Advanced Physics 1/2

Theory of Magnetism

2019, terms 3/4

岡山大学 異分野基礎科学研究所
RESEARCH INSTITUTE FOR INTERDISCIPLINARY SCIENCE, OKAYAMA UNIVERSITY

Harald Jeschke
Research Institute for Interdisciplinary Science, Okayama University

Sources

This lecture is based on the books "Lecture Notes on Electron Correlation and Magnetism" by Patrik Fazekas and "Quantum Theory of Magnetism" by Wolfgang Nolting and Anupuru Ramakanth. In parts, it follows the lecture notes "Theory of Magnetism" by Carsten Timm. Original papers are used as cited in the footnotes.

1. Introduction

1.1 Magnetism as an effect of the electron-electron interaction

Since antiquity it was known that there is an attraction between lodestone (magnetite, Fe_3O_4) and iron. Plato and Aristotle mention permanent magnets. They are also mentioned in Chinese texts of the 4th century BC. Use of a magnetic compass for navigation was first mentioned in a Chinese text dated 1040-1044 AD but a much earlier use is possible. Apparently the magnetic compass was first used for orientation on land, not on the sea.

Thus, magnetism at first referred to the long-range interaction between ferromagnetic macroscopic entities. However, in this course we will focus on the microscopic origin of the magnetic order in solids; one of these orders is ferromagnetism.

Initially, we want to describe the effect of an external magnetic field H on the behavior of a solid. For a weak field, the relevant response function is the susceptibility

$$\chi = \frac{M}{H}$$

with the magnetization density H. The probing field can be space and time dependent. Consequently, if we introduce Fourier components of H and M, we find in general a wave vector and frequency dependent generalized susceptibility $\chi(\vec{q}, \omega)$ which fully characterizes the behavior of the systems in weak fields. Calculating the magnetization requires solving a quantum mechanical eigenvalue problem where the interaction of the external field with the system is added to the microscopic Hamiltonian. The energy scale of this interaction is $\mu_B H$ with the Bohr magneton

$$\mu_B = \begin{cases} \frac{e\hbar}{2mc} \approx 9.27 \cdot 10^{-21} \frac{\mathrm{erg}}{\mathrm{G}} \text{ in cgs units} \\ \frac{e\hbar}{2m} \approx 9.27 \cdot 10^{-24} \frac{\mathrm{J}}{\mathrm{T}} \text{ in SI units}\,, \end{cases} \tag{1.1}$$

and with Gauss (G) and Tesla (T). μ_B is nearly equal to the spin moment of an electron in vacuum.

Now how large is this energy scale $\mu_B H$? In condensed matter physics, a common energy unit is electron Volt $1\,\mathrm{eV} = 1.6 \cdot 10^{-19}\,\mathrm{J}$, and equivalent temperatures are obtained using the Boltzmann constant $k_B = 1.38 \cdot 10^{-23}\,\frac{\mathrm{J}}{\mathrm{K}} = 8.6 \cdot 10^{-5}\,\frac{\mathrm{eV}}{\mathrm{K}}$. This means that $1\,\mathrm{eV}$ corresponds to the temperature $11605\,\mathrm{K}$. A standard laboratory field of $5\,\mathrm{T} = 5 \cdot 10^4\,\mathrm{G}$ then corresponds to an interaction energy of $\mu_B H = 3\,\mathrm{K}$. Comparing to other solid state energy scales, this is rather small: band widths can be of the order $1 - 10\,\mathrm{eV} \sim 10^4 - 10^5\,\mathrm{K}$; Coulomb matrix elements are of similar size, and phonons are characterized by Debye temperatures of $\Theta_D = 100 - 1000\,\mathrm{K}$. Even spin orbit coupling is usually stronger than $\mu_B H$. $\mu_B H$ is of the order of the superconducting T_c of conventional superconductors; it is well known that sufficiently strong magnetic fields can suppress superconductivity.

It would however be wrong to assume that because of the small energy scales, no drastic change in the behavior of solids is to be expected. One possibility is a situation of degeneracies in a system with competing energy scales where a magnetic field can trigger strong effects. Also, if strong correlations lead to very narrow effective bands and therefore very small Fermi energies, as in heavy Fermion materials, laboratory fields can have dramatic effects.

Conventionally, weak magnetism denotes the situation when the magnetization of a material is induced by an external field and vanishes when the field is turned off. Strong magnetism means that a material shows a spontaneous magnetization also in the absence of an external field. Formally, a divergent static susceptibility $\chi(\vec{q}, \omega = 0)$ indicates the onset of magnetic ordering.

While field induced magnetization can be attributed to a weak perturbation, this is not anymore true for the origin of spontaneous magnetic order. Realizable magnetic fields are rather weak, and dipole fields of atomic moments at interatomic distances are even weaker. Let's consider if the magnetism of iron could be due to classical moments that align by sitting in each others fields. In ferromagnetic iron, the moments are $\mu_{at} = 2.1\mu_B$, and nearest neighbor distances are $a = 2.55\,\text{Å} \approx 5a_B$ (with Bohr radius $a_B = \frac{\hbar^2}{me^2}$). The fine structure constant is $\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$, and a Rydberg is $1\,\mathrm{Ry} = \frac{e^4 m}{2\hbar^2} \approx 13.6\,\mathrm{eV}$. Then the dipole-dipole interaction can be estimated as $\frac{\mu_{at}^2}{a^3} = \left(\frac{\mu_{at}}{\mu_B}\right)^2 \left(\frac{a}{a_B}\right)^{-3} \frac{\alpha^2}{2}\,\mathrm{Ry} \approx 10^{-5}\,\mathrm{eV} \approx 0.1\,\mathrm{K}$. This clearly cannot account for a Curie temperature of $1043\,\mathrm{K}$.

Spontaneous magnetic order is, in most cases, a consequence of strong

electron-electron interactions, rather than a secondary effect due to a weak perturbation. The magnetism of strongly magnetic materials arises from large terms in the Hamiltonian. Whenever band theory indicates a metal with narrow conduction bands, magnetism arises in a natural manner (*i.e.* it is one of the leading instabilities). Then the material often turns out to be a magnetic insulator rather than a metal. Magnetic instabilities are closely related to the problem of metal insulator transitions. Not all materials loose their metallicity upon becoming magnetic, though: Other circumstances like band filling play a role. There are the famous ferromagnetic metals Fe, Co, Ni, magnetic rare earth metals Gd and Dy or the ferromagnetic metallic oxide CrO_2.
Magnetism will be discussed here in the context of electron-electron interactions. Once these are strong, systems can become magnetic, they can distort structurally, they can show metal insulator transitions, and they might even become unconventional superconductors. The focus will here be on the microscopic mechanisms of magnetism.

1.2 Magnetic field sources

The phenomena that can be observed experimentally are determined by the available magnetic fields, and new field induced effects are discovered with each progress in available magnetic field strengths. The cgs unit of magnetic field H is Oersted, and conversion to SI is by $10\,\mathsf{Oe} = \frac{10^3}{4\pi}\,\frac{\mathsf{A}}{\mathsf{m}}$. The unit of the field $\mathsf{B} = \mathsf{H} + 4\pi\mathsf{M}$ is the same by value and dimension but it is called Gauss. The SI unit is 1 Tesla $= 10^4$ G. The field of the earth is about 0.5 G. Simple iron based permanent magnets provide a few hundred Gauss, and powerful permanent magnets like samarium-cobalt ($SmCo_5$) or neodymium-iron-boron ($Nd_2Fe_{14}B$) have fields of 3000-4000 G. Large fields for research are produced by electromagnets; fields of 5-30 T are routine. There are limits to the fields that can be produced by currents through coils because the force exerted by the field on the coil eventually exceeds the tensile strength of the material. Resistive heating through the current is another limiting factor. This can be avoided by the use of superconducting coils; however, superconductivity breaks down when the magnetic field exceeds the critical field of the superconductor. Hybrid magnets with resistive electromagnets inside a superconducting magnet can reach higher fields, ~ 45 T. For many experiments, it is sufficient to have an intense field for milliseconds or microseconds; fields up to 100 T can be produced with non-destructive pulsed magnets. Self-destructing pulsed magnets can produce an order of magnitude larger fields, for example very recently 1200 T at ISSP, University of Tokyo, and constitute a very active field of research.

1.3 Some concepts at the example of magnetite

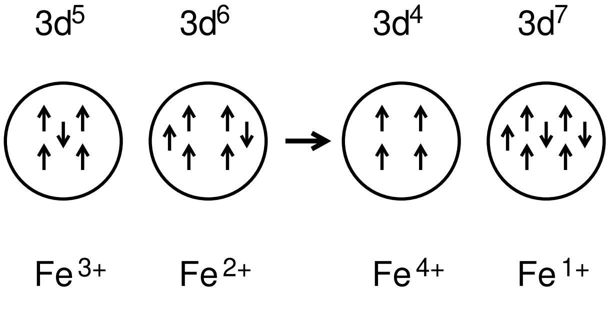
Even though magnetite was known to most ancient civilizations, it is a complicated substance and a topic of research even today. By way of an introduction we will learn about some of the questions that can be asked, but not about all the answers

1.3.1 Charge states

Magnetite is an iron oxide with the formula Fe_3O_4; besides it, there are also the iron oxides FeO and Fe_2O_3 (hematite). We can use them to analyze the ionic bonding: We have divalent iron in $Fe^{2+}O^{2-}$ and trivalent iron in $Fe_2^{3+}O_3^{2-}$. Now in magnetite, both divalent and trivalent iron are present: $Fe^{2+}Fe_2^{3+}O_4^{2-}$. It is called a mixed valent oxide. We can try to imagine what these charge states of iron would mean in terms of a band picture. Fe^{2+} and Fe^{3+} both have partially filled d shells, $3d^5$ for Fe^{3+} and $3d^6$ for Fe^{2+}. Now we can try to think of the resulting band as a tight binding band with nearest neighbor hopping. For a pair of $3d^5$ and $3d^6$ sites, we can think of this hopping event between sites i and j:

$$\left|3d^5\right\rangle_i\left|3d^6\right\rangle_j \to \left|3d^4\right\rangle_i\left|3d^7\right\rangle_j \tag{1.2}$$

This can be realized by a minority spin electron jumping away from a $3d^5$ site:



If such a hopping process would be realized, we would also expect monovalent Fe^+ and tetravalent Fe^{4+} ions to appear. However, this is contrary to our chemical picture that magnetite contains only divalent and trivalent iron; at least, independent electron theory seems to give us a too high fraction of other valences. Mixed valence is rather common in transition metals and rare earths and means that a substance is dominated by two valence states at the exclusion of others. Thus, in magnetite we should somehow restrict hopping to processes that interchange Fe^{2+} and Fe^{3+}:

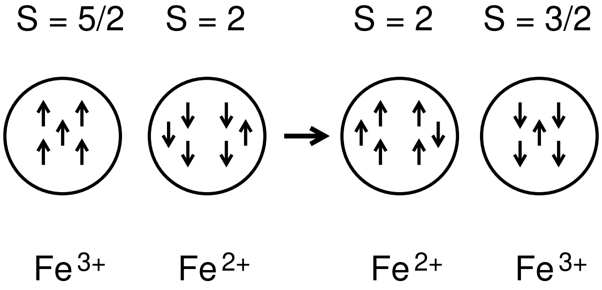
$$\left|3d^5\right\rangle_i \left|3d^6\right\rangle_j \rightarrow \left|3d^6\right\rangle_i \left|3d^5\right\rangle_j \tag{1.3}$$

Restricting hopping in this way is an example of correlated motion rather than the usual band motion of electrons.

1.3.2 Spin states

Even equation 1.3 is still too permissive because we have not yet taken into account the spin state. Fe^{2+} and Fe^{3+} are magnetic ions and will retain the total spin (the magnetic moment) that they would have in free space. First of all, we need to understand why free Fe^{2+} ions have total spin $S = 2$, and Fe^{3+} ions have total spin $S = \frac{5}{2}$. Then we need to study if and how this survives in the solid state.

So let us consider the hopping process



This is consistent with equation 1.3 but we would consider it forbidden because it would lead to a low spin state of Fe^{3+}, $S = \frac{3}{2}$. Physically, the reason to exclude such processes is the same as in the case of the valence restriction above: It would lead from a low energy subspace to a high energy subspace. Later, both the reason for the energetic separation of subspaces and the way to introduce such constraints formally will be discussed.

For the special case of magnetite, it is possible to give a simple form both to the valence constraint and to the restriction to high spin states. Magnetite is ferrimagnetic, with an ordering temperature $T_N = 858$ K. If we consider temperatures of $T < 200$ K, we can approximate the electron motion as hopping on the background of a frozen pattern of spin order. Magnetite has an inverse spinel structure (Figure 1.1), AB_2O_4 with two crystallographically inequivalent iron sites. Below T_N, spins on the two inequivalent lattice sites are polarized in the opposite way. The tetrahedrally coordinated A site has one Fe^{3+} ion, the octahedrally coordinated B site has one Fe^{3+} ion, one Fe^{2+} ion. While the opposite moment Fe^{3+} ions compensate each

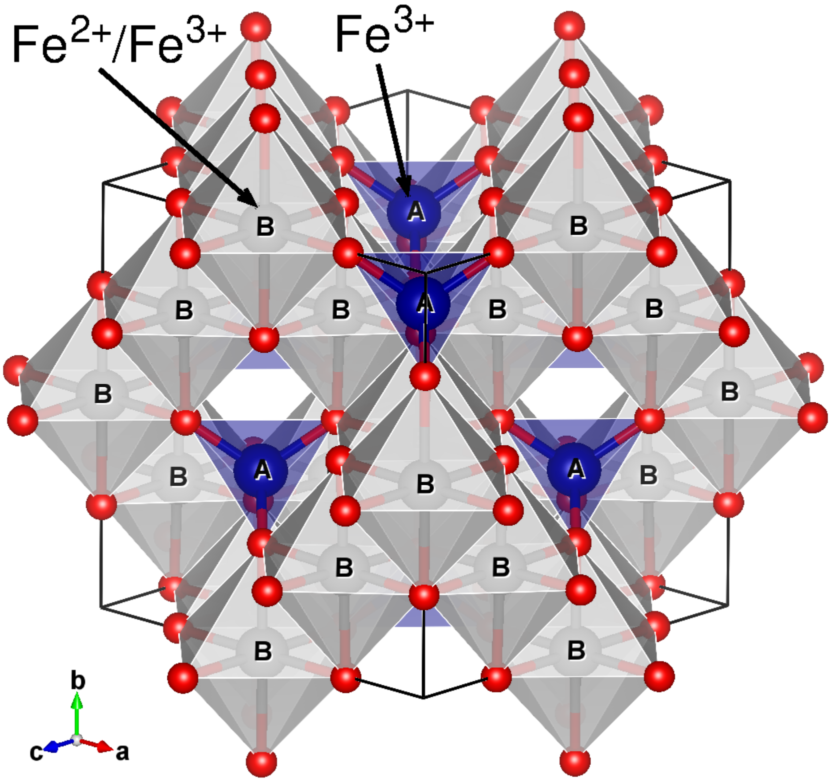
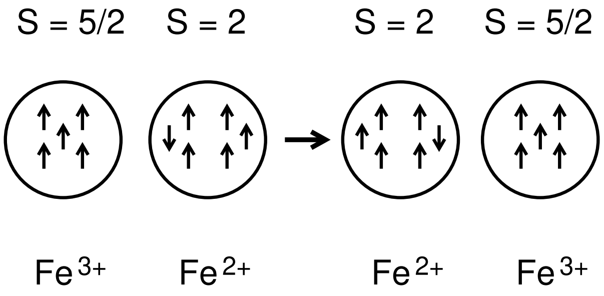


Figure 1.1: Inverse spinel structure AB_2O_4 of Fe_3O_4: Fe^{3+} ions occupy the tetrahedral A sites as well as half of the octahedral B sites; the other B sites are occupied by Fe^{2+} ions.

other, a residual ferrimagnetic moment of Fe^{2+} on the B site remains. At sufficiently low temperature when magnetic ordering is nearly perfect, the only allowed hopping is that of minority spin electrons that switch Fe^{2+} and Fe^{3+} sites:



This means that after fixing the spin order, we can forget about the spin degree of freedom and focus only on the mobile electrons which are now “spin free”. The problem is reduced to nearest neighbor hopping of spin-less Fermions in a half-filled band:

$$H_{\mathrm{hop}} = -t \sum_{\langle ij \rangle} \left(c_i^\dagger c_j + c_j^\dagger c_i \right) \tag{1.4}$$

Here, $c_i^\dagger$ creates an electron at lattice site i, c_i annihilates an electron. t is an energy and represents the hopping amplitude.

1.3.3 Charge order

So far, the problem of 16 d electrons in the chemical unit cell of Fe_3O_4 has been reduced to a half-filled band of spin-less Fermions moving on the the B sublattice of the inverse spinel. For a half-filled band, we would now expect a metallic ground state. However, the resistivity of Fe_3O_4 shown in Figure 1.2 clearly doesn't show a metallic temperature dependence and is rather high at all temperatures.

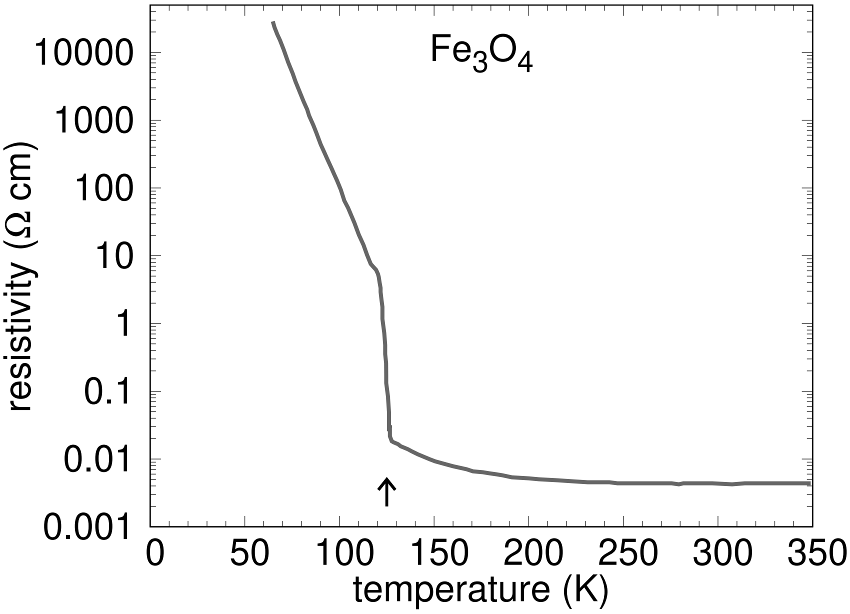


Figure 1.2: Resistivity of Fe_3O_4 as function of temperature[a]. The temperature of the Verwey transition is indicated by an arrow.

In particular, resistivity jumps up by two orders of magnitude when Fe_3O_4 is cooled below the so-called Verwey temperature $T_V \approx 125$ K. Even though it is a transition from a semiconducting to an insulating state, it can be considered an example of a correlation-driven metal-insulator transition. We have accounted for substantial parts of the electron-electron repulsion by restricting the valence states to Fe^{2+} and Fe^{3+} and by assigning definite, maximum spins to these valence states. However, an important part is missing. In the spin-less Fermion model, the electrons are prohibited by the Pauli principle to share the same site but there is so far nothing preventing them from sitting at nearest neighbor sites. However, the associated Coulomb energy is large at $V = \frac{e^2}{a}$ with lattice spacing a. We should add this term to the Hamiltonian:

$$H = H_{\text{hop}} + H_{\text{e-e}} = -t \sum_{\langle ij \rangle} \left(c_i^\dagger c_j + c_j^\dagger c_i \right) + V \sum_{\langle ij \rangle} \hat{n}_i \hat{n}_j \tag{1.5}$$

with site occupation operator $\hat{n}_i = c_i^\dagger c_i$. This is of course a very simplified band model but it contains some essential aspects of the Verwey transition.

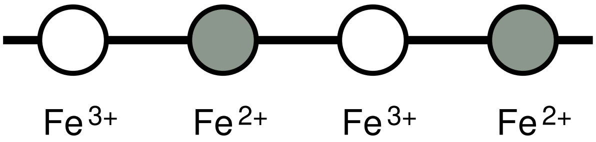
---

[a] V. V. Shchennikov, S. V. Ovsyannikov, J. Phys.: Condens. Matter **21**, 271001 (2009).

It is an example of how model Hamiltonians are devised. The two terms of the Hamiltonian stand for competing tendencies. For small V, the first, kinetic energy term dominates and essentially a half-filled band with little suppression of simultaneous occupation of neighboring sites is obtained. On the other hand, setting $t = 0$ results in

$$H_{\text{e-e}} = V \sum_{\langle ij \rangle} \hat{n}_i \hat{n}_j \,. \tag{1.6}$$

This energy would be minimized by a state in which only every other site is occupied:



This could be called a large amplitude charge density wave. It has also been interpreted as an example of Wigner crystallization, the formation of an electron crystal. This is a prime example of the meaning of a strongly correlated state: The electrons do their utmost to avoid each other in direct space, and it is clear that this is an insulating state. It is also interesting that this state does not have the full translational invariance of the underlying crystal. An important effect of strong electronic correlations is the appearance of symmetry breaking ground states, ground states whose symmetry is lower than that of the Hamiltonian.

2. Magnetism of free atoms and ions

The study of magnetism starts with small systems – atoms, ions and molecules – where many of the mechanisms that give rise to magnetic ordering are already active as soon as there are two interacting electrons. The tendency to form a high spin ground state in a small system is related to a ferromagnetic state in a solid, the tendency to align antiparallel and form a spin singlet is related to antiferromagnetism. On the other hand, when electrons are essentially localized to ions, the properties of the isolated ions are directly relevant also for the solid, as is the case for rare earth ions.

2.1 Hartree approximation for the electron shell

We cannot exactly solve the many-body problem of a nucleus with many electrons. In the simplest non-trivial approximation, the Hartree approximation, a given electron moves in a potential resulting from the nucleus and the average density of the other electrons; it is important that the self-interaction, the interaction of the electron with its own averaged charge density is excluded. The total potential is

$$V_{\text{eff}}(\vec{r}) = -\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r} - \frac{1}{4\pi\epsilon_0}\int d^3r'\,\frac{e\rho_{\vec{r}}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} \tag{2.1}$$

where Z is the atomic number of the nucleus, the electron charge is $-e < 0$, and $\rho_{\vec{r}}(\vec{r}\,') < 0$ is the charge density at $\vec{r}\,'$ of the other electrons if the given electron is at $\vec{r}$. $V_{\text{eff}}(\vec{r})$ is spherically symmetric due to the isotropy of space, but $\rho_{\vec{r}}(\vec{r}\,')$ as a function of $\vec{r}\,'$ is not spherically symmetric except for $\vec{r} = 0$. For the electron at $\vec{r}$ we solve the single-particle Schrödinger equation

$$\left(\frac{p^2}{2m} + V_{\text{eff}}(\vec{r})\right)\psi(\vec{r}) = E\psi(\vec{r}) \tag{2.2}$$

From a separation of variables and using the spherical symmetry, the eigenfunctions are

$$\psi_{nlm}(\vec{r}) = R_{nl}(r)Y_{lm}(\vartheta, \phi) \tag{2.3}$$

with principal quantum number $n = 1, 2, 3, \ldots$, orbital angular momentum quantum number $l = 0, 1, 2, \ldots, n-1$ and magnetic quantum number $m = -l, -l+1, ..., l$. The angular part is the same for any spherically symmetric potential and is given by the spherical harmonics $Y_{lm}(\vartheta, \phi)$. Eigenenergies $\epsilon_{n,l}$ only depend on n, l in the present approximation. Including a factor of 2 from spin $s = 1/2$, the $\epsilon_{n,l}$ are $2(2l+1)$-fold degenerate. Completely full shells comprising all orbitals with a given quantum number n, l have $\langle \sum_i \vec{l_i} \rangle = 0$ and $\langle \sum_i \vec{s_i} \rangle = 0$, *i.e.* vanishing total angular moment because for every electron there is another one with opposite $\langle \vec{l_i} \rangle$, $\langle \vec{s_i} \rangle$. The total magnetic moment of filled shells also vanishes. Therefore, magnetic ions require incompletely filled shells.

In the ground states, we need to fill the Hartree orbitals starting from the lowest in energy. For a shell containing p electrons (with $p < 2(2l+1)$), the number of possibilities of doing this is given by

$$\binom{2(2l+1)}{p} \, .$$

This represents the degeneracy of the many-particle state. For a filled state, we get no degeneracy:

$$\binom{2(2l+1)}{2(2l+1)} = 1 \, .$$

2.2 Beyond the Hartree approximation

The large degeneracy we found is partially lifted by the Coulomb repulsion beyond the Hartree approximation. The Coulomb interaction

$$V_C = \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\vec{r_i} - \vec{r_j}|} \tag{2.4}$$

commutes with the total orbital angular momentum of the shell $\vec{L} = \sum_i \vec{l_i}$ as V_C is spherically symmetric, and with the total spin of the shell $\vec{S} = \sum_i \vec{s_i}$ as well because the Hamiltonian doesn't depend on it:

$$[\vec{S}, H]_- = 0 \, , \quad [\vec{L}, H]_- = 0 \tag{2.5}$$

The Hamiltonian also commutes with $\vec{L}^2$ and with $\vec{S}^2$. Furthermore, the total angular momentum of the shell

$$\vec{J} = \vec{L} + \vec{S} \tag{2.6}$$

commutes with the Hamiltonian

$$[\vec{J}, H]_- = 0\,. \tag{2.7}$$

Physically, this means that

$$L = \sum_{i=1}^{p} m_l^{(i)},\ S = \sum_{i=1}^{p} m_s^{(i)},\ J \tag{2.8}$$

are good quantum numbers; $m_l^{(i)}$ and $m_s^{(i)}$ are the magnetic quantum numbers of the electrons. We can also say that there exists a simultaneous set of eigenstates for the operators

$$H, J^2, J_z, L^2, L_z, S^2, S_z\,,$$

and states can be labeled by the corresponding quantum numbers

$$|\dots\rangle = |J, M_J, L, M_L, S, M_S\rangle \tag{2.9}$$

Here, S is the maximum possible value of $\langle S_z\rangle$, L is the maximum possible value of $\langle L_z\rangle$. For example,

$$\begin{aligned} J^2|\dots\rangle &= \hbar^2 J(J+1)|\dots\rangle \qquad & J_z|\dots\rangle &= \hbar M_J|\dots\rangle \\ J &= |L-S|, \dots, L+S \qquad & -J \leqslant M_J &\leqslant +J \end{aligned} \tag{2.10}$$

The other angular momentum operators act in similar fashion. The energy eigenvalues of the Hamiltonian

$$H|\dots\rangle = E^{(0)}_{JLS}|\dots\rangle \tag{2.11}$$

will depend on J, L, S, but in the absence of a magnetic field, they will be degenerate with respect to M_J, M_L and M_S.

If we apply the raising operator $L^+ = L_x + iL_y$ to $|\psi\rangle$ and obtain a $|\psi'\rangle$ with $M_L' = M_L + 1$, we reach a new state that has the same energy as the old one because $[H, L^+] = 0$. There are $(2L+1)(2S+1)$ states that are connected by $L^\pm$ and by $S^\pm$ ($S^- = S_x - iS_y$, $S^+ = S_x + iS_y$); the $\binom{2(2l+1)}{p}$-fold degenerate states split into multiplets with fixed S and L and degeneracies $(2L+1)(2S+1)$. Typically, energy splitting between multiplets are much larger than 1 eV so that for magnetic problems, we only need to know the ground state multiplet.

2.2.1 Hund's rules for LS coupling

We assume here that the Coulomb interaction H_C is significantly larger than the spin-orbit coupling H_{SO}

$$H_C \gg H_{SO}$$

which leads us to the Russell-Saunders- or LS-coupling. This is relevant for light nuclei.
The empirical Hund's rules tell us how to build the ground-state LS multiplet for given L and S, and for p electrons filled into a shell with orbital angular momentum quantum number l:

1st Hund's rule: The ground state multiplet has the largest possible S. (The maximum S corresponds to the largest possible value of $\langle S_z \rangle$).

$$S = \frac{1}{2}\left[(2l+1) - |2l+1-p|\right]$$

2nd Hund's rule: If the 1st Hund's rule leaves several possibilities, the state with maximum L is lowest in energy:

$$L = S|2l+1-p|$$

(The maximum L corresponds to the largest possible value of $\langle L_z \rangle$).

A short qualitative explanation for the 1st Hund's rule is that the same spin together with the Pauli principle means that the electrons are on average further apart, reducing the Coulomb repulsion. Also the aligned orbital momenta, *i.e.* rotation in the same sense, of the second rule optimizes distance between electrons and reduces Coulomb energy.
The multiplets are labeled as ${}^{2S+1}L$ where a letter is used for L:

$L =$	0	1	2	3	4	5	6	…
$X =$	S	P	D	F	G	H	I	continuing alphabetically

We can count the number of distinct states within the LS multiplets as

$$\sum_{J=|L-S|}^{L+S} (2J+1) = (2S+1)(2L+1) \tag{2.12}$$

Often (not always), these states are ordered according to the

Table 2.1: Atomic term scheme for an f shell occupied by p electrons. ↑ represents spin projection $1/2$, ↓ represents spin projection $-1/2$.

p				m_l							
	3	2	1	0	-1	-2	-3	S	L	J	term
1	↑							$1/2$	3	$5/2$	$^2F_{5/2}$
2	↑	↑						1	5	4	$^3H_4$
3	↑	↑	↑					$3/2$	6	$9/2$	$^4I_{9/2}$
4	↑	↑	↑	↑				2	6	4	$^5I_4$
5	↑	↑	↑	↑	↑			$5/2$	5	$5/2$	$^6H_{5/2}$
6	↑	↑	↑	↑	↑	↑		3	3	0	$^7F_0$
7	↑	↑	↑	↑	↑	↑	↑	$7/2$	0	$7/2$	$^8S_{7/2}$
8	↑↓	↑	↑	↑	↑	↑	↑	3	3	6	$^7F_6$
9	↑↓	↑↓	↑	↑	↑	↑	↑	$5/2$	5	$15/2$	$^6H_{15/2}$
10	↑↓	↑↓	↑↓	↑	↑	↑	↑	2	6	8	$^5I_8$
11	↑↓	↑↓	↑↓	↑↓	↑	↑	↑	$3/2$	6	$15/2$	$^4I_{15/2}$
12	↑↓	↑↓	↑↓	↑↓	↑↓	↑	↑	1	5	6	$^3H_6$
13	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑	$1/2$	3	$7/2$	$^2F_{7/2}$
14	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	0	0	0	$^1S_0$

3rd Hund's rule: For the total angular momentum of the shell,
$\vec{J} = |\vec{L} - \vec{S}|$ in case the shell is less than half filled ($p \leqslant (2l+1)$),
$\vec{J} = \vec{L} + \vec{S}$ in case the shell is more than half filled ($p \geqslant (2l+1)$),

which means

$$J = S|2l - p|\,.$$

Table 2.1 gives the terms for an f shell (*e.g.* 4f, relevant for rare earths), occupied by p electrons.

2.3 Spin-orbit coupling

Experimentally, the magnetic moment of an electron is

$$\vec{m}_s = -g\mu_B \frac{\vec{s}}{\hbar}$$

with $g \approx 2.0023$. This cannot be explained with a classical calculation. A classical estimate of g based on the assumption that an electron is a spinning charged sphere with angular momentum $\vec{s}$ would only lead to

$g = 1$. However, relativistic Dirac quantum theory gives $g = 2$, and the remaining ≈ 0.0023 are due to small corrections arising from the interaction of the electronic charge with the electromagnetic field it generates; this can be calculated precisely with quantum electrodynamics.

We will see now that also for the many-particle states of ions, a relativistic description is necessary. The starting point for the derivation is the non-relativistic $v \ll c$ limit of the Dirac equation which gives rise to the two-component Pauli theory. However, we have to go further in perturbation theory, including terms of the order $\frac{v^2}{c^2}$. This higher order approximation is done formally correctly using the Foldy-Wouthuysen transformation. A static electric potential ϕ of the nucleus is considered. One obtains the following Hamiltonian in spinor space:

$$H = \frac{p^2}{2m} + \phi(\vec{r}) - \frac{p^4}{8m_e^3c^2} + \frac{\hbar^2}{2m_e^2c^2}\frac{1}{r}\frac{\partial\phi}{\partial r}\vec{s}\cdot\vec{l} + \frac{\hbar^2}{8m_e^2c^2}\nabla^2\phi \qquad (2.13)$$

Here, the first two terms are the non-relativistic H_0, the third term is a relativistic correction to the kinetic energy, the fourth term contains the spin-orbit coupling H_{SO}, and the last term is a correction to the potential, known as the Darwin term.

We will now consider the spin-orbit coupling term for the central potential of the nucleus

$$H_{SO} = -\frac{\hbar^2}{2m_e^2c^2}\frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r}\frac{\partial}{\partial r}\frac{1}{r}\vec{s}\cdot\vec{l} = \frac{\hbar^2}{2m_e^2c^2}\frac{Ze^2}{4\pi\epsilon_0}\frac{\vec{s}\cdot\vec{l}}{r^3} = \frac{\mu_0}{4\pi}g\mu_B^2 Z\frac{\vec{s}\cdot\vec{l}}{r^3}\,,$$

where we assume $g = 2$ and use $\mu_B = \frac{e\hbar}{2m_e}$ and $c = \frac{1}{\mu_0\epsilon_0}$.

The operator of spin-orbit coupling for several electrons in an incompletely filled shell is then

$$H_{SO} = \frac{\mu_0}{4\pi}g\mu_B^2 Z\sum_i \frac{\vec{s}_i\cdot\vec{l}_i}{r_i^3}\,. \qquad (2.14)$$

In principle, not only the bare nuclear potential but the full effective potential of the Hartree approximation should be taken into account; however, this is done in practice by replacing the atomic number Z by an effective $Z_{\text{eff}} < Z$. Next, we treat H_{SO} as a weak perturbation to H_0 and evaluate the contribution of the spin-orbit coupling to the energy:

$$E_{SO} := \langle H_{SO}\rangle = \frac{\mu_0}{4\pi}g\mu_B^2 Z\sum_i \left\langle\frac{\vec{s}_i\cdot\vec{l}_i}{r_i^3}\right\rangle\,. \qquad (2.15)$$

For free ions, the radial wave function R_{nl} for all orbitals of the nl shell is the same; therefore

$$E_{SO} = \frac{\mu_0}{4\pi} g\mu_B^2 Z \sum_i \left\langle \frac{1}{r^3} \right\rangle_{nl} \langle \vec{s}_i \cdot \vec{l}_i \rangle . \tag{2.16}$$

Now we call electrons with spin parallel to $\vec{S}$ spin up (↑) and the others spin down (↓). Also, the $\vec{s}_i$ and $\vec{l}_i$ commute. Thus

$$E_{SO} = \frac{\mu_0}{4\pi} g\mu_B^2 Z \sum_i \left\langle \frac{1}{r^3} \right\rangle_{nl} \Bigg(\sum_{\substack{i \\ \text{spin up}}} \left\langle \frac{\vec{S} \cdot \vec{l}_i}{2S} \right\rangle - \sum_{\substack{i \\ \text{spin up}}} \left\langle \frac{\vec{S} \cdot \vec{l}_i}{2S} \right\rangle \Bigg) .$$

We distinguish three cases:

(i) If the shell is less than half filled ($n_{nl} < 2l + 1$), all spins are aligned and the spin down sum doesn't contain any terms:

$$E_{SO} = \frac{\mu_0}{4\pi} g\mu_B^2 Z \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} \langle \vec{S} \cdot \sum_i \vec{l}_i \rangle = \frac{\mu_0}{4\pi} g\mu_B^2 Z \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} \langle \vec{S} \cdot \vec{L} \rangle =: \lambda \langle \vec{S} \cdot \vec{L} \rangle \tag{2.17}$$

with

$$\lambda = \frac{\mu_0}{4\pi} g\mu_B^2 Z \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} .$$

(ii) If the shell is more than half filled ($n_{nl} > 2l + 1$), the spin up sum vanishes because it contains

$$\sum_{m_l=-l}^{l} \langle l m_l | \vec{l} | l m_l \rangle = 0$$

and we obtain

$$E_{SO} = -\frac{\mu_0}{4\pi} g\mu_B^2 Z \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} \langle \vec{S} \cdot \vec{L} \rangle =: \lambda \langle \vec{S} \cdot \vec{L} \rangle \tag{2.18}$$

with

$$\lambda = -\frac{\mu_0}{4\pi} g\mu_B^2 Z \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} .$$

(iii) If the shell is half filled ($n_{nl} = 2l + 1$), both spin up and spin down sums vanish and we get $E_{SO} = 0$. Note that at higher orders in perturbation theory, there is a nonzero contribution.

In summary, the spin-orbit coupling in a free ion behaves, within perturbation theory, like a term $H_{SO} = \lambda\langle \vec{S} \cdot \vec{L} \rangle$ in the Hamiltonian, with $\lambda > 0$ for less than half filled shells and $\lambda < 0$ for more than half filled shells.

As a consequence of spin-orbit coupling H_{SO}, even in the absence of an external magnetic field ($\vec{B}_0 = 0$) $\vec{l}$ and $\vec{s}$ do not commute with the Hamiltonian.

One can show that

$$\left[\vec{l} \cdot \vec{s}, \vec{l}\right]_- = i\hbar\left(\vec{l} \times \vec{s}\right) = -\left[\vec{l} \cdot \vec{s}, \vec{s}\right]_- .$$

On the other hand,

$$\left[\vec{l} \cdot \vec{s}, \vec{j}\right]_- = 0 \quad \text{with } \vec{j} = \vec{l} + \vec{s} .$$

Furthermore,

$$\left[\vec{l} \cdot \vec{s}, \vec{j}^2\right]_- = \left[\vec{l} \cdot \vec{s}, \vec{l}^2\right]_- = \left[\vec{l} \cdot \vec{s}, \vec{s}^2\right]_- = 0 .$$

This means that the energy eigenstates can be classified by j, m_j, l and s (which are good quantum numbers) but not by m_l and m_s. H_{SO} couples, *i.e.* hybridizes the states with different m_l and m_s.

H_{SO} partially lifts the degeneracy of the LS-multiplet (here the doublet as our treatment is for one electron, $j = l \pm 1/2$). Due to

$$\vec{j} = \vec{l} + \vec{s} \curvearrowright 2\left(\vec{l} \cdot \vec{s}\right) = \vec{j}^2 - \vec{l}^2 - \vec{s}^2$$

H_{SO} produces a fine structure of the energy terms

$$E^{(0)}_{nlj} = E^{(0)}_{nl} + \lambda_{nl}\hbar^2[j(j+1) - l(l+1) - s(s+1)] \tag{2.19}$$

with energy $E^{(0)}_{nlj}$ in the absence of spin-orbit coupling. The constant λ_{nl} is

$$\lambda_{nl} = -\frac{e}{2m_e^2c^2}\left\langle nls\middle|\frac{1}{r}\frac{d\phi}{dr}\middle|nls\right\rangle .$$

Thus, terms with $j = l \pm 1/2$ have different energies for $l \neq 0$ while the $2j + 1$-fold degeneracy due to m_j remains.

2.4 Magnetic moments of ions

A problem arises when we want to calculate the magnetic moment of an ion with quantum number S, L, J: Using $g = 2$ for the g factor of the electron, the magnetic moment is

$$\vec{m}_J = \vec{m}_S + \vec{m}_L = -2\mu_B\vec{S} - \mu_B\vec{L} = -\mu_B\left((2\vec{S} + \vec{L}\right) = -\mu_B\left(\vec{J} + \vec{S}\right)$$

But $\vec{m}_J$ does not commute with the Hamiltonian because of the spin-orbit coupling term $\lambda\vec{L}\cdot\vec{S}$ ($\vec{J}$ does commute but $\vec{S}$ does not). Therefore, $\vec{J}$ is a constant of motion but $\vec{m}_J$ is not; we can think of $\vec{S}$ and $\vec{L}$ and thus $\vec{m}_J$ as rotating around the fixed vector $\vec{J}$ (see Figure 2.1).

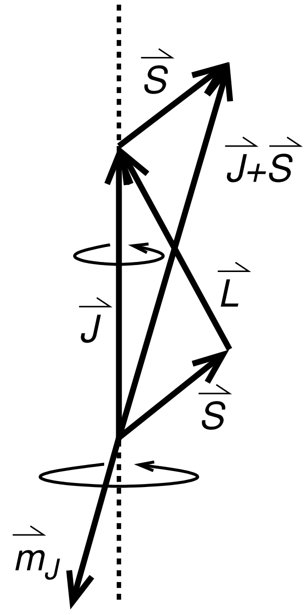


Figure 2.1: Neither $\vec{L}$, $\vec{S}$ nor $\vec{m}_J$ are constants of motion, only $\vec{J}$.

The typical time scale of this rotation should be $\frac{\hbar}{|\lambda|}$. For "slow" experiments like magnetization measurements, only a time averaged $\vec{m}_{\mathrm{obs}}$ will be observable. We can determine it by projecting $\vec{m}_J$ on the direction of the constant $\vec{J}$:

$$\begin{aligned}
\vec{m}_{\mathrm{obs}} &= \frac{(\vec{m}_J\cdot\vec{J})\vec{J}}{\vec{J}\cdot\vec{J}} = -\mu_B\frac{\left[(\vec{J}+\vec{S})\cdot\vec{J}\right]\vec{J}}{\vec{J}\cdot\vec{J}} = -\mu_B\vec{J} - \mu_B\frac{(\vec{S}\cdot\vec{J})\vec{J}}{\vec{J}\cdot\vec{J}} \\
&= -\mu_B\vec{J} + \frac{\mu_B}{2}\frac{\left(\vec{J}-\vec{S}\right)^2 - \vec{J}\cdot\vec{J} - \vec{S}\cdot\vec{S}}{\vec{J}\cdot\vec{J}}\vec{J} \text{ and with } \vec{J}-\vec{S}=\vec{L} \\
\vec{m}_{\mathrm{obs}} &= -\mu_B\vec{J} - \frac{\mu_B}{2}\frac{J(J+1)+S(S+1)-L(L+1)}{J(J+1)}\vec{J} =: -g_J\mu_B\vec{J} \quad (2.20)
\end{aligned}$$

after introducing the Landé g factor

$$g_J = 1 + \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)} \quad (2.21)$$

g_J satisfies $0 \leqslant g_J \leqslant 2$ and can be smaller than the orbital value $g_L = 1$.

3. Diamagnetism

Diamagnetism is defined by $\chi^{\text{dia}} < 0$ where $\chi^{\text{dia}} = \text{constant}$. The simple explanation is that this is an induction effect: The external field induces magnetic dipoles which, according to Lenz's law, are oriented antiparallel to the field and therefore χ is negative. We will see in a moment that strictly speaking, this easily understandable picture is not quite true, as without quantum mechanical effects, there is not even diamagnetism.

Empirically, the effect of diamagnetism is displayed by all materials; however, they are only called diamagnets if no other, stronger type of magnetism like paramagnetism or collective magnetism is present. Examples for diamagnets are almost all organic substances, metals like Bi, Zn and Hg, nonmetals like S, I and Si, and superconductors for $T < T_c$; in fact they are perfect diamagnets: $\chi^{\text{dia}} = -1$ which is called the Meissner-Ochsenfeld effect.

3.1 Bohr-van-Leeuwen theorem

The Bohr-van Leeuwen theorem states: Magnetism is a quantum mechanical effect. Strictly classically, there cannot be either dia-, para- or collective magnetism.

Proof: We assume a solid of identical ions with translational symmetry. Then the magnetization is

$$\vec{M} = \frac{N}{V} \langle \vec{m} \rangle \,,$$

where $\vec{m}$ is the magnetic moment of the individual ion, and N is the number of ions in volume V. The magnetic moment can be related to the energy W of the magnetic system and to the Hamiltonian H:

$$\vec{m} = -\frac{\partial W}{\partial \vec{B}_0} = -\frac{\partial H}{\partial \vec{B}_0} \,, \tag{3.1}$$

where H is the classical Hamiltonian function of a single ion. The classical average can be calculated by

$$\langle \vec{m} \rangle = \frac{1}{Z^*} \int \cdots \int dx_1 \ldots dx_{3N_e} \int dp_1 \ldots dp_{3N_e} \vec{m} e^{-\beta H} \tag{3.2}$$

with the number of electrons N_e and with the classical partition function

$$Z = \frac{Z^*}{N_e! h^{3N_e}}$$

given by

$$Z^* = \int \cdots \int dx_1 \ldots dx_{3N_e} \int dp_1 \ldots dp_{3N_e} e^{-\beta H} \tag{3.3}$$

with the inverse temperature $\beta = \frac{1}{k_B T}$. Then

$$\begin{aligned} \frac{1}{\beta Z} \frac{\partial Z}{\partial \vec{B}_0} &= \frac{1}{\beta Z^*} \frac{\partial Z^*}{\partial \vec{B}_0} \\ &= \frac{1}{\beta Z^*} \int \cdots \int dx_1 \ldots dx_{3N_e} \int dp_1 \ldots dp_{3N_e} e^{-\beta H} \Big(-\beta \frac{\partial H}{\partial \vec{B}_0} \Big) \\ &= \frac{1}{Z^*} \int \cdots \int dx_1 \ldots dx_{3N_e} \int dp_1 \ldots dp_{3N_e} e^{-\beta H} \vec{m} = \langle \vec{m} \rangle \end{aligned} \tag{3.4}$$

Thus,

$$\langle \vec{m} \rangle = \frac{1}{\beta Z} \frac{\partial Z}{\partial \vec{B}_0} . \tag{3.5}$$

If we can show that Z doesn't change when an external field $\vec{B}_0$ is switched on, the theorem is proven. The general form of H in the presence of a magnetic field $\vec{B}_0 = \nabla \times \vec{A}$ is

$$H = \frac{1}{2m} \sum_{i=1}^{N_e} (\vec{p}_i + e\vec{A}_i)^2 + H_1(x_1, \ldots, x_{3N_e}) \tag{3.6}$$

where H_1 represents the electron-electron interactions. Then we can write for the partition function

$$\begin{aligned} Z^* = \int \ldots \int dx_1 \ldots dx_{3N_e} e^{-\beta H_1(x_1, \ldots, x_{3N_e})} \times \\ \times \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dp_1 \ldots dp_{3N_e} e^{-\frac{\beta}{2m} \sum_{i=1}^{N_e} (\vec{p}_i + e\vec{A}_i)^2} \end{aligned} \tag{3.7}$$

As the momentum integration runs from $-\infty$ to ∞, we can substitute $\vec{u}_i = \vec{p}_i + e\vec{A}_i$ without changing the limits of integration:

$$Z^* = \int \ldots \int dx_1 \ldots dx_{3N_e} e^{-\beta H_1(x_1, \ldots, x_{3N_e})} \times$$
$$\times \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dp_1 \ldots dp_{3N_e} e^{-\frac{\beta}{2m} \sum_{i=1}^{N_e} \vec{u}_i^2} \tag{3.8}$$

which obviously is independent of the magnetic field, $Z \neq Z(B_0)$. Thus, the average magnetic moment vanishes in all cases:

$$\langle \vec{m} \rangle \equiv 0 \,.$$

Rigorously, classically, there is no magnetism, and it is best to always argue quantum mechanically.

In matter, we have charged particles in motion which respond to an external magnetic field $\vec{B}_0$. Either the system contains permanent magnetic moments. Then they will order in a field and give rise to collective phenomena like paramagnetism, ferromagnetism, antiferromagnetism or ferrimagnetism. Or, the field itself induces the magnetic moments. This is called diamagnetism and is only observable if no permanent moments are present. We can distinguish between the diamagnetism of insulators which is called Larmor diamagnetism and diamagnetism of itinerant electrons in metals which is called Landau diamagnetism.

3.2 Larmor diamagnetism

As diamagnetism is only observable in a system without other kinds of magnetism, we consider a solid made up out of ions with completely filled shells. For the ground state, we have

$$\vec{J}|0\rangle = \vec{L}|0\rangle = \vec{S}|0\rangle = 0 \tag{3.9}$$

We switch on an external magnetic field given by

$$\vec{B}_0 = \mu_0 \vec{H} = (0, 0, B_0)$$

and look for the response of the system, *i.e.* the field induced magnetic moment, the magnetization. We consider an insulator where all electrons are strictly localized. Then

$$\vec{M}(B_0) = \frac{N}{V} \langle 0 | \vec{m} | 0 \rangle \tag{3.10}$$

for N ions in volume V. We already know that magnetic energies $\approx \mu_{\mathrm{B}} B_0$ are usually so small that the system remains in the ground state $|0\rangle$ in the average. We now need the Hamiltonian in a homogeneous magnetic field from which we obtain the magnetization as $\vec{m} = -\frac{\partial H}{\partial \vec{B}_0}$. We consider only the electrons and ignore the interactions between them and with the nucleus for the time being. We choose the vector potential in the Coulomb gauge so that

$$\vec{B}_0 = \nabla \times \vec{A} \quad \text{with } \nabla \cdot \vec{A} = 0\,. \tag{3.11}$$

This can be achieved by

$$\vec{A} = \frac{1}{2}\vec{B}_0 \times \vec{r}\,.$$

The kinetic energy without field is

$$T_0 = \sum_{i=1}^{n} \frac{p_i^2}{2}$$

for n electrons with charge $-e$ so that $e > 0$. In the field, the canonical momentum $\vec{p}_i$ is different from the mechanical momentum $m\vec{v}_i$:

$$\vec{p}_i = m\vec{v}_i - e\vec{A}(\vec{r}_i)\,. \tag{3.12}$$

Then we have for the kinetic energy

$$T = \frac{1}{2m}\sum_{i=1}^{n} \left(\vec{p}_i + e\vec{A}(\vec{r}_i)\right)^2 = \frac{1}{2m}\sum_{i=1}^{n} \left[\vec{p}_i^{\,2} + e\left(\vec{p}_i \cdot \vec{A}(\vec{r}_i) + \vec{A}(\vec{r}_i) \cdot \vec{p}_i\right) + e^2 \vec{A}^2(\vec{r}_i)\right] \tag{3.13}$$

In general, operators $\vec{p}_i$ and $\vec{A}(\vec{r}_i)$ do not commute, but in Coulomb gauge they do:

$$\vec{p}_i \cdot \vec{A}(\vec{r}_i) = \frac{\hbar}{i}\left(\underbrace{\nabla_i \cdot \vec{A}}_{=0} + \vec{A} \cdot \nabla_i\right) = \vec{A}(\vec{r}_i) \cdot \vec{p}_i\,.$$

Therefore, the kinetic energy becomes

$$T = T_0 + \frac{e}{m}\sum_{i=1}^{n} \vec{A}(\vec{r}_i) \cdot \vec{p}_i + \frac{e^2}{2m}\sum_{i=1}^{n} \vec{A}^2(\vec{r}_i) \tag{3.14}$$

With the field $\vec{B}_0$ in z direction $\vec{B}_0 = (0, 0, B_0)$, the vector potential is

$$\vec{A} = \frac{1}{2}\vec{B}_0 \times \vec{r} = \frac{B_0}{2}(-y, x, 0)\,. \tag{3.15}$$

The scalar product $\vec{A}(\vec{r}_i) \cdot \vec{p}_i$ can be expressed by the orbital angular momentum $\vec{l}_i$ of electron i:

$$\vec{A}(\vec{r}_i) \cdot \vec{p}_i = \frac{B_0}{2}(-y_i p_{ix} + x_i p_{iy}) = \frac{1}{2} B_0 l_{iz} = \frac{1}{2} \vec{B}_0 \cdot \vec{l}_i \,.$$

Then using the total orbital angular momentum $\vec{L} = \sum_{i=1}^{n} \vec{l}_i$, we have

$$T = T_0 + \frac{\mu_B}{\hbar} \vec{L} \cdot \vec{B}_0 + \frac{e^2 B_0^2}{8m} \sum_{i=1}^{n} \left(x_i^2 + y_i^2 \right) . \tag{3.16}$$

In this derivation we have neglected the spin $\vec{S}$ of the electrons; we already know that the field couples not only to $\vec{L}$ but actually to $\vec{J} = \vec{L} + 2\vec{S}$. Thus, two terms in the Hamiltonian contain the magnetic field,

$$H_Z = -\frac{\mu_B}{\hbar}(L_z + 2S_z) B_0 \text{ and } H_{\text{dia}} = \frac{e^2 B_0^2}{8m} \sum_{i=1}^{n} \left(x_i^2 + y_i^2 \right) \tag{3.17}$$

Performing the average for the first term (considering completely filled shells, Equation 3.9) gives

$$\left\langle 0 \middle| \frac{\partial H_Z}{\partial \vec{B}_0} \middle| 0 \right\rangle = 0 \,.$$

We are left with

$$\vec{M}(B_0) = -\frac{N}{V} \left\langle 0 \middle| \frac{\partial H_{\text{dia}}}{\partial \vec{B}_0} \middle| 0 \right\rangle \tag{3.18}$$

Due to the spherical symmetry of the ion (noble gas configuration), we have

$$\sum_{i=1}^{n} \langle 0 | x_i^2 | 0 \rangle = \sum_{i=1}^{n} \langle 0 | y_i^2 | 0 \rangle = \sum_{i=1}^{n} \langle 0 | z_i^2 | 0 \rangle = \frac{1}{3} \sum_{i=1}^{n} \langle 0 | \vec{r}_i^{\,2} | 0 \rangle \,.$$

As x and y components of the magnetization vanish, we find for the z component

$$M(B_0) = -\frac{N e^2}{6mV} B_0 \sum_{i=1}^{n} \langle \vec{r}_i^{\,2} \rangle \,. \tag{3.19}$$

The diamagnetic susceptibility is obtained by differentiating again with respect to B_0:

$$\chi^{\mathrm{dia}} = \mu_0 \left(\frac{\partial M}{\partial \vec{B}_0} \right)_T = -\frac{Ne^2}{6mV} \sum_{i=1}^{n} \langle 0 | \vec{r}_i^{\,2} | 0 \rangle \,, \tag{3.20}$$

where the negative sign indicates diamagnetism. The external field induces a moment whose field is directed opposite to the applied field.
To estimate the order of magnitude of χ^{dia}, we first note that in the literature, usually the molar susceptibility is given:

$$\chi_m^{\mathrm{dia}} = \frac{N_A V}{N} \chi^{\mathrm{dia}} \quad \left[\frac{\mathrm{cm}^3}{\mathrm{mol}} \right] \tag{3.21}$$

with Avogadro number $N_A = 6.022 \cdot 10^{23}\,\mathrm{mol}^{-1}$. The average ion radius

$$\langle r^2 \rangle = \frac{1}{n} \sum_{i=1}^{n} \langle 0 | r_i^2 | 0 \rangle$$

can be expressed in units of the Bohr radius

$$a_B = \frac{4\pi\epsilon_0 \hbar^2}{me^2} = 0.529\,\text{Å}\,.$$

Then the molar susceptibility is

$$\chi_m^{\mathrm{dia}} = -0.995 \cdot 10^{-5} n \left\langle \frac{r^2}{a_B^2} \right\rangle \quad \left[\frac{\mathrm{cm}^3}{\mathrm{mol}} \right], \tag{3.22}$$

where $\left\langle \frac{r^2}{a_B^2} \right\rangle$ is of the order of 1. Thus, χ_m^{dia} is very small and diamagnetism is only observable when it is not shadowed by paramagnetism or collective magnetism.

Table 3.1: Examples of diamagnetic molar susceptibilities χ_m^{dia} in $10^{-6}\mathrm{cm}^3/\mathrm{mol}$.

		He	-1.9	Li^+	-0.7
F^-	-9.4	Ne	-7.2	Na^+	-6.1
Cl^-	-24.2	Ar	-19.4	K^+	-14.6
Br^-	-34.5	Kr	-28.0	Rb^+	-22.0
I^-	-50.6	Xe	-43.0	Cs^+	-35.1

Examples are noble gases and simple ionic crystals like alkali metal halides for which contributions of cations and anions add up.
In Table3.1, the electron number increases within each column, and so does $|\chi_m^{\text{dia}}|$ (see Eq. 3.22). In each row, electron number is the same but the nuclear charge Z increases, increasing the attractive force on the electron shells and thus shrinking the size of the ion $\langle r^2 \rangle$ from left to right; thus, $|\chi_m^{\text{dia}}|$ decreases.

3.3 Landau diamagnetism

There is also a diamagnetic contribution to the susceptibility due to (nearly) free electrons in metals; free electrons in an electron gas lead to charge currents in a magnetic field which generate magnetic moments. This would still be true if the electrons had no spin (no spin magnetic moment). This diamagnetic response, due to the Bohr-van Leeuwen theorem, has to be a quantum-mechanical phenomenon.

3.3.1 Two-dimensional electron gas

As a first step, we consider a two-dimensional electron gas in a uniform magnetic field based on the single-electron Hamiltonian

$$H = \frac{1}{2m}\left[\vec{p} + e\vec{A}(\vec{r})\right]^2 \tag{3.23}$$

with charge $-e$. The Zeeman term which would lead to paramagnetism is neglected here. Without loss of generality, we assume the uniform field to point in z direction, $\vec{B}_0 = (0, 0, B_0)$. We choose the so-called Landau gauge

$$\vec{A}(\vec{r}) = (-B_0 y, 0, 0)$$

which gives

$$\vec{B}_0 = \nabla \times \vec{A} = \begin{pmatrix} \partial_y 0 - \partial_z 0 \\ -\partial_z B_0 y - \partial_x 0 \\ \partial_x 0 + \partial_y B_0 y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix}. \tag{3.24}$$

Thus

$$H = \frac{1}{2m}(p_x - eB_0 y)^2 + \frac{1}{2m}p_y^2 \tag{3.25}$$

we have $[H, p_x] = 0$ as H doesn't contain x; p_x is a constant of motion, and we can replace it by its eigenvalue $\hbar k_x$. If we define $y_0 := \frac{\hbar}{eB_0}k_x$ and

$\omega_c := \frac{eB_0}{m}$ with cyclotron frequency ω_c, we obtain

$$H = \frac{1}{2m}p_y^2 + \frac{m}{2}\omega_c^2(y - y_0)^2\,. \tag{3.26}$$

This Hamiltonian describes a harmonic oscillator with potential minimum shifted to y_0. It has the eigenvalues

$$E_{n,k_x} = \hbar\omega_c\left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, \ldots \tag{3.27}$$

This is a huge degeneracy because the energies do not depend on k_x. The apparent asymmetry between x and y direction in H is gauge dependent and therefore without physical consequence. We could have chosen the vector potential to point in any direction within the xy plane. The isotropy of the two-dimensional space is not broken by the choices of a special gauge. The magnetic field transforms the spectrum of the two-dimensional electron gas

$$E_{\vec{k}} = \frac{\hbar^2}{2m}\left(k_x^2 + k_y^2\right)$$

into a discrete spectrum of Landau levels enumerated by n. For $B_0 = 0$, the density of states is

$$\begin{aligned} D(\varepsilon) &= \int \frac{d^2k}{(2\pi)^2}\delta\left(\varepsilon - \frac{\hbar^2k^2}{2m}\right) = \frac{1}{2\pi}\int_0^\infty dk\, k\delta\left(\varepsilon - \frac{\hbar^2k^2}{2m}\right) \\ &\underbrace{=}_{u=k^2,\, du=2k\,dk} \frac{1}{4\pi}\int_0^\infty du\delta\left(\varepsilon - \frac{\hbar^2u}{2m}\right) = \frac{1}{4\pi}\int_0^\infty dE\frac{2m}{\hbar^2}\delta(\varepsilon - E) \\ &= \frac{m}{2\pi\hbar^2} \quad \text{for } \varepsilon \geqslant 0 \end{aligned} \tag{3.28}$$

Thus, the density of states is constant. For $B_0 > 0$, it is replaced by δ function peaks (Figure 3.1).

We can now determine the degeneracy of the Landau levels. Since the total number of states does not change, the Landau levels must accommodate these states; thus, the degeneracy of the first one, and all others, is (for an electron gas enclosed in a sample with area L^2)

$$N_L = L^2\int_0^{\hbar\omega_c} d\varepsilon\, D(\varepsilon) = \frac{m}{2\pi\hbar^2}\hbar\omega_c L^2 = \frac{m}{2\pi\hbar}\frac{eB_0}{m}L^2 = \frac{eB_0}{h}L^2 \tag{3.29}$$

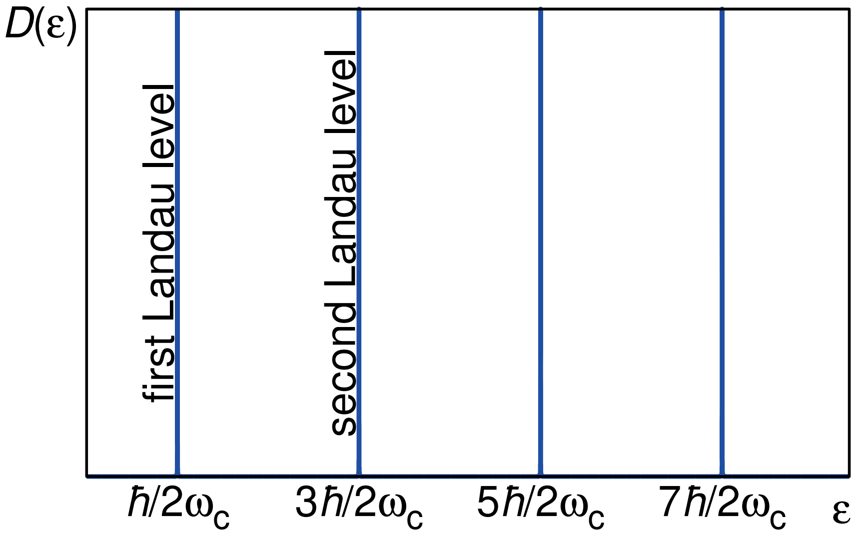


Figure 3.1: Density of states of two-dimensional electron gas at finite magnetic field $B_0 > 0$.

At low temperature, the low energy states are filled successively until all N electrons are accommodated. If

$$N = 2nN_L \quad \text{with } n = 1, 2, \ldots$$

the n lowest Landau levels are completely filled and the others are empty. The factor 2 is due to the two spin directions. In case that $\frac{N}{2N_L}$ is not an integer, the highest Landau level is partially filled. Landau level quantization is one of the key ingredients of the integer quantum Hall effect. For calculating the total energy of N electrons, we define $\lfloor x \rfloor$ as the largest integer smaller or equal to x, and $\underline{n} = \left\lfloor \frac{N}{2N_L} \right\rfloor$; then

$$E = \sum_{n=0}^{\underline{n}-1} 2N_L \hbar\omega_c \left(n + \frac{1}{2}\right) + (N - \underline{n}2N_L)\hbar\omega_c \left(\underline{n} + \frac{1}{2}\right). \tag{3.30}$$

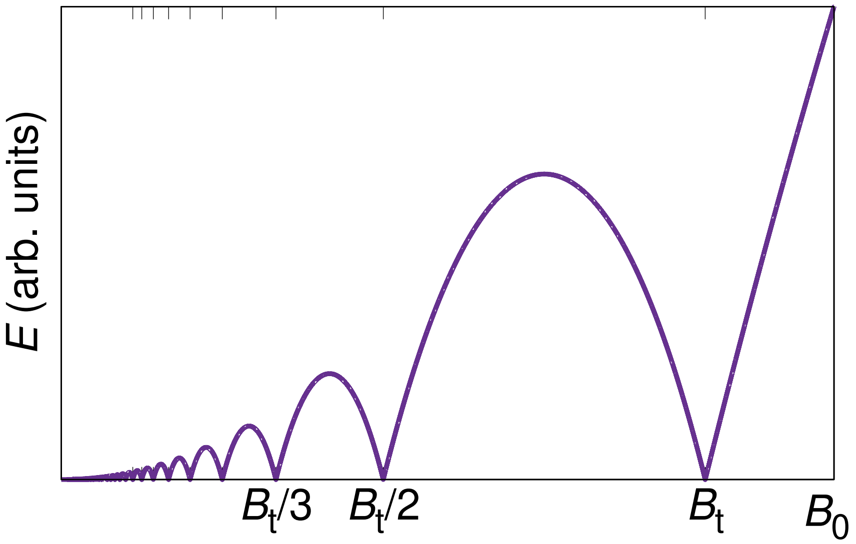
The first term describes the filled Landau levels, and the second the partially filled one. With the filling factor $\nu := \frac{N}{2N_L}$, the energy per electron is

$$\frac{E}{N} = \sum_{n=0}^{\underline{n}-1} \frac{2}{\nu} \hbar\omega_c \left(n + \frac{1}{2}\right) + \left(1 - \frac{2\underline{n}}{\nu}\right) \hbar\omega_c \left(\underline{n} + \frac{1}{2}\right). \tag{3.31}$$

with $\underline{n} = \left\lfloor \frac{\nu}{2} \right\rfloor$. This function is continuous but not everywhere differentiable (see Figure 3.2). B_t is the field for which $\nu = 2$, *i.e.* for which the lowest Landau level is completely filled:

$$\frac{N}{N_L} = \frac{Nh}{eB_tL^2} \overset{!}{=} 2 \curvearrowright B_t = \frac{hN}{2eL^2}$$

Figure 3.2: Energy of the two-dimensional electron gas as function of magnetic field B_0.

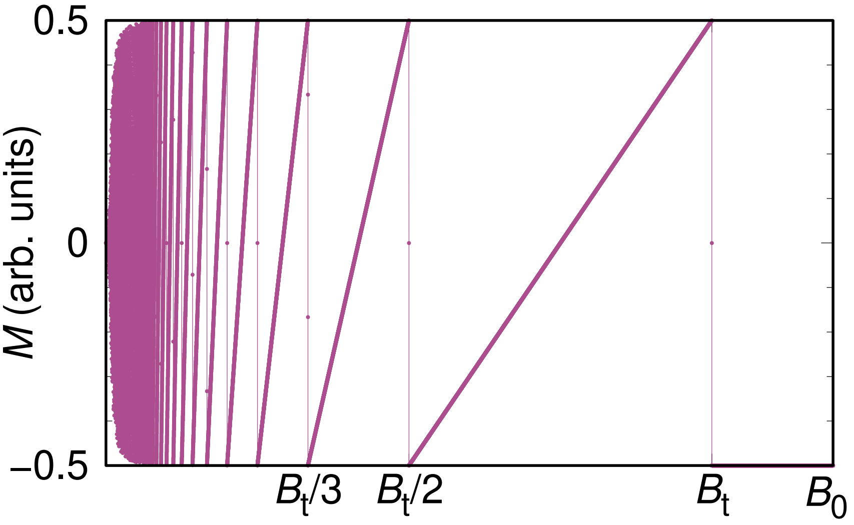


The areal magnetization

$$M = -\frac{1}{L^2}\frac{\partial E}{\partial B_0}$$

shows oscillations that are periodic in $\frac{1}{B_0}$. These are the de Haas-van Alphen oscillations (see Figure 3.3).

Figure 3.3: Magnetization of the two-dimensional electron gas as function of magnetic field B_0.



The limit $\lim_{B_0\to 0} M$ does not exist; neither does the limit $\lim_{B_0\to 0}\chi$ for the susceptibility $\chi = \frac{\partial M}{\partial B_0}$. This unphysical result is due to the assumption of zero temperature $T = 0$. At any $T > 0$, the thermal energy $k_B T$ will be large compared to the energy spacing $\hbar\omega_c$ between the Landau levels for sufficiently small B_0. In this regime, the discreteness of the Landau levels can be neglected, and the energy becomes

$$\begin{aligned}E \overset{N\gg 2N_L}{\cong} \int_0^{\frac{N}{2N_L}} dn\, 2N_L\hbar\omega_c n &= \frac{2eB_0}{h}L^2\frac{\hbar eB_0}{m}\left[\frac{n^2}{2}\right]_0^{\frac{Nh}{2eB_0L^2}} \\ &= \frac{e^2B_0^2}{\pi m}L^2\frac{1}{2}\frac{N^2h^2}{4e^2B_0^2L^4} = \frac{N^2h^2}{8\pi m}\frac{1}{L^2}\,. \end{aligned} \tag{3.32}$$

Thus, the areal magnetization vanishes:

$$M = -\frac{1}{L^2}\frac{\partial E}{\partial B_0} = 0$$

We could have guessed this result: Smearing out the rapid oscillations at small B_0 would lead to $M = 0$. Thus, $\chi = \frac{\partial M}{\partial B_0} = 0$ in this limit; the diamagnetic susceptibility of the two-dimensional electron gas vanishes.

3.3.2 The three-dimensional electron gas

The Hamiltonian of free electrons in three dimensions in the presence of a uniform magnetic field $\vec{B}_0 = B_0\vec{e}_z$ is

$$H = \frac{1}{2m}\left[\vec{p} + e\vec{A}(\vec{r})\right]^2 = \frac{1}{2m}p_y^2 + \frac{m}{2}\omega_c^2(y-y_0)^2 + \frac{1}{2m}p_z^2\,, \tag{3.33}$$

using again the Landau gauge $A = (-B_0y, 0, 0)$. Now we obtain free motion in z direction in addition to shifted harmonic oscillators in the xy plane. The eigenenergies are

$$E_{n,\vec{k}} = \hbar\omega_c\left(n+\frac{1}{2}\right) + \frac{\hbar^2k_z^2}{2m}, \quad n = 0,1,2,\ldots \tag{3.34}$$

The density of states is thus a sum of the densities of states of the one-dimensional electron gas, shifted to the minimum energies $\hbar\omega_c\left(n+\frac{1}{2}\right)$, $n = 0,1,2,\ldots$. The resulting density of states is for one spin direction

$$\begin{aligned}D(\varepsilon) &= \frac{N_L}{L^2}\sum_{n=0}^{\infty}\frac{1}{\pi\hbar}\sqrt{\frac{m}{2\left(\varepsilon-\hbar\omega_c\left(n+\frac{1}{2}\right)\right)}}\,\Theta\left(\varepsilon-\hbar\omega_c\left(n+\frac{1}{2}\right)\right) \\ &= \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\frac{N_L}{L^2}\sum_{n=0}^{\infty}\frac{\Theta\left(\varepsilon-\hbar\omega_c\left(n+\frac{1}{2}\right)\right)}{\sqrt{\varepsilon-\hbar\omega_c\left(n+\frac{1}{2}\right)}} \end{aligned} \tag{3.35}$$

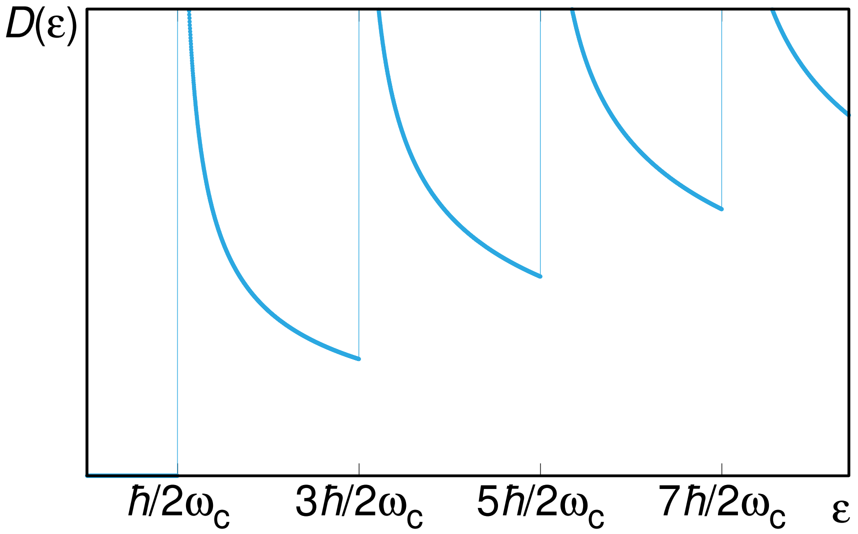


Figure 3.4: Density of states of three-dimensional electron gas at finite magnetic field $B_0 > 0$.

For $k_B T \ll \hbar\omega_c$, the low-energy states are filled up until all electrons are accommodated, as in the case of the two-dimensional electron gas. As before, we expect special features whenever the chemical potential μ reaches $\hbar\omega_c\big(1+\frac{1}{2}\big)$ with $n = 1, 2, \ldots$. Since μ is roughly constant in three dimensions (it has a saw-tooth behavior in two dimensions, with discontinuities at $B_t, B_t/2, B_t/3, \ldots$) while $\hbar\omega_c \propto B_0$, the total energy $E(B_0)$ and therefore the magnetization $M = -\frac{1}{V}\frac{\partial E}{\partial B_0}$ (with volume V) will have features which are periodic in $\frac{1}{B_0}$; these are again de Haas-van Alphen oscillations. They become visible in large magnetic fields.

In the context of conduction electron diamagnetism, we are interested in the susceptibility for small fields. We have $\hbar\omega_c \ll k_B T$, and for a typical metal, we assume $k_B T \ll \mu$. It is useful to define the iterated integrals of the density of states

$$\begin{aligned} P_1(x) &:= \int_0^x d\varepsilon\, D(\varepsilon)\,, \\ P_2(x) &:= \int_0^x d\varepsilon\, P_1(\varepsilon)\,, \end{aligned} \tag{3.36}$$

The total number of electrons is

$$N = 2VP_1(\mu)$$

where the factor 2 accounts for two spin directions. The total energy is

$$E = 2V \int_0^\mu d\varepsilon\, \varepsilon D(\varepsilon)\,.$$

Partial integration gives

$$E = 2V\left[\mu P_1(\mu) - \int_0^{\mu} d\varepsilon\, P_1(\varepsilon)\right] = \mu N - 2VP_2(\mu)\,. \tag{3.37}$$

The explicit expression for $P_2(\mu)$ is

$$P_2(\mu) = \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\frac{N_L}{L^2}\sum_{n=0}^{\infty}\frac{4}{3}\left(\mu - \hbar\omega_c\left(n+\frac{1}{2}\right)\right)^{\frac{3}{2}}\Theta\left(\mu - \hbar\omega_c\left(n+\frac{1}{2}\right)\right)\,. \tag{3.38}$$

Since $\hbar\omega_c \ll \mu$, we can replace the sum over n by a integral. The Poisson summation formula provides the correct expression for this:

$$\sum_{n=0}^{\infty} f\left(n+\frac{1}{2}\right) = \int_0^{\infty} dx\, f(x) + 2\sum_{s=0}^{\infty}(-1)^s\int_0^{\infty} dx\, f(x)\cos(2\pi s x) \tag{3.39}$$

We find

$$\begin{aligned} P_2(\mu) = \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\frac{N_L}{L^2}\frac{4}{3}\Bigg\{ & \int_0^{\frac{\mu}{\hbar\omega_c}-\frac{1}{2}} dx\,(\mu - x\hbar\omega_c)^{\frac{3}{2}} \\ & + 2\sum_{s=1}^{\infty}(-1)^s\int_0^{\frac{\mu}{\hbar\omega_c}-\frac{1}{2}} dx\,(\mu - x\hbar\omega_c)^{\frac{3}{2}}\cos(2\pi s x)\Bigg\} \\ = \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\frac{N_L}{L^2}\frac{4}{3}\Bigg\{ & \frac{2}{5}\frac{\mu^{\frac{5}{2}}}{\hbar\omega_c} - \frac{1}{10\sqrt{2}}(\hbar\omega_c)^{\frac{3}{2}} + 2\sum_{s=1}^{\infty}(-1)^s\frac{3}{8\pi^2}\frac{\hbar\omega_c\sqrt{\mu}}{s^2} \\ & + \text{oscillating terms}\Bigg\} \end{aligned} \tag{3.40}$$

The oscillating terms contain $\cos\left(\frac{2\pi s\mu}{\hbar\omega_c}\right)$ or $\sin\left(\frac{2\pi s\mu}{\hbar\omega_c}\right)$, become rapidly oscillating in the limit $\frac{\mu}{\hbar} \to \infty$ (i.e. $B \to 0$) and are neglected. Then using

$$\sum_{s=0}^{\infty}\frac{(-1)^s}{s^2} = -\frac{\pi^2}{12}$$

we get

$$P_2(\mu) \cong \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\frac{N_L}{L^2}\frac{4}{3}\left\{\frac{2}{5}\frac{\mu^{\frac{5}{2}}}{\hbar\omega_c} - \frac{1}{10\sqrt{2}}(\hbar\omega_c)^{\frac{3}{2}} - \frac{1}{16}\hbar\omega_c\sqrt{\mu}\right\}\,. \tag{3.41}$$

Since $N_L \propto B_0$ and $\hbar\omega_c \propto B_0$, the second term is of order $B_0^{\frac{5}{2}}$ and is not important for the susceptibility in the limit $B_0 \to 0$. The first term is of order B_0^0 and determines the energy in the absence of a magnetic field. Thus

$$E = E(B_0 = 0) + 2V \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}} \frac{N_L}{L^2} \frac{4}{3} \frac{1}{16} \hbar\omega_c \sqrt{\mu} = E(B_0 = 0) + \frac{1}{12\pi^2} V \frac{e^2\sqrt{\mu}}{\sqrt{2m}\hbar} B_0^2 \tag{3.42}$$

and the derivatives are

$$\begin{aligned} M &= -\frac{1}{V}\frac{\partial E}{\partial B_0} = -\frac{1}{6\pi^2}\frac{e^2\sqrt{\mu}}{\sqrt{2m}\hbar} B_0 \,, \\ \chi &= \frac{\partial M}{\partial B_0} = -\frac{1}{6\pi^2}\frac{e^2\sqrt{\mu}}{\sqrt{2m}\hbar} \,. \end{aligned} \tag{3.43}$$

Using the zero field density of states at the chemical potential $D(\mu) = \frac{m}{2\pi^2\hbar^3\sqrt{2m\mu}}$ we get (with $g = 2$)

$$\chi = -\frac{e^2\hbar^2}{6m^2} D(\mu) = -\frac{g^2\mu_B^2}{6} D(\mu) \,. \tag{3.44}$$

We will see later that this value is minus one third the value of the Pauli susceptibility, the paramagnetic response of the free electrons.

4. Paramagnetism

We continue studying the magnetism of materials in the absence of interactions between magnetic moments. Here, we discuss the response of permanent magnetic moments to the external field; this effect is called paramagnetism, and the induced magnetization is parallel to the external field. The permanent moments can originate from partially filled electron shells, for example $3d$ (transition metals), $4f$ (rare earths) or $5f$ (actinides). The moments can also be due to the spin of the itinerant conduction electrons in metals.

4.1 Paramagnetism of the electron gas

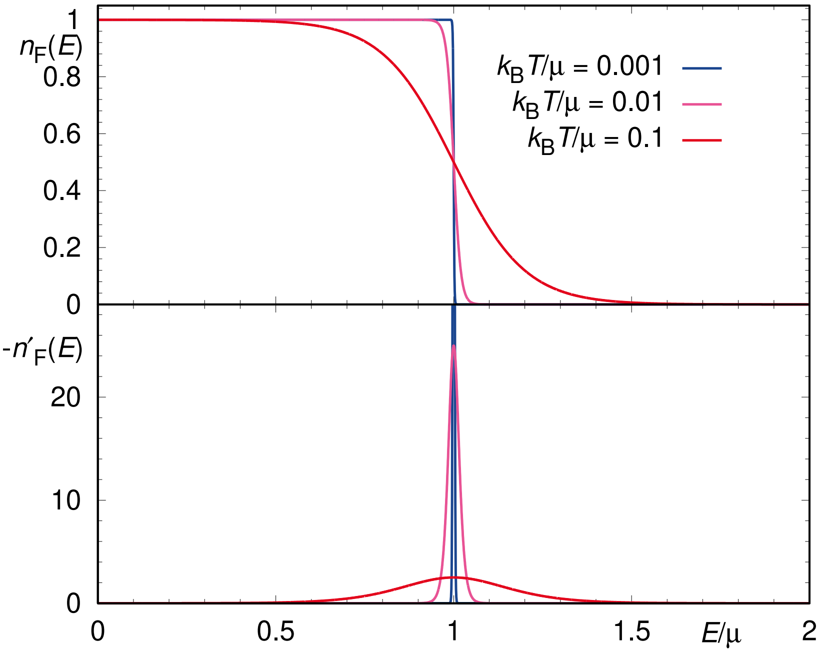


Figure 4.1: Fermi-Dirac distribution function at three different temperatures. Unit of energy is the chemical potential μ.

We already know that electrons possess spin $S = 1/2$ and a spin magnetic moment $m_s = \frac{1}{2} g \mu_B \approx 1 \mu_B$ oriented oppositely to the spin. In a uniform magnetic field, the energy of a free electron is

$$\varepsilon_{\vec{k}\sigma} = \frac{\hbar^2 k^2}{2m} + \sigma \frac{1}{2} g \mu_B B_0 \quad \text{with } \sigma = \uparrow, \downarrow \equiv \pm 1 \tag{4.1}$$

With chemical potential μ, the total energy of the Fermi gas is then

$$E = \sum_{\vec{k}\sigma} \varepsilon_{\vec{k}\sigma} n_F\left(\varepsilon_{\vec{k}\sigma}\right) \tag{4.2}$$

with Fermi-Dirac distribution function (in short, Fermi function)

$$n_F(\varepsilon) = \frac{1}{1 + e^{\frac{\varepsilon - \mu}{k_B T}}} .$$

This is shown in Figure 4.1. The magnetization is given by

$$M = -\frac{g\mu_B}{2V} \sum_{\vec{k}\sigma} \sigma n_F(\varepsilon_{\vec{k}\sigma}) \tag{4.3}$$

and the susceptibility is

$$\begin{aligned} \chi = \left.\frac{\partial M}{\partial B_0}\right|_{B_0=0} &= -\frac{g\mu_B}{2V} \sum_{\vec{k}\sigma} \sigma n_F'\left(\frac{\hbar^2 k^2}{2m}\right) \sigma \frac{g\mu_B}{2} \\ &= -\frac{g^2\mu_B^2}{4V} 2 \sum_{\vec{k}} n_F'\left(\frac{\hbar^2 k^2}{2m}\right) = -\frac{g^2\mu_B^2}{2V} \sum_{\vec{k}} n_F'\left(\frac{\hbar^2 k^2}{2m}\right) \end{aligned} \tag{4.4}$$

$\sum_\sigma$ contributes the factor 2. We can now replace the k summation by an energy integration,

$$\sum_{\vec{k}} \to \int d\varepsilon \, D(\varepsilon) ,$$

using the density of states for one spin direction

$$D(\varepsilon) = \frac{1}{V} \sum_{\vec{k}} \delta\left(\varepsilon - \frac{\hbar^2 k^2}{2m}\right)$$

and find

$$\chi = -\frac{g^2\mu_B^2}{2} \int d\varepsilon \, D(\varepsilon) n_F'(\varepsilon) . \tag{4.5}$$

For $k_B T \ll \mu$, which is typically the case for metals, $n_F(\varepsilon)$ can be approximated by a Heaviside step function $\Theta(\varepsilon)$, and therefore $-n_F'(\varepsilon)$ as a δ function: $n_F'(\varepsilon) \approx -\delta(\varepsilon)$ (see also Figure 4.1). This means

$$\chi = \frac{g^2\mu_B^2 D(\mu)}{2} =: \chi_{\text{Pauli}} . \tag{4.6}$$

This is the so-called Pauli susceptibility which describes the Pauli paramagnetism. The result is valid not only for free electrons but for any kind of dispersion if the appropriate density of states $D(\varepsilon)$ is inserted. As long as $k_B T \ll \mu$, the result is temperature independent. The Pauli susceptibility is related to the Landau susceptibility via

$$\chi_{\text{Landau}} = -\frac{1}{3}\chi_{\text{Pauli}} \,. \tag{4.7}$$

Thus, the total susceptibility of the free electron gas in three dimensions is

$$\chi = \chi_{\text{Pauli}} + \chi_{\text{Landau}} = \frac{2}{3}\chi_{\text{Pauli}} \,. \tag{4.8}$$

Excursion: Sommerfeld expansion

Integrals of the type

$$\int dE\, H(E) n_F(E) \tag{4.9}$$

involving the Fermi function $n_F(E) = \left(e^{\frac{E-\mu}{k_B T}} + 1\right)^{-1}$ often occur in solid state physics, for example in the internal energy of the electron system

$$U = \sum_{l\vec{k}\sigma} \langle n_{l\vec{k}\sigma} \rangle \varepsilon_l(\vec{k}) = N \int dE\, n_F(E) D(E) E \tag{4.10}$$

where $\langle n_{l\vec{k}\sigma} \rangle$ is the average occupation of the single particle states l as function of temperature, given by the Fermi function; $D(E)$ is the density of states per unit cell. The total particle number can be expressed as

$$N_e = \sum_{l\vec{k}\sigma} \langle n_{l\vec{k}\sigma} \rangle = N \int dE\, D(E) n_F(E) \tag{4.11}$$

The Sommerfeld expansion can be applied if the function $H(E)$ is multiply continuously differentiable and integrable and disappears for $E \to -\infty$; then we have (by partial integration)

$$\int_{-\infty}^{\infty} dE\, H(E) n_F(E) = \int_{-\infty}^{\infty} dE\, K(E) \left(-\frac{dn_F}{dE} \right) + \underbrace{\left[K(E) n_F(E) \right]_{-\infty}^{\infty}}_{=0} \tag{4.12}$$

where

$$K(E) = \int_{-\infty}^{E} dE'\, H(E') \tag{4.13}$$

is the antiderivative of $H(E)$. Explicitly, the negative derivative of the Fermi function is

$$-\frac{dn_F}{dE} = \frac{1}{k_b T}\frac{1}{(e^{\frac{E-\mu}{k_B T}}+1)(e^{-\frac{E-\mu}{k_B T}}+1)} \,. \tag{4.14}$$

Examples of this function are shown in Figure 4.1. It is symmetric with respect to the chemical potential μ and falls exponentially on both sides. For $T \to 0$ the derivative becomes a δ function. At finite temperatures, it is only nonzero in a region which grows approximately linearly with temperature. Therefore, the energy integral needs only to be done in a small interval around μ, and the function $K(E)$ is expanded in a Taylor series around μ:

$$K(E) = K(\mu) + \sum_{n=1}^{\infty} \frac{1}{n!}(E-\mu)^n \frac{d^n K(E)}{dE^n}\bigg|_{E=\mu} \tag{4.15}$$

Inserting the integral that we want to calculate, we find

$$\int_{-\infty}^{\infty} dE\, H(E) n_F(E) = \int_{-\infty}^{\mu} dE\, H(E) + \sum_{n=1}^{\infty} \frac{d^{n-1}H(E)}{dE^{n-1}}\bigg|_{E=\mu} \int_{-\infty}^{\infty} dE\, \frac{(E-\mu)^n}{n!}\left(-\frac{dn_F}{dE}\right)$$
$$\text{because } \int_{-\infty}^{\infty} dE \left(-\frac{dn_F}{dE}\right) K(\mu) = \left[-n_F(E)K(\mu)\right]_{-\infty}^{\infty} = K(\mu) \tag{4.16}$$

Thus, the series contains only derivatives of the function we want to integrate, taken at the chemical potential μ as well as integrals that are independent of $H(E)$. As $-\frac{dn_F}{dE}$ is symmetric around μ, integrals over odd powers of $(E-\mu)$ drop out. With the substitution $x = \frac{E-\mu}{k_B T}$ we have

$$\int_{-\infty}^{\infty} dE\, H(E) n_F(E) = \int_{-\infty}^{\mu} dE\, H(E) + \sum_{n=1}^{\infty} a_n (k_B T)^{2n} \frac{d^{2n-1}H(E)}{dE^{2n-1}}\bigg|_{E=\mu}$$
$$\text{with } a_n = \int_{-\infty}^{\infty} dx \frac{x^{2n}}{(2n)!}\frac{1}{(e^x+1)(e^{-x}+1)} \tag{4.17}$$

The a_n can be calculated analytically:

$$a_n = \left(2 - \frac{1}{2^{2(n-1)}}\right)\zeta(2n) \tag{4.18}$$

with the **Riemann zeta function**

$$\zeta(x) = \sum_{m=1}^{\infty} \frac{1}{m^x} = 1 + \frac{1}{2^x} + \frac{1}{3^x} + \dots$$

In particular,

$$a_1 = \zeta(2) = \frac{\pi^2}{6}, \qquad a_2 = \frac{7}{4}\zeta(4) = \frac{7}{4}\frac{\pi^4}{90} = \frac{7\pi^4}{360}$$

Thus, the expansion up to fourth order in $k_B T$ is

$$\begin{aligned}\int_{-\infty}^{\infty} dE\, H(E) n_F(E) = \int_{-\infty}^{\mu} dE\, H(E)& \\ + \frac{\pi^2}{6}(k_B T)^2 \frac{dH(E)}{dE}\bigg|_{E=\mu} &+ \frac{7\pi^4}{360}(k_B T)^4 \frac{d^2H(E)}{dE^2}\bigg|_{E=\mu}\end{aligned} \tag{4.19}$$

Using the Sommerfeld expansion, we find

$$\begin{aligned} Z_e = \frac{N_e}{N} = \int dE\, D(E) n_F(E) = \int_{-\infty}^{\mu} dE\, D(E) + \frac{\pi^2}{6}(k_B T)^2 D'(\mu) + \mathcal{O}(T^4) \\ \frac{U}{N} = \int dE\, n_F(E) D(E) E = \int_{-\infty}^{\mu} dE\, D(E) E + \frac{\pi^2}{6}(k_B T)^2 \big(\mu D'(\mu) + D(\mu)\big) + \mathcal{O}(T^4)\end{aligned} \tag{4.20}$$

4.2 Paramagnetism of localized electrons

Now we discuss the paramagnetism of insulators. Here, the electrons that are responsible for the paramagnetism are strictly localized at fixed lattice points and produce a permanent magnetic moment there. This picture is almost ideally realized in rare earths and their compounds; they are often called 4f systems because of the electron shell that is successively filled. The neutral rare earth atom has the configuration

$$[\mathrm{Xe}]4f^p 6s^2\,,$$

i.e. a noble gas configuration of xenon plus 4f and 6s electrons. In the periodic table, the rare earths start with La (lanthanum) and then add 4f electrons until Lu (lutetium). Usually, the rare earths are trivalent, *i.e.* $(RE)^{3+}$, and they give away the 6s electrons and one 4f electron. The

partially filled 4f shell is situated inside the xenon core and is strongly screened by the fully filled $5s^2 5p^6$ shells which lie outside the xenon core. This strongly localizes the magnetic moments of the 4f electrons to their respective lattice points.
Such a system can be described by a very simple model: We assume that in volume V, there are N identical independent ions; we only care about the magnetic moments produced by these ions. We can assume that intra-atomic correlations are so strong that the localized magnetic moment is determined by the atomic Hund's rules. The temperature and field dependence of the magnetization M is essentially an atomic problem.

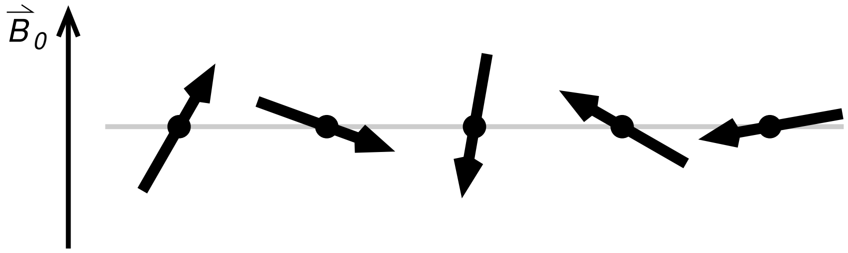


Figure 4.2: Paramagnetic local moments in an external magnetic field.

We now assume LS-coupling which is valid for not too strong spin-orbit coupling. The distances between different LS multiplets are so large that transitions among them are improbable and we can assume that quantum numbers L and S belonging to the squares L^2 and S^2 of the total angular momentum and total spin quantum numbers are good quantum numbers. The magnetic moment $\vec{m}_j$ of the ion at lattice site j is given by

$$\vec{m}_j = -\frac{\mu_B}{\hbar}\left(\vec{L}_j + 2\vec{S}_j\right) = -\frac{\mu_B}{\hbar}\left(\vec{J}_j + \vec{S}_j\right). \tag{4.21}$$

In the following, we disregard the minus sign in the definition of the magnetic moment. The model Hamiltonian is

$$H = \sum_{j=1}^{N}\left(H_0^{(j)} + H_{SO}^{(j)} - \vec{m}_j \cdot \vec{B}_0\right) = \sum_{j=1}^{N} H_1^{(j)} \tag{4.22}$$

As we are restricting the discussion to a single LS multiplet, $H_0^{(j)}$ which determines the coarse structure of the terms is not important. The spin orbit coupling of each ion $H_{SO}^{(j)}$ determines the fine structure of the terms. The last expression in the bracket represents the Zeeman energy, and its relative strength in comparison to $H_{SO}^{(j)}$ is decisive for the calculation of the magnetization

$$\vec{M} = n\langle m\rangle \quad \text{with } n = \frac{N}{V}.$$

Here, the bracket $\langle \ldots \rangle$ means quantum mechanical expectation value and thermal averaging; in general

$$\langle A \rangle = \frac{1}{Z} \mathrm{Tr}\Big(A e^{-\beta H} \Big)$$

with canonical partition function $Z = \mathrm{Tr}\Big(e^{-\beta H} \Big) = Z_1^N$ which factorizes for the model (4.22) into a product of single-particle partition functions

$$Z_1 = \mathrm{Tr}\Big(e^{-\beta H_1} \Big)$$

because the moments do not interact. β is the inverse temperature $\beta = \frac{1}{k_B T}$. If we now apply a homogeneous magnetic field $\vec{B}_0$ in z direction, the x and y components of the magnetization vanish, and we find for $M_z = M$

$$M = n \frac{1}{Z_1} \mathrm{Tr}\Big(m e^{-\beta H_1} \Big) = k_B T n \frac{\partial}{\partial B_0} \ln Z_1 \tag{4.23}$$

Determining the single-particle partition function Z_1 already solves the problem; however, we can calculate the trace only for particular situations. M is influenced by the thermal energy $k_B T$, by the spin-orbit interaction $H_{SO} = \Lambda(\gamma, LS) \vec{L} \cdot \vec{S}$ (γ stands for other quantum numbers) and the magnetic field $H_z = -\frac{\mu_B}{\hbar}(J_z + S_z) B_0$. Only if there are orders of magnitude between these three terms can the partition function Z_1 be determined easily.

Weak spin-orbit interaction.- This is the case $\hbar^2 \Lambda(\gamma, LS) \vec{L} \ll k_B T$. Furthermore, we have to distinguish small and large field compared to the temperature of interest.

a) $\hbar^2 \Lambda(\gamma, LS) \vec{L} \ll k_B T, \mu_B B_0$

We can assume that we have the so-called normal Zeeman effect

$$E_{\gamma LSJM_J} = E^{(0)}_{\gamma LS} - (M_L + 2M_S) \mu_B B_0 \,,$$

and all terms of the LS multiplet are occupied with almost equal probability. M_L and M_S are "still good" quantum numbers, but J is not a good quantum number. $E^{(0)}_{\gamma LS}$ are eigenenergies of H_0 without field and without spin-orbit coupling. We calculate the partition function in energy representation

$$Z_1 = e^{-\beta E^{(0)}_{\gamma LS}} \sum_{M_L=-L}^{+L} \sum_{M_S=-S}^{+S} e^{\beta \mu_B B_0 (M_L + 2M_S)} \tag{4.24}$$

We focus on the lowest LS multiplet as the prefactor will be very small for the higher ones. With the notation $b = \beta\mu_B B_0 > 0$ we calculate

$$\begin{aligned}\sum_{M_L=-L}^{+L} e^{bM_L} &= e^{bL}\sum_{n=0}^{2L}(e^{-b})^n = e^{bL}\frac{1-e^{-b(2L+1)}}{1-e^{-b}} \\ &= e^{bL}\frac{e^{b/2}-e^{-b/2-2Lb}}{e^{b/2}-e^{-b/2}} = \frac{e^{b(L+1/2)}-e^{-b(L+1/2)}}{e^{b/2}-e^{-b/2}} \\ &= \frac{\sinh\big(b(L+1/2)\big)}{\sinh\left(\frac{b}{2}\right)} \end{aligned} \tag{4.25}$$

The same calculation also gives

$$\sum_{M_S=-S}^{+S} e^{bM_S} = \frac{\sinh\big(b(2S+1)\big)}{\sinh b}\,. \tag{4.26}$$

Therefore, the partition function is

$$Z_1 = e^{-\beta E^{(0)}_{\gamma LS}}\frac{\sinh\big(\beta\mu_B B_0(L+1/2)\big)}{\sinh\left(\frac{1}{2}\beta\mu_B B_0\right)}\frac{\sinh\big(\beta\mu_B B_0(2S+1)\big)}{\sinh(\beta\mu_B B_0)}\,. \tag{4.27}$$

With the magnetization of a paramagnet (4.23), we now need to differentiate the partition function with respect to the field:

$$\frac{\partial}{\partial B_0}\ln Z_1 = \frac{1}{Z_1^{(L)}}\frac{\partial Z_1^{(L)}}{\partial B_0} + \frac{1}{Z_1^{(S)}}\frac{\partial Z_1^{(S)}}{\partial B_0}\,. \tag{4.28}$$

We do the first term explicitly:

$$\begin{aligned}\frac{1}{Z_1^{(L)}}\frac{\partial Z_1^{(L)}}{\partial B_0} &= \frac{\sinh\left(\frac{b}{2}\right)}{\sinh\big(b(L+1/2)\big)}\left\{\frac{\sinh\left(\frac{b}{2}\right)\beta\mu_B(L+1/2)\cosh\big(b(L+1/2)\big)}{\sinh^2\left(\frac{b}{2}\right)}-\right. \\ &\qquad\left. - \frac{\sinh\big(b(L+1/2)\big)\frac{1}{2}\beta\mu_B\cosh\left(\frac{b}{2}\right)}{\sinh^2\left(\frac{b}{2}\right)}\right\} \\ &= \beta\mu_B(L+1/2)\coth\big(b(L+1/2)\big) - \frac{1}{2}\beta\mu_B\coth\left(\frac{b}{2}\right)\end{aligned} \tag{4.29}$$

Now we introduce a function which is central to the theory of magnetism, the Brillouin function:

$$B_D(x) = \frac{2D+1}{2D}\coth\left(\frac{2D+1}{2D}x\right) - \frac{1}{2D}\coth\left(\frac{x}{2D}\right) \tag{4.30}$$

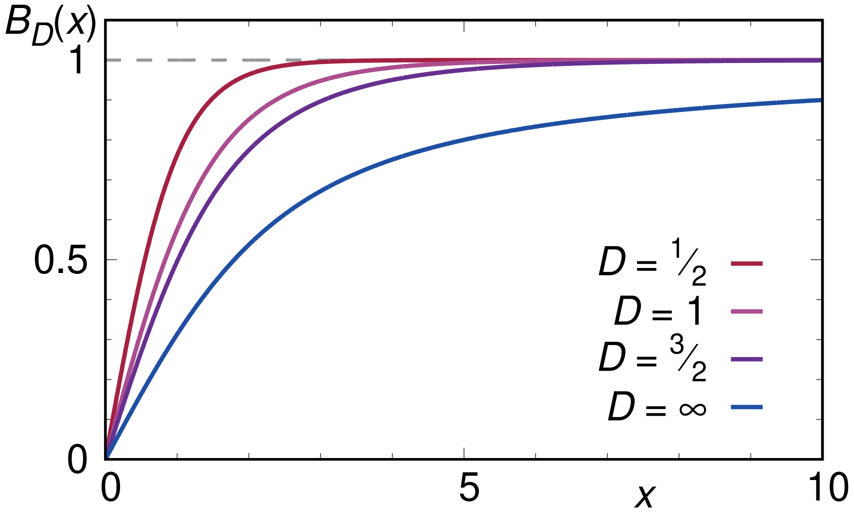


Figure 4.3: Brillouin function $B_D(x)$ as function of x for several parameters D.

Using this function, we can write the magnetization of Eq. (4.23) in the form

$$M(T, B_0) = n\mu_B\big[LB_L(\beta\mu_B B_0 L) + 2SB_S(2\beta\mu_B B_0 S)\big] \tag{4.31}$$

A few general properties of the Brillouin function are:

1. $D = 1/2$: In this special case,
$$B_{1/2}(x) = \tanh x\,.$$
2. $D \to \infty$: In this limit, the Brillouin function reduces to the Langevin function $L(x)$:
$$B_{D\to\infty}(x) = L(x) = \coth x - \frac{1}{x}\,.$$
This function appears in the classical treatment of paramagnetism if the quantization of the orbital angular momentum is neglected (moments can assume any angle in space).
3. Small x: By expanding $\coth x$, one finds
$$B_D(x) = \frac{D+1}{3D}x - \frac{D+1}{3D}\frac{2D^2+2D+1}{30D^2}x^3 + \dots\,, \tag{4.32}$$
and, in particular
$$B_D(0) = 0\,.$$
Due to this property, according to (4.31), the magnetization of a paramagnet is zero if either $B_0 = 0$ or $T = \infty$. Physically, this means that a paramagnet does not possess spontaneous magnetization.

4. $B_D(-x) = -B_D(x)$: This means that if the direction of the magnetic field is reversed, the direction of the magnetization is also reversed.

5. $\lim_{x\to\infty} B_D(x) = 1$: The magnetization shows saturation for $B_0 \to \infty$ or for $T \to 0$. This means that all the moments are oriented parallel to the field: $M \to M_0 = n\mu_B(L{+}2S)$. The high temperature behavior of the magnetization is interesting. With the precondition we are discussing here $\hbar^2\Lambda(\gamma, LS) \ll \mu_B B_0 \ll k_B T$ or $\beta\mu_B B_0 \ll 1$, the argument of the Brillouin function is small, and the expansion (4.32) can be terminated after the linear term:

$$M \approx \frac{n\mu_0\mu_B^2}{3k_BT} B_0\big[L(L+1) + 4S(S+1)\big] \tag{4.33}$$

The susceptibility

$$\chi = \mu_0\left(\frac{\partial M}{\partial B_0}\right)_T$$

then shows a characteristic $\frac{1}{T}$ behavior, which is called the Curie law:

$$\chi(T) = \frac{C_1}{T}\,. \tag{4.34}$$

C_1 is the so-called Curie constant which is given by

$$C_1 = \frac{n\mu_0\mu_B^2}{3k_B}\big[L(L+1) + 4S(S+1)\big]\,.$$

A purely classical calculation would have given a similar high-temperature behavior:

$$\chi_{\text{cl}}(T) = \frac{C_{\text{cl}}}{T}, \quad C_{\text{cl}} = \frac{n\mu_0\mu^2}{3k_B}\,, \tag{4.35}$$

where μ is the magnetic moment. In analogy, one therefore defines

$$\mu_{\text{eff}} = \mu_B p_{\text{eff}}, \quad p_{\text{eff}} = \sqrt{L(L+1) + 4S(S+1)} \tag{4.36}$$

with the effective number of Bohr magnetons p_{eff}.

b) So far, we assumed that both thermal energy and field energy are large compared to the spin-orbit interaction. If we now allow the spin-orbit coupling energy to be of the same order of magnitude as the magnetic energy but still smaller than the thermal energy

$$\hbar^2\Lambda(\gamma, LS), \mu_B B_0 \ll k_B T$$

the calculation becomes more complicated because the spin-orbit coupling term enters the partition function. However, after considering all terms, the result for the magnetization is

$$M = \frac{n\mu_0\mu_B^2}{3k_BT}B_0\big[L(L+1) + 4S(S+1)\big] \tag{4.37}$$

which is the same as the high temperature limit of case a), Eq. (4.33).

Strong spin-orbit interaction.- $\hbar^2\Lambda(\gamma, LS) \gg k_BT, \mu_BB_0$. This case is different; it is often discussed as Langevin paramagnetism and occurs for $4f$ systems in moderate fields. Rather than a uniform distribution over the fine structure terms of the LS multiplets, only the lowest term is occupied to a certain degree. J is still a "good" quantum number, and non-diagonal terms of S^z only play a marginal role; this region is called the anomalous Zeeman effect, and the energies are

$$E_{\gamma LSJM_j} = E^{(0)}_{\gamma LSJ} + g_J(L, S)M_J\mu_BB_0 \tag{4.38}$$

with Landé g factor from Eq. (2.21). Then, the partition function is

$$Z_1 = e^{-\beta E^{(0)}_{\gamma LSJ}} \sum_{M_J=-J}^{+J} e^{-\beta g_J M_J \mu_B B_0}\,. \tag{4.39}$$

Only the energetically most favorable J value, according to Hunds' third rule, has to be taken into account. The partition function is calculated as shown above for weak spin-orbit coupling, Eq. (4.27):

$$Z_1 = e^{-\beta E^{(0)}_{\gamma LSJ}} \frac{\sinh\big(\beta g_J\mu_BB_0(J + 1/2)\big)}{\sinh\big(\frac{1}{2}\beta g_J\mu_BB_0\big)} \tag{4.40}$$

This gives us the magnetization

$$M = M_0B_J\big(\beta g_J J\mu_BB_0\big) \tag{4.41}$$

with saturation magnetization

$$M_0 = nJg_J\mu_B\,. \tag{4.42}$$

As before, the susceptibility follows from differentiating with respect to the field B_0. The high temperature behavior is again the Curie law (for $\beta\mu_BB_0 \ll 1$)

$$\chi = \frac{C}{T} \quad \text{with } C = n\mu_0\frac{p_{\text{eff}}^2}{3k_B}\mu_B^2 \tag{4.43}$$

where in the Curie constant C now a different effective magneton number is found:

$$\mathsf{p}_{\text{eff}} = \mathsf{g}_{\mathrm{J}}\sqrt{\mathsf{J}(\mathsf{J}+1)}\,. \tag{4.44}$$

The Curie law is experimentally very well confirmed, and the order of magnitude of χ_{Langevin} is normally much larger than the Pauli magnetism of the conduction electrons:

$$\frac{\chi_{\text{Pauli}}}{\chi_{\text{Langevin}}} = \frac{9}{2}\frac{1}{\mathsf{g}_{\mathrm{J}}^2\mathsf{J}(\mathsf{J}+1)}\frac{\mathsf{k}_{\mathrm{B}}\mathsf{T}}{\epsilon_{\mathrm{F}}} \tag{4.45}$$

5. Exchange interaction

The dipole interaction between the magnetic moments of the electrons is much too weak for supporting magnetic order of materials at high temperatures. For explaining the observed magnetism, we need to find a strong interaction between the electrons. We might think that an interaction depending explicitly on the spin (the magnetic moments) of the electrons is needed; no such interaction is known though. In 1928, Werner Heisenberg realized that the responsible interaction is the Coulomb repulsion between electrons; this is a strong interaction but does not explicitly depend on the spin. Spin selectivity is due to quantum mechanics, in particular the Pauli principle: Two electrons with parallel or antiparallel spin behave differently even though the fundamental interaction is the same; the spatial wave function $\psi(\vec{r}_1, \vec{r}_2)$ has to be antisymmetric and symmetric, respectively. One consequence is that two electrons with parallel spin cannot be in the same place. In order to discuss how the Coulomb interaction term leads to an exchange interaction of the spins, we first write this interaction in second quantization; we will introduce this very useful representation first.

5.1 Occupation number representation for fermions

So far, we have discussed the Hamiltonian in so-called first quantization:

$$H = H_0 + H_1 \, , H_0 = \sum_{i=1}^{N_e} h_i = \sum_{i=1}^{N_e} \frac{\vec{P}_i^2}{2m} + \sum_{i=1}^{N_e} V(\vec{r}_i) \, , H_1 = \sum_{i<j} u(\vec{r}_i - \vec{r}_j) \tag{5.1}$$

where $u(\vec{r}_i - \vec{r}_j)$ is either the bare Coulomb potential

$$u(\vec{r}_i - \vec{r}_j) = \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \tag{5.2}$$

or an effective, screened interaction. First quantization implies that the antisymmetry of the wave function has to be taken into account by working with Slater determinants which is rather cumbersome. Therefore, many-body calculations in solid state theory are usually performed in so-called second quantization, using the occupation number representation.

Slater determinant

Let's assume that we can solve the single particle problem exactly, *i.e.* for electron i we have

$$h_i|k_\alpha\rangle^{(i)} = \varepsilon_{k_\alpha}|k_\alpha\rangle^{(i)} \tag{5.3}$$

or in position representation

$$h_i\varphi_{k_\alpha}(\vec{r}_i) = \left(\frac{\vec{p}_i^{\,2}}{2m} + V(\vec{r}_i)\right)\varphi_{k_\alpha}(\vec{r}_i) = \varepsilon_{k_\alpha}\varphi_{k_\alpha}(\vec{r}_i) \tag{5.4}$$

with a complete set of single particle quantum numbers $k_\alpha = (l, \vec{k}, \sigma)$ of Bloch states. The Pauli principle now implies that only the part of the product space of N_e single particle Hilbert spaces is realized which consists of the particle indices of totally antisymmetric wave functions. A basis of the N_e particle Hilbert space are the Slater determinants which we can compose of single particle states as follows:

$$\begin{aligned}|\psi_{k_1,\cdots,k_{N_e}}(1\cdots N_e)\rangle &= \frac{1}{\sqrt{N_e!}}\sum_{P\in S_{N_e}}(-1)^{\chi_P}|k_{P(1)}\rangle^{(1)}\cdots|k_{P(N_e)}\rangle^{(N_e)}\\ &= \frac{1}{\sqrt{N_e!}}\det\begin{pmatrix}|k_1\rangle^{(1)} & \cdots & |k_1\rangle^{(N_e)}\\ \vdots & & \vdots\\ |k_{N_e}\rangle^{(1)} & \cdots & |k_{N_e}\rangle^{(N_e)}\end{pmatrix} = \frac{1}{\sqrt{N_e!}}\det\left(|k_\alpha\rangle^{(i)}\right)\end{aligned} \tag{5.5}$$

where P are the elements of the permutation group S_{N_e} of N_e elements, and χ_P is the character of the permutation (number of transpositions, which lead to the permutation). The product state

$$|k_1\rangle^{(1)}|k_2\rangle^{(2)}\cdots|k_{N_e}\rangle^{(N_e)}$$

means that particle 1 is in state k_1, particle 2 in state k_2 and so on; but as the particles are not distinguishable, it has to be irrelevant which particle is in state k_1, k_2 etc. Therefore, we have to sum over all possible permutations. The Slater determinants are a suitable basis for the N_e particle Hilbert space $\mathcal{H}_A(N_e)$, even if not all states of this Hilbert space correspond to a single Slater determinant.

Fock space

The basis of $\mathcal{H}_A(N_e)$ which is described by Slater determinants can also be written down in occupation number representation by writing down how many of the indistinguishable N_e particles are in state k_α; however, the sum over all occupation numbers has to be N_e. We can get rid of this restriction if we do not work in the N_e particle Hilbert space but instead in Fock space

$$\mathcal{H}_{A,\mathsf{Fock}} = \mathcal{H}_A(0) \oplus \mathcal{H}_A(1) \oplus \cdots \oplus \mathcal{H}_A(N_e) \oplus \cdots \tag{5.6}$$

which is defined as direct sum over the Hilbert spaces for all possible particle numbers. If we allow an arbitrary number of (identical) particles in the Hilbert space, then this product space is called Fock space. We can now define operators that "ascend" and "descend" between segments of Fock space with different particle numbers. These operators create and annihilate particles; therefore, they are called creation and annihilation operators. They play a central role in many serious calculations within quantum mechanics. Fock space is always explicitly or implicitly used for grand canonical treatments. In the following, we note the most important relations for fermions and bosons; therefore, we use N for the number of particles.

Starting point is the representation of N particle states. Let's assume a discrete, ordered single particle basis is given: $|1\rangle$, $|2\rangle$, …, where i in $|i\rangle$ stands for a set of single particle quantum numbers $(lk\sigma)_i$. The normalization is $\langle i|j\rangle = \delta_{ij}$. All N particle states can be represented by superposition of

$$P_\pm\big(|r_1\rangle|r_2\rangle\dots|r_N\rangle\big) \tag{5.7}$$

where P_+ symmetrizes for bosons and P_- antisymmetrizes for fermions. Explicitly, we have

$$P_-\big(|r_1\rangle|r_2\rangle\dots|r_N\rangle\big) = \frac{1}{\sqrt{N!}}\sum_{P\in S_N}(-1)^P|r_{P(1)}\rangle|r_{P(2)}\rangle\dots|r_{P(N)}\rangle \tag{5.8}$$

and

$$P_+\big(|r_1\rangle|r_2\rangle\dots|r_N\rangle\big) = \frac{1}{\sqrt{N!n_1!n_2!\dots}}\sum_{P\in S_N}|r_{P(1)}\rangle|r_{P(2)}\rangle\dots|r_{P(N)}\rangle \tag{5.9}$$

where P runs over all permutations and n_i is the number of single particle states $|i\rangle$ in the product.

An equivalent characterization of the (basis) states is possible in occupation number representation:

$$\big|\{n\}\big\rangle \equiv |n_1, n_2, \dots\rangle := P_\pm\big(\underbrace{|1\rangle\dots|1\rangle}_{n_1 \text{ times}}\underbrace{|2\rangle\dots|2\rangle}_{n_2 \text{ times}}\dots\big) \tag{5.10}$$

(Obviously, for fermions we have $n_i \in \{0,1\}$). $\big|\{n\}\big\rangle$ thus stands for a complete set of occupation numbers for all single particle states.
A further step for the efficient representation is the introduction of particle creation and annihilation operators: $c_i^\dagger$, c_i. We do this for fermions here and cite the result for bosons below. We define c_i and $c_i^\dagger$ by their effect on the basis states as follows:

$$\begin{aligned} c_i\big|\{n\}\big\rangle = c_i|\dots n_i\dots\rangle &= (-1)^{\sum_{j=1}^{i-1} n_j}|\dots(n_i-1)\dots\rangle \\ c_i^\dagger\big|\{n\}\big\rangle = c_i^\dagger|\dots n_i\dots\rangle &= (-1)^{\sum_{j=1}^{i-1} n_j}|\dots(n_i+1)\dots\rangle \end{aligned} \tag{5.11}$$

Concerning the notation, $c_i^\dagger$ is indeed the adjunct operator for c_i:

$$\begin{aligned} \big\langle\{m\}\big|c_i\big|\{n\}\big\rangle &= \begin{cases} (-1)^{\sum_{j<i} n_j}, & \text{if } m_i = n_i - 1\,, m_j = n_j \text{ for } j \neq i\,, \\ 0, & \text{otherwise} \end{cases} \\ \big\langle\{n\}\big|c_i^\dagger\big|\{m\}\big\rangle &= \begin{cases} (-1)^{\sum_{j<i} m_j}, & \text{if } n_i = m_i + 1\,, n_j = m_j \text{ for } j \neq i\,, \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{5.12}$$

Remark: c_i maps totally antisymmetric N particle states to totally antisymmetric $(N-1)$ particle states; $c_i^\dagger$ acts “in the opposite way”: Creation and annihilation, respectively.
For the particle operators the important anticommutation relations are valid:

$$[c_i, c_j]_+ = c_i c_j + c_j c_i = 0 \tag{5.13}$$

and also

$$[c_i^\dagger, c_j^\dagger]_+ = c_i^\dagger c_j^\dagger + c_j^\dagger c_i^\dagger = 0 \tag{5.14}$$

Verification: we assume $i < j$ (otherwise, the labels can be exchanged, $i = j$ is obvious):

$$\begin{aligned} c_i c_j\big|\{n\}\big\rangle &= c_i(-1)^{\nu_j}|\dots(n_j-1)\dots\rangle \\ &= (-1)^{\nu_i+\nu_j}|\dots(n_i-1)\dots(n_j-1)\dots\rangle \\ c_j c_i\big|\{n\}\big\rangle &= c_j(-1)^{\nu_i}|\dots(n_i-1)\dots\rangle \end{aligned}$$

$$= \ (-1)^{\nu_j-1+\nu_i}|\ldots(n_i-1)\ldots(n_j-1)\ldots\rangle \tag{5.15}$$

Here we abbreviated $\nu_i = \sum_{j=1}^{i-1} n_j$. Thus, the second equation of (5.15) has an extra minus sign, and if we add both equations, we find $c_i c_j + c_j c_i = 0$. Furthermore,

$$[c_i, c_j^\dagger]_+ = \delta_{ij} \tag{5.16}$$

Justification: We assume $i < j$ (otherwise, we again relabel); as before, we find

$$\begin{aligned} c_i c_j^\dagger \big|\{n\}\big\rangle &= (-1)^{\nu_i+\nu_j}|\ldots(n_i-1)\ldots(n_j+1)\ldots\rangle \\ c_j^\dagger c_i \big|\{n\}\big\rangle &= (-1)^{\nu_j-1+\nu_i}|\ldots(n_i-1)\ldots(n_j+1)\ldots\rangle \end{aligned} \tag{5.17}$$

i.e. for $i \neq j$ c_i and $c_j^\dagger$ anticommute. Now let $i = j$:

$$\begin{aligned} c_i c_i^\dagger \big|\{n\}\big\rangle &= \begin{cases} \big|\{n\}\big\rangle, & \text{if } n_i = 0\,, \\ 0, & \text{if } n_i = 1\,, \end{cases} \\ c_i^\dagger c_i \big|\{n\}\big\rangle &= \begin{cases} 0, & \text{if } n_i = 0\,, \\ \big|\{n\}\big\rangle, & \text{if } n_i = 1\,, \end{cases} \end{aligned} \tag{5.18}$$

From the sum of these two equations, we obtain

$$(c_i c_i^\dagger + c_i^\dagger c_i)|\{n\}\rangle = \big|\{n\}\big\rangle \curvearrowright [c_i, c_i^\dagger]_+ = 1\,.$$

If we now define the vacuum as

$$|0\rangle := |00\ldots0\ldots\rangle \tag{5.19}$$

then we have

$$\big|\{n\}\big\rangle = (c_1^\dagger)^{n_1}(c_2^\dagger)^{n_2}\ldots|0\rangle, \tag{5.20}$$

so that for N particle states, we obtain

$$c_{r_1}^\dagger c_{r_2}^\dagger \ldots c_{r_N}^\dagger|0\rangle, \tag{5.21}$$

Therefore, we have the correspondence, for example for two particles

$$\big|\psi_{k_1k_2}(12)\big\rangle = \frac{1}{\sqrt{2}}\begin{vmatrix} |k_1\rangle^{(1)} & |k_1\rangle^{(2)} \\ |k_2\rangle^{(1)} & |k_2\rangle^{(2)} \end{vmatrix} \ \leftrightarrow\ c_1^\dagger c_2^\dagger|0\rangle \tag{5.22}$$

$$\big|\psi_{k_1k_2}(12)\big\rangle = -\big|\psi_{k_2k_1}(12)\big\rangle \ \leftrightarrow\ c_2^\dagger c_1^\dagger|0\rangle = -c_1^\dagger c_2^\dagger|0\rangle \tag{5.23}$$

Both for the Slater determinant in first quantization and in second quantization, the antisymmetry of the wave function is guaranteed.
In particular for fermions, this has the consequence

$$c_1^\dagger c_1^\dagger|0\rangle = -c_1^\dagger c_1^\dagger|0\rangle = 0\,,$$

two fermions cannot have the same quantum numbers.
For bosons, we briefly note the definitions

$$\begin{aligned} c_i|\ldots n_i\ldots\rangle &= \sqrt{n_i}|\ldots(n_i-1)\ldots\rangle \\ c_i^\dagger|\ldots n_i\ldots\rangle &= \sqrt{n_i+1}|\ldots(n_i+1)\ldots\rangle \end{aligned} \tag{5.24}$$

as well as the commutation relations

$$\begin{aligned} &[c_i, c_j] = [c_i^\dagger, c_j^\dagger] = 0 \\ &[c_i, c_j^\dagger] = \delta_{ij} \end{aligned} \tag{5.25}$$

Particle number operator

From the anticommutation (commutation) relations, we can also conclude that the operator $n_i = c_i^\dagger c_i$ (for bosons: $n_i = b_i^\dagger b_i$) is the particle number operator. For fermions, we have

$$\begin{aligned} n_i|\ldots(n_i=0)\ldots\rangle &= c_i^\dagger c_i|\ldots(n_i=0)\ldots\rangle = 0 \\ n_i|\ldots(n_i=1)\ldots\rangle &= c_i^\dagger c_i c_i^\dagger|\ldots(n_i=0)\ldots\rangle \\ &= c_i^\dagger\big(1-c_i^\dagger c_i\big)|\ldots(n_i=0)\ldots\rangle = |\ldots(n_i=1)\ldots\rangle\,. \end{aligned} \tag{5.26}$$

For bosons,

$$n_i\big(b_i^\dagger\big)^N|0\rangle = N\,\big(b_i^\dagger\big)^N|0\rangle\,,$$

as can easily been shown recursively.

Change of basis

If we change from one single particle basis $|i\rangle$ to another, $|\tilde{i}\rangle$, by inserting a unit operator

$$\tilde{c}_j^\dagger\,|0\rangle \;\equiv\; |\tilde{j}\rangle \;=\; \sum_i |i\rangle\langle i|\tilde{j}\rangle = \sum_i \langle i|\tilde{j}\rangle\, c_i^\dagger\,|0\rangle\,, \tag{5.27}$$

the operators c_i turn into operators $\tilde{c}_j$ according to

$$\boxed{\tilde{c}_j^\dagger \;=\; \sum_i \langle i|\tilde{j}\rangle\, c_i^\dagger\,, \qquad \tilde{c}_j \;=\; \sum_i \langle\tilde{j}|i\rangle\, c_i}\,. \tag{5.28}$$

Operators in second quantization

So far, we have only considered a discrete single particle basis $|i\rangle$. We will now define field operators $\hat{\Phi}(\vec{x})$ over the single particle states corresponding to the position operator, and analogously field operators $c_{\vec{k}}$ that correspond to the momentum operator:

$$\begin{aligned}
\hat{\Phi}(\vec{x}) &= \sum_i \langle\vec{x}|i\rangle c_i \\
c_{\vec{k}} &= \sum_i \langle\vec{k}|i\rangle c_i = \int d^3x \langle\vec{k}|\vec{x}\rangle \sum_i \langle\vec{x}|i\rangle c_i = \frac{1}{(2\pi)^{3/2}} \int d^3x\, e^{-i\vec{k}\cdot\vec{x}} \hat{\Phi}(\vec{x})
\end{aligned} \tag{5.29}$$

Here, the unit operator $\mathbb{1} = \int d^3x\, |\vec{x}\rangle\langle\vec{x}|$ was introduced. Furthermore,

$$\hat{\Phi}(\vec{x}) = \int d^3k \sum_i \langle\vec{x}|\vec{k}\rangle\langle\vec{k}|i\rangle c_i = \int d^3k \langle\vec{x}|\vec{k}\rangle c_{\vec{k}} = \frac{1}{(2\pi)^{3/2}} \int d^3k\, e^{i\vec{k}\cdot\vec{x}} c_{\vec{k}} \tag{5.30}$$

Field operators corresponding to the position operator $\hat{\Phi}(\vec{x})$ ($\hat{\Phi}^\dagger(\vec{x})$) annihilate (create) a particle at position $\vec{x}$ (while annihilation/creation operators c_i, $c_i^\dagger$ do this for particles in a certain single particle state). The anticommutation relations

$$\begin{aligned}
\left[\hat{\Phi}(\vec{x}), \hat{\Phi}(\vec{y})\right]_+ &= \left[\hat{\Phi}^\dagger(\vec{x}), \hat{\Phi}^\dagger(\vec{y})\right]_+ = 0, \\
\left[\hat{\Phi}(\vec{x}), \hat{\Phi}^\dagger(\vec{y})\right]_+ &= \sum_{i,j} \langle\vec{x}|i\rangle\langle j|\vec{y}\rangle \left[c_i, c_j^\dagger\right]_+ \\
&= \sum_i \langle\vec{x}|i\rangle\langle i|\vec{y}\rangle = \langle\vec{x}|\vec{y}\rangle = \delta^3(\vec{x}-\vec{y})
\end{aligned} \tag{5.31}$$

and analogously for $c_{\vec{k}}$ are valid.

Remark. For systems with a finite volume V one defines $\hat{\Phi}(\vec{x})$ and $c_{\vec{k}}$ as above; $\vec{x}$ then varies continuously over V and $\vec{k}$ is quantized with corresponding volume $d^3k = \frac{(2\pi)^3}{V}$. Transformation from $\hat{\Phi}(\vec{x})$ to $c_{\vec{k}}$ is done using $\langle x|k\rangle = \frac{1}{\sqrt{V}} e^{i\vec{k}\vec{x}}$. The (anti-) commutation relations are valid as noted above, and the δ symbol for operators $c_{\vec{k}}$, $c_{\vec{k}}^\dagger$ becomes a Kronecker delta.

Representation of states

Let $|\psi\rangle$ be an N particle state; then we can write $|\psi\rangle$ as

$$|\psi\rangle = \frac{1}{N!}\int d^3x_1 \ldots d^3x_N \psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)\hat{\Phi}^\dagger(\vec{x}_1)\ldots\hat{\Phi}^\dagger(\vec{x}_N)|0\rangle \quad (5.32)$$

with a function $\psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)$ which is chosen totally (anti-)symmetric without loss of generality. We remember that

$$P_-\big(|\vec{x}_1\rangle|\vec{x}_2\rangle\ldots|\vec{x}_N\rangle\big) =: |\vec{x}_1\vec{x}_2\ldots\vec{x}_N\rangle = \hat{\Phi}^\dagger(\vec{x}_1)\hat{\Phi}^\dagger(\vec{x}_2)\ldots\hat{\Phi}^\dagger(\vec{x}_N)|0\rangle \quad (5.33)$$

Apparently, the function $\psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)$ is the wave function in second quantization.
We can also see by applying (anti-) commutation rules, that $\langle\vec{y}_1\vec{y}_2\ldots\vec{y}_N|\psi\rangle = \psi(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_N)$:

$$\begin{aligned}\langle\vec{y}_1\vec{y}_2\ldots\vec{y}_N|\psi\rangle &= \frac{1}{N!}\int d^3x_1\ldots d^3x_N\psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)\times \\ &\qquad \times\langle 0|\hat{\Phi}(\vec{y}_N)\ldots\hat{\Phi}(\vec{y}_1)\hat{\Phi}^\dagger(\vec{x}_1)\ldots\hat{\Phi}^\dagger(\vec{x}_N)|0\rangle \\ &= \psi(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_N) \quad (5.34)\end{aligned}$$

This presents the explicit relationship between first and second quantization.

So far, we have only considered one species of particles. We will for example consider interactions between electrons (fermions) and phonons (bosons). The corresponding states will then be created by products of fermionic and bosonic creation operators, for example

$$c^\dagger_{\vec{k}_1} c^\dagger_{\vec{k}_2} b^\dagger_{\vec{q}_1} b^\dagger_{\vec{q}_2} b^\dagger_{\vec{q}_3}|0\rangle \quad (5.35)$$

as a state of two fermions and three bosons with momenta $\vec{k}_1$, $\vec{k}_2$, and $\vec{q}_1$, $\vec{q}_2$, $\vec{q}_3$. Here, different fermions anticommute, and bosons among each other as well as bosons and fermions commute.

Operators in occupation number representation

In the description of many particle systems, so-called one particle and two particle operators appear. In first quantization, we can write a one particle operator for an N_e particle system as

$$A^{(1)} = \sum_{i=1}^{N_e} A^{(1)}(\vec{r}_i)\,. \quad (5.36)$$

It is added up from contributions, each of which is only acting on one of the N_e particles. Examples are the kinetic energy and the external potential of the single particle Hamiltonian, the current operator or the particle density operator. Two particle operators in first quantization are of the form

$$A^{(2)} = \frac{1}{2}\sum_{i\neq j} A^{(2)}(\vec{r}_i, \vec{r}_j)\,. \tag{5.37}$$

All addends simultaneously act on two different particles. An example is the Coulomb interaction. Now, we express the one particle operator by creation and annihilation operators by introducing the unit operators $\mathbb{1} = \sum_\alpha |k_\alpha\rangle^{(i)}\,{}^{(i)}\langle k_\alpha|$:

$$\begin{aligned} A^{(2)} &= \sum_{i=1}^{N_e}\sum_{\alpha=1}^{\infty} |k_\alpha\rangle^{(i)}\,{}^{(i)}\langle k_\alpha|A^{(1)}(\vec{r}_i)\sum_{\beta=1}^{\infty}|k_\beta\rangle^{(i)}\,{}^{(i)}\langle k_\beta| \\ &= \sum_{\alpha,\beta=1}^{\infty}\langle k_\alpha|A^{(1)}(\vec{r})|k_\beta\rangle\sum_{i=1}^{N_e}|k_\alpha\rangle^{(i)}\,{}^{(i)}\langle k_\beta| \end{aligned} \tag{5.38}$$

Here we used that the matrix element of $A^{(1)}$ with respect to the single particle states is not anymore dependent on the particle index:

$${}^{(i)}\langle k_\alpha|A^{(1)}(\vec{r}_i)|k_\beta\rangle^{(i)} = \int d^3r_i\,\varphi^*_{k_\alpha}(\vec{r}_i)A^{(1)}(\vec{r}_i)\varphi_{k_\beta}(\vec{r}_i) = \langle k_\alpha|A^{(1)}(\vec{r})|k_\beta\rangle \tag{5.39}$$

We have

$$\sum_{i=1}^{N_e}|k_\alpha\rangle^{(i)}\,{}^{(i)}\langle k_\beta| = c^+_{k_\alpha}c_{k_\beta} \tag{5.40}$$

because the operator $|k_\alpha\rangle^{(i)}\,{}^{(i)}\langle k_\beta|$, applied to an N_e particle state only gives a non-zero result if the one particle state k_β is contained in the manyparticle state; in this case, the state $|k_\beta\rangle$ in this manyparticle state is replaced by $|k_\alpha\rangle$; this corresponds to the annihilation of a particle in the state k_β and the creation of a particle in the state k_α.
Thus, the single particle operator in occupation number representation is

$$A^{(1)} = \sum_{\alpha,\beta=1}^{\infty} A^{(1)}_{k_\alpha,k_\beta}c^\dagger_{k_\alpha}c_{k_\beta} \quad \text{with } A^{(1)}_{k_\alpha,k_\beta} = \langle k_\alpha|A^{(1)}(\vec{r})|k_\beta\rangle \tag{5.41}$$

The one particle operator is expressed as a linear combination over all possible pairs of creation and annihilation operators, with matrix elements of the single particle operator with respect to the single particle states as coefficients. Correspondingly, we have for the two particle operator, if we insert unit operators:

$$\begin{aligned} A^{(2)} &= \frac{1}{2}\sum_{i\neq j}\sum_{\alpha,\beta,\gamma,\delta} |k_\alpha\rangle^{(i)}|k_\beta\rangle^{(j)}\ {}^{(i)}\langle k_\alpha|\ {}^{(j)}\langle k_\beta|A^{(2)}(\vec{r}_i,\vec{r}_j)|k_\gamma\rangle^{(j)}|k_\delta\rangle^{(i)}\ {}^{(j)}\langle k_\gamma|\ {}^{(i)}\langle k_\delta| \\ &= \frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta} A^{(2)}_{k_\alpha k_\beta, k_\gamma k_\delta} \sum_{i\neq j} |k_\alpha\rangle^{(i)}|k_\beta\rangle^{(j)}\ {}^{(j)}\langle k_\gamma|\ {}^{(i)}\langle k_\delta| \end{aligned} \tag{5.42}$$

with the two particle matrix element

$$\begin{aligned} A^{(2)}_{k_\alpha k_\beta, k_\gamma k_\delta} &= {}^{(i)}\langle k_\alpha|\ {}^{(j)}\langle k_\beta|A^{(2)}(\vec{r}_i,\vec{r}_j)|k_\gamma\rangle^{(j)}|k_\delta\rangle^{(i)} \\ &= \int d^3r_i d^3r_j\ \varphi^*_{k_\alpha}(\vec{r}_i)\varphi^*_{k_\beta}(\vec{r}_j)A^{(2)}(\vec{r}_i,\vec{r}_j)\varphi_{k_\gamma}(\vec{r}_j)\varphi_{k_\delta}(\vec{r}_i) \end{aligned} \tag{5.43}$$

which is again independent of the particle indices because in the calculation of the matrix element in position representation, there is an integral over these indices. If we apply the operator $\sum_{i\neq j}|k_\alpha\rangle^{(i)}|k_\beta\rangle^{(j)}\ {}^{(j)}\langle k_\gamma|\ {}^{(i)}\langle k_\delta|$ on a manybody state we only obtain something nonzero if in this many body state, the single particle states k_γ and k_δ are occupied. In this case the operator replaces these states by k_α and k_β; then, in the manybody state k_γ, k_δ are unoccupied and k_α, k_β are occupied; this corresponds to

$$\sum_{i\neq j}|k_\alpha\rangle^{(i)}|k_\beta\rangle^{(j)}\ {}^{(j)}\langle k_\gamma|\ {}^{(i)}\langle k_\delta| = c^\dagger_{k_\alpha}c^\dagger_{k_\beta}c_{k_\gamma}c_{k_\delta} \tag{5.44}$$

Thus, the two particle operator $A^{(2)}$ is in occupation number representation

$$\begin{aligned} A^{(2)} &= \frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta} A^{(2)}_{k_\alpha k_\beta, k_\gamma k_\delta} c^\dagger_{k_\alpha}c^\dagger_{k_\beta}c_{k_\gamma}c_{k_\delta} \\ A^{(2)}_{k_\alpha k_\beta, k_\gamma k_\delta} &= \int d^3r_i d^3r_j\ \varphi^*_{k_\alpha}(\vec{r}_i)\varphi^*_{k_\beta}(\vec{r}_j)A^{(2)}(\vec{r}_i,\vec{r}_j)\varphi_{k_\gamma}(\vec{r}_j)\varphi_{k_\delta}(\vec{r}_i) \end{aligned} \tag{5.45}$$

All operators become linear combinations of creation and annihilation operators with coefficients which are given by the matrix element of the corresponding operator with respect to the single particle states.

The operators can also be expressed using the representation of free field operators:

$$A^{(1)} = \int d^3r\, \hat{\Phi}^\dagger(\vec{r})A^{(1)}(\vec{r})\hat{\Phi}(\vec{r})$$
$$A^{(2)} = \frac{1}{2}\int d^3r d^3r'\, \hat{\Phi}^\dagger(\vec{r})\hat{\Phi}^\dagger(\vec{r}')A^{(2)}(\vec{r},\vec{r}')\hat{\Phi}(\vec{r}')\hat{\Phi}(\vec{r}) \tag{5.46}$$

This can be checked by introducing the expansion of the field operators in a single particle basis. The name second quantization is now due to the fact that in these relationships, also the wave functions are replaced by operators.

Especially the Hamiltonian (5.1) can now be written in occupation number representation:

$$H = H_0 + H_1 = \sum_\alpha \varepsilon_{k_\alpha} c^\dagger_{k_\alpha} c_{k_\alpha} + \frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta} u_{k_\alpha k_\beta, k_\gamma k_\delta} c^\dagger_{k_\alpha} c^\dagger_{k_\beta} c_{k_\gamma} c_{k_\delta} \tag{5.47}$$

where the single particle basis of the single particle Hamiltonian was used. We have

$$u_{k_\alpha k_\beta, k_\gamma k_\delta} = \int d^3r d^3r'\, \varphi^*_{k_\alpha}(\vec{r})\varphi^*_{k_\beta}(\vec{r}')u(\vec{r}-\vec{r}')\varphi_{k_\gamma}(\vec{r}')\varphi_{k_\delta}(\vec{r}) \tag{5.48}$$

The use of a different single particle basis is also possible; however, the single particle component of H is then not diagonal anymore.

5.2 Direct ferromagnetic exchange interaction

The Coulomb interaction is

$$H_{\text{Coulomb}} = \frac{1}{2}\frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2 \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} \tag{5.49}$$

In second quantized notation, $\rho(\vec{r})$ is the operator of the charge density

$$\rho(\vec{r}) = -e\sum_\sigma \psi^\dagger_\sigma(\vec{r})\psi_\sigma(\vec{r})$$

with spin orientation $\sigma =\uparrow,\downarrow$ and the field operator $\psi_\sigma(\vec{r})$. This operator satisfies the anticommutation relations

$$\left[\psi_{\sigma_1}(\vec{r}_1), \psi^+_{\sigma_2}(\vec{r}_2)\right]_+ \equiv \psi_{\sigma_1}(\vec{r}_1)\psi^+_{\sigma_2}(\vec{r}_2) + \psi^+_{\sigma_2}(\vec{r}_2)\psi_{\sigma_1}(\vec{r}_1) = \delta_{\sigma_1\sigma_2}\delta(\vec{r}_1 - \vec{r}_2)$$
$$\left[\psi_{\sigma_1}(\vec{r}_1), \psi_{\sigma_2}(\vec{r}_2)\right]_+ = 0 \qquad \left[\psi^+_{\sigma_1}(\vec{r}_1), \psi^+_{\sigma_2}(\vec{r}_2)\right]_+ = 0$$

(5.50)

as it is a Fermionic field. We find

$$\begin{aligned}H_{\text{Coulomb}} &= \frac{1}{2}\frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2 \sum_{\sigma_1\sigma_2} \psi^\dagger_{\sigma_1}(\vec{r}_1)\psi_{\sigma_1}(\vec{r}_1)\frac{e^2}{|\vec{r}_1-\vec{r}_2|}\psi^\dagger_{\sigma_2}(\vec{r}_2)\psi_{\sigma_2}(\vec{r}_2)\\ &= \frac{1}{2}\frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2 \sum_{\sigma_1\sigma_2} \psi^\dagger_{\sigma_1}(\vec{r}_1)\psi^\dagger_{\sigma_2}(\vec{r}_2)\frac{e^2}{|\vec{r}_1-\vec{r}_2|}\psi_{\sigma_2}(\vec{r}_2)\psi_{\sigma_1}(\vec{r}_1)\\ &\quad + \frac{1}{2}\frac{1}{4\pi\epsilon_0}\int d^3r_1 \sum_{\sigma_1}\frac{e^2}{|\vec{r}_1-\vec{r}_1|}\psi^\dagger_{\sigma_1}(\vec{r}_1)\psi_{\sigma_1}(\vec{r}_1)\end{aligned} \tag{5.51}$$

The last singular term is unphysical; in fact, the field operators should be written in normal order $(\psi^\dagger\psi^\dagger\psi\psi)$ from the start. ψ can be expanded into any orthonormal set of single particle wave functions. A set of Wannier functions, *i.e.* orthonormal functions $\phi_{\vec{R}m}(\vec{r})$ localized at the ion position $\vec{R}$ is advantageous; m includes all orbital quantum numbers except for the spin σ. We also introduce spinors

$$\chi_\uparrow = \begin{pmatrix}1\\0\end{pmatrix}, \chi_\downarrow = \begin{pmatrix}0\\1\end{pmatrix}$$

which are eigenvectors of

$$s^z = \frac{1}{2}\sigma^z = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

with eigenvalues $\pm\frac{1}{2}$. Then the field operator is

$$\psi_\sigma(\vec{r}) = \sum_{\vec{R}m} a_{\vec{R}m\sigma}\phi_{\vec{R}m}\chi_\sigma \tag{5.52}$$

with Fermionic annihilation operators $a_{\vec{R}m\sigma}$ which satisfy $[a_{\vec{R}m\sigma}, a^\dagger_{\vec{R}m\sigma}]_+ = \delta_{\vec{R}\vec{R}'}\delta_{mm'}\delta_{\sigma\sigma'}$ etc. The Coulomb interaction becomes

$$\begin{aligned}H_{\text{Coulomb}} &= \frac{1}{2}\sum_{\vec{R}_1m_1}\cdots\sum_{\vec{R}_4m_4}\int d^3r_1 d^3r_2 \phi^*_{\vec{R}_1m_1}(\vec{r}_1)\phi^*_{\vec{R}_2m_2}(\vec{r}_2)\frac{e^2}{4\pi\epsilon_0|\vec{r}_1-\vec{r}_2|}\times\\ &\quad\times \phi_{\vec{R}_3m_3}(\vec{r}_2)\phi_{\vec{R}_4m_4}(\vec{r}_1)\sum_{\sigma_1\sigma_2}\chi^\dagger_{\sigma_1}\chi^\dagger_{\sigma_2}\chi_{\sigma_2}\chi_{\sigma_1}a^\dagger_{\vec{R}_1m_1\sigma_1}a^\dagger_{\vec{R}_2m_2\sigma_2}a_{\vec{R}_3m_3\sigma_2}a_{\vec{R}_4m_4\sigma_1}\end{aligned}$$

$$(5.53)$$

The scalar product of the spinors are simply $\chi^\dagger_{\sigma_1}\chi_{\sigma_1} = \chi^\dagger_{\sigma_2}\chi_{\sigma_2} = 1$. We define the integral

$$\left\langle \vec{R}_1 m_1, \vec{R}_2 m_2 \middle| \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} \middle| \vec{R}_3 m_4, \vec{R}_4 m_4 \right\rangle := \\ \int d^3r_1 d^3r_2 \phi^*_{\vec{R}_1 m_1}(\vec{r}_1)\phi^*_{\vec{R}_2 m_2}(\vec{r}_2) \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} \phi_{\vec{R}_3 m_3}(\vec{r}_2)\phi_{\vec{R}_4 m_4}(\vec{r}_1) \quad (5.54)$$

and obtain

$$H_{\text{Coulomb}} = \frac{1}{2}\sum_{\vec{R}_1 m_1} \cdots \sum_{\vec{R}_4 m_4} \left\langle \vec{R}_1 m_1, \vec{R}_2 m_2 \middle| \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} \middle| \vec{R}_3 m_4, \vec{R}_4 m_4 \right\rangle \times \\ \times \sum_{\sigma_1\sigma_2} a^\dagger_{\vec{R}_1 m_1 \sigma_1} a^\dagger_{\vec{R}_2 m_2 \sigma_2} a_{\vec{R}_3 m_3 \sigma_2} a_{\vec{R}_4 m_4 \sigma_1} \quad (5.55)$$

5.2.1 On-site Coulomb interaction

We first consider the contribution of $\vec{R}_1 = \vec{R}_2 = \vec{R}_3 = \vec{R}_4 \equiv \vec{R}$ and drop the index $\vec{R}$ where it doesn't create confusion. In general, the quantum numbers $m_1, \ldots, m_4$ in the integral $\left\langle m_1, m_2 \middle| \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} \middle| m_3, m_4 \right\rangle$ can all be different and still lead to a nonzero integral. However, in treating H_{Coulomb} as a perturbation, a nonzero first-order contribution requires a pairing of creation and annihilation operators $a^\dagger$, a for each orbital. This requires $m_1 = m_4$ and $m_2 = m_3$ or $m_1 = m_3$, $m_2 = m_4$. This leads to the direct Coulomb integrals

$$K_{m_1 m_2} := \left\langle m_1, m_2 \middle| \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} \middle| m_2, m_1 \right\rangle \\ = \int d^3r_1 d^3r_2 |\phi_{m_1}(\vec{r}_1)|^2 \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} |\phi_{m_2}(\vec{r}_2)|^2 \quad (5.56)$$

and the exchange integrals

$$J_{m_1 m_2} := \left\langle m_1, m_2 \middle| \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} \middle| m_1, m_2 \right\rangle \\ = \int d^3r_1 d^3r_2 \phi^*_{m_1}(\vec{r}_1)\phi^*_{m_2}(\vec{r}_2) \frac{e^2}{4\pi\epsilon_0|\vec{r}_1 - \vec{r}_2|} \phi_{m_1}(\vec{r}_2)\phi_{m_2}(\vec{r}_1)\,, \quad (5.57)$$

so-called because m_1 and m_2 are exchanged in the last factor compared to the direct integrals. To first order we obtain

$$\begin{aligned} H_{\text{Coulomb}} \approx \frac{1}{2}\sum_{\vec{R}}\sum_{m_1 m_2}\sum_{\sigma_1\sigma_2}\Big(&K_{m_1 m_2} a^\dagger_{\vec{R}m_1\sigma_1} a^\dagger_{\vec{R}m_2\sigma_2} a_{\vec{R}m_2\sigma_2} a_{\vec{R}m_1\sigma_1} \\ &+ J_{m_1 m_2} a^\dagger_{\vec{R}m_1\sigma_1} a^\dagger_{\vec{R}m_2\sigma_2} a_{\vec{R}m_1\sigma_2} a_{\vec{R}m_2\sigma_1}\Big) \end{aligned} \tag{5.58}$$

(double counting of contributions from $m_1 = m_2$ and $\sigma_1 = \sigma_2$ is not a problem because the terms contain $a_{\vec{R}m_1\sigma_1} a_{\vec{R}m_1\sigma_1} = 0$). Therefore

$$\begin{aligned} H_{\text{Coulomb}} \approx &\frac{1}{2}\sum_{\vec{R}}\sum_{m_1 m_2}\sum_{\sigma_1\sigma_2}\Big(K_{m_1 m_2} a^\dagger_{\vec{R}m_1\sigma_1} a_{\vec{R}m_1\sigma_1} a^\dagger_{\vec{R}m_2\sigma_2} a_{\vec{R}m_2\sigma_2} \\ &- J_{m_1 m_2} a^\dagger_{\vec{R}m_1\sigma_1} a_{\vec{R}m_1\sigma_2} a^\dagger_{\vec{R}m_2\sigma_2} a_{\vec{R}m_2\sigma_1}\Big) + \text{irrelevant potential terms} \end{aligned} \tag{5.59}$$

We now define the number operators $n_{\vec{R}m} := \sum_\sigma a^\dagger_{\vec{R}m\sigma} a_{\vec{R}m\sigma}$ and the spin operators $s^\alpha_{\vec{R}m} := \sum_{\sigma\sigma'} a^\dagger_{\vec{R}m\sigma} \frac{1}{2}\sigma^\alpha_{\sigma\sigma'} a_{\vec{R}m\sigma'}$ with Pauli matrices σ^α, $\alpha = x, y, z$. After some algebra,

$$\begin{aligned} &\sum_{\sigma_1\sigma_2} a^\dagger_{\vec{R}m_1\sigma_1} a_{\vec{R}m_1\sigma_2} a^\dagger_{\vec{R}m_2\sigma_2} a_{\vec{R}m_2\sigma_1} \\ &\qquad = \frac{1}{2} n_{\vec{R}m_1} n_{\vec{R}m_2} + 2 s^z_{\vec{R}m_1} s^z_{\vec{R}m_2} + s^+_{\vec{R}m_1} s^-_{\vec{R}m_2} + s^-_{\vec{R}m_1} s^+_{\vec{R}m_2} \\ &\qquad = \frac{1}{2} n_{\vec{R}m_1} n_{\vec{R}m_2} + 2 \vec{s}_{\vec{R}m_1} \cdot \vec{s}_{\vec{R}m_2} \end{aligned} \tag{5.60}$$

Therefore, we obtain

$$\begin{aligned} H_{\text{Coulomb}} \approx \sum_{\vec{R}} \frac{1}{2} \sum_{m_1 m_2} \Big\{ &\Big(K_{m_1 m_2} - \frac{1}{2} J_{m_1 m_2}\Big) n_{\vec{R}m_1} n_{\vec{R}m_2} \\ &- 2 J_{m_1 m_2} \vec{s}_{\vec{R}m_1} \cdot \vec{s}_{\vec{R}m_2} \Big\}. \end{aligned} \tag{5.61}$$

The first term is the m site Coulomb interaction. From the definition (5.56), we can read off immediately that $K_{m_1 m_2} > 0$. To show that $J_{m_1 m_2} \geqslant 0$, we use the Fourier transform

$$\frac{1}{|\vec{r}_1 - \vec{r}_2|} = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot(\vec{r}_1 - \vec{r}_2)} \frac{4\pi}{k^2}\,. \tag{5.62}$$

Then,

$$
\begin{aligned}
J_{m_1 m_2} &= \int \frac{d^3k}{(2\pi)^3} \frac{e^2}{\epsilon_0 k^2} \underbrace{\int d^3r_1 \phi^*_{m_1}(\vec{r}_1)\phi_{m_2}(\vec{r}_1) e^{i\vec{k}\cdot\vec{r}_1}}_{=:I(\vec{k})} \underbrace{\int d^3r_2 \phi_{m_1}(\vec{r}_2)\phi^*_{m_2}(\vec{r}_2) e^{-i\vec{k}\cdot\vec{r}_2}}_{=:I^*(\vec{k})} \\
&= \int \frac{d^3k}{(2\pi)^3} \frac{e^2}{\epsilon_0 k^2} |I(\vec{k})|^2 \geqslant 0\,.
\end{aligned}
\tag{5.63}
$$

We also show that $K_{m_1 m_2} \geqslant J_{m_1 m_2}$:

$$
\begin{aligned}
K_{m_1 m_2} - J_{m_1 m_2} &= \frac{1}{2}\big(K_{m_1 m_2} + K_{m_2 m_1} - J_{m_1 m_2} - J_{m_2 m_1}\big) \\
&= \frac{1}{2} \int d^3r_1 d^3r_2 \frac{e^2}{4\pi\epsilon_0 |\vec{r}_1 - \vec{r}_2|} \underbrace{\big[\phi^*_{m_1}(\vec{r}_1)\phi^*_{m_2}(\vec{r}_2) - \phi^*_{m_2}(\vec{r}_1)\phi^*_{m_1}(\vec{r}_2)\big]}_{=:f(\vec{r}_1,\vec{r}_2)} \times \\
&\qquad \times \underbrace{\big[\phi_{m_2}(\vec{r}_2)\phi_{m_1}(\vec{r}_1) - \phi_{m_1}(\vec{r}_2)\phi_{m_2}(\vec{r}_1)\big]}_{=:f^*(\vec{r}_1,\vec{r}_2)} \\
&= \frac{1}{2} \int d^3r_1 d^3r_2 \frac{e^2}{4\pi\epsilon_0 |\vec{r}_1 - \vec{r}_2|} |f(\vec{r}_1, \vec{r}_2)|^2 \geqslant 0\,.
\end{aligned}
\tag{5.64}
$$

Therefore, the corrected Coulomb term $K_{m_1 m_2} - \frac{1}{2} J_{m_1 m_2} > 0$ is reduced but still repulsive. The new finding is a spin-spin interaction of the form

$$
-J_{m_1 m_2} \vec{s}_{\vec{R}m_1} \cdot \vec{s}_{\vec{R}m_2}
$$

with $J_{m_1 m_2} \geqslant 0$. This interaction prefers parallel alignment of the spins, *i.e.* it is a ferromagnetic interaction. Thus, we have derived the first Hund's rule: The total spin of electrons in a partially filled shell of one ion tends to be maximal.

Note that all terms containing $J_{m_1 m_2}$ are quantum mechanical in origin; they appear because we have written the density $\rho = -e\psi^\dagger\psi$ as a bilinear form in the field operator which made unconventional pairings of the orbital indices possible. There is no analogy in classical physics. For a single relevant orbital $\phi(\vec{r})$, we get

$$\begin{aligned}
H_{\text{Coulomb}} &\approx \frac{1}{2}\sum_{\vec{R}}\int d^3r_1 d^3r_2 \phi^*(\vec{r}_1)\phi^*(\vec{r}_2)\frac{e^2}{4\pi\epsilon_0|\vec{r}_1-\vec{r}_2|}\phi(\vec{r}_2)\phi(\vec{r}_1)\times \\
&\qquad\times\sum_{\sigma_1\sigma_2} a^\dagger_{\vec{R}\sigma_1}a^\dagger_{\vec{R}\sigma_2}a_{\vec{R}\sigma_2}a_{\vec{R}\sigma_1} \\
&= \frac{1}{2}\sum_{\vec{R}}\int d^3r_1 d^3r_2 \phi^*(\vec{r}_1)\phi^*(\vec{r}_2)\frac{e^2}{4\pi\epsilon_0|\vec{r}_1-\vec{r}_2|}\phi(\vec{r}_2)\phi(\vec{r}_1)\times \\
&\qquad\times\left(a^\dagger_{\vec{R}\uparrow}a^\dagger_{\vec{R}\downarrow}a_{\vec{R}\downarrow}a_{\vec{R}\uparrow} + a^\dagger_{\vec{R}\downarrow}a^\dagger_{\vec{R}\uparrow}a_{\vec{R}\uparrow}a_{\vec{R}\downarrow}\right) \\
&= \sum_{\vec{R}}\underbrace{\int d^3r_1 d^3r_2 \phi^*(\vec{r}_1)\phi^*(\vec{r}_2)\frac{e^2}{4\pi\epsilon_0|\vec{r}_1-\vec{r}_2|}}_{=:U} a^\dagger_{\vec{R}\uparrow}a^\dagger_{\vec{R}\downarrow}a_{\vec{R}\downarrow}a_{\vec{R}\uparrow} \\
&= \sum_{\vec{R}} U a^\dagger_{\vec{R}\uparrow}a^\dagger_{\vec{R}\downarrow}a_{\vec{R}\downarrow}a_{\vec{R}\uparrow},
\end{aligned} \tag{5.65}$$

where U is the famous Hubbard U.

5.2.2 Ion-ion exchange interaction

If we allow the ionic sites $\vec{R}_1, \ldots, \vec{R}_4$ to be different, most of the calculation of the previous section remains unchanged; we just have to treat $\vec{R}_i$ as another quantum number besides m_i. Here, we restrict ourselves to a model with a single, non-degenerate (apart from spin) orbital per site. Then, we can drop the orbital quantum numbers m_i. We again assume the orbitals on different sites to have negligible overlap, *i.e.* they are orthogonal. This time, we will have a first order contribution if $\vec{R}_1 = \vec{R}_4$ and $\vec{R}_2 = \vec{R}_3$ or $\vec{R}_1 = \vec{R}_3$ and $\vec{R}_2 = \vec{R}_4$. In complete analogy to the previous section we obtain

$$H_{\text{Coulomb}} \approx \frac{1}{2}\sum_{\vec{R}_1\vec{R}_2}\left\{\left(K_{12}-\frac{1}{2}J_{12}\right)n_1n_2 - 2J_{12}\vec{s}_1\cdot\vec{s}_2\right\} \tag{5.66}$$

where

$$K_{12} \equiv K_{\vec{R}_1\vec{R}_2} := \int d^3r_1 d^3r_2 \left|\phi_{\vec{R}_1}(\vec{r}_1)\right|^2 \frac{e^2}{4\pi\epsilon_0|\vec{r}_1-\vec{r}_2|} \left|\phi_{\vec{R}_2}(\vec{r}_2)\right|^2$$

$$J_{12} \equiv J_{\vec{R}_1\vec{R}_2} := \int d^3r_1 d^3r_2 \phi^*_{\vec{R}_1}(\vec{r}_1)\phi^*_{\vec{R}_2}(\vec{r}_2) \frac{e^2}{4\pi\epsilon_0|\vec{r}_1-\vec{r}_2|} \phi_{\vec{R}_1}(\vec{r}_2)\phi_{\vec{R}_2}(\vec{r}_1)\,. \tag{5.67}$$

If we make the assumption (to be relaxed later) than in an ionic crystal the charge $-en_i$ does not fluctuate much, we can consider the electron number $n_i = 1$ (otherwise there is no spin) and find for the interaction if we disregard a constant

$$H_{\text{exc}} = -\sum_{\vec{R}_1\vec{R}_2} J_{12}\vec{s}_1 \cdot \vec{s}_2\,. \tag{5.68}$$

By the argument given above, $J_{12} \geqslant 0$. thus, the Coulomb repulsion between electrons in orthogonal orbitals always leads to a ferromagnetic exchange interaction. The physical interpretation is that electrons with parallel spins cannot occupy the same orbital; therefore, they avoid the strong intra-orbital Coulomb repulsion. Thus, their energy is lower than for antiparallel spins.

5.3 Kinetic antiferromagnetic exchange interaction

Previously, we neglected charge fluctuations but as this is not usually a good approximation even for ionic crystals, we now go one step further. In an independent-electron or band picture, the hybridization between orbitals of different ions leads to charge fluctuations and thus allows electrons to tunnel or hop from one ion to another. While studying the effect of the hybridization, we neglect the non-local (inter-ionic) Coulomb repulsion; we already know that it leads to ferromagnetic exchange. The model describing the competition between the kinetic energy and the on-site Coulomb repulsion for a single relevant orbital is the Hubbard model:

$$H = \sum_{\vec{R}\vec{R}'\sigma} t(\vec{R}-\vec{R}')a^\dagger_{\vec{R}'\sigma}a_{\vec{R}\sigma} + U\sum_{\vec{R}} a^\dagger_{\vec{R}\uparrow}a_{\vec{R}\uparrow}a^\dagger_{\vec{R}\downarrow}a_{\vec{R}\downarrow}\,. \tag{5.69}$$

As a toy model, we consider the case of a dimer:

$$H = -t\sum_\sigma \left(a^\dagger_{1\sigma}a_{2\sigma}+a^\dagger_{2\sigma}a_{1\sigma}\right) - \mu\sum_\sigma \left(a^\dagger_{1\sigma}a_{1\sigma}+a^\dagger_{2\sigma}a_{2\sigma}\right) + U\sum_{i=1}^{2} a^\dagger_{i\uparrow}a_{i\uparrow}a^\dagger_{i\downarrow}a_{i\downarrow}\,.$$

(5.70)

The dimension of the Fock space is $4^2 = 16$ because each site can be in one of four states (empty $|0\rangle$, spin up $|\uparrow\rangle$, spin down $|\downarrow\rangle$, and doubly occupied $|\uparrow\downarrow\rangle$). The sectors with $0, 1, 2, 3, 4$ electrons have $1, 4, 6, 4, 1$ states, respectively. We consider the two electron sector which corresponds to a six dimensional Hilbert space. In this space, the chemical potential is an irrelevant constant. We choose as basis vectors $\Big\{|\uparrow\downarrow, 0\rangle,\ |0, \uparrow\downarrow\rangle,\ |\uparrow, \downarrow\rangle,$ $|\downarrow, \uparrow\rangle, |\uparrow, \uparrow\rangle, |\downarrow, \downarrow\rangle\Big\}$. In this basis, the Hamiltonian is a 6×6 matrix

$$H' = \begin{pmatrix} U & 0 & t & -t & 0 & 0 \\ 0 & U & t & -t & 0 & 0 \\ t & t & 0 & 0 & 0 & 0 \\ -t & -t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} . \tag{5.71}$$

We can simplify H' by transforming from $|\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle$ to $(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle)/\sqrt{2}$, $(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle)/\sqrt{2}$ and obtain

$$H' = \begin{pmatrix} U & 0 & \sqrt{2}t & 0 & 0 & 0 \\ 0 & U & \sqrt{2}t & 0 & 0 & 0 \\ \sqrt{2}t & \sqrt{2}t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.72}$$

with eigenenergies

$$\begin{aligned} & U, \frac{1}{2}\left(U \pm \sqrt{U^2 + 16t^2}\right) \text{ in the first sector} \\ & 0, 0, 0 \text{ in the second sector} \end{aligned} \tag{5.73}$$

The latter eigenenergies correspond to the spin triplet, the states $(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle)/\sqrt{2}$, $|\uparrow, \uparrow\rangle$ and $|\downarrow, \downarrow\rangle$.

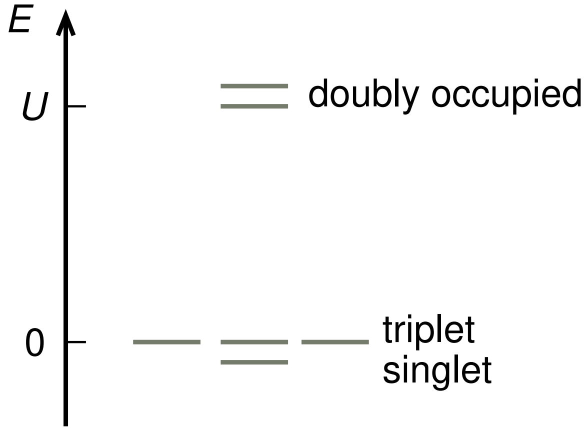
We are interested in the ionic systems for which t should be small, $t \ll U$. Then the first sector contains two very large energies

$$U \quad \text{and} \quad \frac{1}{2}\left(U + \sqrt{U^2 + 16t^2}\right) \approx U + \frac{4t^2}{U} ,$$

and one small energy

$$\frac{1}{2}\left(U - \sqrt{U^2 + 16t^2}\right) \approx -\frac{4t^2}{U} < 0\,. \tag{5.74}$$

For $\frac{U}{|t|} \to \infty$, the corresponding eigenstate approaches $(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle)/\sqrt{2}$, *i.e.* the spin singlet. For finite U, it has some admixture of doubly occupied states. The spectrum looks like this:



We find that the singlet $(S = 0)$ is lower in energy than the triplet $(S = 1)$, *i.e.* there is an antiferromagnetic interaction. This results from the lowering of the kinetic energy for antiparallel spins. For parallel spins, the hopping is blocked by the Pauli principle which is why t does not even appear in the eigenenergies of the triplet. Therefore, this mechanism is called kinetic exchange. An example is the H_2 molecule which has a singlet ground state. To compare this model to an interacting pair of spins $s_1 = s_2 = \frac{1}{2}$, we write

$$\begin{aligned}
H_{\text{eff}} &= -J\vec{s}_1 \cdot \vec{s}_2 = -\frac{J}{2}\left[\vec{S}\cdot\vec{S} - \vec{s}_1\cdot\vec{s}_1 - \vec{s}_2\cdot\vec{s}_2\right] \quad \text{with } \vec{S} = \vec{s}_1 + \vec{s}_2 \\
&= -\frac{J}{2}\Big[S\Big(S+\frac{1}{2}\Big)\underbrace{-\frac{3}{4}-\frac{3}{4}}_{\text{const}}\Big] = \text{const } -\frac{J}{2}S\Big(S+\frac{1}{2}\Big) \\
&= \text{const} \begin{cases} +0 \text{ for } S = 0 \\ -J \text{ for } S = 1 \end{cases}
\end{aligned} \tag{5.75}$$

By comparing this with Eq. (5.74), we read off

$$J = -\frac{4t^2}{U} \quad \text{for } U \gg t\,. \tag{5.76}$$

An analogous result holds for the Hubbard model on a lattice, not only for a dimer. The result is for a lattice at half filling and in the limit $U \gg t$

$$H_{\text{eff}} = -J \sum_{\langle ij \rangle} \vec{s}_i \cdot \vec{s}_j \quad \text{with } J = -\frac{4t^2}{U} \tag{5.77}$$

where $\sum_{\langle ij \rangle}$ runs over all nearest neighbor bonds, and each bond is counted only once ($\langle ij \rangle$ and $\langle ji \rangle$ are the same bond and enter the sum only once).