Chapter 8

Magnetism

8.1 Introduction

This part of the lecture deals with the magnetic properties and behavior of solids. Most solids are generally considered to be "non-magnetic", which is a loose way of saying that these solids become magnetized only in the presence of an applied magnetic field (dia- and paramagnetism). We will see that in most cases these effects are very weak, and the magnetization is lost, as soon as the external field is removed. Much more interesting (also from a technological point of view) are those materials, which not only have a large magnetization, but also retain it even after removal of the external field. Such materials are called *permanent magnets* (ferromagnetism), and the property that like (unlike) poles of permanent magnets repel (attract) each other has already been known to (and "technologically" exploited in form of compasses by) the Ancient Greeks and Chinese over 2000 years ago. Since then, the importance of magnets has risen continuously, and they now play a crucial role in many modern technologies. Indeed, the revolution in information technology that has taken place over the last few decades owes as much to the developments in magnetic storage of information as it does to the ubiquitous silicon chip.

Magnetic fields act on moving electric charges, and very roughly one may understand the behavior of solids in magnetic fields as a consequence of the presence of "moving charges" in form of the (bound) core and/or (quasi-free) valence electrons of the material. Both types of electrons possess a spin, which in a trivialized classical picture can be viewed as a rotating charge, i.e., a current. Bound electrons furthermore exhibit an orbital momentum, which adds another source of current (again in a simplistic classical picture). These "microscopic currents" react in two general ways to an applied magnetic field: First, a so-called induction current according to the Lenz rule is initiated, which yields a magnetic field opposite to the external one (diamagnetism). Second, the "molecular magnets" represented by the electron currents align to the external field and reinforce it (paramagnetism). If these effects happen for each "molecular magnet" independently, they remain small. Corresponding magnetic properties of solids may then be understood from the individual behavior of the solid's constituents, i.e., atoms/ions (insulators, semiconductors) or atoms/ions and free electrons (conductors), cf. section 8.3. The much stronger ferromagnetism, on the other hand, arises from a collective behavior of the "molecular magnets" in the solid. In section 8.4 we will first discuss the source for such an interaction, before we move on to simple models treating either a possible coupling of localized moments or itinerant ferromagnetism.

8.2 Macroscopic Electrodynamics

Before we plunge into the microscopic sources of magnetism in solids, let's first recap a few definitions from macroscopic electrodynamics. In vacuum we have $\mathbf{E}$ and $\mathbf{B}$ as the electric field (unit: V/m) and the magnetic flux density (unit: Tesla = Vs/m^2), respectively. Note that both are vector quantities, and in principle they are also functions of place and time. Since we will only deal with uniform constant fields in this lecture, this dependence will be dropped though. Inside macroscopic media, both fields can be affected by the charges and currents present in the material, yielding the two new net fields, $\mathbf{D}$ (electric displacement, unit: As/m^2) and $\mathbf{H}$ (magnetic field, unit: A/m). For the formulation of macroscopic electrodynamics (the Maxwell equations in particular) we therefore need so-called *constitutive relations* between the external (applied) and internal (effective) fields:

$$\mathbf{D} = \mathbf{D}(\mathbf{E}, \mathbf{B}), \qquad \mathbf{H} = \mathbf{H}(\mathbf{E}, \mathbf{B}). \tag{8.1}$$

In general, this functional dependence can be written in form of a multipole expansion series. In most materials, however, only the first (dipole) contribution in the series needs to be considered. These dipole terms are called $\mathbf{P}$ (el. polarization) and $\mathbf{M}$ (magnetization), and we can write

$$\mathbf{H} = (1/\mu_o)\mathbf{B} - \mathbf{M} + \ldots \tag{8.2}$$

$$\mathbf{D} = \epsilon_o \mathbf{E} + \mathbf{P} + \ldots, \tag{8.3}$$

where $\epsilon_o = 8.85 \cdot 10^{-12}$ As/Vm and $\mu_o = 4\pi \cdot 10^{-7}$ Vs/Am are the dielectric constant and permeability of vacuum, respectively. $\mathbf{P}$ and $\mathbf{M}$ depend on the applied field, which we can formally write down in a general Taylor expansion in $\mathbf{E}$ and $\mathbf{B}$. If the applied fields are not too strong, one can truncate this series after the first term (linear response), i.e., the induced polarization/magnetization is then simply proportional to the applied field. We will see below that external magnetic fields we can generate at present in the laboratory are indeed weak compared to microscopic magnetic fields, i.e., the assumption of a magnetic linear response is often well justified. As a note aside, current high-intensity lasers may, however, bring us easily out of the linear response regime for the electric polarization. Corresponding phenomena are treated in the field of non-linear optics.

In the linear response regime, we can thus write for the α Cartesian coordinate of the induced magnetization and electric polarization

$$M_\alpha = (1/\mu_o) \sum_\beta \chi^{\mathrm{mag}}_{\alpha\beta} B_\beta \tag{8.4}$$

$$P_\alpha = \epsilon_o \sum_\beta \chi^{\mathrm{el}}_{\alpha\beta} E_\beta, \tag{8.5}$$

where $\chi^{\mathrm{el/mag}}_{\alpha\beta}$ is the dimensionless electric/magnetic susceptibility tensor. In simple materials (on which we will focus in this lecture), the linear response is often isotropic in space and parallel to the applied field. The susceptibility tensors then reduce to diagonal form and we can simplify to

$$\mathbf{M} = (1/\mu_o)\chi^{\mathrm{mag}}\, \mathbf{B} \tag{8.6}$$

$$\mathbf{P} = \epsilon_o \chi^{\mathrm{el}}\, \mathbf{E}. \tag{8.7}$$

These equations are in fact often directly written as defining equations for the dimensionless susceptibility *constants* in solid state textbooks. We recall, however, that the approximations/restrictions we have employed to arrive at them are that of dipole moment linear response in isotropic media. For this case, the constitutive relations between external and internal field reduce to

$$\mathbf{D} = \epsilon_o \epsilon \, \mathbf{E} \tag{8.8}$$

$$\mathbf{H} = \frac{1}{\mu_o \mu} \, \mathbf{B}, \tag{8.9}$$

where $\epsilon = 1 + \chi^{\rm el}$ is the dielectric constant, and $\mu = 1 - \chi^{\rm mag}$ the permeability of the medium. For $\chi^{\rm mag} < 0$, we have $\mu > 1$, which implies that the effective field in the solid is smaller than the applied one. The "molecular magnets" tend therefore to screen the exterior field, and such a system is called *diamagnetic*. For the opposite case ($\chi^{\rm mag} > 0$, $\mu < 1$), the external field is reinforced in the solid, and we talk about *paramagnetism*.

The process of screening or reinforcing the external field will obviously require work, i.e., the energy of the system is changed (magnetic energy). Assume therefore that we have a solid of volume V, which we bring into a magnetic field B (which for simplicity we take to be along one dimension only). From electrodynamics we know that the energy of this magnetic field is

$$E^{\rm mag} = (1/2BH)V \qquad \Rightarrow \quad 1/V dE^{\rm mag} = 1/2BdH + 1/2HdB. \tag{8.10}$$

From Eq. (8.9) we find $HdB = BdH$, which leads to

$$1/V dE^{\rm mag} = BdH = (\mu_o H + \mu_o M)dH = B_o dH + \mu_o M dH. \tag{8.11}$$

In the first term, we have realized that $\mu_o H = B_o$ is just the field in the vacuum, i.e., without the solid. This term describes therefore simply the energy change connected with the presence of the field, if no solid was present. Only the second term has really to do with the system we are interested in, and describes the energy change due to the reaction of the solid to the applied field. The energy change of the solid itself is therefore

$$dE^{\rm mag}_{\rm solid} = -\mu_o M V dH. \tag{8.12}$$

Recalling that the energy of a dipole with magnetic moment m in a magnetic field is $E = -mB$, we see that the approximations leading to Eq. (8.9) mean nothing else but assuming that the homogeneous solid is build up of a constant density of "molecular dipoles", i.e., the magnetization M is the (average) dipole moment density.

Rearranging Eq. (8.12), we arrive finally at an expression that is a special case of a frequently employed alternative *definition* of the magnetization

$$M(H) = -\frac{1}{\mu_o V} \left. \frac{\partial E(H)}{\partial H} \right|_{S,V} , \tag{8.13}$$

and in turn of the susceptibility

$$\chi^{\rm mag}(H) = \left. \frac{\partial M(H)}{\partial H} \right|_{S,V} = -\frac{1}{\mu_o V} \left. \frac{\partial^2 E(H)}{\partial H^2} \right|_{S,V} . \tag{8.14}$$

At finite temperatures, it is straightforward to generalize these definitions to

$$M(H,T) = -\frac{1}{\mu_o V}\left.\frac{\partial F(H,T)}{\partial H}\right|_{S,V} \qquad \chi^{\mathrm{mag}}(H,T) = -\frac{1}{\mu_o V}\left.\frac{\partial^2 F(H,T)}{\partial H^2}\right|_{S,V}. \tag{8.15}$$

While the derivation via macroscopic electrodynamics is conceptually most important, we will see that these last two equations will be much more useful for the actual quantum-mechanical computation of the magnetization and susceptibility of real systems. All we have to do, is to derive an expression for the energy of the system as a function of the applied external field. The first and second derivatives with respect to the field yield then M and χ^{mag}, respectively.

8.3 Magnetism of Atoms and Free Electrons

We had already discussed that magnetism arises out of the "microscopic currents" connected with the orbital and spin momentum of the electrons. Each electron represents therefore a kind of "microscopic magnet", and the question in solids with their huge numbers of electrons is obviously how all these micromagnets couple together to produce the macroscopic magnetization. If we take a tight-binding type of viewpoint and approach the solid from the extreme of an ensemble of atoms, it is reasonable to address this formidable issue first by evaluating how the electrons couple within one atom, before we worry about how all these "atomic magnets" couple with each other. Since both orbital and spin momentum of all bound electrons of an atom will contribute to its magnetic behavior, it will be useful to first recall how the total electronic angular momentum of an atom is determined (section 8.3.1), before we turn to the effect of a magnetic field on the atomic Hamiltonian (section 8.3.2). In metals, we have in addition the presence of free conduction electrons, the magnetism of which will be addressed in section 8.3.3. Finally, we will discuss in section 8.3.4, which aspects of the magnetism of solids we can already understand just on the basis of these results on atomic magnetism, i.e., in the limit of a vanishing coupling between the resulting "atomic magnets".

8.3.1 Total Angular Momentum of Atoms

In the atomic shell model, the possible quantum states for the electrons are labeled as nl_{m_l}, where n is the *principle quantum number*, l the *orbital quantum number*, and m_l the *orbital magnetic quantum number*. For any given n, defining a so-called "shell", the orbital quantum number l can take only integer values between 0 and $(n-1)$. For $l = 0, 1, 2, 3\ldots$, we generally use the letters s, p, d, f and so on. Within such a "subshell" of defined nl, the orbital magnetic number m_l may have the $(2l+1)$ integer values between $-l$ and $+l$. As a last quantum number, the spin magnetic quantum number m_s takes values of $-1/2$ and $+1/2$. For the example of the first two shells this leads therefore to two $1s$, two $2s$, and six $2p$ states.

When an atom has more than one electron, the *Pauli exclusion principle* states that each quantum state specified by the set of four quantum numbers nl_{m_l,m_s} can only be occupied

el.	$m_l =$	2	1	0	-1	-2	S	$L = \lvert\sum m_l\rvert$	J	Symbol
1		$\downarrow$					$1/2$	2	$3/2$	${}^2D_{3/2}$
2		$\downarrow$	$\downarrow$				1	3	2	${}^3F_2$
3		$\downarrow$	$\downarrow$	$\downarrow$			$3/2$	3	$3/2$	${}^4F_{3/2}$
4		$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$		2	2	0	${}^5D_0$
5		$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$5/2$	0	$5/2$	${}^6S_{5/2}$
6		$\downarrow\uparrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	2	2	4	${}^5D_4$
7		$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow$	$\downarrow$	$\downarrow$	$3/2$	3	$9/2$	${}^4F_{9/2}$
8		$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow$	$\downarrow$	1	3	4	${}^3F_4$
9		$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow$	$1/2$	2	$5/2$	${}^2D_{5/2}$
10		$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow\uparrow$	$\downarrow\uparrow$	0	0	0	${}^1S_0$

Table 8.1: Ground states of ions with partially filled d-shells ($l = 2$), as constructed from Hund's rules.

by one electron. In all but the very heavy ions, spin-orbital coupling is not too strong. In this case, the total orbital and spin angular momentum for a given subshell,

$$L = \sum m_l \qquad \text{and} \qquad S = \sum m_s, \tag{8.16}$$

are good quantum numbers (Russel-Saunders coupling). If a subshell is completely filled, it is easy to verify that its L and S are zero. The total electronic angular momentum of the atom $J = L + S$, is therefore determined by the outermost partially filled shell. How the electrons are distributed among the states of this last partially filled shell is finally summarized in the three Hund's rules:

1. The states are occupied so that as many electrons as possible (within the limitations of the Pauli exclusion principle) have their spins aligned parallel to one another, i.e., so that the value of S is as large as possible.

2. When determined how the spins are assigned, then the electrons occupy states such that the value of L is a maximum.

3. The total angular momentum J is obtained by combining L and S as follows:

   - if the subshell is less than half filled (i.e., if the number of electrons is $< 2l+1$), then $J = L - S$;
   - if the subshell is more than half filled (i.e., if the number of electrons is $> 2l+1$), then $J = L + S$;
   - if the subshell is exactly half filled (i.e., if the number of electrons is $= 2l + 1$), then $L = 0, J = S$.

The first two Hund's rules are determined purely by electrostatic energy considerations (e.g. electrons with equal spins are farther away from each other on account of the exchange-hole). Only the third rule, giving the final splitting within the LS multiplet follows from spin-orbit coupling. The notation for the *ground-state multiplet*, obtained by these rules, indicates the total angular momentum as a capital letter ($L = 0, 1, 2, 3, \ldots =$

$S, P, D, F, \ldots$). Attached to this is the spin as a superprefix $(2S+1)$, and the total angular momentum, J, as a subscript, i.e., ${}^{(2S+1)}L_J$. The rules are in fact easier to apply than their description might suggest at first glance. Table 8.1 gives the example for the filling and notation of a d-shell.

8.3.2 General Derivation of Atomic Susceptibilities

Having determined the total angular momentum of an atom, we now turn to how the presence of a uniform magnetic field $\mathbf{B}$ (taken along the z-axis) modifies the electronic Hamiltonian of an atom $H^e = T^e + V^{e-ion} + V^{e-e}$. Note that the focus on H^e means that we neglect the effect of $\mathbf{B}$ on the nuclear motion and spin. This is in general justified by the much greater mass of the nuclei (rendering the nuclear contribution to the atomic magnetic moment very small), but it would of course be crucial if we were to address e.g. nuclear magnetic resonance (NMR) experiments. Since V^{e-ion} and V^{e-e} are not affected by the magnetic field, we are first left with the effect on the kinetic energy operator.

From classical electrodynamics we know that in the presence of a magnetic field, we have to replace all momenta $\mathbf{p}$ $(= i\hbar\nabla)$ by the canonic momenta $\mathbf{p} \to \mathbf{p} + e\mathbf{A}$. For a uniform magnetic field (along the z-axis), a suitable vector potential $\mathbf{A}$ is

$$\mathbf{A} = -\frac{1}{2}\left(\mathbf{r} \times \mathbf{B}\right), \tag{8.17}$$

for which it is straightforward to verify that it fulfills the conditions $\mathbf{B} = (\nabla \times \mathbf{A})$ and $\nabla \cdot \mathbf{A} = 0$. The total kinetic energy operator is then written

$$T^e(\mathbf{B}) = \frac{1}{2m}\sum_k \left[\mathbf{p}_k + e\mathbf{A}\right]^2 = \frac{1}{2m}\sum_k \left[\mathbf{p}_k - \frac{e}{2}\left(\mathbf{r}_k \times \mathbf{B}\right)\right]^2 \quad . \tag{8.18}$$

Expanding the square leads to

$$\begin{aligned} T^e(\mathbf{B}) &= \sum_k \left[\frac{p_k^2}{2m} + \frac{e}{2m}\mathbf{p}_k \cdot (\mathbf{B} \times \mathbf{r}_k) + \frac{e^2}{8m}\left(\mathbf{r}_k \times \mathbf{B}\right)^2\right] \\ &= T_o^e + \sum_k \frac{e}{2m}\left(\mathbf{r}_k \times \mathbf{p}_k\right) \cdot \mathbf{B} + \frac{e^2}{8m}B^2 \sum_k \left(x_k^2 + y_k^2\right) \quad . \end{aligned} \tag{8.19}$$

If we exploit that the total electronic orbital momentum operator $\mathbf{L}$ can be expressed as $\hbar\mathbf{L} = \sum_k (\mathbf{r}_k \times \mathbf{p}_k)$, and introduce the Bohr magneton $\mu_B = (e\hbar/2m) = 0.579 \cdot 10^{-4}\,\mathrm{eV/T}$, the variation of the kinetic energy operator due to the magnetic field can be further simplified to

$$\Delta T^e(\mathbf{B}) = \mu_B \mathbf{L} \cdot \mathbf{B} + \frac{e^2}{8m}B^2 \sum_k \left(x_k^2 + y_k^2\right) \quad . \tag{8.20}$$

If this was the only effect of the magnetic field on H^e, one can, however, show that the magnetization in thermal equilibrium must always vanish (Bohr-van Leeuwen theorem). There must therefore be another contribution, and it turns out that this is the interaction energy of the field with each electron spin. Such a term is not obvious in the non-relativistic electronic Hamiltonian we had started out with, but follows naturally in

the non-relativistic limit of the Dirac equation. As a note aside, one also obtains in this limit another term, the spin-orbit coupling, but we will neglect this contribution here. The new interaction energy term has the form

$$\Delta H^{\rm spin}(\mathbf{B}) = g_0 \mu_B B \mathbf{S}_z \,, \quad \text{where} \quad \mathbf{S}_z = \sum_k \mathbf{s}_{z,k} \quad . \tag{8.21}$$

Here, g_0 is the so-called electronic *g-factor* (i.e., the g-factor for one electron spin),

$$\begin{aligned} g_0 &= 2\left[1 + \frac{\alpha}{2\pi} + O(\alpha^2) + \ldots\right] \,, \quad \text{where} \quad \alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137} \quad , \\ &= 2.0023\ldots, \end{aligned} \tag{8.22}$$

which can usually be taken as just 2.

The total field-dependent Hamiltonian is therefore,

$$\Delta H^e(\mathbf{B}) = \Delta T^e(\mathbf{B}) + \Delta H^{\rm spin}(\mathbf{B}) \;=\; \mu_B \left(\mathbf{L} + g_0 \mathbf{S}\right) \cdot \mathbf{B} \;+\; \frac{e^2}{8m} B^2 \sum_k \left(x_k^2 + y_k^2\right). \tag{8.23}$$

We will see below that the energy shifts produced by Eq. (8.23) are generally quite small on the scale of atomic excitation energies, even for the highest presently attainable laboratory field strengths. Perturbation theory can thus be used to calculate the changes in the energies connected with this modified Hamiltonian. Remember that magnetization and susceptibility were the first and second derivative of the field-dependent energy with respect to the field, which is why we want to obtain the energy as a function of $\mathbf{B}$. Since we lateron want the second derivative, terms to second order in $\mathbf{B}$ must be retained. The general result of the corresponding second order perturbation theory for the energy of the *non-degenerate* nth level is,

$$E_n \to E_n + \Delta E_n(\mathbf{B}); \quad \Delta E_n = < n|\Delta H^e(\mathbf{B})|n > + \sum_{n' \neq n} \frac{| < n|\Delta H^e(\mathbf{B})|n' > |^2}{E_n - E'_n} \quad . \tag{8.24}$$

Substituting Eq. (8.23) into the above, and retaining terms to quadratic in $\mathbf{B}$, we arrive at the basic equation for theories of the magnetic susceptibility of atoms,

$$\begin{aligned} \Delta E_n &= \mu_B \mathbf{B} \cdot < n|\mathbf{L} + g_0 \mathbf{S}|n > \; + \frac{e^2}{8m} B^2 < n| \sum_k \left(x_k^2 + y_k^2\right) |n > \\ &+ \sum_{n' \neq n} \frac{| < n|\mu_B \mathbf{B} \cdot (\mathbf{L} + g_0 \mathbf{S})|n' > |^2}{E_n - E_{n'}} \quad . \end{aligned} \tag{8.25}$$

Since we are mostly interested in the ground state |0> of the atom, we will now proceed to evaluate the magnitude of the three terms contained in Eq. (8.25) and their corresponding susceptibilities for |0>. This will bring us also a feeling of what is actually the physics behind the various contributions.

Second term: Larmor/Langevin Diamagnetism

To obtain an estimate for the magnitude of this term, we assume spherically symmetric wavefunctions ($< 0|\sum_k (x_k^2+y_k^2)|0> \approx 2/3 < 0|\sum_k r_k^2|0>$). This allows us to approximate the energy shift in the atomic ground state due to the second term as

$$\Delta E_0^{\rm dia} \approx \frac{e^2B^2}{12m}\sum_k < 0|r_k^2|0> \sim \frac{e^2B^2}{12m}\, Z\,\bar{r}^2_{\rm atom} \quad , \tag{8.26}$$

where Z is the total number of electrons in the atom (resulting from the sum over the k electrons in the atom), and $\bar{r}^2_{\rm atom}$ is the mean square atomic radius. If we take $Z \sim 30$ and $\bar{r}^2_{\rm atom}$ of the order of Å^2, we find $\Delta E_0^{\rm dia} \sim 10^{-9}$ eV even for fields of the order of Tesla. This contribution is therefore rather small, which we find similarly expressed, when we make a similar order of magnitude estimate for the susceptibility,

$$\chi^{\rm mag,dia} = -\frac{\mu_o}{V}\frac{\partial^2 E_0}{\partial B^2} = -\frac{\mu_o e^2 Z \bar{r}^2_{\rm atom}}{6mV} \sim -10^{-4}. \tag{8.27}$$

With $\chi^{\rm mag,dia} < 0$, this term represents a diamagnetic contribution (so-called Larmor or Langevin diamagnetism), i.e., the associated magnetic moment tends to screen the exterior field. We have thus identified the term which we qualitatively expected to arise out of the induction currents initiated by the applied magnetic field.

We will see below that this diamagnetic contribution will only determine the overall magnetic susceptibility of the atom, when the other two terms in Eq. (8.25) vanish. This is only the case for atoms with $\mathbf{J}|0> = \mathbf{L}|0> = \mathbf{S}|0>$, i.e., for atoms or ions with all electronic shells filled (e.g. the noble gas atoms). In this case, the ground state $|0>$ is indeed non-degenerate, which justified our use of Eq. (8.25) to calculate the energy level shift, and from chapter 6 on cohesion we recall that the first excited state is much higher in energy. In all but the highest temperatures, there is then also a negligible probability of the atom being in any but its ground state in thermal equilibrium. This means that the diamagnetic susceptibility will be largely temperature independent (otherwise it would have been necessary to use Eq. (8.15) for the derivation of $\chi^{\rm mag,dia}$).

Third term: Van Vleck Paramagnetism

By taking the second derivative of the third term in Eq. (8.25), we obtain for its susceptibility

$$\chi^{\rm mag,vleck} = \frac{2\mu_o\mu_B^2}{V}\sum_n \frac{|< 0|(\mathbf{L_z} + g_0\mathbf{S_z})|n>|^2}{E_n - E_0}. \tag{8.28}$$

Note that we have reversed the order in the denominator as compared to Eq. (8.25), which induced the change of sign of $\chi^{mag,vleck}$. Since the energy of any excited state will necessarily be higher than the ground state energy ($E_n > E_0$), $\chi^{\rm mag,vleck} > 0$. The third term represents therefore a paramagnetic contribution (so-called Van Vleck paramagnetism), which is connected to field-induced electronic transitions. If the electronic ground state is non-degenerate (and only for this case, does the above formula hold), it is precisely the normally quite large separation between electronic states which makes this term similarly small as the diamagnetic one. Van Vleck paramagnetism plays therefore only a role for

atoms or ions with shells that are one electron short of being half filled (which is the only case when the third term in Eq. (8.25) does not vanish, while the first term does). The magnetic behavior of such atoms or ions is determined by a balance between Larmor diamagnetism and Van Vleck paramagnetism. Both terms are small, and tend even to cancel each other out.

First term: Paramagnetism

For any atom with $J \neq 0$ (which is the case for most atoms), the first term does not vanish and will then completely dominate the magnetic behavior. Anticipating this result, we will neglect the first and third term in Eq. (8.25) for the moment. Atoms with $J \neq 0$ have a $(2J+1)$-fold degenerate ground state, which implies that the simple form of Eq. (8.25) can not be applied. Instead, the $\alpha = 1, \ldots, (2J+1)$ energy shifts within the ground state subspace are given by

$$\Delta E_{0,\alpha} \;=\; \mu_B B \sum_{\alpha'=1}^{(2J+1)} < 0\alpha|\mathbf{L_z} + g_0\mathbf{S_z}|0\alpha' > = \; \mu_B B \sum_{\alpha'=1}^{(2J+1)} V_{\alpha,\alpha'}, \tag{8.29}$$

where we have again used that the magnetic field goes along the z-axis, and have defined the interaction matrix $V_{\alpha,\alpha'}$. The energy shifts are therefore obtained by diagonalizing the interaction matrix within the ground state subspace. This is a standard problem in atomic physics (see e.g. Ashcroft/Mermin), and one finds that the basis diagonalizing this matrix are the states of defined J and J_z,

$$< JLS, J_z|\mathbf{L_z} + g_0\mathbf{S_z}|JLS, J_z' > = \; g(JLS)J_z \; \delta_{J_z,J_{z'}}. \tag{8.30}$$

where JLS are the quantum numbers defining the atomic ground state $|0>$ in the shell model, and $g(JLS)$ is the so-called *Landé g-factor*. Approximating $g_o \approx 2$, this factor is obtained as

$$g(JLS) \;=\; \frac{3}{2} \;+\; \frac{1}{2}\left[\frac{S(S+1) - L(L+1)}{J(J+1)}\right] \quad . \tag{8.31}$$

If we insert Eq. (8.30) into Eq. (8.29), we obtain

$$\Delta E_{JLS,J_z} \;=\; g(JLS)\mu_B J_z \; B \quad , \tag{8.32}$$

i.e., the magnetic field splits the degenerate ground state into $(2J+1)$ equidistant levels separated by $g(JLS)\mu_B B$ (i.e., $-g(JLS)\mu_B JB, -g(JLS)\mu_B(J-1)B, \ldots, +g(JLS)\mu_B(J-1)B, +g(JLS)\mu_B JB$). This is the same effect, as the magnetic field would have on a magnetic dipole with magnetic moment

$$\mathbf{m}_{\text{atom}} = -g(JLS)\mu_B \; \mathbf{J}. \tag{8.33}$$

The first term in Eq. (8.25) can therefore be interpreted as the expected paramagnetic contribution due to the alignment of the "microscopic magnet" connected with the total angular momentum of the atom. This happens, of course, only when we have unpaired electrons in partially filled shells of the atom ($J \neq 0$).

Before we proceed to derive the paramagnetic susceptibility that arises from this contribution, we note that this identification of $\mathbf{m}_{\text{atom}}$ via Eq. (8.25) serves nicely to illustrate the

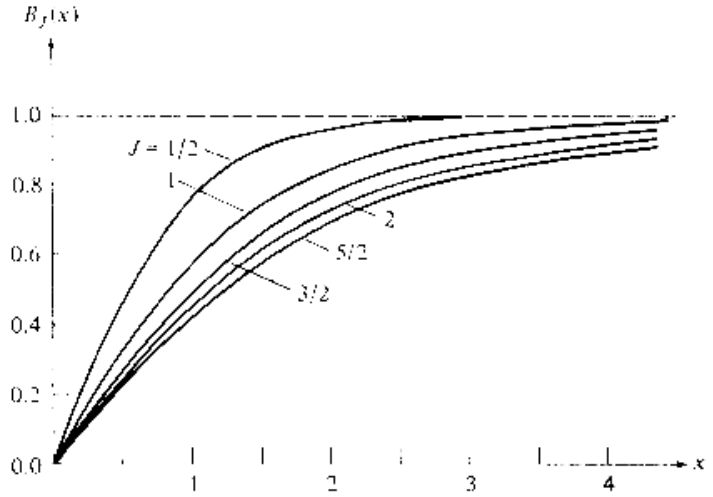

Figure 8.1: Plot of the Brillouin function $B_J(x)$ for various values of the spin J (the total angular momentum).

difference between phenomenological and so-called first-principles theories. We have seen that the simple Hamiltonian $H^{\text{spin}} = -\mathbf{m}_{\text{atom}} \cdot \mathbf{B}$ would equally well describe the splitting of the $(2J+1)$ ground state levels. If we had only known this splitting, e.g. from experiment (*Zeeman effect*), we could have simply used H^{spin} as a phenomenological, so-called "spin Hamiltonian". By fitting to the experimental splitting, we could even have obtained the magnitude of the magnetic moment for individual atoms (which due to the different total angular momentum is of course different for different species). Studying empirically all the fitted magnitudes, we might even have discovered that the splitting is connected to the total angular momentum. The virtue of the first-principles derivation used in this lecture, on the other hand, is that this prefactor is directly obtained, without any fitting to experimental data. This allows us not only to unambiguously check our theory by comparison with the experimental data (which is not possible in the phenomenological theory, where the fitting can often easily cover up for deficiencies in the assumed ansatz). We also obtain directly the dependency on the total angular momentum, and can even predict the properties of atoms which have not (yet) been measured. Admittedly, the first-principles theory will be harder to accomplish, and we will for example see later on several other "spin Hamiltonians", which describe the observed collective magnetic properties of solids very well, but where an all-embracing first-principles derivation is still lacking.

Paramagnetic Susceptibility: Curie's Law

The separation of the $(2J+1)$ split ground state levels of a paramagnetic atom in a magnetic field is $g(JLS)\mu_B B$. Recalling that $\mu_B = 0.579 \cdot 10^{-4}\,\text{eV/T}$, we see that the splitting is only of the order of $10^{-4}\,\text{eV}$, even for fields of the order of Tesla. This is obviously small compared to $k_B T$ for realistic temperatures. At finite temperature, a significant population not only of the lowest lying level can therefore be expected, which requires the use of Eq. (8.15) for the evaluation of the magnetic susceptibility. This means that first the free energy (or equivalently the partition function) of the $(2J+1)$ level system must be obtained. Defining $\eta = (g(JLS)\mu_B B)/(k_B T)$ (i.e., the fraction of magnetic versus

thermal energy), the partition function turns out as

$$\begin{aligned} q &= \sum_{J_z=-J}^{J} e^{-\eta J_z} = \frac{e^{-\eta J} - e^{\eta(J+1)}}{1-e^{\eta}} \\ &= \frac{e^{-\eta(J+1/2)} - e^{\eta(J+1/2)}}{e^{-\eta/2} - e^{\eta/2}} = \frac{\sinh\left[(J+1/2)\eta\right]}{\sinh\left[\eta/2\right]} \quad . \end{aligned} \tag{8.34}$$

With this, one finds for the magnetization

$$\begin{aligned} M(T) &= -\frac{1}{V}\frac{\partial F}{\partial B} = -\frac{1}{V}\frac{\partial(-k_B T \ln q)}{\partial B} \\ &= \frac{k_B T}{V}\frac{\partial}{\partial B}\left[\ln\left(\sinh[(J+1/2)\eta]\right) - \ln\left(\sinh[\eta/2]\right)\right] = \frac{g(JLS)\mu_B J}{V}\, B_J(\eta), \end{aligned} \tag{8.35}$$

where in the last step, we have defined the so-called *Brillouin function*

$$B_J(\eta) = \frac{1}{J}\left\{(J+\frac{1}{2})\coth\left[(J+\frac{1}{2})\eta/J\right] - \frac{1}{2}\coth\left[\frac{\eta}{2J}\right]\right\} \quad . \tag{8.36}$$

As shown in Fig. 8.1, $B_J \to 1$ for $\eta \gg 1$, in which case Eq. (8.35) simply tells us that all atomic magnets with momentum $\mathbf{m}_{\rm atom}$, cf. Eq. (8.33), have aligned to the external field, and the magnetization reaches its maximum value of $M = m_{\rm atom}/V$, i.e., the magnetic moment density.

We had, however, discussed above, that $\eta = (g(JLS)\mu_B B)/(k_B T)$ will rather be much smaller than unity for normal field strengths. At all but the lowest temperatures, the other limit for the Brillouin function, i.e., $B_J(\eta \to 0)$, will therefore be much more relevant. This small η-expansion is

$$B_J(\eta \ll 1) \approx \frac{J+1}{3}\eta + O(\eta^3) \quad . \tag{8.37}$$

In this limit, we obtain finally for the paramagnetic susceptibility

$$\chi^{\rm mag,para}(T) = \mu_o \frac{\partial M}{\partial B} = \frac{\mu_o \mu_B^2 g(JLS)^2 J(J+1)}{3Vk_B}\,\frac{1}{T} \quad . \tag{8.38}$$

With $\chi^{\rm mag,para} > 0$, we have now confirmed our already made assessment, that the first term in Eq. (8.25) represents indeed a paramagnetic contribution. The variation of the susceptibility of paramagnetic atoms ($J \neq 0$) inversely with temperature is known as *Curie's law*, $\chi^{\rm mag,para} = C_{\rm Curie}/T$ with $C_{\rm Curie}$ the Curie constant. Such a dependence is characteristic for a system with "molecular magnets" whose alignment is favored by the applied field, and opposed by thermal disorder. Again, we see that the carried out first-principles derivation allows not only to show the validity range of this empirical law (remember that it only holds in the small η-limit), but also for the determination of the Curie proportionality constant in terms of fundamental system properties.

If we take the volume again as of an atomic order of magnitude ($V \sim$ Å^3), Eq. (8.38) allows the straightforward estimate that $\chi^{\rm mag,para} \sim 10^{-2}$ at room temperature. As already mentioned several times, the paramagnetic contribution is thus even at room temperature orders of magnitude larger than the diamagnetic or the Van Vleck paramagnetic one.

Nevertheless, it is with $\chi^{\mathrm{mag,para}} \ll 1$ still quite small (compared to electric susceptibilities, which are of the order of unity). We will discuss the consequences of this minute value in section 8.3.4, but before we have to evaluate a last remaining, additional source of magnetic behavior in metals, i.e., the free electrons.

8.3.3 Susceptibility of the Free Electron Gas

Having determined the magnetic properties of electrons when strongly bound within ions, it is valuable to also consider the opposite extreme and examine their properties as they move about nearly free in a metal. Similar to the case of bound electrons in atoms, there are essentially two major terms in the response of free fermions to an external magnetic field: One results from the kind of "alignment" of the spins (which is a rough way of visualizing the quantum mechanical action of the field on the spins) and is called *Pauli paramagnetism*. The other arises from the orbital moments created by induced circular motions of the electrons. It thus tends to screen the exterior field and is known as *Landau diamagnetism*.

Let us first try to analyze the origin of Pauli paramagnetism and assess its order of magnitude like we have also done for all atomic magnetic effects. For this we consider again the free electron gas, i.e., N non-interacting electrons in a volume V. In chapter 2 we had already derived that the density of states (DOS) per volume for this model is

$$N(\epsilon) \;=\; \frac{1}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2}\sqrt{\epsilon} \quad . \tag{8.39}$$

In the electronic ground state, all states are filled up according to the Fermi distribution

$$\frac{N}{V} \;=\; \int_0^\infty N(\epsilon)\; f(\epsilon,T)\; d\epsilon, \tag{8.40}$$

yielding in total the correct electron density N/V uniquely characterizing our electron gas. For sufficiently low temperatures, the Fermi-distribution can be approximated by a step function, so that simply all states up to the Fermi level are occupied. This Fermi level is conveniently specified in terms of the Wigner-Seitz radius r_s that we had already defined in chapter 6

$$\epsilon_F \;=\; \frac{50.1\,\mathrm{eV}}{\left(\frac{r_s}{a_B}\right)} \quad . \tag{8.41}$$

The DOS at the Fermi level can in turn equally be given as a mere function of r_s

$$N(\epsilon_F) \;=\; \left(\frac{1}{20.7\,\mathrm{eV}^3}\right)\left(\frac{r_s}{a_B}\right)^{-1} \quad . \tag{8.42}$$

Up to now we have not explicitly considered the spin degree of freedom, which leads to the double occupation of each electronic state with a spin up and a spin down electron. Such an explicit consideration becomes necessary, however, when we now want to address the effect of an external field on the free electron gas. This is accomplished by defining spin up and spin down DOSs, $N^\uparrow$ and $N^\downarrow$, in an analog manner as in the case of the spin-unresolved DOS before. In the absence of a magnetic field, the ground state of the free

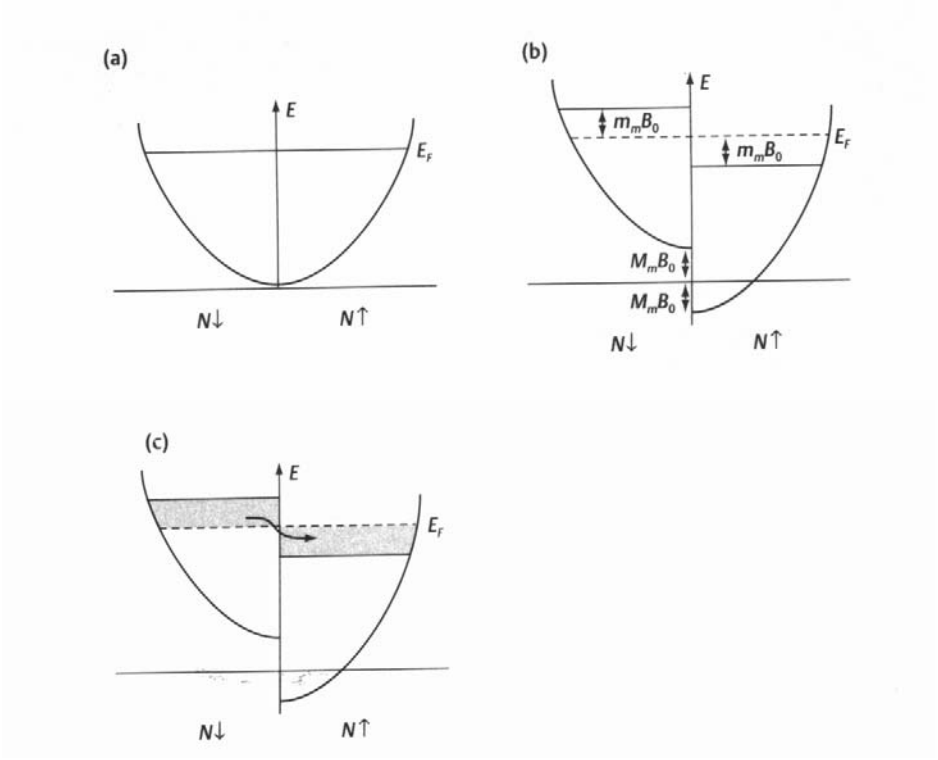

Figure 8.2: Free electron density of states and the effect of an external field: N↑ corresponds to the number of electrons with spins aligned parallel and N↓ to the number of electrons with spins antiparallel to the z-axis. (a) No magnetic field, (b) a magnetic field, $\mathbf{B}_0$ is applied along the z-direction, (c) as a result of (b) some electrons with spins antiparallel to the field (shaded region on left) change their spin and move into available lower energy states with spins aligned parallel to the field.

electron gas has equal numbers of spin-up and spin-down electrons, and the degenerate spin-DOS are therefore simply half-copies of the original total DOS defined in Eq. (8.39)

$$N^{\uparrow}(\epsilon) \;=\; N^{\downarrow}(\epsilon) \;=\; \frac{1}{2}N(\epsilon) \quad , \tag{8.43}$$

cf. the graphical representation in Fig. 8.2a.

An external field $\mathbf{B}$ will now interact differently with spin up and spin down states. Since the electron gas is fully isotropic we can always choose the spin axis defining up and down spin to go along the direction of the applied field, and thus reduce the problem to one dimension. If we focus only on the interaction of the spin with the field and neglect the orbital response for the moment, the effect boils down to an equivalent of Eq. (8.21), i.e., we have

$$\Delta H^{\text{spin}}(\mathbf{B}) \;=\; \mu_B g_o \mathbf{B}\cdot\mathbf{s} \;=\; +\mu_B B \quad \text{(for spin up)} \tag{8.44}$$

$$\;=\; -\mu_B B \quad \text{(for spin down)} \tag{8.45}$$

where we have approximated the electronic g-factor g_o with 2. The effect is therefore simply to shift the energies of spin up and spin down electrons, which are either aligned

parallel or antiparallel to the external field. This can be expressed via the spin DOSs

$$N^{\uparrow}(\epsilon) = \frac{1}{2}N(\epsilon - \mu_B B) \tag{8.46}$$

$$N^{\downarrow}(\epsilon) = \frac{1}{2}N(\epsilon + \mu_B B) \quad , \tag{8.47}$$

or again graphically by shifting the two parabolas against each other as in Fig. 8.2b. If we now fill the electronic states up according to the Fermi-distribution, a different number of up and down electrons is obtained and our electron gas exhibits a net magnetization due to the applied field. Since the states aligned parallel to the field are favored, the magnetization will enhance the exterior field and a paramagnetic contribution results.

For $T \approx 0$ both parabolas are essentially filled up to a sharp energy, and the net magnetization is simply given by the dark grey shaded region in Fig. 8.2c. Even for fields of the order of Tesla, the energetic shift of the states is $\mu_B B \sim 10^{-4}\,\mathrm{eV}$, i.e., small on the scale of electronic energies. The shaded region representing our net magnetization can therefore be well approximated by a simple rectangle of height $\mu_B B$ and width $\frac{1}{2}N(\epsilon_F)$, i.e., the Fermi level DOS of the unperturbed system without applied field. We then have

$$N^{\uparrow} = \frac{N}{2} + \frac{1}{2}N(\epsilon_F)\mu_B B \tag{8.48}$$

up electrons per volume and

$$N^{\downarrow} = \frac{N}{2} - \frac{1}{2}N(\epsilon_F)\mu_B B \tag{8.49}$$

down electrons per volume. Since each electron contributes a magnetic moment of μ_B to the magnetization, we have

$$M = (N^{\uparrow} - N^{\downarrow})\mu_B = N(\epsilon_F)\mu_B^2 B \tag{8.50}$$

(remember that the magnetization was the dipole magnetic density, and can in our uniform system therefore be written as dipole moment per electron times DOS). This yields then the susceptibility as

$$\chi^{\mathrm{mag,Pauli}} = \mu_o \frac{\partial M}{\partial B} = \mu_o \mu_B^2 N(\epsilon_F) \quad . \tag{8.51}$$

Using Eq. (8.42) for the Fermi level DOS of the free electron gas, this can be rewritten as a pure function of the Wigner-Seitz radius r_s

$$\chi^{\mathrm{mag,Pauli}} = 10^{-6}\left(\frac{2.59}{r_s/a_B}\right) \quad . \tag{8.52}$$

Recalling that the Wigner-Seitz radius of most metals was of the order of 3-5 a_B, this expression reveals that the Pauli paramagnetic contribution ($\chi^{\mathrm{mag,Pauli}} > 0$) from a free electron gas has the minute size characteristic of diamagnetic susceptibilities of atoms, and is thus in distinct contrast to the strinkingly larger paramagnetic susceptibilities of atoms. The reason for this is that for the latter it is only thermal disorder at finite temperatures which prevents their complete alignment to an external field. For an electron gas, on the other hand, it is the Pauli exclusion principle that opposes such an alignment

by forcing the electrons to occupy energetically higher lying states when aligned. It is also for this reason, why the free electron paramagnetism does not show a linear temperature dependence like the Curie law in Eq. (8.38). The characteristic temperature for the electron gas is not the thermal one, but the Fermi temperature. One could therefore cast the Pauli susceptibility in a Curie law form, but with a fixed temperature of order T_F playing the role of T. Since $T_F \gg T$, the Pauli susceptibility is then hundreds of times smaller and almost always completely negligible.

Turning to the second effect in form of Landau diamagnetism, we note that its calculation would require solving the full quantum mechanical free electron problem along similar lines as done in the last section for the atoms. The field-dependant Hamilton operator would then similarly yield a term trying to screen the applied field. Taking the second derivative of the resulting total energy with respect to $\mathbf{B}$, one would obtain for the diamagnetic susceptibility

$$\chi^{\text{mag,Landau}} = -\frac{1}{3}\chi^{\text{mag,Pauli}} \quad , \tag{8.53}$$

i.e., another term that is equally vanishingly small as the paramagnetic response of the free electron gas. In particular, because this term is so small and the derivation does not yield new insights compared to the one already undertaken for the atomic case, we refer to e.g. the book by M.P. Marder, *Condensed Matter Physics* (Wiley, 2000) for a proper derivation of this result.

8.3.4 Atomic Magnetism in Solids

In our first approach to understand the magnetic behavior of solids we took the tight-binding viewpoint of a solid as a mere ensemble of atoms. This led us to discuss in the preceding sections the magnetic properties of isolated atoms (or equivalently ions). As an additional source of magnetism in solids we looked at free (delocalized) electrons. In both cases we found two major sources of magnetic response: a paramagnetic one resulting from the alignment of existing "microscopic magnets" (either total angular momentum in the case of atoms, or spin in case of free electrons), and a diamagnetic one due to induction currents trying to screen the exterior field.

- Atoms (bound electrons):

- Paramagnetism	$\chi^{\text{mag,para}}$	$\approx 1/T \sim 10^{-2}\,(\text{RT})$	$\Delta E_0^{\text{para}} \sim 10^{-4}\,\text{eV}$
- Larmor diamagnetism	$\chi^{\text{mag,dia}}$	$\approx \text{const.} \sim -10^{-4}$	$\Delta E_0^{\text{dia}} \sim 10^{-9}\,\text{eV}$

- Free electrons:

- Pauli paramagnetism	$\chi^{\text{mag,Pauli}}$	$\approx \text{const.} \sim 10^{-6}$	$\Delta E_0^{\text{Pauli}} \sim 10^{-4}\,\text{eV}$
- Landau diamagnetism	$\chi^{\text{mag,Landau}}$	$\approx \text{const.} \sim -10^{-6}$	$\Delta E_0^{\text{Landau}} \sim 10^{-4}\,\text{eV}$

If the coupling between the different sources of magnetism is small, the magnetic behavior of a solid would be given by a simple superposition of all sources. In the case of isolators, this would for example mean that the magnetic moment connected with each atom/ion

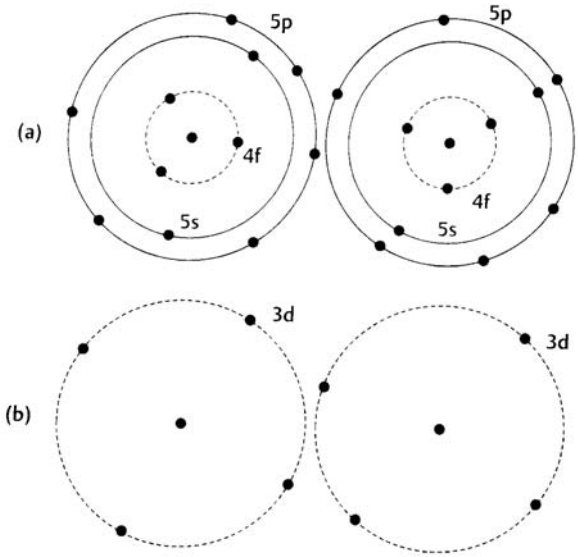

Figure 8.3: (a) In the rare earth metals the incomplete $4f$ electronic subshell is located within the $5s$ and $5p$ subshells, so that the $4f$ electrons are not strongly affected by neighboring atoms. (b) In the transition metals, here the iron group, the $3d$ electrons are the outermost electrons, and so interact strongly with other nearby atoms.

does not change appreciably when condensed in the solid phase, and the total susceptibility is just the sum of all atomic susceptibilities. And this is indeed what one finds primarily when going through the periodic table:

Insulators: When $J = 0$, the paramagnetic contribution vanishes. If in addition $L = S = 0$ (closed-shell), the response is purely diamagnetic as in the case of noble gas solids or simple ionic crystals like the alkali halides (recall from the discussion on cohesion that the latter can be viewed as closed-shell systems!). Otherwise, the response results as a balance between Van Vleck paramagnetism and Larmor diamagnetism. In all cases, the effects are very small, though, and in excellent quantitative agreement with the values following from the theoretical derivation of the last section. More frequent is the situation $J \neq 0$, where the response is then dominated by the paramagnetic term. This is realized, when insulating solids contain immersed rare earth (RE) or transition metal (TM) ions (with partially filled f or d shells, respectively). Such solids are indeed found to obey a Curie law, i.e., they exhibit susceptibilities that scale inversely with temperature. For the RE case, even the magnitude of the measured Curie constant corresponds very well to the theoretical one given by Eq. (8.38). For insulating solids containing TM ions, on the other hand, the measured Curie constant can only be understood by assuming $L = 0$, while S is still given by the Hund's rules. This is denoted as *quenching of the orbital angular momentum* and is due to a general phenomenon known as crystal field splitting: As illustrated by Fig. 8.3, the f orbitals of a RE atom are located deep inside the filled s and p shells and are thus not significantly affected by the crystalline environment. On the contrary, the d shells of a TM atom belong to the outermost valence shells and are correspondingly strongly influenced. As a result, the Hund's rules are partially invalidated. One can derive new rules for any given crystal environment and symmetry and thus bring measured and

calculated Curie constants in agreement. The important thing to notice here is, that although the solid perturbs the magnitude of the "microscopic magnet" connected with the partially filled shell, all individual moments inside the solid are still virtually decoupled from each other (i.e., the magnetic moment in the solid corresponds to an atom that has simply adapted to a new symmetry situation). This is why still a Curie law results, and only the proportionality constant is changed.

Semiconductors: Covalent materials have only partially filled shells, so they could be expected to have a finite magnetic moment. However, covalent bonds form through a pair of electrons with opposite spin, and hence the net orbital angular momentum is zero. Covalent materials like Si (when they are not doped!!) exhibit therefore only a vanishingly small diamagnetic response.

Metals: In metals, the delocalized electrons add a new contribution to the magnetic behavior. In simple metals like the alkalis, this new contribution is in fact the only remaining one. As discussed in the chapter on cohesion, such metals can be viewed as closed-shell ion cores and a free electron glue formed from the valence electrons. The closed-shell ion cores have $J = L = S = 0$ and exhibit therefore only a negligible diamagnetic response. The magnetic behavior is then dominated by the Pauli paramagnetism of the conduction electrons. Measured susceptibilities are indeed of the order as expected from Eq. (8.52), and the quantitative differences arise from the exchange-correlation effects that were not treated in our free electron model. More to come on metals below ...

8.4 Magnetic Order: Permanent Magnets

The punchline of the theory we have developed up to now is, that the magnetic properties of the overwhelming fraction of solids is completely comprehensible within the picture of atomic and free electron magnetism. The paramagnetic or diamagnetic response is in all cases very small, at least when compared to electric susceptibilities, which are of the order of unity. This is by the way also the reason, why typically only the electric effects are discussed, when considering the interaction of electromagnetic radiation with solids in many spectroscopic methods. At this stage, we could therefore close the chapter on magnetism of solids as something really unimportant, if, well if there weren't a few elemental and compound solids (formed of or containing transition metal or rare earth atoms) that exhibit a completely different magnetic behavior: Huge susceptibilities and in particular a magnetization that does not vanish when the external field is removed. Since such so-called *ferromagnetic effects* are totally incomprehensible within the hitherto developed theory, they must originate from a strong coupling of the individual magnetic moments present in these solids, which is what we will now look into in more detail.

8.4.1 Ferro-, Antiferro-, and Ferrimagnetism

The simple theory of paramagnetism in solids assumes that the discrete sources of magnetic moment (ionic shells of non-zero angular momentum in insulators or the conduction

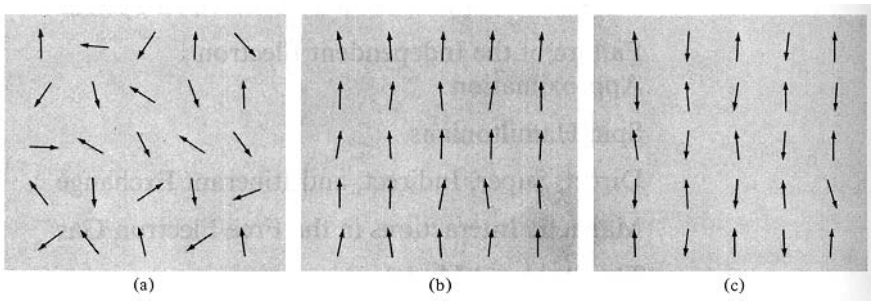

Figure 8.4: Schematic illustration of the distribution of directions of the local magnetic moments connected to "molecular magnets" in a solid. (a) random thermal disorder in a paramagnetic solid with inconsequential magnetic interactions, (b) complete alignment either in a paramagnetic solid due to a strong external field or in a ferromagnetic solid below its critical temperature due to magnetic interactions. (c) Example of antiferromagnetic ordering below the critical temperature.

electrons in simple metals) do not interact with one another. In the absence of an external field, the individual magnetic moments are then thermally disordered at any finite temperature, i.e., they point in random directions yielding a zero net moment for the solid as a whole, cf. Fig. 8.4a. In such cases, an alignment can only be caused by an applied external field, which leads to an ordering of all magnetic moments at sufficiently low temperatures as schematized in Fig. 8.4b. However, a similar effect could also be obtained by a coupling between the different magnetic moments, i.e., an interaction that would e.g. favor a parallel alignment. It is hereby important to realize that already quite short-ranged interactions, for example only between nearest neighbors, could well lead to an ordered structure as shown in Fig. 8.4b. Such interactions are often generically denoted as *magnetic interaction*, although this should not be misunderstood as implying that the source of interaction is really magnetic in nature (e.g. a magnetic dipole-dipole interaction, which in fact is not the reason for the ordering as we will see below).

Solids that due to some magnetic interaction can exhibit an ordered magnetic structure in the absence of an applied external field are called *ferromagnets* (or permanent magnets), and their resulting (often nonvanishing) magnetic moment is known as *spontaneous magnetization* M_s. The complexity of magnetically ordered states exceeds hereby by far the simple parallel alignment case shown in Fig. 8.4b. Another common case is one where the individual local moments sum to zero total moment, and no spontaneous magnetization is present to reveal the microscopic ordering. Such magnetically ordered states are classified as *antiferromagnetic* and one possible realization is shown in Fig. 8.4c. If magnetic moments of different magnitude are present in the solid, and not all local moments have a positive component along the direction of spontaneous magnetization, one finally talks about *ferrimagnets* ($M_\mathrm{s} \neq 0$), which are thus somewhere intermediate between ferro- and antiferromagnetic materials. Some of the many types of magnetic ordering that are possible within these categories are illustrated schematically in Fig. 8.5, but it should be stressed that many magnetic structures are so complex that it is better to describe them explicitly rather than using any of the above three categories. However, such complexities are well beyond the scope of our introductory level lecture, and we will henceforth primar-

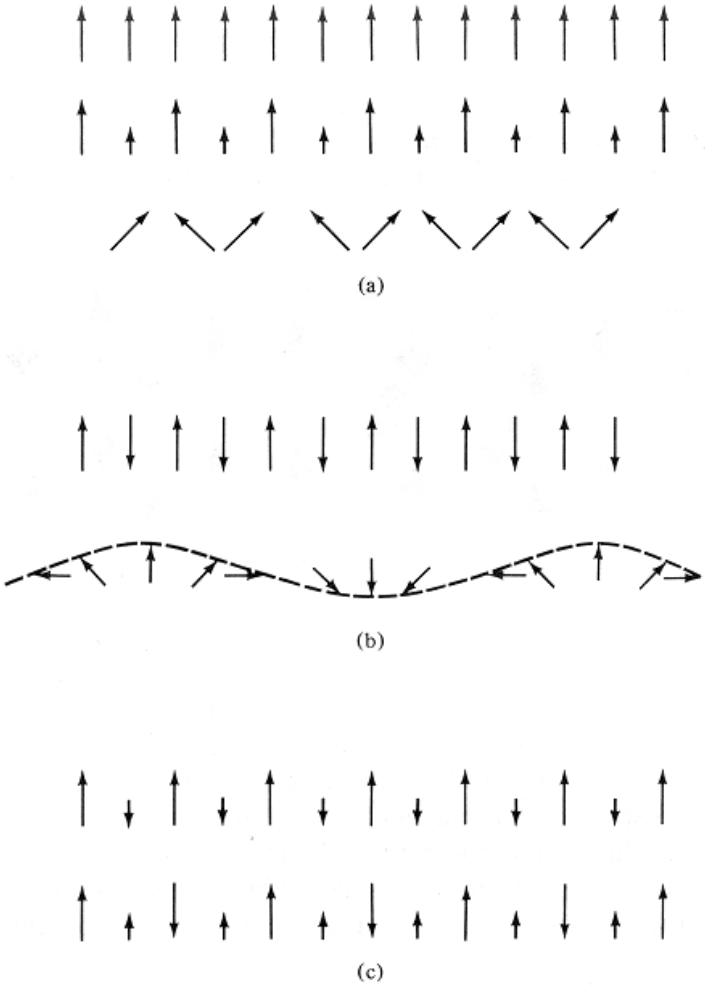

Figure 8.5: Typical magnetic orderings already possible in a simple linear array of spins. (a) ferromagnetic, (b) antiferromagnetic, and (c) ferrimagnetic.

ily focus on the comparably simple (but still annoyingly complex) case of ferromagnetic ordering.

8.4.2 Interaction Versus Thermal Disorder: Curie-Weiss Law

Even in permanent magnets, the magnetic ordering does not prevail at all temperatures. Above a critical temperature T_c, the thermal energy can overcome the ordering tendency induced by the magnetic interaction, the ordering vanishes and the solid (often) behaves like a simple paramagnet. In ferromagnets, T_c is known as the *Curie temperature*, and in antiferromagnets as *Néel temperature* (sometimes denoted with T_N). Note that T_c depends strongly on an applied field (both on strength and direction!). If the external field is parallel to the direction of spontaneous magnetization, T_c is typically the higher, the higher the external field, since both ordering tendencies support each other. Quite a range of complex phenomena can arise for other field directions, for which then competition between both ordering tendencies sets in.

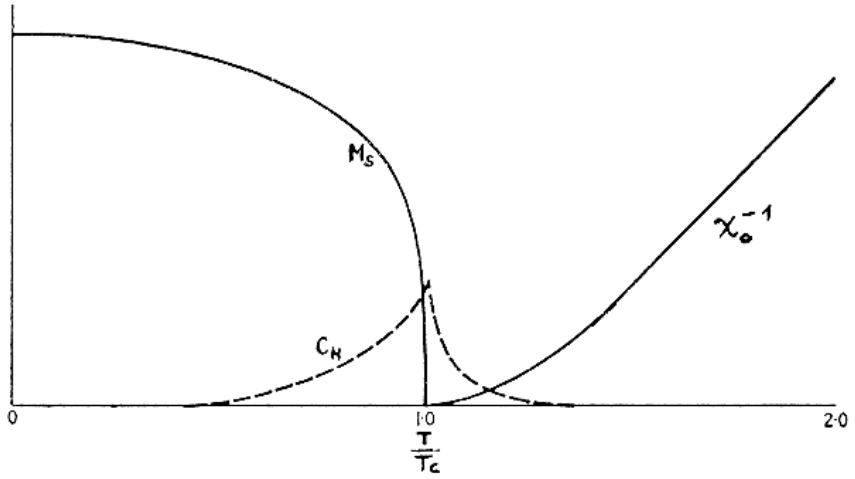

Figure 8.6: Typical temperature dependence of the magnetization M_s, the specific heat c_V and the zero-field susceptibility χ_o of a ferromagnet. The temperature scale is normalized to the critical Curie temperature T_c.

At the moment we are, however, more interested in the ferromagnetic properties of solids in the absence of external fields, i.e., the existence of a finite spontaneous magnetization. The gradual loss of order with increasing temperature is reflected by a continuous drop of $M_\mathrm{s}(T)$ as illustrated in Fig. 8.6. Just below T_c, a power law dependence is typically observed

$$M_\mathrm{s}(T) \;\sim\; (T_c - T)^{\beta} \qquad (\text{for } \; T \to T_c^{-}) \quad , \tag{8.54}$$

with a critical exponent β somewhere around 1/3. Coming from the high temperature side, the onset of ordering is equally signaled by the zero-field susceptibility, which is found to diverge as T drops to values very close to T_c

$$\chi_o(T) \;=\; \chi(T)|_{B=0} \;\sim\; (T - T_c)^{-\gamma} \qquad (\text{for } \; T \to T_c^{+}) \quad , \tag{8.55}$$

with γ around 4/3. This indicates already that the solid undergoes quite some dramatic changes at the critical temperature, which is also reflected by a divergence in other fundamental quantities like the zero-field specific heat

$$c_V(T)|_{B=0} \;\sim\; (T - T_c)^{-\alpha} \qquad (\text{for } \; T \to T_c) \quad , \tag{8.56}$$

with α around 0.1. The reason for this divergence is that by an infinitesimal change in the temperature, a macroscopic magnetization is created due to the onset of long-range order (alignment) among the magnetic moments. Precisely because of these pronounced changes happening over a very small temperature range, the critical behavior (of this second order phase transition) is most difficult to handle theoretically. It has therefore become common practice that theories of magnetism are preferentially benchmarked against how well (or bad) they reproduce the experimentally measured *critical exponents* α, β and γ.

Further, away from the critical temperature, the properties of the magnetic solid become more normal. Particularly, for much higher temperatures than T_c, one very often finds a

	$\bar{m}_s$ (in μ_B)	m_{atom} (in μ_B)	T_c (in K)	Θ_c (in K)
Fe	2.2	6 (4)	1043	1100
Co	1.7	6 (3)	1394	1415
Ni	0.6	5 (2)	628	650
Eu	7.1	7	289	108
Gd	8.0	8	302	289
Dy	10.6	10	85	157

Table 8.2: Magnetic quantities of some elemental ferromagnets: The saturation magnetization at $T = 0\,\text{K}$ is given in form of the average magnetic moment $\bar{m}_s$ per atom, while critical temperature T_c and Curie-Weiss temperature Θ_c are given in K. For comparison also the magnetic moment m_{atom} of the corresponding isolated atoms is listed (the value in brackets results for orbital angular momentum quenching in the case of the transition metals).

simple paramagnetic behavior, with a susceptibility that obeys the so-called *Curie-Weiss law*, cf. Fig. 8.6,

$$\chi_o(T) \;\sim\; (T - \Theta_c)^{-1} \qquad (\text{for } \; T \gg T_c) \quad , \tag{8.57}$$

and Θ_c is called the Curie-Weiss temperature. Since γ in Eq. (8.55) is almost never exactly 1, Θ_c does not necessarily coincide with the Curie temperature T_c. Mostly, γ is of the order of unity, though, so that the two are still of the same order of magnitude. This can for example be seen from the values for some representative elemental ferromagnetic materials given in Table 8.2, where also the saturated values of the spontaneous magnetization at $T \to 0\,\text{K}$ are listed.

In theoretical studies one typically converts this measured $M_s(T \to 0\,\text{K})$ directly into an average magnetic moment $\bar{m}_s$ per atom (in units of μ_B) by simply dividing through the atomic density in the solid. In this form, one can compare directly with the corresponding value for the isolated atoms, obtained as $m_{\text{atom}} = g(JLS)\mu_B J$, cf. Eq. (8.33). As apparent from Table 8.2, the two values agree very well for the rare earth metals, but differ substantially for the transition metals. Recalling the discussion on the quenching of the orbital angular momentum of TM atoms immersed in insulating solids in section 8.3.4, one could again try to ascribe this difference to the symmetry lowering of the interacting d orbitals in the solid. Yet, even using $L = 0$ (and thus $J = S$, $g(JLS) = 2$), the agreement does not become much better, cf. Table 8.2. In the RE ferromagnets, the saturation magnetization appears therefore as a simple result of a perfect alignment of the magnetic moments connected to each atom in the solid, where these magnetic moments have retained pretty much their atomic character. The TM ferromagnets (Fe, Co, Ni), on the other hand, seem to have a much more complicated magnetic structure that is not well described by individual atomic properties.

As a first summary of our discussion on ferromagnetism (or ordered magnetic states in general) we are therefore led to conclude that it must arise out of an (yet unspecified) interaction between magnetic moments existing in a solid. This interaction effects an alignment of the magnetic moments, yielding a large net magnetization even in the absence of an external field. We are thus certainly far outside of the linear response regime that has served

as basis for the dia- and paramagnetic theories in the first sections of this chapter. Considering the alignment, the interaction has apparently the same effect as an applied field on a paramagnet, which explains somehow the similar form of the Curie-Weiss law compared to the Curie law of paramagnetic solids: Interaction and external field favor alignment, and are opposed by thermal disorder. Below the critical temperature the alignment is perfect, and far above T_c the competition between aligning and disordering tendencies is equivalent to the situation in a paramagnet (just the origin of the temperature scale has shifted).

Sources for the interacting magnetic moments can be either delocalized electrons (Pauli paramagnetism) or partially filled atomic shells (Paramagnetism). This makes it roughly plausible why ferromagnetism is a phenomenon specific to some metals, and in particular metals with a high magnetic moment due to only partially filled d or f shells (i.e., transition metals, rare earths, as well as their compounds). However, why only some, and not all of the TMs and REs exhibit ferromagnetic behavior is something that we can not yet rationalize. From the comparison of the saturated value of the spontaneous magnetization with the atomic magnetic moments, it seems that the ferromagnetism of RE metals can be understood on the basis of a coupling of localized magnetic moments (due to the inert partially filled f shells). The situation is more complicated for the TM ferromagnets, where the picture of a simple coupling between atomic-like moments fails. We will see below that this so-called *itinerant ferromagnetism* arises out of a subtle mixture of the delocalized s electrons and the more localized, but nevertheless not inert d orbitals.

8.4.3 Phenomenological Theories of Ferromagnetism

After this first overview of phenomena arising out of magnetic order, let us now try to establish a proper theoretical understanding. Unfortunately, the theory of ferromagnetism is one of the less well developed of the fundamental areas of solid state physics (a bit better for the localized ferromagnetism of the REs, worse for the itinerant ferromagnetism of the TMs). The underlying complex mixing of single electron and many-particle effects, as well as collective and local aspects makes it difficult to project the topic down onto simple models appropriate for an introductory level lecture. Correspondingly, we will not be able to present and discuss a full-blown *ab initio* theory like in the preceding chapters. Instead we will first consider simple phenomenological theories and concentrate afterwards only on some selected more refined treatments, which will enable us to address some of the fundamental questions not amenable to a phenomenological theory (like the involved material parameters).

Molecular (Mean) Field Theory

The simplest phenomenological theory of ferromagnetism is the *molecular field theory* due to Weiss (1906). We had seen in the preceding section that the effect of the presumed interaction between the discrete sources of magnetic moment in the solid is very similar to the one of an applied external field (leading to alignment). For each magnetic moment, the net effect of the interaction with all other moments can therefore be represented as a kind of internal field created by all other moments (and the magnetic moment itself). Obviously, this averages over all interactions and condenses them into one effective field,

i.e., the molecular field theory is a typical representative of a *mean field theory*, under which name it is also known. Since this effective internal, or so-called *molecular field* arises as the average result of all interactions, it is reasonable to assume that it scales somehow with M, i.e., the overall magnetization (density of magnetic moments!). Without knowing anything about the microscopic origin of the magnetic interaction and the resulting field, the ansatz of Weiss is then simply to postulate a linear scaling with M, leading to

$$B^{\mathrm{mol}} \;=\; \mu_o \lambda M \quad , \tag{8.58}$$

where λ is the so-called molecular field constant. For simplicity, we will consider here only the model case of an isotropic solid, so that internal and external field can go along the same direction, allowing us to write all formulae as scalar relations.

Since in this mean field theory the molecular field has an effect that is equivalent to (and thus indistinguishable from) an applied external field, the susceptibility will follow a Curie law above the critical temperature ($\chi^{\mathrm{mag,para}} = C/T$) and we arrive at

$$\begin{aligned} M \;=\; \frac{1}{\mu_o}\chi^{\mathrm{mag}} B \;&=\; \frac{1}{\mu_o}\chi^{\mathrm{mag}}(B^{\mathrm{ext}} + B^{\mathrm{mol}}) \\ &=\; \frac{1}{\mu_o}\frac{C}{T}(B^{\mathrm{ext}} + \mu_o \lambda M) \quad . \end{aligned} \tag{8.59}$$

Solving for the magnetization, this yields

$$M \;=\; \frac{1}{\mu_o}\,\frac{C}{(T - \lambda C)} B^{\mathrm{ext}} \quad , \tag{8.60}$$

and in turn for the susceptibility

$$\chi_o(T) \;=\; \mu_o \left.\frac{\partial M}{\partial B^{\mathrm{ext}}}\right|_{B^{\mathrm{ext}}=0} \;=\; \frac{C}{T - \lambda C} \qquad (\text{for } \; T > T_c) \quad . \tag{8.61}$$

Identifying $\lambda = \Theta_c/C$, we are therefore led to the Curie-Weiss law, i.e., in other words, the form of the Curie-Weiss law is compatible with an effective molecular field that scales linearly with the magnetization. Given the experimental observation of Curie-Weiss like scaling, the mean field ansatz of Weiss seems therefore reasonable. However, in detail we see already now that this simple theory fails, as it predicts the linear scaling for all $T > T_c$, i.e., within molecular field theory $\Theta_c = T_c$ and $\gamma = 1$ (this latter critical exponent is by the way characteristic for any mean field theory).

Given the qualitative plausibility of the theory, let us nevertheless use it to derive a first order of magnitude estimate for this (yet unspecified) molecular field. This phenomenological approach has thus the completely reverse strategy compared to an *ab initio* theory. In the latter, we would have analyzed the microscopic origin of the magnetic interaction and derived the molecular field as a suitable average, which in turn would have brought us into a position to predict materials properties like the saturation magnetization and critical temperature. Now, we do the reverse: We will use the experimentally known values of M_s and T_c to obtain a first estimate of the size of the molecular field, which might help us lateron to identify the microscopic origin of the magnetic interactions (about which we

still don't know anything). For the derivation, recall that the Curie constant C was given by Eq. (8.35) as

$$C \;=\; \frac{N\mu_o\mu_B^2 g(JLS)^2 J(J+1)}{3k_B V} \;\approx\; \left(\frac{N}{V}\right)\frac{\mu_o m_{\rm atom}^2}{3k_B} \quad , \tag{8.62}$$

where we have exploited that $J(J+1) \sim J^2$, allowing us to identify the atomic magnetic moment $m_{\rm atom} = \mu_B g(JLS)J$. With this, it is straightforward to arrive at the following estimate of the molecular field at $T = 0\,\mathrm{K}$,

$$B^{\rm mol}(0\,\mathrm{K}) \;=\; \mu_o\lambda M_s(0\,\mathrm{K}) \;= \mu_o\frac{T_c}{C}\left(\frac{N}{V}\bar{m}_s\right) \approx \frac{3k_B T_c \bar{m}_s}{m_{\rm atom}^2} \tag{8.63}$$

When discussing the content of Table 8.2 we had already seen that for most ferromagnets, $m_{\rm atom}$ is at least of the same order of $\bar{m}_s$, so that plugging in the numerical constants we arrive at the following order of magnitude estimate

$$B^{\rm mol} \;\sim\; \frac{[5T_c \text{ in K}]}{[m_{\rm atom} \text{ in } \mu_B]}\ \text{Tesla} \quad . \tag{8.64}$$

Looking at the values listed in Table 8.2, one realizes that molecular fields are of the order of some 10^3 Tesla, which is at least one order of magnitude more than the strongest magnetic fields that can currently be produced in the laboratory. As take-home message we therefore keep in mind that the magnetic interactions must be quite strong to yield such a high internal molecular field.

Molecular field theory can also be employed to understand the temperature behavior of the zero-field spontaneous magnetization. As shown in Fig. 8.6 $M_s(T)$ declines smoothly from its saturation value at $T = 0\,\mathrm{K}$ to zero at the critical temperature. Since the molecular field produced by the magnetic interaction is indistinguishable from an applied external one, the variation of $M_s(T)$ must be equivalent to the one of a paramagnet, just with the external field replaced by the molecular one. Using the results from section 8.3.2 for atomic paramagnetism, we therefore obtain directly a behavior equivalent to Eq. (8.35)

$$M_s(T) \;=\; M_s(0\,\mathrm{K})\; B_J\!\left(\frac{g(JLS)\mu_B B^{\rm mol}}{k_B T}\right) \quad , \tag{8.65}$$

i.e., the temperature dependence is entirely given by the Brillouin function $B_J(\eta)$ defined in Eq. (8.36). And, in fact, this function exhibits exactly the kind of qualitative behavior just mentioned: It declines smoothly to zero as sketched in Fig. 8.6. Similar to the reproduction of the Curie-Weiss law, mean field theory provides us therefore with a simple rationalization of the observed phenomenon, but fails again on a quantitative level (let alone that it does of course not yield the value of the material parameters). The functional dependence of $M_s(T)$ in both limits $T \to 0\,\mathrm{K}$ and $T \to T_c^-$ is incorrect, with e.g.

$$M_s(T \to T_c^-) \;\sim\; (T_c - T)^{1/2} \tag{8.66}$$

and thus $\beta = 1/2$ instead of the observed values around $1/3$. In the low temperature limit, the Brillouin function declines exponentially, which is also in gross disagreement to the measured deviation by an amount proportional to $T^{3/2}$ (known as *Bloch* $T^{3/2}$ *law*). The

failure to properly describe the critical exponents stems in general from the inability of any mean field theory to account for the fluctuations connected with a phase transition. The wrong low temperature limit, on the other hand, is due to the non-existence of a particular set of low-energy excitations called spin-waves (or magnons) in a mean field approach, and we will come back to this point below.

Heisenberg Hamiltonian

Another class of phenomenological theories is called *spin Hamiltonians*. Such theories are based on the perception that magnetic order arises out of the coupling between discrete microscopic magnetic moments. If these microscopic moments are for example connected with the partially filled f-shells of RE atoms, one can consider a discrete lattice model of the solid, where each atomic site i in the lattice exhibits a magnetic moment $\mathbf{m}_i$. A coupling between two different sites i and j in the lattice can then be described by a term

$$H_{i,j}^{\text{coupling}} \;=\; -\frac{J_{ij}}{\mu_B^2}\,\mathbf{m}_i \cdot \mathbf{m}_j \quad . \tag{8.67}$$

Here, J_{ij} is known as the exchange coupling constant between the two sites (unit: eV). Due to the scalar product of the two magnetic moment vectors, the coupling can take any value between $-J_{ij}m_i m_j/\mu_B^2$ (parallel alignment) and $+J_{ij}m_i m_j/\mu_B^2$ (antiparallel alignment) depending on the relative orientation of the two moment vectors. Note that in this form, there is therefore no absolute dependence on any spatial coordinate. If this is required (e.g. in a non-isotropic crystal with a preferred magnetization axis), additional terms need to be introduced to the Hamiltonian.

When generalizing to the whole solid, one first has to decide between which lattice sites a coupling should be present. Recall that this is a phenomenological theory, so that there is nothing that would tell us which interactions are important. Instead, we have to simply choose some interactions (e.g. between nearest, or up to next-nearest neighbors) and leave their strengths J_{ij} as unspecified variables. When one then calculates through the model and sees what observables it yields, one can employ this to fit the J_{ij} to experimental data and thereby hope to extract some first information about the order of magnitude of the latter microscopic quantity, and of course, whether the choice of couplings was reasonable at all. This choice is commonly guided by some model conceptions about the microscopic origin of the coupling. For the localized ferromagnetism of the rare earths one typically assumes a quite short-ranged interaction, which could either be between directly neighboring atoms (*direct exchange*), or sometimes across non-magnetic atoms in more complex structures (*superexchange*). As illustrated in Fig. 8.7 this conveys somehow the image that the interaction arises out of an overlap of electronic wavefunctions. However, alternatively, one could also think of an interaction mediated by the glue of conduction electrons (*indirect exchange*), cf. Fig. 8.7c, but without a proper microscopic theory one cannot distinguish between these possibilities.

Having decided on the couplings, the total Hamiltonian describing the whole solid may in its simplest form assumed to be the mere pairwise sum of all the interactions between

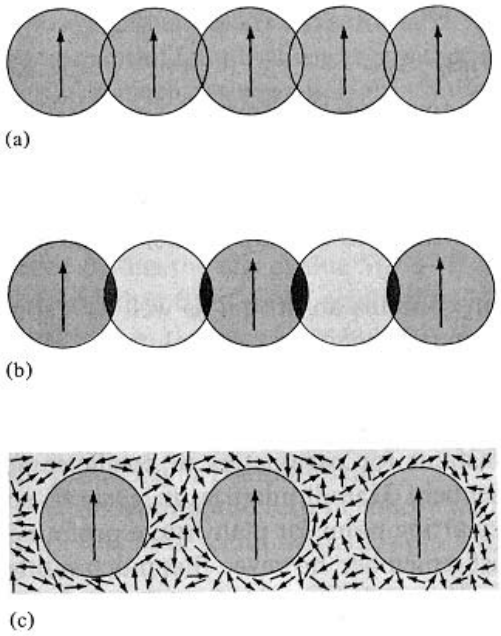

Figure 8.7: Schematic illustrations of possible coupling mechanisms between localized magnetic moments: (a) direct exchange between neighboring atoms, (b) superexchange mediated by non-magnetic ions, and (c) indirect exchange mediated by interactions with the conduction electron glue.

different sites in the lattice

$$H^{\text{Heisenberg}} = \sum_{i,j=1}^{M} H_{i,j}^{\text{coupling}} = -\sum_{i,j=1}^{M} \frac{J_{ij}}{\mu_B^2} \mathbf{m}_i \cdot \mathbf{m}_j \quad . \tag{8.68}$$

Similar to the discussion in chapter 6 on cohesion, such a pairwise summation is only justified, when the interaction between two sites is not altered appreciably by the presence of the magnetic moments at other sites nearby. Since we don't know anything about the microscopic origin of the interaction, this is quite hard to validate. In general, we may expect that it could be necessary that all magnetic ions are far enough apart from each other (so that e.g. the overlap of their electronic wavefunctions is small). As already stated, spin Hamiltonians are therefore more suitable for the description of the ferromagnetism of the rare earths with their inert partially filled f-shells.

The particular spin Hamiltonian of Eq. (8.68) is known as *Heisenberg Hamiltonian* and is the starting point for many theories of ferromagnetism. Although it appears quite easy, it is in reality quite complex to solve. This was particularly so in the past, when one could only try to tackle the problem analytically. With the real Heisenberg Hamiltonian defying such an analytical solution, a range of even further simplified spin Hamiltonians were first addressed alternatively. Examples of this are one- or two-dimensional Heisenberg Hamiltonians (possibly only with nearest neighbor interaction) or the *Ising model*, in which no longer a continuous relative orientation between two magnetic moments is considered, but only the completely parallel and antiparallel alignment. These models opened up a whole field of research of its own in the statistical mechanics community, that became more and

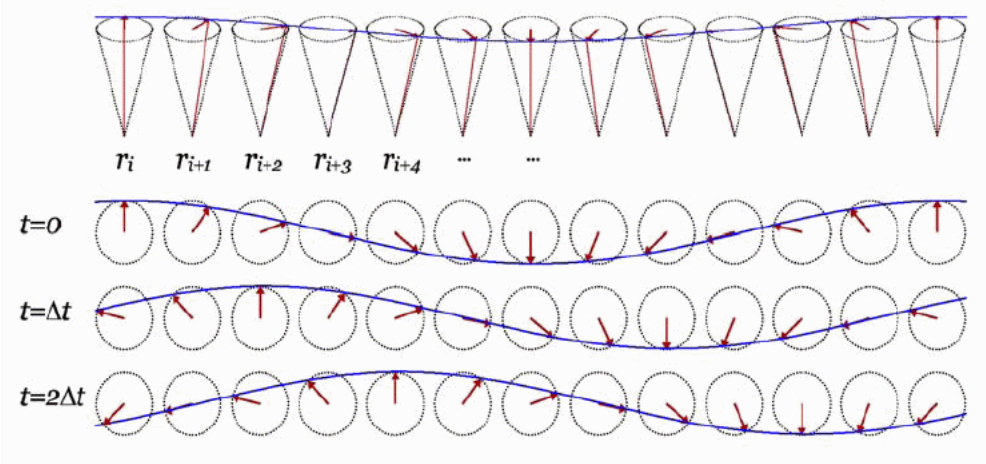

Figure 8.8: Schematic representation of the directional alignment of the magnetic moments in a one-dimensional spin-wave.

more detached from the original magnetic problem, but was and is more driven by the curiosity to solve the generic models themselves. In the meanwhile even the real Heisenberg Hamiltonian can be very efficiently tackled with powerful numerical techniques like Monte Carlo simulations, although they can of course not reach the beauty of a proper analytical solution.

If we concentrate on the case of an elemental solid, where all atoms are equivalent and exhibit the same magnetic moment $\mathbf{m}$, inspection of Eq. (8.68) reveals that $J > 0$ describes ferromagnetic and $J < 0$ antiferromagnetic coupling. For a ferromagnet, the ground state is hereby obviously given by a perfect alignment of all moments in one direction $|\uparrow\uparrow\uparrow \ldots \uparrow\uparrow>$. One is inclined to believe that the lowest energy excitations in such a system would correspond to a simple spin flip $|\uparrow\downarrow\uparrow \ldots \uparrow\uparrow>$. Surprisingly, the understanding arising out of the detailed study of Heisenberg Hamiltonians is that not only are such states no eigenstates to the real problem, but that the proper low energy excitations have a much more complicated wave-like form, with only slight directional changes between directly neighboring magnetic moments as illustrated in Fig. 8.8. At second thought it is then intuitive, that such waves will have only an infinitesimal higher energy than the ground state, if only their wave vector is infinitely long. Such spin-waves are called *magnons*, and one can indeed show that they are what was missing in the previous mean field approach, i.e., contrary to molecular field theory the Heisenberg Hamiltonian properly accounts for the Bloch $T^{3/2}$-scaling of $M_s(T)$ for $T \to 0\,\mathrm{K}$. Such spin wave solutions can in fact be expected quite generally, whenever there is a direction associated with the local order that can vary spatially in a continuous manner at a cost in energy that becomes small as the wavelength of the variation becomes very long. Magnonic solutions can therefore neither exist in a mean field approach nor in the discrete Ising model, but should be a characteristic property of real systems. The existence of spin-waves and properties that are very well accounted for by the Heisenberg model has indeed been experimentally confirmed by inelastic scattering experiments with spin-polarized neutrons.

The Heisenberg model does also a very good job in accounting for other characteris-

tic properties of ferromagnets. The high-temperature susceptibility follows a Curie-Weiss type law, and upon approaching the critical temperature the $1/T$-scaling receives gradually increasing corrections. Correspondingly, Monte Carlo simulations of Heisenberg ferromagnets yield critical exponents that are in very good agreement with the experimental data. One can exploit the data on susceptibility or spontaneous magnetization, in order to fit the exchange constant J_{ij}. The values obtained for nearest neighbor interaction are typically of the order of 10^{-2} eV for the TM ferromagnets and $10^{-3} - 10^{-4}$ eV for the RE ferromagnets. These quite high values (compared to the energy splittings in dia- and paramagnetism) give us already a first impression that the microscopic origin of ferromagnetic coupling must be rather strong.

This can equally be obtained by deriving the above discussed molecular field theory as an approximation to the Heisenberg model. This is achieved by rewriting Eq. (8.68) in the following way

$$\begin{aligned} H^{\text{Heisenberg}} &= -\sum_{i,j=1}^{M} \frac{J_{ij}}{\mu_B^2}\, \mathbf{m}_i \cdot \mathbf{m}_j \\ &= -\sum_{i=1}^{M} \mathbf{m}_i \cdot \left(\sum_{j=1}^{M} \frac{J_{ij}}{\mu_B^2} \mathbf{m}_j \right) = -\sum_{i=1}^{M} \mathbf{m}_i \cdot \mathbf{B}^{\text{mol}} \quad . \end{aligned} \tag{8.69}$$

In the last step, we have realized that the term in brackets has the units of Tesla, and can therefore be perceived as a magnetic field $\mathbf{B}^{\text{mol}}$ arising from the coupling of moment i with all other moments in the solid. However, up to now we have not gained anything by this rewriting, because this effective magnetic field is different for any state (i.e., directional orientation of all moments). The situation can only be simplified, by assuming a mean field point of view and approximating this effective field by its thermal average $< \mathbf{B}^{\text{mol}} >$. For this one obtains

$$< \mathbf{B}^{\text{mol}} > = \sum_{j=1}^{M} \frac{J_{ij}}{\mu_B^2} \; < \mathbf{m}_j > = \frac{\sum_{j=1}^{M} J_{ij}}{\mu_B^2} \left(\frac{N}{V} \right) \mathbf{M}_s(T) \; \sim \; \lambda \mathbf{M}_s(T) \quad , \tag{8.70}$$

since the mean magnetic moment at site i in an elemental crystal is simply given by the macroscopic magnetization (dipole moment density) divided by the density of sites. As apparent from Eq. (8.70) we arrive therefore indeed at a molecular mean field that is linear in the magnetization as assumed by Weiss in his molecular field theory. In addition to that, we can now even relate the molecular field to the more microscopic quantity J_{ij}. If we assume for simplicity, only nearest neighbor interaction in a fcc metal, this relation follows from Eq. (8.70) as

$$J \; = \; J_{ij}^{NN} \; = \; \frac{\mu_B^2 B^{\text{mol}}}{12 \bar{m}_s} \quad . \tag{8.71}$$

Using the saturation magnetizations listed in Table 8.2 and the estimate of molecular fields of the order of 10^3 Tesla obtained above, one obtains precisely exchange constants of the order of 10^{-2} eV for the TM or 10^{-3}–10^{-4} eV for the RE ferromagnets as mentioned before.

8.4.4 Microscopic Origin of Magnetic Interaction

Having reproduced a large fraction of the properties characteristic for ferromagnetism with the phenomenological models, the crucial insight that we are now still lacking and that must come from an *ab initio* theory is the microscopic mechanism that can lead to the derived rather high values of J, i.e., the origin of the quite strong magnetic interaction. The natural choice somehow also (falsely) conveyed by the name "magnetic interaction" is to consider the direct dipolar interaction energy of two magnetic dipoles $\mathbf{m}_1$ and $\mathbf{m}_2$, separated by a distance $\mathbf{r}$. For this macroscopic electrodynamics gives us

$$J^{\text{mag dipole}} = \frac{1}{r^3}\left[\mathbf{m}_1 \cdot \mathbf{m}_2 - 3(\mathbf{m}_1 \cdot \hat{\mathbf{r}})(\mathbf{m}_2 \cdot \hat{\mathbf{r}})\right] \quad , \tag{8.72}$$

Glossing over the angular dependence, this gets simplified by assuming the two dipoles to be oriented parallel to each other and perpendicular to the vector between them,

$$J^{\text{mag dipole}} \approx \frac{1}{r^3}\, m_1 m_2 \quad . \tag{8.73}$$

Inserting magnetic moments of the order of μ_B, cf. Table 8.2, and a typical distance of 2 Å between atoms in solids, this yields $J^{\text{mag dipole}} \sim 10^{-4}$ eV, which is thus 1-2 orders of magnitude too small for the exchange constants determined via the phenomenological models.

Having ruled out magnetic dipole interaction as source for the coupling of magnetic moments in ferromagnets, one is only left with electrostatic (Coulomb) interactions. And these are in fact the reason behind the coupling, i.e., to be more precise it is them together with electronic interaction due to the *Pauli exclusion principle*. Loosely speaking, this exclusion principle keeps electrons with parallel spins apart, and so reduces their Coulomb repulsion. The resulting difference in energy between the parallel spin configuration and the antiparallel one is the sought *exchange energy*. This is favorable for ferromagnetism only in exceptional circumstances because the increase in kinetic energy associated with parallel spins usually outweighs the favorable decrease in potential energy. The difficulty behind understanding magnetic ordering lies therefore in the subtlety that one needs to go beyond the independent electron (one particle) picture to grasp the reason for the interaction between the individual magnetic moments. In particular because this point is quite intricate, let us further illustrate how the Pauli principle and Coulomb interaction can lead to magnetic effects by looking at two simple and opposite model cases: two localized magnetic moments and the free electron gas.

Exchange Interaction between two Localized Moments

The most simple realization of two localized magnetic moments is the familiar problem of the hydrogen molecule. Here the two magnetic moments are given by the spins of the two electrons (1) and (2), which belong to the two H ions A and B. If there was only one ion and one electron, we would simply face the hydrogen atom Hamiltonian h_o,

$$h_o|\phi> = \epsilon_o|\phi> \quad , \tag{8.74}$$

with the ground state solution $|\phi> = |\mathbf{r}, \sigma>$ depending on the position $\mathbf{r}$ and spin σ of the orbiting electron. However, in case of the molecule we also have to consider the

interaction between the two electrons and two ions themselves, as well as between ions and electrons. This leads to the new Hamiltonian

$$H|\Phi> = (h_o(A) + h_o(B) + h_{\rm int})|\Phi> \quad , \tag{8.75}$$

where the two-electron wave function $|\Phi>$ is now a function of the positions of the ions A and B, the two electrons (1) and (2), and their two spins σ_1, σ_2. Since H has no explicit spin dependence, all spin operators commute with H and the total two-electron wave function must separate into the form

$$|\Phi> = |\Psi_{\rm orb}>|\chi_{\rm spin}> \quad , \tag{8.76}$$

i.e., an orbital part and a pure spin part. As a suitable basis for the latter, we can choose the eigenfunctions of $\mathbf{S^2}$ and S_z, so that the spin part of the wave function is given by

$$\begin{array}{lll} |\chi_{\rm spin}>_{\rm S} = 2^{-1/2}\,(|\uparrow\downarrow> - |\downarrow\uparrow>) & S=0 & S_z=0 \\ |\chi_{\rm spin}>_{\rm T1} = |\uparrow\uparrow> & S=1 & S_z=1 \\ |\chi_{\rm spin}>_{\rm T2} = 2^{-1/2}\,(|\uparrow\downarrow> + |\downarrow\uparrow>) & S=1 & S_z=0 \\ |\chi_{\rm spin}>_{\rm T3} = |\downarrow\downarrow> & S=1 & S_z=-1 \quad . \end{array}$$

The Fermion character of the two electrons (and the Pauli exclusion principle) manifests itself at this stage in the general postulate that the two-electron wave function must change sign when interchanging the two indistinguishable electrons. Since $|\chi_{\rm spin}>_{\rm S}$ changes sign upon interchanging the two electron spins, while the three triplet wave functions $|\chi_{\rm spin}>_{\rm T}$ do not, only the following combinations with the orbital wave function part are then possible

$$\begin{aligned} |\Phi>_{\rm singlet} &= |\Psi_{\rm orb,sym}> \ |\chi_{\rm S}> \\ |\Phi>_{\rm triplet} &= |\Psi_{\rm orb,asym}> \ |\chi_{\rm T}> \quad , \end{aligned} \tag{8.77}$$

where $|\Psi_{\rm orb,sym}(1,2)> = |\Psi_{\rm orb,sym}(2,1)>$ and $|\Psi_{\rm orb,asym}(1,2)> = -|\Psi_{\rm orb,asym}(2,1)>$. Despite the spin-independent Hamiltonian, there is thus a strict correlation between the spatial symmetry of the orbital wave function and the total spin of the system state.

An obvious solution to the H_2 problem is given, when the two ions are infinitely far apart. In this case $<\Phi|h_{\rm int}|\Phi> = 0$, and the total wave function must separate into a product of the one-electron wave functions $|\phi>$ representing the solution to the aforementioned isolated H atom problem. Maintaining the symmetry requirement of Eq. (8.77), the only two suitably symmetrized possibilities are then

$$\begin{aligned} |\Psi^{\infty}_{\rm orb,sym}> &= 2^{-1/2}\,[\,|\phi(1A)>|\phi(2B)> + |\phi(2A)>|\phi(1B)>\,] \\ |\Psi^{\infty}_{\rm orb,asym}> &= 2^{-1/2}\,[\,|\phi(1A)>|\phi(2B)> - |\phi(2A)>|\phi(1B)>\,] \quad , \end{aligned} \tag{8.78}$$

where $|\phi(1A)>$ denotes for example that electron (1) is in orbit around ion A. These two wave functions correspond to the ground state of H_2 in either a spin singlet or a spin triplet configuration, and it is straightforward to verify that both states exhibit the degenerate energy $2\epsilon_o$.

Having analyzed the problem in this (academic) limit, let us consider what happens if we bring the two atoms closer together. At some distance, the two electron densities will

slowly start to overlap. Since the atomic wave functions $|\phi(1A)>$ and $|\phi(1B)>$ (and equivalently $|\phi(2A)>$ and $|\phi(2B)>$ for the other electron) were determined for the separate atoms, they are not necessarily orthogonal to each other, and the wave function overlap gives then rise to a non-vanishing so-called *overlap integral*

$$S = <\phi(1A)|\phi(1B)><\phi(2A)|\phi(2B)> = |<\phi(1A)|\phi(1B)>|^2 \quad . \tag{8.79}$$

If we restrict ourselves to the onset of overlap at still quite a large distance, it is reasonable to assume that despite the non-vanishing overlap integral the delocalization of the two electrons is not yet appreciable enough to conceive system states where both electrons are located at the same ion. In this *Heitler-London approximation*, the solutions to the two-electron problem are then still of the form as Eq. (8.78), just the normalization has changed due to the non-vanishing overlap integral

$$\begin{aligned} |\Psi^{\rm HL}_{\rm orb,sym}> &= (2+2S)^{-1/2}\,[\,|\phi(1A)>|\phi(2B)> + |\phi(2A)>|\phi(1B)>\,] \quad (8.80)\\ |\Psi^{\rm HL}_{\rm orb,asym}> &= (2-2S)^{-1/2}\,[\,|\phi(1A)>|\phi(2B)> - |\phi(2A)>|\phi(1B)>\,] \quad . \end{aligned}$$

Due to this change, the singlet ground state energy $E_{\rm S} = <\Phi|H|\Phi>_{\rm singlet}$ and the triplet ground state energy $E_{\rm T} = <\Phi|H|\Phi>_{\rm triplet}$ begin to deviate, and one obtains

$$E_{\rm T} - E_{\rm S} = 2\,\frac{CS-A}{1-S^2} \quad , \tag{8.81}$$

where

$$\begin{aligned} C &= <\phi(1A)|<\phi(2B)|\,H_{\rm int}\,|\phi(2B)>|\phi(1A)> \quad \text{(Coulomb integral)} \quad (8.82)\\ A &= <\phi(1A)|<\phi(2B)|\,H_{\rm int}\,|\phi(2A)>|\phi(1B)> \quad \text{(Exchange integral)}\,. \quad (8.83) \end{aligned}$$

C is therefore simply the Coulomb interaction of the two electrons with the two ions and themselves, while A arises out of the exchange of the indistinguishable particles. If this was possible without any sign change (e.g. for two bosons), then $A = CS$ and the singlet-triplet energy splitting would vanish. However, it is precisely the Fermion character of the two electrons that via the Pauli exclusion principle yields $A \neq CS$ and in turn a finite energy difference between singlet ground state (antiparallel spin alignment) and triplet ground state (parallel spin alignment) of the H_2 molecule. In fact, the triplet ground state results here as less favorable, because it is combined with the asymmetric orbital wave function: The latter must have a node exactly midway between the two atoms, meaning that there is no electron density at precisely that point where they could screen the Coulomb repulsion between the two atoms most efficiently. The coupling between the two localized magnetic moments is therefore due to electrostatic interaction mixed with the Pauli exclusion principle: What matters for the alignment is the effort by the electrons to choose a spin state that minimizes the Coulomb repulsion between them once Fermi statistics are taken into account. For this reason, magnetic interaction is also often denoted as *exchange interaction* (and one talks about an *exchange splitting* of the energy levels).

Evaluating the two integrals C and A in Eq. (8.83) using the one-electron wave functions of the H atom problem, the energy splitting $(E_{\rm T} - E_{\rm S})$ can be determined. If one realizes that the two level problem of the H_2 molecule can be equally described by the Heisenberg

Hamiltonian of Eq. (8.68), one has therefore found a means to determine the exchange constant J entering the latter. The Heitler-London approximation sketched in this section can be generalized to more than two magnetic moments, and is then the gateway to computing the exchange constants for ensuing Heisenberg Hamiltonian based studies of the (anti)ferromagnetism of solids. However, it works only in the very limit of highly localized magnetic moments, where the wave function overlap is just good enough to provide an exchange interaction, but not yet highly delocalized electrons (for which the wave function ansatz of Eq. (8.80) is unreasonable). With this understanding of the interaction of localized moments, also the frequently successful use of only nearest-neighbor interaction in Heisenberg Hamiltonians becomes plausible. As a note aside, in the extreme limit of a highly screened Coulomb interaction that only operates at each lattice site itself (and doesnŠt even penetrate to the nearest neighbor lattice sites anymore), magnetic behavior can also be described by a so-called *Hubbard model*, which discusses electron hopping on a discrete lattice and a repulsive interaction U once two electrons of opposite spin happen to occupy the same site.

Exchange Interaction in the Homogeneous Electron Gas

The conclusion from the localized moment model is thus that feeding Heitler-London derived exchange constants into Hubbard or Heisenberg models provides a good semi-quantitative basis to describe the ferromagnetism of localized moments as is typical for the rare earth solids. The other extreme of the more *itinerant ferromagnetism* of the transition metals can not be described with it, though. For this, the more appropriate picture is that of delocalized electrons, the exchange interaction of which is conceptually most amenably studied for the homogeneous electron gas, i.e., treating the ions as a jellium background. We had looked into this model already in section 8.3.3 in the independent electron approximation (zero electron-electron interaction) and found that the favorable alignment of the spins in an external field is opposed by the increased kinetic energy connected with the necessary occupation of energetically higher lying states. Due to this competition the net effect was weak and since there is no driving force for a spin alignment without applied field only a (Pauli) paramagnetic behavior was obtained.

If electron-electron interaction is considered, on the other hand, there is a new source for alignment due to the exchange interaction, i.e., the exchange-hole mediated effect to reduce the repulsion of parallel spin states. This new contribution supports the ordering tendency of an applied field, and could (if strong enough) maybe even favor ordering in the absence of an applied field. To estimate whether the latter is possible, we consider the electron gas in the Hartree-Fock approximation as the minimal theory accounting for exchange interaction. In the chapter on cohesion (and in exercise 19) the energy per electron was for this approximation already obtained without explicitly considering spins

$$(E/N)_{\text{jellium}} \;=\; T_s \;+\; E^{\text{XC}} \;\overset{HF}{\approx}\; \frac{30.1\,\text{eV}}{\left(\frac{r_s}{a_B}\right)^2} \;-\; \frac{12.5\,\text{eV}}{\left(\frac{r_s}{a_B}\right)} \quad , \tag{8.84}$$

i.e., compared to the independent electron model of section 8.3.3, which only contains the first kinetic energy term, a second (attractive) term arises from exchange. To discuss magnetic behavior we now have to extend this description to the spin-polarized case. This

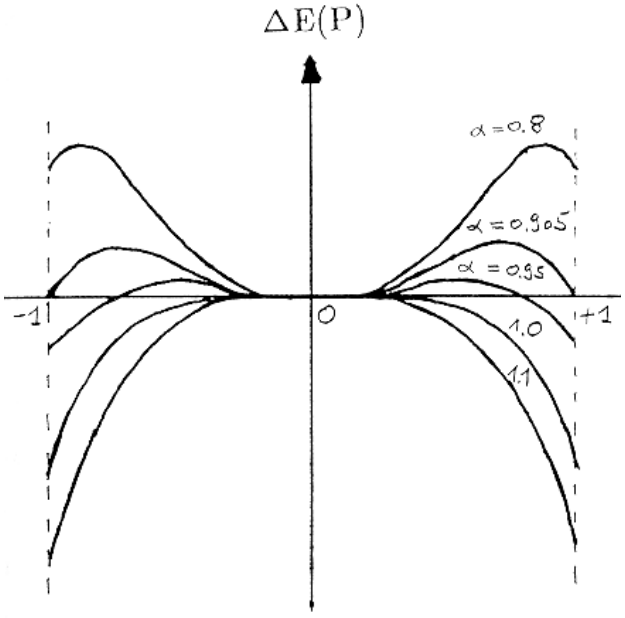

Figure 8.9: Plot of $\Delta E(P)$, cf. Eq. (8.89), as a function of the polarization P and for various values of α. A ferromagnetic solution is found for $\alpha > 0.905$.

is done in most straightforward manner by rewriting the last expression as a function of the electron density (exploiting the definition of the Wigner-Seitz radius as $r_s/a_B = \left(\frac{3}{4\pi}(V/N)\right)^{1/3}$)

$$E_{\text{jellium,HF}}(N) \;=\; N\left\{78.2\,\text{eV}\left(\frac{N}{V}\right)^{2/3} \;-\; 20.1\,\text{eV}\left(\frac{N}{V}\right)^{1/3}\right\} \quad . \tag{8.85}$$

With this, the obvious generalization to a spin-resolved gas is

$$\begin{aligned} E^{\text{spin}}_{\text{jellium,HF}}(N^\uparrow, N^\downarrow) \;&=\; E_{\text{jellium,HF}}(N^\uparrow) \;+\; E_{\text{jellium,HF}}(N^\downarrow) \;= \\ =\; &\left\{78.2\,\text{eV}\left[\left(\frac{N^\uparrow}{V}\right)^{2/3} + \left(\frac{N^\downarrow}{V}\right)^{2/3}\right] \;-\; 20.1\,\text{eV}\left[\left(\frac{N^\uparrow}{V}\right)^{1/3} + \left(\frac{N^\downarrow}{V}\right)^{1/3}\right]\right\} \quad , \end{aligned} \tag{8.86}$$

where $N^\uparrow$ is the number of electrons with spin up, and $N^\downarrow$ the number of electrons with spin down ($N^\uparrow + N^\downarrow = N$). More elegantly, the effects of spin-polarization can be discussed by introducing the polarization

$$P \;=\; \frac{N^\uparrow - N^\downarrow}{N} \quad , \tag{8.87}$$

which possesses the obvious limits $P = \pm 1$ for complete spin alignment and $P = 0$ for a non-spinpolarized gas. Exploiting the relations $N^\uparrow = \frac{N}{2}(1+P)$ and $N^\downarrow = \frac{N}{2}(1-P)$, Eq. (8.86) can be cast into the form

$$\begin{aligned} &E_{\text{jellium,HF}}(N, P) \;= \\ &NT\left\{\frac{1}{2}\left[(1+P)^{5/3} + (1-P)^{5/3}\right] \;-\; \frac{5}{4}\alpha\left[(1+P)^{4/3} + (1-P)^{4/3}\right]\right\} \quad , \end{aligned} \tag{8.88}$$

with $\alpha = 0.10(V/N)^{1/3}$. Ferromagnetism, i.e., a non-vanishing spontaneous magnetization at zero field, can occur when a system configuration with $P \neq 0$ leads to a lower energy than the non-polarized case, $\Delta E(N,P) = E(N,P) - E(N,0) < 0$. Plugging in the expression for the total energy in the Hartree-Fock approximation we thus arrive at

$$\Delta E(N,P) = NT\left\{\frac{1}{2}\left[(1+P)^{5/3} + (1-P)^{5/3} - 2\right] - \frac{5}{4}\alpha\left[(1+P)^{4/3} + (1-P)^{4/3} - 2\right]\right\}. \quad (8.89)$$

Figure 8.9 shows this function in dependence of the polarization P for various α. Apparently, the condition $\Delta E(N,P) < 0$ can only be fulfilled for

$$\alpha > \alpha_c = \frac{2}{5}(2^{1/3}+1) \approx 0.905 \quad , \quad (8.90)$$

in which case the lowest energy state is then always given by a completely polarized gas ($P = 1$). Using the definition of α, this condition can be converted into a maximum electron density below which the gas possesses a ferromagnetic solution. In units of the Wigner-Seitz radius this is

$$\left(\frac{r_s}{a_B}\right) \gtrsim 5.45 \quad , \quad (8.91)$$

i.e., only electron gases with a density that is low enough to yield Wigner-Seitz radii above $5.6\,a_B$ would be ferromagnetic. Recalling that $1.8 < r_s < 5.6$ for real metals, only Cs would just result as a ferromagnet in this Hartree-Fock approximation.

Already this is obviously not in agreement with reality, where Cs is paramagnetic and Fe/Co/Ni are ferromagnetic. However, the situation becomes even worse, when we start to improve the theoretical description by considering further electron exchange-correlation effects beyond Hartree-Fock. The prediction of the paramagnetic-ferromagnetic phase transition has a long history in homogeneous electron gas theory, and the results vary dramatically with each higher level of exchange-correlation included. The maximum density below which ferromagnetism can occur is, however, consistently obtained much lower as in Eq. (8.91), and the currently most quantitative work predicts ferromagnetism only for electron gases with densities below $(r_s/a_B) \approx 50 \pm 2$ [F.H. Zhong, C. Lin, and D.M. Ceperley, Phys. Rev. B **66**, 036703 (2002)]. Although Hartree-Fock theory allows us therefore to qualitatively understand the reason behind itinerant ferromagnetism of delocalized electrons, it overestimates the effect of the exchange interaction by one order of magnitude in r_s. Yet, even the more refined treatment of electron exchange and correlation does not yield a good description of the itinerant magnetism of real metals, none of which (with $1.8 < r_s < 5.6$) should be ferromagnetic at all. The reason for this lies now not in still present deficiencies of the exchange-correlation description, but in the jellium-type model itself: Completely delocalized electrons as in simple metals can not lead to ferromagnetism, this is correctly obtained. That Fe/Co/Ni do exhibit ferromagnetic behavior results primarily from their partially filled *d*-electron shell, and the latter is not well described by a free electron model as we had already seen in the chapter on cohesion. In order to finally understand the itinerant ferromagnetism of the transition metals, we therefore need to combine the exchange interaction obtained by looking at the jellium model with something that we have not yet considered at all: band structure effects.

8.4.5 Band Consideration of Ferromagnetism

In chapter 3.7 we had seen that an efficient treatment of electron correlation together with explicit consideration of the ionic positions is possible within density-functional theory (DFT). Here, we can write for the total energy of the problem

$$E/N \;=\; T_s[n] + E^{\mathrm{ion-ion}}[n] + E^{\mathrm{el-ion}}[n] + E^{\mathrm{Hartree}}[n] + E^{\mathrm{XC}}[n] \quad , \tag{8.92}$$

where all terms are uniquely determined functions of the electron density $n(\mathbf{r})$. The most frequently employed approximation to the exchange-correlation term is the local density approximation (LDA), which gives the XC-energy at each point $\mathbf{r}$ as the one of the homogeneous electron gas with equivalent density $n(\mathbf{r})$. Recalling our procedure of the last section, it is straightforward to generalize this to the so-called spin-polarized DFT, with

$$E/N \;=\; T_s[n^{\uparrow}, n^{\downarrow}] + E^{\mathrm{ion-ion}}[n] + E^{\mathrm{el-ion}}[n] + E^{\mathrm{Hartree}}[n] + E^{\mathrm{XC}}[n^{\uparrow}, n^{\downarrow}] \quad . \tag{8.93}$$

$n^{\uparrow}(\mathbf{r})$ and $n^{\downarrow}(\mathbf{r})$ are then the spin up and down electron densities, respectively. It is only necessary to consider them separately in the kinetic energy term (since there might be a different number of up- and down- electrons) and in the exchange-correlation term. For the latter, one proceeds in an analog manner as before and obtains it in the *local spin density approximation* (LSDA) from the XC-energy of the spin-polarized homogeneous electron gas that we have just discussed before. Introducing the magnetization density

$$m(\mathbf{r}) \;=\; \left(n^{\uparrow}(\mathbf{r}) - n^{\downarrow}(\mathbf{r})\right)\mu_B \quad , \tag{8.94}$$

one can write the total energy as a pure function of $n(\mathbf{r})$ and $m(\mathbf{r})$. Since the integral of the magnetization density over the complete unit cell yields the total magnetic moment, self-consistent solutions of the coupled Kohn-Sham equations can then either be obtained under the constraint of a specified magnetic moment (*fixed spin moment (FSM) calculations*), or directly yielding that magnetic moment which minimizes the total energy of the system.

Already in the LSDA this approach is very successful and yields indeed a non-vanishing magnetic moment for Fe, Co, and Ni. Resulting spin-resolved density of states (DOS) for Fe and Ni are exemplified in Fig. 8.10. After self-consistency is reached, these three ferromagnetic metals exhibit a larger number of electrons in one spin direction (called majority spin) than in the other (called minority spin). Due to this the effective potential seen by up and down electrons is different, and correspondingly the Kohn-Sham eigenlevels are obtained differently (which upon filling the states up to a common Fermi-level yields the different number of up and down electrons that we had started with). Quantitatively one obtains in the LSDA the following average magnetic moments per atom, $\bar{m}_s = 2.1\ \mu_B$ (Fe), $1.6\ \mu_B$ (Co), and $0.6\ \mu_B$ (Ni), which compare excellently with the experimental saturated spontaneous magnetizations at $T = 0\,\mathrm{K}$ listed in Table 8.2. We can therefore state that DFT-LSDA reproduces quantitatively the itinerant ferromagnetism of the elemental transition metals. This is nice, but at this stage somewhat unsatisfying, because we still donŠt *understand* why it is just these three elements for which the exchange interaction is large enough to outweigh the increased kinetic energy and give rise to a net magnetic

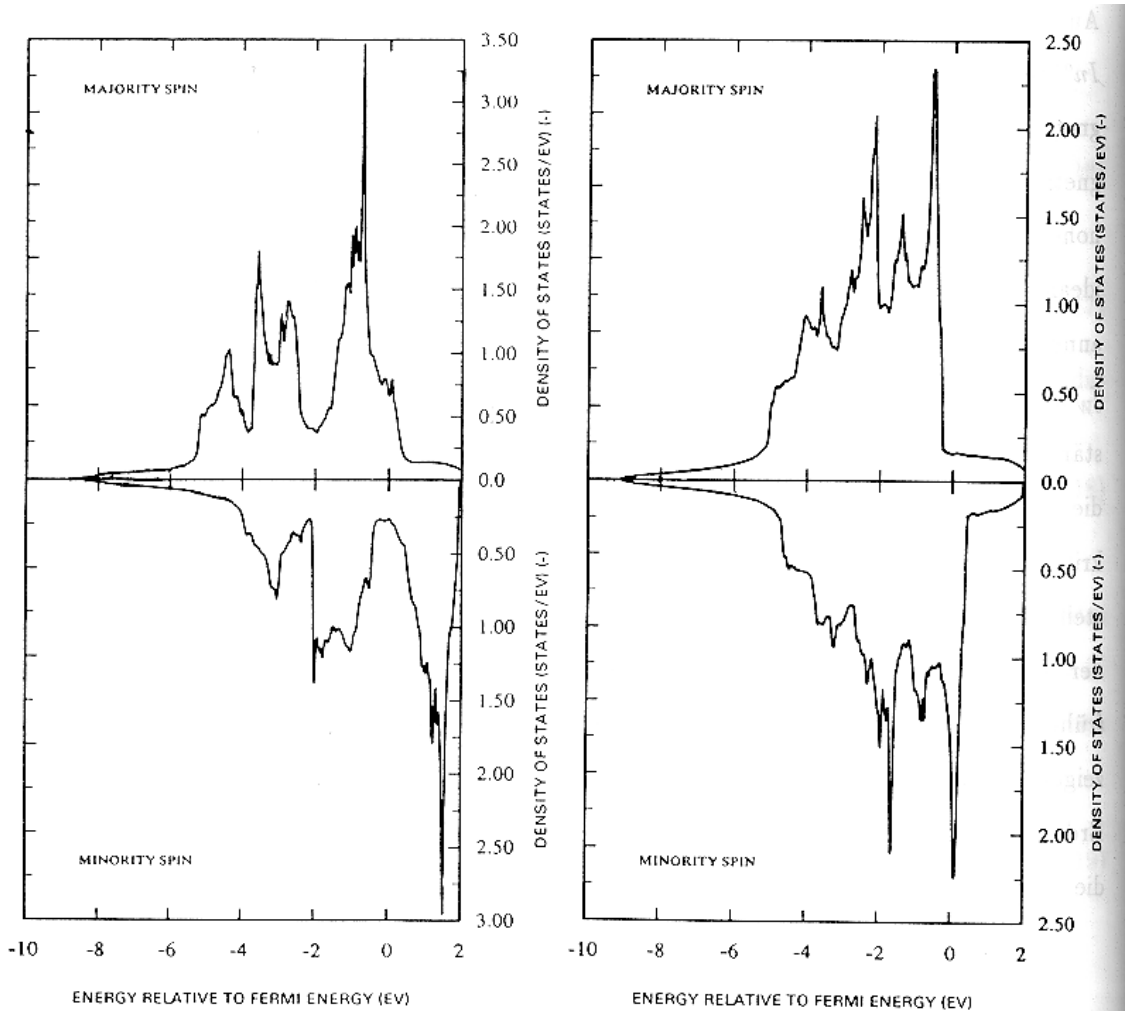

Figure 8.10: Spin-resolved density of states (DOS) for bulk Fe and Ni in DFT-LSDA [from V.L. Moruzzi, J.F. Janak, and A.R. Williams, *Calculated electronic properties of metals*, Pergamon Press (1978)].

moment.

Stoner Model of Itinerant Ferromagnetism

For such an understanding, we notice after a closer look that the DOS of majority and minority spins in Fig. 8.10 is different, but mostly due to a kind of constant shift between both channels. There are some changes in the precise structure of the DOS, but to first order it is really only such a constant shift that distinguishes the eigenstates of up and down electrons. This can be rationalized by realizing that the magnetization density (i.e., difference between up and down electron density) is significantly smaller than the total electron density itself. For the XC-correlation potential one can therefore write a Taylor expansion to first order in the magnetization density as

$$V_{\rm XC}^{\uparrow\downarrow}(\mathbf{r}) \;=\; \frac{\delta E_{\rm XC}^{\uparrow\downarrow}[n(\mathbf{r}),m(\mathbf{r})]}{\delta n(\mathbf{r})} \;\approx\; V_{\rm XC}^{o}(\mathbf{r}) \pm \tilde{V}(\mathbf{r}) m(\mathbf{r}) \quad , \tag{8.95}$$

where $V^o_{\rm XC}(\mathbf{r})$ is the XC-potential in the non-magnetic case. If the electron density is only slowly varying, one may assume the difference in up and down XC-potential (i.e., the term $\tilde{V}(\mathbf{r})m(\mathbf{r})$) to also vary slowly. In the *Stoner model*, one approximates it therefore with a constant term given by

$$V^{\uparrow\downarrow}_{\rm XC,Stoner}(\mathbf{r}) \;=\; V^o_{\rm XC}(\mathbf{r}) \pm \frac{1}{2} IM \quad , \tag{8.96}$$

where the proportionality constant I is called the Stoner parameter (or exchange integral), and $M = \int_{unit-cell} m(\mathbf{r}) d\mathbf{r}$ (and thus $M = (\bar{m}_s/\mu_B)\times$ number of elements in the unit cell!). Since the only difference between the XC-potential seen by up and down electrons is then a constant shift (in total: IM), the wave functions in the Kohn-Sham equations are not changed compared to the non-magnetic case. The only effect the spin-polarized XC-potential has is therefore to shift the eigenvalues $\epsilon^{\uparrow\downarrow}_i$ by a constant to lower or higher values

$$\epsilon^{\uparrow\downarrow}_i \;=\; \epsilon^o_i \;\pm\; 1/2IM \quad . \tag{8.97}$$

The band structure and in turn the spin DOS are respectively shifted compared to the non-magnetic case,

$$N^{\uparrow\downarrow}(E) \;=\; \sum_i \int_{\rm BZ} \delta(E - \epsilon^{\uparrow\downarrow}_{\mathbf{k},i}) d\mathbf{k} \;=\; N^o(E \pm 1/2IM) \quad , \tag{8.98}$$

i.e., the Stoner approximation corresponds exactly to our initial observation of a constant shift between up and down spin DOS.

Yet, what have we gained with this formalized writing of our initial observation that the self-consistent spin DOS differs for up and down electrons only by a constant shift? So far nothing, really, but one can employ this Stoner model to arrive at a simple condition that tells us by looking at an unmagnetic calculation, whether the material is likely to exhibit a ferromagnetic ground state or not. What one starts out with in this approach are therefore the results of the non-magnetic calculation, namely the non-magnetic DOS $N^o(E)$ and the total number of electrons N. In the magnetic case, N remains the same, but what enters as a new free variable is the possible magnetization M. In the Stoner model the whole effect of a possible spin-polarization is contained in the parameter I, so that one has to relate the sought quantity M with I to arrive at a conclusion about which I will yield a non-vanishing M (and thereby a ferromagnetic ground state). We therefore proceed by noticing that N and M are given by integrating over all occupied states of the shifted spin DOSs up to the Fermi-level

$$N \;=\; \int_{\epsilon_F} \left[N^o(E + 1/2IM) + N^o(E - 1/2IM) \right] dE \tag{8.99}$$

$$M \;=\; \int_{\epsilon_F} \left[N^o(E + 1/2IM) - N^o(E - 1/2IM) \right] dE \quad . \tag{8.100}$$

Since $N^o(E)$ and N are fixed, these two equations completely determine the remaining free variables of the magnetic calculation, ϵ_F and M. However, the two equations are coupled (i.e., they depend on the same variables), so that one has to solve them self-consistently. The resulting M is therefore characterized by self-consistently fulfilling the relation

$$M \;=\; \int_{\epsilon_F(M)} \left[N^o(E + 1/2IM) - N^o(E - 1/2IM) \right] dE \;\equiv\; F(M) \tag{8.101}$$

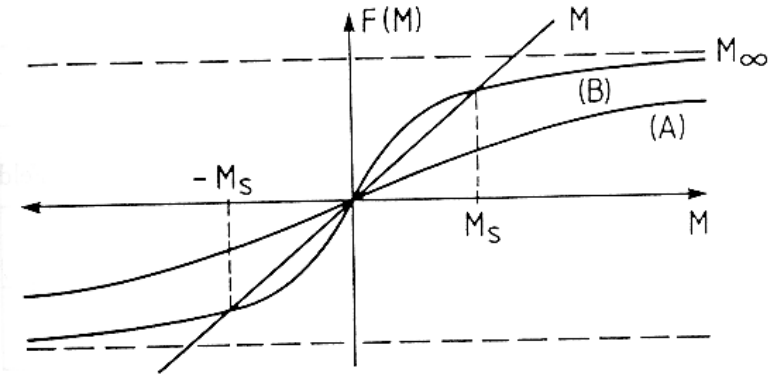

Figure 8.11: Graphical solution of the relation $M = F(M)$ of Eq. (8.101) for two characteristic forms of $F(M)$.

As expected M is a pure function $F(M)$ of the Stoner parameter I, the DOS $N^o(E)$ of the non-magnetic material, and the filling of the latter (determined by ϵ_F and in turn by N). To really quantify the resulting M, we would need an explicit knowledge of $N^o(E)$ at this point. Yet, even without such a knowledge we can derive some universal properties of the functional dependence $F(M)$ that simply follow from its form and the fact that $N^o(E) > 0$,

$$\begin{array}{ll} F(M) = F(-M) & F(0) = 0 \\ F(\pm\infty) = \pm M_\infty & F'(M) > 0 \quad . \end{array}$$

Here, M_∞ corresponds to the saturation magnetization at complete spin polarization, when all majority spin states are occupied and all minority spin states are empty. The structure of $F(M)$ must therefore be quite simple: monotonously rising from a constant value at $-\infty$ to a constant value at $+\infty$, and in addition exhibiting mirror symmetry. Two possible generic forms of $F(M)$ that comply with these requirements are sketched in Fig. 8.11. In case (A), there is only the trivial solution $M = 0$ to the relation $M = F(M)$ of Eq. (8.101), i.e., only the non-magnetic state results. In contrast, in case (B) there are three possible solutions, namely apart from the always existing possibility $M = 0$, now also solutions with a finite spontaneous magnetization $M = M_s \neq 0$. The difference between the two cases arises solely from the slope of $F(M)$ at the origin: If the slope is larger than 1, the monotonous function $F(M)$ must necessarily cross the line M another time, when saturating at some stage to a constant value. A sufficient criterion for the existence of ferromagnetic solutions with finite magnetization M_s is therefore $F'(0) > 1$. Taking the derivative of Eq. (8.101), one finds that $F'(0) = IN^o(\epsilon_F)$, and arrives finally at the famous *Stoner criterion for ferromagnetism*

$$IN^o(\epsilon_F) > 1 \quad . \tag{8.102}$$

A sufficient condition for a material to exhibit ferromagnetism is therefore to have a high density of states at the Fermi-level and a large Stoner parameter. The latter is a material constant that can be computed within DFT (by taking the functional derivative of $V_{xc}^{\uparrow\downarrow}(\mathbf{r})$ with respect to magnetization; see e.g. J.F. Janak, Phys. Rev. B **16**, 255 (1977)). Loosely

speaking it measures the localization of the wave functions at the Fermi-level. In turn it represents the strength of the exchange interaction in the solid per electron, whereas $N^o(\epsilon_F)$ tells how many electrons are able to participate in the ordering. Consequently, the derived Stoner criterion has a very intuitive form: strength per electron times number of electrons participating. If this exceeds a critical threshold, the ferromagnetic state results favorable.

Figure 8.12 shows the quantities entering the Stoner criterion calculated within DFT-LSDA for all elements. Gratifyingly, only the three elements Fe, Co and Ni that are indeed ferromagnetic in reality fulfill the condition. With the understanding of the Stoner model we can now, however, also rationalize why it is only these three elements and not the others: For ferromagnetism we need a high number of states at the Fermi-level, such that only transition metals come into question. In particular, $N^o(\epsilon_F)$ is always highest for the very late TMs with an almost filled d-band. In addition, also a strong localization of the wave functions at the Fermi-level is required to yield a high Stoner parameter. And here, the compact $3d$ orbitals outwin the more delocalized ones of the $4d$ and $5d$ elements. Combining the two demands, only Fe, Co, and Ni remain.

Yet, it is interesting to see that also Ca, Sc and Pd come quite close to fulfilling the Stoner criterion. This is in particular relevant, if we consider other situations than the bulk: The DOSs of solids are in general quite structured, but scale to first order inversely with the band width W: In other words, the narrower the band, the higher the number of states per energy within the band. Just as much as by a stronger localization, W is also decreased by a lower coordination. In the atomic limit, the band width is zero and the Stoner criterion always fulfilled. Correspondingly, all atoms with partially filled shells have a non-vanishing magnetic moment, given by Hund's rules. Somewhere intermediate are surfaces and thin films, where the atoms have a reduced coordination compared to the bulk. Due to the then narrower bands, such systems can exhibit a larger magnetic moment or even a change of their magnetic properties (i.e., when they are non-magnetic in the bulk case). The magnetism of surfaces and thin films is therefore substantially richer than the bulk magnetism discussed in this introductory lecture. Most importantly, it is also the basis of many technological applications.

8.5 Magnetic Domain Structure

With the results from the last section we can understand the ferromagnetism of the RE metals as a consequence of the exchange coupling between localized moments due to the partially filled (and mostly atomic-like) f-shells. The itinerant ferromagnetism of the TMs Fe, Co and Ni, on the other hand, comes from the exchange interaction of the still localized, but no longer atomic-like d orbitals together with the delocalized s-electron density. We can treat the prior localized magnetism using discrete Heisenberg or Hubbard models and exchange constants from Heitler-London type theories (or LDA+U type approaches), while the latter form of magnetism is more appropriately described within band theories like spin-DFT-LSDA. Despite this quite comprehensive understanding, an apparently contradictory feature of magnetic materials is that a ferromagnet does not

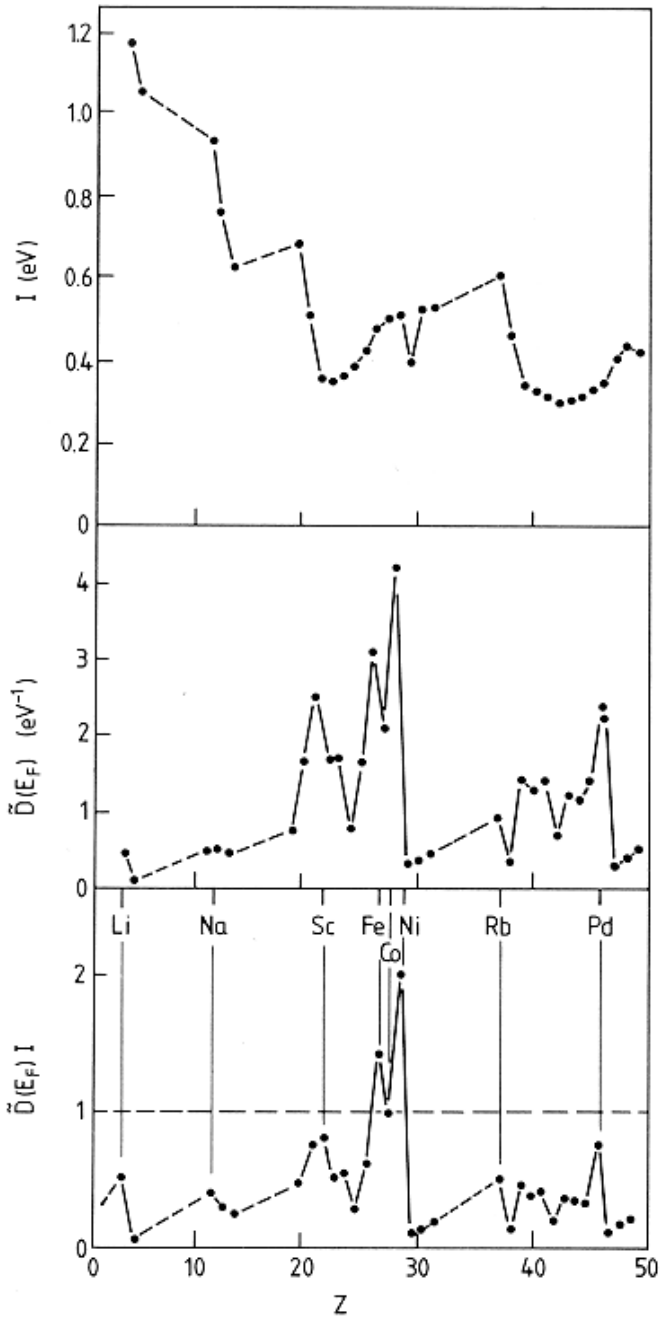

Figure 8.12: Trend over all elements of the Stoner parameter I, of the non-magnetic DOS at the Fermi-level (here denoted as $\tilde{D}(E_F)$) and of their product. The Stoner criterion is only fulfilled for the elements Fe, Co, and Ni [from H. Ibach and H. Lüth, *Solid State Physics*, Springer (1990)].

always show a macroscopic net magnetization, even when the temperature is well below the Curie temperature. In addition it is hard to grasp, how the magnetization can be affected by external fields, when we know that the molecular internal fields are vastly larger.

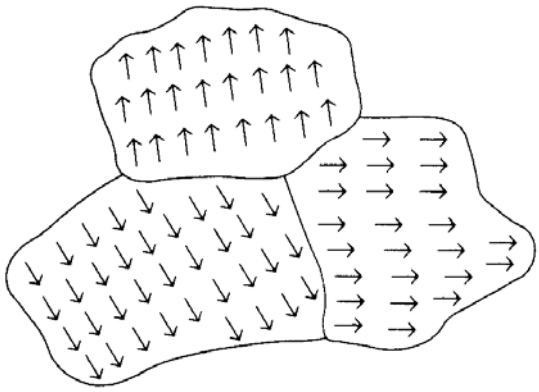

Figure 8.13: Schematic illustration of magnetic domain structure.

The explanation for these puzzles lies in the fact that two rather different forces operate between the magnetic moments of a ferromagnet. At very short distance (of the order of atomic spacings), magnetic order is induced by the powerful, but quite short-ranged exchange forces. Simultaneously, at distances large compared with atomic spacings, the magnetic moments still interact as magnetic dipoles. We stressed in the beginning of the last section that the latter interaction is very weak; in fact between nearest neighbors it is orders of magnitude smaller than the exchange interaction. On the other hand, the dipolar interaction is rather long-ranged, falling off only as the inverse cube of the separation. Consequently, it can still diverge, if one sums over all interactions between a large population that is uniformly magnetized. In order to accommodate these two competing interactions (short range alignment, but divergence if ordered ensembles become too large), ferromagnets organize themselves into *domains* as illustrated in Fig. 8.13. The dipolar energy is then substantially reduced, since due to the long-range interaction the energy of *every spin* drops. This bulk-like effect is opposed by the unfavorable exchange interactions with the nearby spins in the neighboring misaligned domains. Because, however, the exchange interaction is short-ranged, it is only the spins near the domain boundaries that will have their exchange energies raised, i.e., this is a surface effect. Provided that the domains are not too small, domain formation will be favored in spite of the vastly greater strength of the exchange interaction: Every spin can lower its (small) dipolar energy, but only a few (those near the domain boundaries) have their (large) exchange energy raised.

Whether or not a ferromagnet exhibits a macroscopic net magnetization, and how this magnetization is altered by external fields, has therefore less to do with creating alignment (or ripping it apart), but with the domain structure and how its size and orientational distribution is changed by the field. In this respect it is important to notice that the boundary between two domains (known as the domain wall or *Bloch wall*) has often not the (at first thought intuitive) abrupt structure. Instead a spread out, gradual change as shown in Fig. 8.14 is less costly in exchange energy. All an external field has to do then is to slightly affect the relative alignments (not flip one spin completely), in order to induce

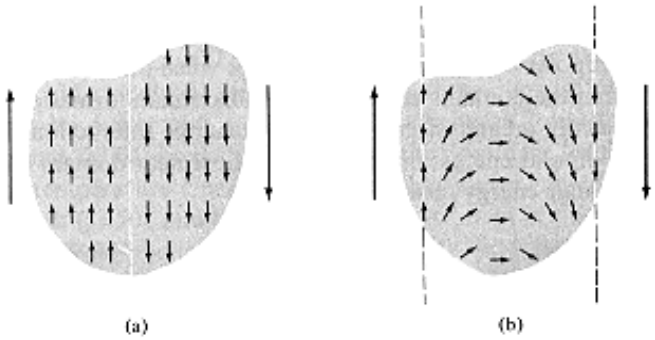

Figure 8.14: Schematic showing the alignment of magnetic moments at a domain wall with (a) an abrupt boundary and (b) a gradual boundary. The latter type is less costly in exchange energy.

a smooth motion of the domain wall, and thereby increase the size of favorably aligned domains. The way how the comparably weak external fields can "magnetize" a piece of unmagnetized ferromagnet is therefore to rearrange and reorient the various domains, cf. Fig. 8.15. The corresponding domain wall motion can hereby be reversible, but it may also well be that it is hindered by crystalline imperfections (e.g. a grain boundary), through which the wall will only pass when the driving force due to the applied field is large. When the aligning field is removed, these defects may prevent the domain walls from returning to their original unmagnetized configuration, so that it becomes necessary to apply a rather strong field in the opposite direction to restore the original situation. The dynamics of the domains is thus quite complicated, and their states depend in detail upon the particular history that has been applied to them (which fields, how strong, which direction).

As a consequence, ferromagnets always exhibit so-called *hysteresis* effects, when plotting the magnetization of the sample against the external magnetic field strength as shown in Fig. 8.16. Starting out with an initially unmagnetized crystal (at the origin), the applied external field induces a magnetization, which reaches a saturation value M_s when all dipoles are aligned. On removing the field, the irreversible part of the domain boundary dynamics leads to a slower decrease of the magnetization, leaving the so-called *remnant magnetization* M_R at zero field. One needs a reverse field (opposite to the magnetization of the sample) to further reduce the magnetization, which finally reaches zero at the so-called *coercive field* B_c.

Depending on the value of this coercive field, one distinguishes between *soft and hard magnets* in applications: A soft magnet has a small B_c and exhibits therefore only a small hysteresis loop as illustrated in Fig. 8.17. Such materials are employed, whenever one needs to reorient the magnetization frequently and without wanting to resort to strong fields to do so. Typical applications are kernels in engines, transformators, or switches and examples of soft magnetic materials are Fe or Fe-Ni, Fe-Co alloys. The idea of a hard magnet which requires a huge coercive field, cf. Fig. 8.17, is to maximally hinder the demagnetization of a once magnetized state. A major field of application is therefore magnetic storage with Fe- or Cr-Oxides, schematically illustrated in Fig. 8.18. The medium (tape or disc) is covered by many small magnetic particles each of which behaves

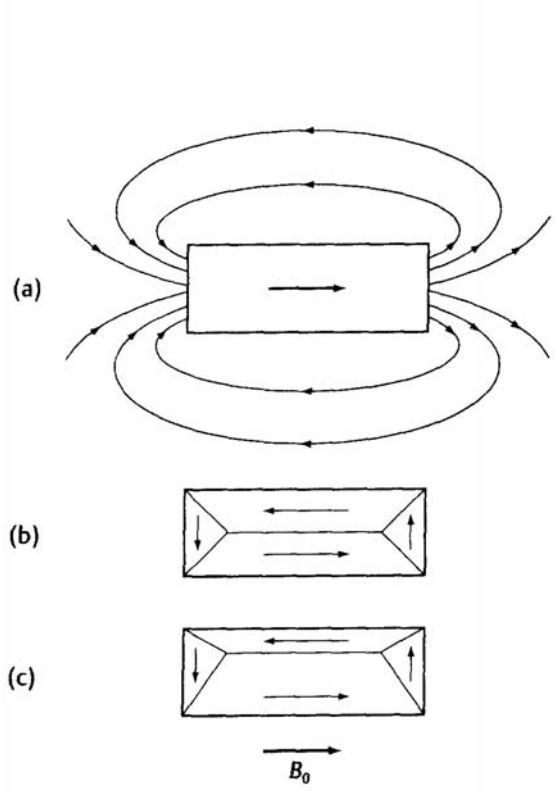

Figure 8.15: (a) If all of the magnetic moments in a ferromagnet are aligned they produce a magnetic field which extends outside the sample. (b) The net macroscopic magnetization of a crystal can be zero, when the moments are arranged in (randomly oriented) domains. (c) Application of an external magnetic field can move the domain walls so that the domains which have the moments aligned parallel to the field are increased in size. In total this gives rise to a net magnetization.

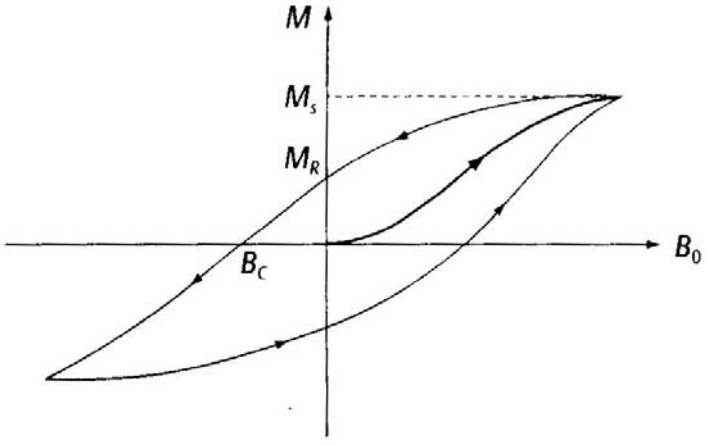

Figure 8.16: Typical hysteresis loop for a ferromagnet.

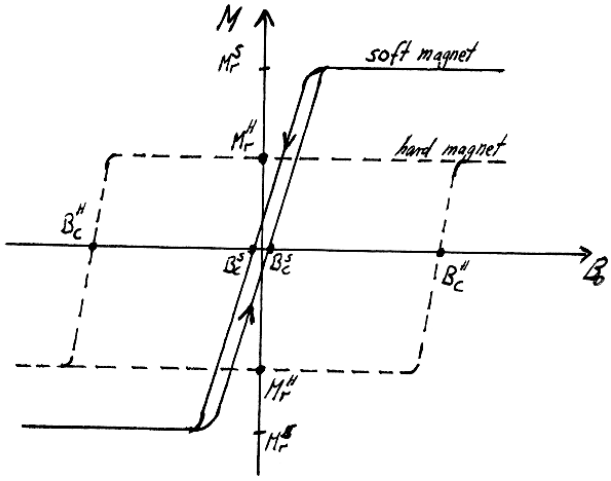

Figure 8.17: Hysteresis loops for soft and hard magnets.

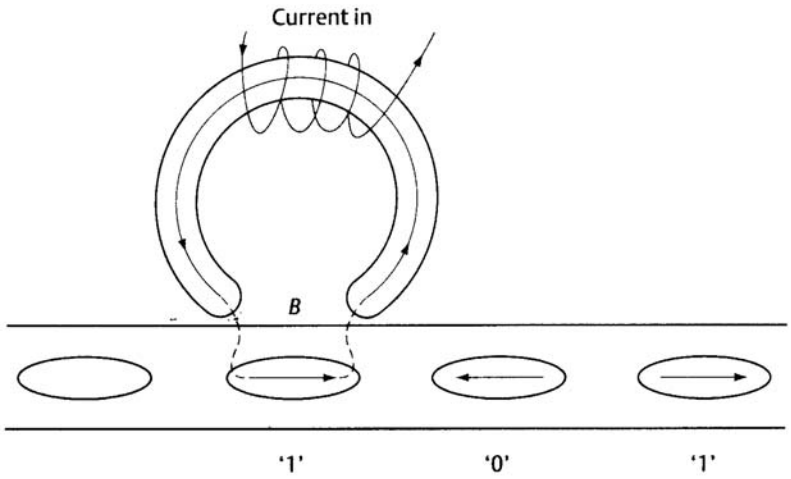

Figure 8.18: Illustration of the process of recording information onto a magnetic medium. The coil converts an electrical signal into a magnetic field in the write head, which in turn affects the orientation of the magnetic particles on the magnetic tape or disc. Here, there are only two possible directions of magnetization of the particles, so the data is in binary form and the two states can be identified as 0 and 1.

as a single domain. A magnet – write head – records the information onto this medium by orienting the magnetic particles along particular directions. Reading information is then simply the reverse: The magnetic field produced by the magnetic particles induces a magnetization in the read head. In magnetoresistive materials, the electrical resistivity changes in the presence of a magnetic field so that it is possible to directly convert the magnetic field information into an electrical response. In 1988 a structure consisting of extremely thin, i.e., three atom-thick-alternative layers of Fe and Co was found to have a magnetoresistance about a hundred times larger than that of any natural material. This effect is known as *giant magnetoresistance* and is in the meanwhile exploited in todays disc drives.