}{\partial S}\right)_{\mathbf{F}} \left(\frac{\partial S}{\partial T}\right)_{\mathbf{F}} = T \cdot \left(\frac{\partial S}{\partial T}\right)_{\mathbf{F}} ,$$

has been used.

B. Derivation of the Mean–Field in the Ising Model

The explicit expression for the mean–field H_0 can be derived with statistical mechanics by minimising the parameter Φ of the Bogoliubov inequality (1.58),

$$\Phi = F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle \ ,$$

where the average is with respect to the ensemble defined by $\mathcal{H}_0$.

Using a mean–field ansatz as a *trial* Hamiltonian,

$$\mathcal{H}_0 = -\left(H_0 + g_{\mathrm{J}}\,\mu_{\mathrm{B}}\,H\right)\sum_i S_i \ , \tag{B.1}$$

together with the Ising–Hamiltonian (1.61) one derives

$$\langle \mathcal{H} - \mathcal{H}_0 \rangle = -J\sum_{\langle i,j \rangle} \langle S_i \rangle \langle S_j \rangle + H_0 \sum_i \langle S_i \rangle \ , \tag{B.2}$$

which for a system with N spins and $\aleph$ *nearest–neighbours* that are *invariant* to space translations can be summarised to

$$\begin{aligned} \langle S_i \rangle = \langle S_j \rangle = \langle S \rangle \ , \\ \Longrightarrow \qquad \langle \mathcal{H} - \mathcal{H}_0 \rangle = N \langle S \rangle \left(-\frac{J\,\aleph\,\langle S \rangle}{2} + H_0 \right) \ . \end{aligned} \tag{B.3}$$

The *free energy* F_0 for the mean–field can be derived calculating the *partition sum* $\mathcal{Z}$ of the ensemble

$$\begin{aligned} \mathcal{Z} = \mathcal{Z}(T,H,N) = \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \exp\Big(-\beta\,\mathcal{H}_0 \Big) = \\ = \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \exp\Big(\beta\left(H_0 + g_{\mathrm{J}}\,\mu_{\mathrm{B}}\,H\right) \sum_i S_i \Big) = \\ = \Big[2\,\cosh\Big(\beta\left(H_0 + g_{\mathrm{J}}\,\mu_{\mathrm{B}}\,H\right) \Big) \Big]^N \ , \end{aligned} \tag{B.4}$$

where $\beta = (k_{\mathrm{B}}\,T)^{-1}$ and the expression for the hyperbolic cosine

$$\cosh(x) \doteq \frac{1}{2} \sum_{k=\pm 1} \exp(k\,x) \ ,$$

has been used. Thus the free energy reads

$$\begin{aligned} F_0 = -k_{\mathrm{B}}\,T \ln \mathcal{Z} = \\ = -N\,k_{\mathrm{B}}\,T\,\ln\Big[2\,\cosh\Big(\beta\left(H_0 + g_{\mathrm{J}}\,\mu_{\mathrm{B}}\,H\right) \Big) \Big] \ . \end{aligned} \tag{B.5}$$

The first derivative of the free energy with respect to the (magnetic) mean–field H_0 yields an average magnetisation M_0 established by the average value of the spins at each site

$$\begin{aligned} M_0 = N\,\langle S \rangle = -\left(\frac{\partial F_0}{\partial H_0} \right) = N\,k_{\mathrm{B}}\,T\,\frac{2\,\beta\,\sinh\left(\beta\,H_0 \ldots\right)}{2\,\cosh\left(\beta\,H_0 \ldots\right)} = N\,\tanh\Big(\beta\,H_0 \ldots \Big) \\ \Longrightarrow \qquad \langle S \rangle = \tanh\Big(\beta\left(H_0 + g_{\mathrm{J}}\,\mu_{\mathrm{B}}\,H\right) \Big) \ . \end{aligned} \tag{B.6}$$

The latter is entered into (B.3) and yields together with (B.5) a *self–consistent* expression for the mean–field H_0

$$F_{\mathrm{MF}} = \min_{H_0} \Phi$$

$$\frac{\partial}{\partial H_0}\Big[- N\, k_{\mathrm{B}}\, T\, \ln\Big(2\, \cosh\big(\beta\, H_0 \ldots\big)\Big) - N \frac{J\, \aleph}{2}\, \tanh^2\big(\beta\, H_0 \ldots\big) + N\, H_0\, \tanh\big(\beta\, H_0 \ldots\big)\Big] = 0$$

$$\Longrightarrow \qquad H_0 = J\, \aleph\, \tanh\big(\beta\, H_0 \ldots\big)\,. \tag{B.7}$$

Equivalently, comparison with (B.6) shows that $H_0 = J\, \aleph\, \langle S \rangle$ and the self–consistent equation for the mean–field magnetisation reads

$$\langle S \rangle = \tanh\Big(\beta\, J\, \aleph\, \langle S \rangle \ldots\Big)\,, \tag{B.8}$$

and hence the mean–field free energy is

$$F_{\mathrm{MF}} = -N\, k_{\mathrm{B}}\, T\, \ln\Big[2\, \cosh\Big(\beta\, H_0 \ldots\Big)\Big] + \frac{M_0\, H_0}{2}\,. \tag{B.9}$$

C. Susceptibility and Correlation Functions

Consider an arbitrary system with a Hamiltonian $\mathcal{H}_0$. This Hamiltonian is modified as soon as an *external inhomogeneous field* $\vec{\boldsymbol{B}}(\vec{\boldsymbol{r}})$ is present,

$$\mathcal{H} = \mathcal{H}_0 - \int \mathrm{d}\vec{\boldsymbol{r}}\, \phi(\vec{\boldsymbol{r}})\, \vec{\boldsymbol{B}}(\vec{\boldsymbol{r}})\,, \tag{C.1}$$

where $\phi(\vec{\boldsymbol{r}})$ is the system variable that couples linearly to the field (spins or magnetisation). The partition function reads

$$\mathcal{Z} = \mathrm{tr}_\phi \left[\exp\left(-\beta\, \mathcal{H}_0 + \beta \int \mathrm{d}\vec{\boldsymbol{r}}\, \phi(\vec{\boldsymbol{r}})\, \vec{\boldsymbol{B}}(\vec{\boldsymbol{r}}) \right) \right], \tag{C.2}$$

and the free energy is derived by $F = -\frac{1}{\beta} \ln \mathcal{Z}$, so that one can define a *generalised isothermal susceptibility* as the second order functional derivative of the free energy

$$\chi(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}\prime) = -\frac{\delta^2 F}{\delta B(\vec{\boldsymbol{r}})\, \delta B(\vec{\boldsymbol{r}}\prime)}\,. \tag{C.3}$$

The two–point *correlation function* is defined as

$$\begin{aligned} \Gamma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}\prime) &\doteq \Big\langle \big(\phi(\vec{\boldsymbol{r}}) - \langle\phi(\vec{\boldsymbol{r}})\rangle\big)\, \big(\phi(\vec{\boldsymbol{r}}\prime) - \langle\phi(\vec{\boldsymbol{r}}\prime)\rangle\big) \Big\rangle \\ &\equiv \big\langle \phi(\vec{\boldsymbol{r}})\, \phi(\vec{\boldsymbol{r}}\prime) \big\rangle - \big\langle \phi(\vec{\boldsymbol{r}}) \big\rangle \big\langle \phi(\vec{\boldsymbol{r}}\prime) \big\rangle\,, \end{aligned} \tag{C.4}$$

so that (C.3) can be evaluated to

$$\begin{aligned} \chi(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}\prime) &= \frac{1}{\beta} \frac{\delta^2 \ln \mathcal{Z}}{\delta B(\vec{\boldsymbol{r}})\, \delta B(\vec{\boldsymbol{r}}\prime)} = \\ &= \frac{1}{\beta} \left(\frac{1}{\mathcal{Z}} \frac{\delta^2 \mathcal{Z}}{\delta B(\vec{\boldsymbol{r}})\, \delta B(\vec{\boldsymbol{r}}\prime)} - \frac{1}{\mathcal{Z}} \frac{\delta \mathcal{Z}}{\delta B(\vec{\boldsymbol{r}})} \cdot \frac{1}{\mathcal{Z}} \frac{\delta \mathcal{Z}}{\delta B(\vec{\boldsymbol{r}}\prime)} \right) = \\ &= \beta \left(\big\langle \phi(\vec{\boldsymbol{r}})\, \phi(\vec{\boldsymbol{r}}\prime) \big\rangle - \big\langle \phi(\vec{\boldsymbol{r}}) \big\rangle \big\langle \phi(\vec{\boldsymbol{r}}\prime) \big\rangle \right) \end{aligned}$$

$$\implies \quad \chi(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}\prime) = \beta\, \Gamma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}\prime)\,. \tag{C.5}$$

A system with translational invariance can use relative paths $\Gamma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}\prime) \to \Gamma(|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}\prime|) = \Gamma(r)$ and the total susceptibility reads

$$\chi = \int \mathrm{d}\vec{\boldsymbol{r}}\, \chi(r) \equiv \beta \int \mathrm{d}\vec{\boldsymbol{r}}\, \Gamma(r)\,. \tag{C.6}$$

The latter equation sets the response χ in relation to an external perturbation $\vec{\boldsymbol{B}}$ of the system with the fluctuations in equilibrium. It could therefore be interpreted as a special case of the *fluctuation–dissipation theorem*.

D. Hartree–Fock Theory

In order to solve the (interaction–free) many–body Hamiltonian,

$$\hat{\mathcal{H}}^0 = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\vec{\boldsymbol{\nabla}}_i^{\,2} + \sum_{i=1}^{N} u(\vec{\boldsymbol{r}}_i) = \hat{T} + \hat{U}\,, \tag{D.1}$$

Douglas Hartree suggested a product ansatz for the general wave function which is built up by the eigenstates of the single particle Schrödinger equations (*orbitals*),

$$\hat{h}^0\,\phi_i = \left(-\frac{\hbar^2}{2m}\vec{\boldsymbol{\nabla}}_i^{\,2} + u(\vec{\boldsymbol{r}}_i)\right)\phi_i = \varepsilon^0{}_i\,\phi_i \quad \Longleftrightarrow \quad \Psi_{\mathrm{H}}(\vec{\boldsymbol{r}}_1,\ldots,\vec{\boldsymbol{r}}_N) = \prod_{i=1}^{N}\phi_i(\vec{\boldsymbol{r}}_i)\,. \tag{D.2}$$

However, the latter is not suitable for fermions (but indeed for bosons) because wave function Ψ_{H} is not strictly antisymmetric with respect to the exchange of two (indistinguishable) fermions with one another. For a correct description of the behaviour of electrons in atoms or molecules the Pauli exclusion principle must be fulfilled. Vladimir Fock improved the ansatz of Hartree by applying an anti–symmetry operator $\mathcal{A}$ to the orbitals in (D.2),

$$\begin{aligned}\Psi_{\mathrm{HF}}(\vec{\boldsymbol{r}}_1,\ldots,\vec{\boldsymbol{r}}_N) &= \sqrt{N!}\,\mathcal{A}\,\Psi_{\mathrm{H}}(\vec{\boldsymbol{r}}_1,\ldots,\vec{\boldsymbol{r}}_N) = \\ &= \frac{1}{\sqrt{N!}}\begin{vmatrix}\phi_1(\vec{\boldsymbol{r}}_1) & \ldots & \phi_1(\vec{\boldsymbol{r}}_N)\\ \vdots & \ddots & \vdots\\ \phi_N(\vec{\boldsymbol{r}}_1) & \ldots & \phi_N(\vec{\boldsymbol{r}}_N)\end{vmatrix}\quad .\end{aligned} \tag{D.3}$$

The latter is called *Slater determinant*, which is an elegant way of expressing the desired antisymmetric state for fermions. Taking the *exchange and correlation* of the (indistinguishable) electrons into account, we derive a Hamiltonian similar to (2.7). The variation of the energy reads

$$\delta\left\{\langle\Psi_{\mathrm{HF}}|\,\hat{\mathcal{H}}\,|\Psi_{\mathrm{HF}}\rangle - \sum_i \varepsilon_i\Big(\langle\phi_i|\phi_i\rangle - 1\Big)\right\} = 0 \quad . \tag{D.4}$$

Note that the orbitals taken from (D.2) account for the presence of other electrons only in an average manner — the Hartree–Fock method thus is a *mean–field theory*. They appear as constraints in (D.4) to grant that the single states are all orthonormal. It is sufficient to project the expectation value in (D.4) onto real space,

$$\begin{aligned}\langle\Psi_{\mathrm{HF}}|\,\hat{\mathcal{H}}\,|\Psi_{\mathrm{HF}}\rangle \;=\; &\sum_i \int \mathrm{d}^3r'\;\phi_i^*(\vec{\boldsymbol{r}}\,')\left[-\frac{\hbar^2}{2m}\vec{\boldsymbol{\nabla}}_i^{\,2} + u(\vec{\boldsymbol{r}}\,')\right]\phi_i(\vec{\boldsymbol{r}}\,') + \\ &+ \underbrace{\frac{1}{2}\sum_{i\neq j}\iint \mathrm{d}^3r'\,\mathrm{d}^3r''\;\phi_i^*(\vec{\boldsymbol{r}}\,')\phi_j^*(\vec{\boldsymbol{r}}\,'')\frac{1}{|\vec{\boldsymbol{r}}\,''-\vec{\boldsymbol{r}}\,'|}\phi_i(\vec{\boldsymbol{r}}\,')\phi_j(\vec{\boldsymbol{r}}\,'')}_{=V_{\mathrm{H}}} - \\ &- \underbrace{\frac{1}{2}\sum_{i\neq j}\iint \mathrm{d}^3r'\,\mathrm{d}^3r''\;\phi_i^*(\vec{\boldsymbol{r}}\,')\phi_j^*(\vec{\boldsymbol{r}}\,'')\frac{1}{|\vec{\boldsymbol{r}}\,''-\vec{\boldsymbol{r}}\,'|}\phi_i(\vec{\boldsymbol{r}}\,'')\phi_j(\vec{\boldsymbol{r}}\,')}_{=V_{\mathrm{F}}}\quad ,\end{aligned} \tag{D.5}$$

and only vary over the imaginary part ϕ_k^*,

$$\frac{\delta}{\delta\phi_k^*(\vec{\boldsymbol{r}})}\left\{ \langle\Psi_{\mathrm{HF}}|\,\hat{\mathcal{H}}|\Psi_{\mathrm{HF}}\rangle \text{ (D.5)} - \sum_i \varepsilon_i \left(\int \mathrm{d}^3r'\,|\phi_i(\vec{\boldsymbol{r}}\,')|^2 - 1 \right) \right\} = 0 \quad . \tag{D.6}$$

As a final result one derives the N *Hartree–Fock–equations*,

$$\begin{aligned} \underbrace{\left(-\frac{\hbar^2}{2m}\vec{\boldsymbol{\nabla}}_k^2 + u(\vec{\boldsymbol{r}}) \right)}_{\hat{h}^0} \phi_k(\vec{\boldsymbol{r}}) + \underbrace{\sum_i \int \mathrm{d}^3r'\,|\phi_i(\vec{\boldsymbol{r}}\,')|^2 \frac{1}{|\vec{\boldsymbol{r}}-\vec{\boldsymbol{r}}\,'|}\phi_k(\vec{\boldsymbol{r}})}_{\text{Coulomb–Term (Hartree)}} - \\ \underbrace{- \sum_i \int \mathrm{d}^3r'\,\phi_i^*(\vec{\boldsymbol{r}}\,') \frac{1}{|\vec{\boldsymbol{r}}-\vec{\boldsymbol{r}}\,'|}\phi_k(\vec{\boldsymbol{r}}\,')\phi_i(\vec{\boldsymbol{r}})}_{\text{Exchange–Term (Fock)}} = \varepsilon_k\phi_k(\vec{\boldsymbol{r}}) \quad . \end{aligned} \tag{D.7}$$

The latter expression is the *exchange term*, which is only derived by the antisymmetric nature of the orbitals. Obviously, the exchange of the electrons results in a lower total energy. The other terms could also be calculated by solely using the Hartree–ansatz (D.2). In order to retrieve a general expression of (D.7), the *Fock operator* $\hat{F}$ is defined as

$$\begin{aligned} \hat{F} = \hat{h}^0 + \sum_i \Big(\langle\phi_i|\hat{w}|\phi_i\rangle \;-\; |\phi_i\rangle\,\langle\phi_i|\,\hat{w} \Big) = \hat{h}^0 + \hat{\jmath} - \hat{k}\,, \\ \text{with } \hat{w} = \frac{1}{|\vec{\boldsymbol{r}}-\vec{\boldsymbol{r}}\,'|} \\ \Longrightarrow \qquad \boxed{\hat{F}\,|\phi_k\rangle = \varepsilon_k\,|\phi_k\rangle} \quad . \end{aligned} \tag{D.8}$$

The Fock operator for an electron k is set up by the wave functions of all the other electrons. Hence, the Fock equations can only be solved iteratively in terms of a *self–consistent field method*. A similar example of this method has been shown at the beginning of Sec.2.3. In case of the free electron gas the Fermi vector can be expressed by the density of states. Slater used the exchange term of the Hartree–Fock approximation to calculate the exchange potential for metals (2.43).

E. Heisenberg Representation

Starting from the Schrödinger equation, $i\hbar\partial_t |\psi\rangle = \mathcal{H} |\psi\rangle$, with a *time–independent* Hamiltonian one can define a unitary *time–evolution operator*

$$|\psi(t)\rangle = \mathcal{T}(t) |\psi(0)\rangle \ , \quad \text{with } \mathcal{T}(t) \doteq \mathrm{e}^{-i\frac{\mathcal{H}}{\hbar}t} \ . \tag{E.1}$$

An arbitrary operator X must yield the same expectation value x in both representations, so that also all state vectors can be constrained to rigid basis of time–independent wave functions and a Heisenberg operator is transformed by

$$x = \langle\psi(t)|X|\psi(t)\rangle = \langle\psi(0)|\mathcal{T}^\dagger(t)\, X\, \mathcal{T}(t)|\psi(0)\rangle \implies X(t) = \mathcal{T}^\dagger(t)\, X\, \mathcal{T}(t) \ . \tag{E.2}$$

As the state $|\psi(t)\rangle$ is not stationary it is possible that also the expectation value is evolving with time. Sometimes also an operator *itself* can explicitly have a time dependency — even in the Schrödinger picture, for example a time–varying potential — which could have an effect on the average value, but will be neglected here. In order to give the rate of change $\frac{\mathrm{d}}{\mathrm{d}t}x = \dot{x}$ one defines an operator $\dot{X}$ with

$$\dot{X} |\psi\rangle = \dot{x} |\psi\rangle \ .$$

Note that $\dot{X}$ is not a time derivative in a classical sense, but indeed the definition of a *new* operator. One finds with (E.2)

$$\langle\psi(0)|\dot{X}(t)|\psi(0)\rangle = \dot{x} = \frac{\mathrm{d}}{\mathrm{d}t}\, x = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\langle\psi(0)|X(t)|\psi(0)\rangle \Big) \ , \tag{E.3}$$

which yields the *Heisenberg equation*

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\Big(\langle\psi(0)|X(t)|\psi(0)\rangle \Big) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\Big(\langle\psi(t)|X|\psi(t)\rangle \Big) = \\
&= \Big(\partial_t \langle\psi(t)|\Big) X |\psi(t)\rangle + \langle\psi(t)| X \Big(\partial_t |\psi(t)\rangle\Big) + \langle\psi(t)| \overbrace{(\partial_t X)}^{\text{(neglected)}} |\psi(t)\rangle = \\
&= \frac{i}{\hbar} \langle\psi(t)|\mathcal{H}\, X|\psi(t)\rangle - \frac{i}{\hbar} \langle\psi(t)|X\, \mathcal{H}|\psi(t)\rangle = \frac{i}{\hbar} \langle\psi(t)|\, [\mathcal{H}, X]\, |\psi(t)\rangle \equiv \\
&\equiv \frac{i}{\hbar} \langle\psi(0)|\mathcal{T}^\dagger\, \mathcal{H}\, X\, \mathcal{T}|\psi(0)\rangle - \frac{i}{\hbar} \langle\psi(0)|\mathcal{T}^\dagger\, X\, \mathcal{H}\, \mathcal{T}|\psi(0)\rangle = \frac{i}{\hbar} \langle\psi(0)|\, [\mathcal{H}, X(t)]\, |\psi(0)\rangle = \\
&\implies \qquad \boxed{\dot{X}(t) = \frac{i}{\hbar}\, [\mathcal{H}, X(t)]} \quad .
\end{aligned} \tag{E.4}$$

This equation of motion in the Heisenberg representation is the equivalence of the Schrödinger equation. The expression for the partial time–derivative of the state vector derives either directly from the Schrödinger equation, or equivalently from the derivative of the exponential function of the time–evolution operator. It was assumed that X is not implicitly dependent on time ($\partial_t X \to 0$), and that the time–evolution operator commutes with the Hamiltonian ($[\mathcal{H}, \mathcal{T}] = 0$) which is proved by the series expansion of the exponential function[1].

[1] Rigorously, the Hamiltonian has to be a bound operator for this proof to be true. It takes a more elaborate effort to show that the commutator relation also holds for an unbound Hamiltonian (which is usually the case).

F. Partition Sum of a Phonon System

Consider a system of $\tilde{N}$ phonons. The probability $\mathcal{P}(n_{\mathbf{k}})$ of finding $n_{\mathbf{k}}$ phonons in a state of frequency $\omega(\vec{\boldsymbol{k}}s)$ is

$$\mathcal{P}(n_{\mathbf{k}}) = \frac{1}{Z_{\mathbf{k}}}\, \mathrm{e}^{-\beta\,\varepsilon(n_{\mathbf{k}})} = \frac{1}{Z_{\mathbf{k}}}\, \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}}\left(n_{\mathbf{k}}+\frac{1}{2}\right)}\,, \tag{F.1}$$

where $\beta = (k_{\mathrm{B}}\,T)^{-1}$ is the inverse temperature and the normalisation factor $Z_{\mathbf{k}}$ is the *partition function* of the single modes $\mathbf{k}$ with energy $\varepsilon(n_{\mathbf{k}})$ from (3.35) in the case of $D = 3$. For better readability the indices $\bullet(\vec{\boldsymbol{k}}s) \to \bullet_{\mathbf{k}}$ have been unified to only one subscript. The chemical potential is zero ($\mu = 0$) as phonons can be created and annihilated in an unlimited way, so that the free energy in equilibrium will always be minimal. The probability of the whole system being in a state of total energy $E(\tilde{N})$ is just the product of the single–modes probabilities,

$$\mathcal{P}(\tilde{N}) = \sum_{\{n_{\mathbf{k}}\}_{\mathrm{B}}}^{\tilde{N}} \prod_{\mathbf{k}} \mathcal{P}(n_{\mathbf{k}}) = \frac{1}{\mathcal{Z}} \sum_{\{n_{\mathbf{k}}\}_{\mathrm{B}}}^{\tilde{N}} \mathrm{e}^{-\beta\,\sum_{\mathbf{k}}\hbar\omega_{\mathbf{k}}\left(n_{\mathbf{k}}+\frac{1}{2}\right)} = \frac{1}{\mathcal{Z}} \sum_{\{n_{\mathbf{k}}\}_{\mathrm{B}}}^{\tilde{N}} \mathrm{e}^{-\beta\,E(\tilde{N})}\,, \tag{F.2}$$

where the sum is to be taken over all combinations of occupation numbers $\{n_{\mathbf{k}}\}_{\mathrm{B}}$ taking any possible (natural) value within boson statistics (B) and under the restriction that

$$\tilde{N} = \sum_{\mathbf{k}} n_{\mathbf{k}}\,. \tag{F.3}$$

The *total partition function* $\mathcal{Z}$ of the system is defined as

$$\mathcal{Z} = \prod_{\mathbf{k}} Z_{\mathbf{k}}\,, \tag{F.4}$$

which again gives rise that the system is similar to $(3N_u\,N)$ *independent harmonic oscillators*. Under the condition that

$$\sum_{\tilde{N}=0}^{\infty} \mathcal{P}(\tilde{N}) = 1 \quad \Longleftrightarrow \quad \mathcal{Z} = \sum_{\tilde{N}=0}^{\infty} \sum_{\{n_{\mathbf{k}}\}_{\mathrm{B}}}^{\tilde{N}} \prod_{\mathbf{k}} \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}}\left(n_{\mathbf{k}}+\frac{1}{2}\right)}\,,$$

follows an expression for the partition function. Now, since the first sum is over all total number values, the restriction of the combination sum can be replaced by

$$\sum_{\tilde{N}=0}^{\infty} \sum_{\{n_{\mathbf{k}}\}_{\mathrm{B}}}^{\tilde{N}} \ldots \quad \Longleftrightarrow \quad \sum_{n_{\mathbf{k}_1}} \cdots \sum_{n_{\mathbf{k}_N}} \ldots\,,$$

because the combination of *independent* summations over every individual occupation number $n_{\mathbf{k}}$ on the right side contains exactly the same terms as on the left (note the difference of N allowed $\mathbf{k}$ values, and the total number $\tilde{N}$ of phonons in the system). Hence, the *partition function of a*

phonon system can equivalently be written as

$$\begin{aligned}\mathcal{Z} &= \sum_{\tilde{N}=0}^{\infty} \sum_{\{n_{\mathbf{k}}\}_{\mathrm{B}}}^{\tilde{N}} \prod_{\mathbf{k}} \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}}\left(n_{\mathbf{k}}+\frac{1}{2}\right)} \equiv \\ &\equiv \sum_{n_{\mathbf{k}_1}} \cdots \sum_{n_{\mathbf{k}_N}} \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}_1}\left(n_{\mathbf{k}_1}+\frac{1}{2}\right)} \cdot \ldots \cdot \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}_N}\left(n_{\mathbf{k}_N}+\frac{1}{2}\right)} = \\ &= \left(\mathrm{e}^{-\frac{1}{2}\,\beta\hbar\omega_{\mathbf{k}_1}} \sum_{n_{\mathbf{k}_1}} \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}_1}\,n_{\mathbf{k}_1}}\right) \cdot \ldots \cdot \left(\mathrm{e}^{-\frac{1}{2}\,\beta\hbar\omega_{\mathbf{k}_N}} \sum_{n_{\mathbf{k}_N}} \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}_N}\,n_{\mathbf{k}_N}}\right) = \\ &= \prod_{\mathbf{k}} \left(\sum_{n_{\mathbf{k}}} \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}}\left(n_{\mathbf{k}}+\frac{1}{2}\right)}\right) = \prod_{\mathbf{k}} Z_{\mathbf{k}}\,,\end{aligned} \tag{F.5}$$

whereby one identifies the single–modes partition functions to be

$$Z_{\mathbf{k}} = \sum_{n_{\mathbf{k}}} \mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}}\left(n_{\mathbf{k}}+\frac{1}{2}\right)}\,. \tag{F.6}$$

The sum in (F.5) and (F.6) can be interpreted as an infinite geometric series,

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}\,, \quad |q| < 1\,.$$

The partition function thus finally reads

$$\mathcal{Z} = \prod_{\mathbf{k}} \frac{\mathrm{e}^{-\frac{1}{2}\beta\,\hbar\omega_{\mathbf{k}}}}{1-\mathrm{e}^{-\beta\,\hbar\omega_{\mathbf{k}}}} = \prod_{\mathbf{k}} \left[2\,\sinh\left(\tfrac{1}{2}\beta\,\hbar\omega_{\mathbf{k}}\right)\right]^{-1}. \tag{F.7}$$

Average Phonon Number

From the standard expression

$$\langle n_{\mathbf{k}}\rangle = \sum_{n_{\mathbf{k}}} n_{\mathbf{k}}\,\mathcal{P}(n_{\mathbf{k}})\,, \tag{F.8}$$

entering (F.1) with (F.6) as the according probability one can find the average occupation number of a phonon state $\mathbf{k}$. Setting $x_{\mathbf{k}} = \beta\,\hbar\omega_{\mathbf{k}}$ it can be shown that,

$$\begin{aligned}\langle n_{\mathbf{k}}\rangle &= \frac{1}{\sum\limits_{n_{\mathbf{k}}} \mathrm{e}^{-x_{\mathbf{k}}\,n_{\mathbf{k}}}} \sum_{n_{\mathbf{k}}} n_{\mathbf{k}}\;\mathrm{e}^{-x_{\mathbf{k}}\,n_{\mathbf{k}}} = -\frac{1}{\sum\limits_{n_{\mathbf{k}}} \mathrm{e}^{-x_{\mathbf{k}}\,n_{\mathbf{k}}}} \sum_{n_{\mathbf{k}}} \frac{\partial}{\partial x_{\mathbf{k}}}\,\mathrm{e}^{-x_{\mathbf{k}}\,n_{\mathbf{k}}} = \\ &= -\frac{1}{\sum\limits_{n_{\mathbf{k}}} \mathrm{e}^{-x_{\mathbf{k}}\,n_{\mathbf{k}}}}\,\frac{\partial}{\partial x_{\mathbf{k}}}\left(\sum_{n_{\mathbf{k}}} \mathrm{e}^{-x_{\mathbf{k}}\,n_{\mathbf{k}}}\right) = -\frac{\partial}{\partial x_{\mathbf{k}}} \ln\left(\sum_{n_{\mathbf{k}}} \mathrm{e}^{-x_{\mathbf{k}}\,n_{\mathbf{k}}}\right) = \\ &= -\frac{\partial}{\partial x_{\mathbf{k}}} \ln\left(\frac{1}{1-\mathrm{e}^{-x_{\mathbf{k}}}}\right) = +\frac{\partial}{\partial x_{\mathbf{k}}} \ln\left(1-\mathrm{e}^{-x_{\mathbf{k}}}\right) = \frac{\mathrm{e}^{-x_{\mathbf{k}}}}{1-\mathrm{e}^{-x_{\mathbf{k}}}}\end{aligned}$$

$$\Longrightarrow \qquad \langle n_{\mathbf{k}}\rangle = \frac{1}{\mathrm{e}^{\beta\,\hbar\omega_{\mathbf{k}}}-1}\quad, \tag{F.9}$$

which yields the expected Bose–Einstein distribution. Note how any zero–point energy cancels out and does not contribute. It follows equivalently with (F.2), (F.3), and (F.7) that the average total number of phonons is

$$\langle \tilde{N} \rangle = \sum_{\tilde{N}} \tilde{N}\, \mathcal{P}(\tilde{N}) = \dots$$

$$\Longrightarrow \qquad \langle \tilde{N} \rangle = \sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle \tag{F.10}$$

Bibliography

[1] P. C. Hohenberg and W. Kohn, Phys. Rev. **136** 864 (1964)

[2] W. Kohn and L. J. Sham, Phys. Rev. A **140** 1133 (1965)

[3] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13** 4274 (1976)

[4] L. Hedin and B. I. Lundqvist, Journal of Physics C **4.14** 2064–2083 (1971)

[5] J. P. Perdew, Phys. Rev. Lett. **55** 1665 (1985)

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

[7] J. P. Perdew and Y. Wang, Phys. Rev. B **45** 13244 (1992)

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

[9] J. P. Perdew, K. Burke and Y. Wang, Phys. Rev. B **54** 16533 (1996)

[10] M. Ernzerhof and J. P. Perdew, J. Chem. Phys. **109** 3313 (1998)

[11] J. Tao, J. P. Perdew *et al.*, Phys. Rev. Lett. **91** 146401 (2003)

[12] J. Sun, M. Marsman *et al.*, Phys. Rev. B **84** 035117 (2011)

[13] J. P. Perdew, M. Ernzerhof and K. Burke, J. Chem. Phys. **105** 9982 (1996)

[14] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **110** 5029 (1999)

[15] C. Adamo and V. Barone, J. Chem. Phys. **110** 6158 (1999)

[16] J. Heyd, G. E. Scuseria and M. Ernzerhof, J. Chem. Phys. **118** 8207 (2003)

[17] J. Heyd and G. E. Scuseria, J. Chem. Phys. **121** 1187 (2004)

[18] G. Kresse, M. Marsman and J. Furthmüller, 2017, URL: http://www.vasp.at

[19] G. Kresse, M. Marsman and J. Furthmüller, VASP *the* GUIDE, Vienna, 2016

[20] G. Kresse and J. Furthmüller, Phys. Rev. B **54** 11169 (1996)

[21] G. Kresse and J. Furthmüller, Comp. Mat. Sci. **6** 15–50 (1996)

[22] J. Harris and R. O. Jones, Phys. Rev. Lett. **41** 191 (1978)

[23] D. Vanderbilt, Phys. Rev. B **41** 7892 (1990)

[24] G. Kresse and J. Hafner, Journal of Physics: Condensed Matter **6** 8245 (1994)

[25] P. E. Blöchl, Phys. Rev. B **50** 17953 (1994)

[26] G. Kresse and D. Joubert, Phys. Rev. B **59** 1758 (1999)

[27] C. Kittel, *Introduction to Solid State Physics*, 8th ed., John Wiley & Sons, Inc, 2005, 674 pp.

[28] F. Schwabl, *Quantenmechanik für Fortgeschrittene*, vol. 2: *(QM II)*, 5th ed., 2, Springer–Lehrbuch, Springer–Verlag Berlin Heidelberg, 2008, 412 pp.

[29] A. Togo and I. Tanaka, Scr. Mater. **108** 1–5 (2015)

[30] K. Parlinski, Z. Q. Li and Y. Kawazoe, Phys. Rev. Lett. **78** 4063 (1997)
[31] G. Shirane, R. Nathans and C. Chen, Phys. Rev. A **134.6** 1547 (1964)
[32] J. Kouvel and C. Hartelius, Journal of Applied Physics **33.3** 1343 (1962)
[33] C. Kittel, Phys. Rev. **120** 335 (1960)
[34] O. N. Mryasov, Phase Transitions **78** 197–208 (2005)
[35] O. N. Mryasov, U. Nowak *et al.*, EPL, (Europhysics Letters) **69.5** 805 (2005)
[36] J. Barker and R. Chantrell, Phys. Rev. B **92** 094402 (2015)
[37] S. Polesya, S. Mankovsky *et al.*, Phys. Rev. B **93** 024423 (2016)
[38] L. Sandratskii and P. Mavropoulos, Phys. Rev. B **83** 174408 (2011)
[39] V. Moruzzi and P. Marcus, Phys. Rev. B **46.5** 2864 (1992)
[40] D. Suess, T. Schrefl *et al.*, Applied Physics Letters **87.1** 012504 (2005)
[41] D. Suess, Applied Physics Letters **89.11** 113105 (2006)
[42] D. Suess, C. Vogler *et al.*, Journal of Applied Physics **117** 163913 (2015)
[43] A. Datta and X. Xu, IEEE Transactions on Magnetics **53.12** 1 (3102105 2017)
[44] A. Q. Wu, Y. Kubota *et al.*, IEEE Transactions on Magnetics **49.2** 779 (2013)
[45] X. Wang, K. Gao *et al.*, IEEE Transactions on Magnetics **49.2** 686 (2013)
[46] M. Wolloch, M. Gruner *et al.*, Phys. Rev. B **94** 174435 (2016)
[47] U. Aschauer, R. Braddell *et al.*, Phys. Rev. B **94** 014109 (2016)
[48] J. B. Staunton, R. Banerjee *et al.*, Phys. Rev. B **89** 054427 (2014)
[49] J. Ricodeau and D. Melville, Journal of Physics, : Metal Physics **2** 337–350 (1972)
[50] D. W. Cooke, F. Hellman *et al.*, Phys. Rev. Lett. **109** 255901 (2012)