Theory of Magnetism

International Max Planck Research School for Dynamical Processes in
Atoms, Molecules and Solids

Winter Semester 2009–2010
Revised September 19, 2019

Contents

Chapter 1

Introduction: What is magnetism?

It has been known since antiquity that "loadstone" (magnetite, Fe_3O_4) and iron attract each other. Plato (428/427–348/347 B.C.) and Aristotle mention permanent magnets. They are also mentioned in Chinese texts from the 4th century B.C. The earliest mention of a magnetic compass used for navigation is from a Chinese text dated 1040–1044 A.D., but it may have been invented there much earlier. It was apparently first used for orientation on land, not at sea.

Thus magnetism at first referred to the long-range interaction between ferromagnetic bodies. Indeed, the present course will mainly address *magnetic order* in solids, of which ferromagnetism is the most straight-forward case. This begs the question of what it is that is ordering in a ferromagnet.

Oersted (1819) found that a compass needle is deflected by a current-carrying wire in the same way as by a pemanent magnet. This and later experiments led to the notion that the magnetization of a permanent magnet is somehow due to pemanent currents of electrons. Biot, Savart, and Ampère established the relationship of the magnetic induction and the current that generates it. As we know, Maxwell essentially completed the classical theory of electromagnetism.

1.1 The Bohr-van Leeuwen theorem

Can we understand ferromagnetism in terms of electron currents in the framework of Maxwellian classical electrodynamics? For N classical electrons with positions $\mathbf{r}_i$ and moments $\mathbf{p}_i$, the partition function is

$$Z \propto \int \prod_i d^3 r_i d^3 p_i \, \exp\big(-\beta H(\mathbf{r}_1, \dots; \mathbf{p}_1, \dots)\big), \tag{1.1}$$

where H is the classical Hamilton function,

$$H = \frac{1}{2m} \sum_i \big(\mathbf{p}_i + e\mathbf{A}(\mathbf{r}_i)\big)^2 + V(\mathbf{r}_1, \dots). \tag{1.2}$$

Here, $\mathbf{A}(\mathbf{r})$ is the vector potential related to the magnetic induction $\mathbf{B}$ through $\mathbf{B} = \nabla \times \mathbf{A}$. $\mathbf{B}$ is presumably due to the currents (through the Biot-Savart or Ampère-Maxwell laws), in addition to a possible external magnetic field. The electron charge is $-e$.

But now we can substitute $\mathbf{p}_i \to \tilde{\mathbf{p}}_i = \mathbf{p}_i + e\mathbf{A}(\mathbf{r}_i)$ in the integrals,

$$Z \propto \int \prod_i d^3 r_i d^3 \tilde{p}_i \, \exp\Big[-\beta\Big(\frac{1}{2m}\sum_i \tilde{\mathbf{p}}_i^2 + V\Big)\Big]. \tag{1.3}$$

Thus we have eliminated the vector potential $\mathbf{A}$ from the partition function.

With the free energy

$$F = -\frac{1}{\beta} \ln Z, \tag{1.4}$$

this leads to the magnetization

$$\mathbf{M} = -\frac{\partial F}{\partial \mathbf{B}} = 0. \tag{1.5}$$

This is called the *Bohr-van Leeuwen theorem*.

What have we shown? We cannot obtain equilibrium ferromagnetism in a theory that

(a) is classical *and*

(b) assumes the magnetic field to be due to currents alone.

Which of these two assumptions is to blame? Most books treat them as essentially equivalent. The usually offered solution is that we require an *intrinsic* magnetic moment carried by the electrons, which is then attributed to an intrinsic angular momentum, the *spin*. Text books on quantum mechanics usually claim that at least non-integer spins are only possible in quantum mechanics—there is some controversy on this, though. Most books thus drop (b) by assuming electrons to carry spin and state that (a) is then broken.

1.2 The electron spin and magnetic moment

We know, however, that electrons carry spin $S = 1/2$ (e.g., from Stern-Gerlach-type experiments) and that they carry a magnetic moment. In the following, we review the relation between angular momenta and magnetic moments.

(a) Orbital motion: We know since the days of Oersted that moving charges generate magnetic fields. We consider an electron of charge $-e$ moving in a circle of radius R with constant angular velocity $\boldsymbol{\omega}$.
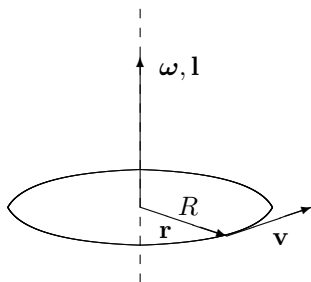
Its angular momentum is

$$\mathbf{l} = \mathbf{r} \times m_e \mathbf{v} = m_e \mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}) = m_e r^2 \boldsymbol{\omega} - m_e \underbrace{(\mathbf{r} \cdot \boldsymbol{\omega})}_{=0} \mathbf{r} = m_e R^2 \boldsymbol{\omega} = \text{const.}$$

if center in origin (1.6)

The magnetic field, on the other hand, is very complicated and time-dependent. However, if the period $T = 2\pi/\omega$ is small on the relevant experimental timescale, we can consider the averaged field. Since the Maxwell equations are linear, this is the magnetic field of the averaged current.

The averaged current is

$$I = \frac{-e}{T} = -\frac{e\omega}{2\pi}. \tag{1.7}$$

The induction $\mathbf{B}$ can now be obtained from the Biot-Savart law,

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} I \int \frac{d\mathbf{l}' \times \Delta\mathbf{r}}{\Delta r^3}. \tag{1.8}$$
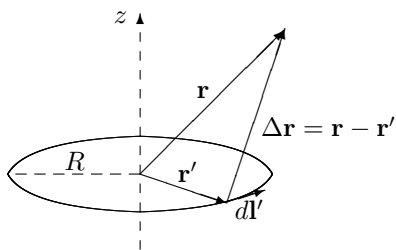
The Biot-Savart law can be rewritten as

$$\mathbf{B}(\mathbf{r}) = -\frac{\mu_0}{4\pi} I \int d\mathbf{l}' \times \nabla \frac{1}{\Delta r} = +\frac{\mu_0}{4\pi} I \nabla \times \int d\mathbf{l}' \frac{1}{\Delta r} \stackrel{!}{=} \nabla \times \mathbf{A}. \tag{1.9}$$

Obviously, we can choose

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} I \int d\mathbf{l}' \frac{1}{\Delta r}. \tag{1.10}$$

The induction $\mathbf{B}$ or the vector potential $\mathbf{A}$ can be evaluated in terms of elliptic integrals, see Jackson's book. We here consider only the limit $R \ll \Delta r$ (the far field). To that end, we perform a multipole expansion of $1/\Delta r$,

$$\frac{1}{\Delta r} = \frac{1}{|\mathbf{r} - \mathbf{r}'|} \cong \frac{1}{r} + \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|^3}\right)_{\mathbf{r}'=0} \cdot \mathbf{r}' = \frac{1}{r} + \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}\Big|_{\mathbf{r}'=0} \cdot \mathbf{r}'$$
$$= \frac{1}{r} + \frac{\mathbf{r} \cdot \mathbf{r}'}{r^3}. \tag{1.11}$$

4

Writing unit vectors with a hat, $\hat{\mathbf{a}} = \mathbf{a}/a$, we obtain

$$
\begin{aligned}
\mathbf{A}(\mathbf{r}) &\cong \frac{\mu_0}{4\pi} I \int_0^{2\pi} d\phi' R \, \hat{\boldsymbol{\phi}}' \left(\cancel{\frac{1}{r}} + \frac{\mathbf{r} \cdot \mathbf{r}'}{r^3} \right) \\
&= \frac{\mu_0}{4\pi} I \frac{R^2}{r^2} \int_0^{2\pi} d\phi' \, \hat{\boldsymbol{\phi}}' \, \underbrace{(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')}_{\text{unit vectors}} \\
&= \frac{\mu_0}{4\pi} I \frac{R^2}{r^2} \int_0^{2\pi} d\phi' \left(-\sin\phi' \, \hat{\mathbf{x}} + \cos\phi' \, \hat{\mathbf{y}} \right) \\
&\quad \times \left(\sin\theta\cos\phi\cos\phi' + \sin\theta\sin\phi\sin\phi' + 0 \right) \\
&= \frac{\mu_0}{4\pi} I \frac{R^2}{r^2} \sin\theta \int_0^{2\pi} d\phi' \left(-\sin\phi\sin^2\phi' \, \hat{\mathbf{x}} + \cos\phi\cos^2\phi' \, \hat{\mathbf{y}} \right) \\
&= \frac{\mu_0}{4\pi} I \frac{R^2}{r^2} \sin\theta \left(-\sin\phi \, \hat{\mathbf{x}} + \cos\phi \, \hat{\mathbf{y}} \right) \\
&= \frac{\mu_0}{4\pi} I \frac{R^2}{r^2} \sin\theta \, \hat{\boldsymbol{\phi}} = \frac{\mu_0}{4\pi} I \frac{R^2}{r^2} \hat{\mathbf{z}} \times \hat{\mathbf{r}}.
\end{aligned}
\tag{1.12}
$$

We define

$$
\mathbf{m}_l := \pi I R^2 \hat{\mathbf{z}},
\tag{1.13}
$$

which we will interpret in a moment. Then

$$
\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \mathbf{m}_l \times \frac{\hat{\mathbf{r}}}{r^2}.
\tag{1.14}
$$

The induction is then

$$
\begin{aligned}
\mathbf{B} = \nabla \times \mathbf{A} &= \frac{\mu_0}{4\pi} \nabla \times \left(\mathbf{m}_l \times \frac{\mathbf{r}}{r^3} \right) \\
&= \frac{\mu_0}{4\pi} \left[\mathbf{m}_l \left(\nabla \cdot \frac{\mathbf{r}}{r^3} \right) - (\mathbf{m}_l \cdot \nabla) \frac{\mathbf{r}}{r^3} \right] \\
&= \frac{\mu_0}{4\pi} \frac{3(\mathbf{m}_l \cdot \mathbf{r})\mathbf{r} - r^2 \mathbf{m}_l}{r^5}.
\end{aligned}
\tag{1.15}
$$

This is, not surprisingly, the field of a magnetic dipole. We can thus identify $\mathbf{m}_l$ with the magnetic dipole moment of the current loop.

The magnetic (dipole) moment is thus

$$
\mathbf{m}_l = \pi I R^2 \hat{\mathbf{z}} = -\frac{1}{2} e\omega R^2 \hat{\mathbf{z}} = -\frac{1}{2} e R^2 \boldsymbol{\omega}.
\tag{1.16}
$$

Compare the orbital angular momentum of the electron: We had found $\mathbf{l} = m_e R^2 \boldsymbol{\omega}$. We obtain the relation

$$
\mathbf{m}_l = -\frac{e}{2m_e} \mathbf{l}.
\tag{1.17}
$$

Thus the magnetic moment of the loop is antiparallel to the orbital angular momentum. In quantum mechanics, orbital angular momentum is quantized in units of $\hbar$, therefore we define the *Bohr magneton*

$$
\mu_B := \frac{e\hbar}{2m_e}
\tag{1.18}
$$

and write

$$
\mathbf{m}_l = -\mu_B \frac{\mathbf{l}}{\hbar}.
\tag{1.19}
$$

(b) Spin: We have seen that the magnetic field due to the orbital motion is unlikely to lead to magnetic ordering. We also know from many experiments that electrons carry a magnetic moment $\mathbf{m}_s$ and angular momentum $\mathbf{s}$ even if they move in a straight line or are in an atomic s-state ($l = 0$). Classically, it would be natural to attribute the intrinsic magnetic moment to a spinning charged sphere. However, if the charge and mass distribution of this sphere were identical, we would again get

$$
\mathbf{m}_s \stackrel{?}{=} -\mu_B \frac{\mathbf{s}}{\hbar},
\tag{1.20}
$$

whereas one finds experimentally

$$
\mathbf{m}_s = -g\mu_B \frac{\mathbf{s}}{\hbar}
\tag{1.21}
$$

5

with $g \cong 2.0023$, in good approximation *twice* the expected moment. There is no natural classical explanation for $g \approx 2$.

On the other hand, the relativistic *Dirac quantum theory* does give $g = 2$. The solutions of the Dirac equations are 4-component-vector functions ("Dirac spinors"). In the non-relativistic limit $v \ll c$, two of these components become small and for the other two (a "Pauli spinor") one obtains the *Pauli equation*

$$\left\{ \left[\frac{1}{2m_e} (\mathbf{p} + e\mathbf{A})^2 - e\phi \right] \mathbf{1} + \frac{e}{m_e} \mathbf{s} \cdot \mathbf{B} \right\} |\psi\rangle = (E - m_e c^2)|\psi\rangle, \tag{1.22}$$

$$\uparrow$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{in spinor space}$$

where $\mathbf{p}$ is the momentum operator and $\mathbf{s}$ is an angular momentum operator satisfying

$$\mathbf{s} \cdot \mathbf{s} = \frac{1}{2}\left(\frac{1}{2} + 1\right)\hbar^2 = \frac{3}{4}\hbar^2. \tag{1.23}$$

Writing the Zeeman term as $-\mathbf{m}_s \cdot \mathbf{B}$, we find

$$\mathbf{m}_s = -\frac{e}{m_e}\mathbf{s} = -2\mu_B \frac{\mathbf{s}}{\hbar} \overset{!}{=} -g\mu_B \frac{\mathbf{s}}{\hbar} \quad \text{with} \quad g = 2. \tag{1.24}$$

The interaction of the electronic charge with the electromagnetic field it generates leads to small corrections ("anomalus magnetic moment"), which can be evaluated within QED to very high accurancy. One finds

$$g = 2 + \frac{\alpha}{\pi} + \mathcal{O}(\alpha^2) \tag{1.25}$$

with the fine structure constant

$$\alpha = \frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c} = \mu_0 \frac{e^2 c}{2h} \approx \frac{1}{137}. \tag{1.26}$$

The relevant leading Feynman diagrams are
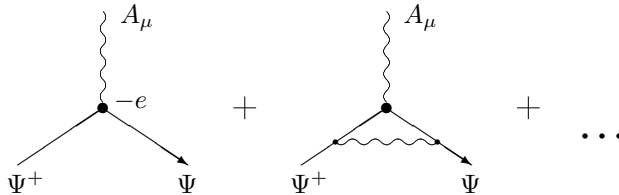
In quantum physics, it is common to write angular momenta as dimensionless quantities by drawing out a factor of $\hbar$. Thus we replace $\mathbf{s} \to \mathbf{s}\hbar$, $\mathbf{s}/\hbar \to \mathbf{s}$, $\mathbf{m}_s \to \mathbf{m}_s = -g\mu_B \mathbf{s}$ etc. We use this convention from now on.

The derivation of the Pauli equation gives $\mathbf{s} \cdot \mathbf{s} = 3/4$. We know from introductory quantum mechanics that any angular momentum operator $\mathbf{L}$ has to satisfy the commutation relations $[L_k, L_l] = i \sum_m \epsilon_{klm} L_m$, $[L_k, L^2] = 0$, which define the spin algebra su(2). (These commutation relations are to a large extend determined by the commutation relations of rotations in three-dimensional space—a purely classical concept. This does of course not fix the value of $\hbar$.) This implies that L^2 can have the eigenvalues $l(l+1)$ with $l = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$ and L_k, $k = x, y, z$, can than have the eigenvalues $m = -l, -l+1, \ldots, l-1, l$.

Thus the Dirac theory, and consequently the Pauli theory, describe particles with spin quantum number $S = 1/2$. As noted, experiments show this to be the correct value for electrons. It is very useful to introduce a representation of the spin algebra for $S = 1/2$. The common but by no means necessary choice are the *Pauli matrices*. We write

$$s_k = \frac{\sigma_k}{2} \tag{1.27}$$

with the Pauli matrices

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.28}$$

One easily checks that $[s_x, s_y] = is_z$ etc. are satisfied. Also,

$$s^2 = \frac{1}{4}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2) = \frac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.29}$$

is proportional to the unit matrix and thus commutes with everything. More generally, one can find $(2l+1)$-dimensional representations for total angular momentum $l \geq 1/2$.

1.3 Dipole-dipole interaction

We have established that electrons in solids carry magnetic moments. Quantum theory (and even QED) was needed to understand the size of the magnetic moment but not its existence. Now we know from electrodynamics that magnetic dipoles interact. The field generated by a magnetic moment $\mathbf{m}_1$ at the origin is

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{3(\mathbf{m}_1 \cdot \mathbf{r})\mathbf{r} - r^2 \mathbf{m}_1}{r^5}. \tag{1.30}$$

The energy of another moment $\mathbf{m}_2$ at $\mathbf{r}$ is then

$$V_{\text{dip}} = -\mathbf{m}_2 \cdot \mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{3(\mathbf{m}_1 \cdot \mathbf{r})(\mathbf{m}_2 \cdot \mathbf{r}) - r^2 \mathbf{m}_1 \cdot \mathbf{m}_2}{r^5}. \tag{1.31}$$

This is clearly symmetric in $\mathbf{m}_1$ and $\mathbf{m}_2$.

Can this dipole-dipole interaction explain magnetic order as we observe it? If it does, we expect the critical temperature T_c, below which the order sets in, to be of the order of the strongest dipolar interaction, which is the one between nearest neighbors. This follows within mean-field theory, as we shall see, and is plausible in general, since the dipolar interaction sets the only obvious energy scale.

A rough estimate is obtained by identifying the nearest neighbor separation with the lattice constant a and writing

$$k_B T_c \sim z \frac{\mu_0}{4\pi} \frac{m_s^2}{a^3} = z \frac{\mu_0}{4\pi} \frac{g^2 \mu_B{}^2}{4a^3} \cong \frac{\mu_0}{4\pi} \frac{e^2 \hbar^2}{4 m_e^2 a^3}, \tag{1.32}$$

where z is the number of nearest neighbors. In the last step we have used $g \cong 2$. Taking $z = 8$ and $a = 2.49\,\text{Å}$ (as appropriate for bcc iron), we get $T_c \sim 0.3\,\text{K}$. But actually iron becomes ferromagnetic at $1043\,\text{K}$. Clearly the dipole-dipole interaction is too weak to explain this.

Interestingly, it is thought that the dipolar interaction can lead to ferromagnetic long-range order on some crystal lattices, including bcc and fcc, even though the interaction is highly anisotropic. This has been predicted by Luttinger and Tisza in 1946. Experimentally, this type of order seems to be realized in $Cs_2NaR(NO_2)_6$ with R = Nd, Gd, Dy, Er.

In any case, we have to search for another, much stronger interaction. As we will discuss in chapter 4, this will turn out to be the Coulomb interaction in conjunction with the Pauli priniciple. Thus quantum mechanics is required for magnetic ordering at high temperatures but not for magnetic ordering per se.

Chapter 2

Magnetism of free atoms and ions

In this chapter we will review the magnetic properties of single atoms and ions. We will from now on subsume atoms under ions.

2.1 The electron shell: Hartree approximation

The nucleus and electrons making up an ion form a complicated many-particle system that we cannot hope to solve exactly. In the simplest non-trivial approximation, the *Hartree approximation*, we assume that a given electron moves in a potential resulting from the nucleus and from the *averaged* charge density due to the other electrons. The word "other" is actually important here. It would certainly be incorrect to include the interaction between an electron and its own averaged charge density. In a solid with of the order of 10^{23} electrons, the correction is negligible but in an ion with a few tens of electrons it is not.

The total potential is thus

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

where Z is the atomic number of the nucleus, the charge of an electron is $-e < 0$, and $\rho_{\mathbf{r}}(\mathbf{r}') < 0$ is the charge density at $\mathbf{r}'$ of the other electrons if the given electron is at $\mathbf{r}$. Due to the isotropy of space, $V_{\text{eff}}(\mathbf{r})$ has spherical symmetry, whereas $\rho_{\mathbf{r}}(\mathbf{r}')$ as a function of $\mathbf{r}'$ does not (unless $\mathbf{r} = 0$). For the given electron, we solve the single-particle Schrödinger equation

$$\underbrace{\left(\frac{1}{2m}p^2 + V_{\text{eff}}(\mathbf{r})\right)}_{H}\psi(\mathbf{r}) = E\,\psi(\mathbf{r}). \tag{2.2}$$

Due to spherical symmetry, the resulting eigenfunctions are

$$\psi_{nlm}(\mathbf{r}) = \underbrace{R_{nl}(r)}_{\text{radial part}}\ \underbrace{Y_{lm}(\theta,\phi)}_{\text{angular part}} \tag{2.3}$$

with $n = 1, 2, 3, \ldots$, $l = 0, 1, 2, \ldots, n-1$, and $m = -l, -l+1, \ldots, l$, as one finds from a separation ansatz. The angular part is identical for any spherically symmetric potential and is given by the spherical harmonics $Y_{lm}(\theta,\phi)$. In the present approximation, the eigenenergies ϵ_{nl} only depend on n, l. Thus ϵ_{nl} is $2(2l+1)$-fold degenerate, including a factor of 2 from the spin $s = 1/2$.

Completely filled shells (made up of all orbitals with the same quantum numbers n, l) have $\langle\sum_i \mathbf{l}_i\rangle = 0$ and $\langle\sum_i \mathbf{s}_i\rangle = 0$, i.e., vanishing total angular momentum, since for each electron there is another one with opposite $\langle\mathbf{l}_i\rangle$, $\langle\mathbf{s}_i\rangle$. Clearly, the total magnetic moment of filled shells also vanishes. Thus magnetic ions require incompletely filled shells.

In the ground state, the Hartree orbitals are filled from the lowest in energy up. If a shell contains n_{nl} electrons there are

$$\binom{2(2l+1)}{n_{nl}} \tag{2.4}$$

possible ways to distribute these electrons, which gives the degeneracy of the many-particle state. Note that for a filled shell we get

$$\binom{2(2l+1)}{2(2l+1)} = 1, \tag{2.5}$$

i.e., no degeneracy.

2.2 Beyond Hartree

The degeneracy found in the previous subsection is partially lifted by the Coulomb repulsion beyond the Hartree approximation. Note that the Coulomb interaction

$$V_c = \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{i\neq j} \frac{e^2}{|\mathbf{r_i} - \mathbf{r_j}|} \tag{2.6}$$

commutes with the total orbital angular momentum (of the shell) $\mathbf{L} := \sum_i \mathbf{l_i}$, since V_c is spherically symmetric, and with the total spin (of the shell) $\mathbf{S} := \sum_i \mathbf{s_i}$ and of course also with $\mathbf{L}^2$ and $\mathbf{S}^2$. $\mathbf{L}$ and $\mathbf{S}$ also commute since they describe completely different degrees of freedom. Thus it is possible to classify the $\binom{2(2l+1)}{n_{nl}}$ many-particle states in terms of quantum numbers L, m_L, S, m_S.

If we now have a state $|\psi\rangle$ with quantum number $m_L < L$ then we can apply the raising operator $L^+ := L_x + iL_y$ to $|\psi\rangle$ and obtain $|\psi'\rangle \propto L^+|\psi\rangle$ with $m'_L = m_L + 1$. However, since $[H, L^+] = 0$, this new state has the same energy as the old one. Since there are $(2L+1)(2S+1)$ states that are connected by $L^\pm$ and $S^\pm$ ($L^- := L_x - iL_y$ etc.), the $\binom{2(2l+1)}{n_{nl}}$-fold degenrate state splits into *multiplets* with fixed L and S and degeneracies $(2L+1)(2S+1)$. Typical energy splittings between multiplets ate of the order of 10eV. The ground-state multiplet is found from the empirical *Hund rules*:

- 1$^{\text{st}}$ *Hund rule*: The ground state multiplet has the maximum possible S. (The maximum S equals the largest possible value of $\langle S_z \rangle$.)

- 2$^{\text{nd}}$ *Hund rule*: If the first rule leaves several possibilities, the state with maximum L is lowest in energy. (The maximum L equals the largest possible value of $\langle L_z \rangle$.)

These rules hold in most cases but not always. We will return to their origin later. A short qualitative explanation can be given as follows:

- For the 1$^{\text{st}}$ rule: same spin and the Pauli principle result in the electrons being further apart, which leads to lower Coulomb repulsion.

- For the 2$^{\text{nd}}$ rule: large L means that the electrons have aligned orbital angular momenta, i.e., rotate in same direction. They are thus further apart, which leads to lower Coulomb repulsion.

The multiplets ate labeled as ^{2S+1}L, where L is denoted by a letter according to

0	1	2	3	4	5	6	...
S	P	D	F	G	H	I	...

2.3 Spin-orbit coupling

We have already seen that a relativistic description is required to understand the magnetic moment of the electron. We will see now that the same holds for the many-particle states of ions.

By taking the non-relativistic limit of the Dirac equation, one arrives at the Pauli equation mentioned above. By including the next order in v/c one obtains additional terms. Technically, this is done using the so-called *Foldy-Wouthuysen transformation*. We only consider the case of a static electric potential. The book by Messiah (vol. 2) contains a clear disscussion. In this case one obtains the Hamiltonian, in spinor space,

$$H = \underbrace{\frac{p^2}{2m_e} + V(\mathbf{r})}_{H_0,\text{ non-relativistic}} - \underbrace{\frac{p^4}{8m_e^3 c^2}}_{\substack{\text{correction to}\\\text{kinetic energy}}} + \underbrace{\frac{\hbar^2}{2m_e^2 c^2}\frac{1}{r}\frac{\partial V}{\partial r}\mathbf{s}\cdot\mathbf{l}}_{\text{spin-orbit coupling}} + \underbrace{\frac{\hbar^2}{8m_e^2 c^2}\nabla^2 V}_{\text{correction to potential}}. \tag{2.7}$$

The last term is also called the "Darwin term". The only term relevant for us is the spin-orbit coupling

$$H_{\text{SO}} = \frac{\hbar^2}{2m_e^2 c^2}\frac{1}{r}\frac{\partial V}{\partial r}\mathbf{s}\cdot\mathbf{l}. \tag{2.8}$$

For the Coulomb potential of the nucleus,

$$H_{\text{SO}} = -\frac{\hbar^2}{2m_e^2 c^2}\frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r}\frac{\partial}{\partial r}\frac{1}{r}\mathbf{s}\cdot\mathbf{l} = \frac{\hbar^2}{2m_e^2 c^2}\frac{Ze^2}{4\pi\epsilon_0}\frac{\mathbf{s}\cdot\mathbf{l}}{r^3} = \frac{\mu_0}{4\pi}g\mu_B^2 Z\frac{\mathbf{s}\cdot\mathbf{l}}{r^3}, \quad \text{assuming } g = 2. \tag{2.9}$$

For several electrons in an incompletely filled shell, the operator of spin-orbit coupling is

$$H_{\text{SO}} = \frac{\mu_0}{4\pi} g \mu_B^2 Z \sum_i \frac{\mathbf{s}_i \cdot \mathbf{l}_i}{r_i^3}. \tag{2.10}$$

In principle, we should include not only the nuclear potential but the full effective potential of the Hartree approximation. In practice, this is expressed by replacing the atomic number Z by an effective one, $Z_{\text{eff}} < Z$.

We now evaluate the contribution of spin-orbit coupling to the energy, treating H_{SO} as a weak pertubation to H_0. Then

$$E_{\text{SO}} := \langle H_{\text{sO}} \rangle = \frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \sum_i \left\langle \frac{\mathbf{s}_i \cdot \mathbf{l}_i}{r_i^3} \right\rangle. \tag{2.11}$$

For free ions, the radial wave function $R_{nl}(r)$ is the same for all orbitals comprising a shell. Thus

$$E_{\text{SO}} = \frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \left\langle \frac{1}{r^3} \right\rangle_{nl} \sum_i \langle \mathbf{s}_i \cdot \mathbf{l}_i \rangle. \tag{2.12}$$

We now call the electrons with spin parallel to $\mathbf{S}$ "spin up" ($\uparrow$) and the others "spin down" ($\downarrow$). Furthermore, $\mathbf{s}_i$ and $\mathbf{l}_i$ commute. We can thus replace, *in the expectation value*, $\mathbf{s}_i$ by $\mathbf{S}/2S$ for spin up and by $-\mathbf{S}/2S$ for spin down, respectively. (Note that $\mathbf{s}_i$ has magnitude $1/2$.) Thus

$$E_{\text{SO}} = \frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \left\langle \frac{1}{r^3} \right\rangle_{nl} \left(\sum_{\substack{i \\ \text{spin up}}} \frac{\langle \mathbf{S} \cdot \mathbf{l}_i \rangle}{2S} - \sum_{\substack{i \\ \text{spin down}}} \frac{\langle \mathbf{S} \cdot \mathbf{l}_i \rangle}{2S} \right). \tag{2.13}$$

We have three cases:

- If the shell is less than half filled, $n_{nl} < 2l + 1$, all spins are aligned and the spin-down sum does not contain any terms. Then

$$E_{\text{SO}} = \frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} \left\langle \mathbf{S} \cdot \sum_i \mathbf{l}_i \right\rangle = \frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} \langle \mathbf{S} \cdot \mathbf{L} \rangle =: \lambda \langle \mathbf{L} \cdot \mathbf{S} \rangle \tag{2.14}$$

  with

$$\lambda = \frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S}. \tag{2.15}$$

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

$$\sum_{m_l=-l}^{l} \langle l m_l | \mathbf{l} | l m_l \rangle = 0 \tag{2.16}$$

  and we obtain

$$E_{\text{SO}} = -\frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S} \langle \mathbf{S} \cdot \mathbf{L} \rangle =: \lambda \langle \mathbf{L} \cdot \mathbf{S} \rangle \tag{2.17}$$

  with

$$\lambda = -\frac{\mu_0}{4\pi} g \mu_B^2 Z_{\text{eff}} \left\langle \frac{1}{r^3} \right\rangle_{nl} \frac{1}{2S}. \tag{2.18}$$

- If the shell is half filled, $n_{nl} = 2l + 1$, both the spin-up and the spin-down sum vanish and we get $E_{\text{SO}} = 0$. Note that one does find a contribution at higher order in pertubation theory.

We have found that the spin-orbit coupling in a free ion behaves, within pertubation theory, like a term $H_{\text{SO}} = \lambda \mathbf{L} \cdot \mathbf{S}$ in the Hamiltonian, where $\lambda > 0$ ($\lambda < 0$) for less (more) then half filled shells. This *LS-coupling* splits the $(2L + 1)(2S + 1)$-fold degeneracy.

We introduce the total angular momentum operator $\mathbf{J} := \mathbf{L} + \mathbf{S}$. We can then write

$$H_{\text{SO}} = \lambda \mathbf{L} \cdot \mathbf{S} = \frac{\lambda}{2} \left[(\mathbf{L} + \mathbf{S})^2 - L^2 - S^2 \right] = \frac{\lambda}{2} \left[J^2 - L^2 - S^2 \right]. \tag{2.19}$$

The full Hamiltonian including H_{SO} commutes with $\mathbf{J}$, and thus with $\mathbf{J}^2$. It does not commute with $\mathbf{L}$ or $\mathbf{S}$ because of H_{SO} but it does commute with $\mathbf{L}^2$ and $\mathbf{S}^2$. Therefore, we can replace $\mathbf{J}^2, \mathbf{L}^2, \mathbf{S}^2$ by their eigenvalues,

$$H_{\text{SO}} \rightarrow \frac{\lambda}{2} \left[J(J+1) - L(L+1) - S(S+1) \right]. \tag{2.20}$$

As we know from quantum mechanics, J can assume the values

$$J = |L - S|, |L - S| + 1, \ldots, L + S. \tag{2.21}$$

Due to H_{SO}, the energy depends not only on L, S but also on J. Since $\mathbf{J}$ commutes with the Hamiltonian, the energy does not depend on the magnetic quantum number $m_J = -J, -J+1, \ldots, J$. The $(2L+1)(2S+1)$-fold degenrate multiplet is thus split into multiplets with fixed J and degenracies $2J + 1$. The sign of λ decides on the ground-state multiplet. This is the...

3rd *Hund rule*: Among the low-energy multiplets with S, L given by the first two rules, the ground-state multiplet has

- minimum J, i.e., $J = |L - S|$, for less than half filled shells,

- maximum J, i.e., $J = L + S$, for more than half filled shells.

For half filled shells we have $L = 0$ and thus $J = S$ anyway.

The notation is now extended to include J: ${}^{2S+1}L_J$, where L is again written as the corresponding letter.

Example: What is the ground state of Nd^{3+}? The electron configuration is $[\mathrm{Xe}]4f^3$. Hund 1: $S \to$ max, thus $S = 3/2$. Hund 2: $L \to$ max ...
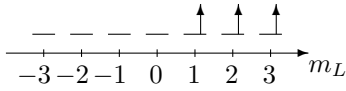
thus $L = 3 + 2 + 1 = 6$, symbol I. Hund 3: less than half filled, $J \to$ min, thus $J = |L - S| = 6 - \frac{3}{2} = \frac{9}{2}$. Thus we obtain a ${}^4I_{9/2}$ multiplet.

2.4 Magnetic moments of ions

When we want to calculate the magnetic moment of an ion with quantum numbers S, L, J, we encounter a problem: Taking the g-factor of the electron spin to be $g = 2$, the magnetic moment is

$$\mathbf{m}_J = \mathbf{m}_S + \mathbf{m}_L = -2\mu_B\mathbf{S} - \mu_B\mathbf{L} = -\mu_B(2\mathbf{S} + \mathbf{L}) = -\mu_B(\mathbf{J} + \mathbf{S}). \tag{2.22}$$

But $\mathbf{m}_J$ does not commute with the Hamiltonian because of the spin-orbit coupling term $\lambda\mathbf{L} \cdot \mathbf{S}$. ($\mathbf{J}$ does commute but $\mathbf{S}$ does not.) Thus $\mathbf{m}_J$ is not a constant of motion, whereas $\mathbf{J}$ is. We can think of $\mathbf{S}$ and $\mathbf{L}$ and thus $\mathbf{m}_J$ as rotating around the fixed vector $\mathbf{J}$:
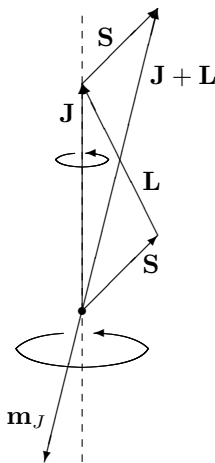
The typical timescale of this rotation should be $h/|\lambda|$. For "slow" experiments like magnetization measurements, only the time-averaged moment $\mathbf{m}_{\mathrm{obs}}$ will be observable. To find it, we project $\mathbf{m}_J$ onto the direction of the constant $\mathbf{J}$:

$$\mathbf{m}_{\mathrm{obs}} = \frac{(\mathbf{m}_J \cdot \mathbf{J})\mathbf{J}}{\mathbf{J} \cdot \mathbf{J}} = -\mu_B\frac{[(\mathbf{J} + \mathbf{S}) \cdot \mathbf{J}]\,\mathbf{J}}{\mathbf{J} \cdot \mathbf{J}} = -\mu_B\mathbf{J} - \mu_B\frac{(\mathbf{S} \cdot \mathbf{J})\mathbf{J}}{\mathbf{J} \cdot \mathbf{J}}$$
$$= -\mu_B\mathbf{J} + \frac{\mu_B}{2}\frac{(\mathbf{J} - \mathbf{S})^2 - \mathbf{J} \cdot \mathbf{J} - \mathbf{S} \cdot \mathbf{S}}{\mathbf{J} \cdot \mathbf{J}}\mathbf{J}. \tag{2.23}$$

Thus since $\mathbf{J} - \mathbf{S} = \mathbf{L}$,

$$\mathbf{m}_{\mathrm{obs}} = -\mu_B \mathbf{J} - \frac{\mu_B}{2} \frac{J(J+1) + S(S+1) - L(L+1)}{J(J+1)} \mathbf{J} =: -g_J \mu_B \mathbf{J}, \qquad (2.24)$$

where we have introduced the *Landé g-factor*

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \qquad (2.25)$$

Note that g_J satifies $0 \leq g_J \leq 2$. It can actually be smaller than the orbital value of unity.

2.5 The nuclear spin and magnetic moment

Protons and neutrons are both spin-1/2 fermions. Both carry magnetic moments, which might be surprising for the neutron since it does not carry a net charge (the neutrinos, also spin-1/2 fermions, do not have magnetic moments). But the neutron, like the proton, consists of charged quarks. Due to this substructure, the g-factor of the proton and the neutron are not close to simple numbers:

- Proton: $\mathbf{m}_p = \underbrace{5.5856}_{=g_p} \mu_N \mathbf{s}$

  with the nuclear magneton $\mu_N := e\hbar/2m_p$ and the proton mass m_p. It is plausible that in the typical scale the electron mass should be replaced by the proton mass. μ_B is by a factor of nearly 2000 larger than μ_N, showing that nuclear magnetic moments are typically small. Incidentally, this suggests that the nuclear moments do not carry much of the magnetization in magnetically ordered solids.

- Neutron: $\mathbf{m}_n = \underbrace{-3.8261}_{=g_n} \mu_N \mathbf{s}$ (note that this is antiparallel to $\mathbf{s}$).

In nuclei consisting of several nucleons, the total spin $\mathbf{I}$ has contibutions from the proton and neutron spins and from the orbital motion of the protons and neutrons. The nuclei also have a magnetic moment $\mathbf{m}_N$ consisting of the spin magnetic moments of protons and neutrons and the orbital magnetic moments of the protons only. The neutrons do not produce orbital currents. The orbital g-factors are thus $g_p^l = 1$ and $g_n^l = 0$.

Due to the different relevant g-factors, the instantaneous magnetic moment is not alligned with $\mathbf{I}$. In analogy to Sec. 2.4, only the averaged moment $\mathbf{m}_N$ parallel to $\mathbf{I}$ is observable, leading to distinct muclear g-factors. We thus have

$$\mathbf{m}_N = g_N \mu_N \mathbf{I} \qquad (2.26)$$

with nucleaus-specific g_N. Typically, $|g_N|$ is of the order of 1 to 10. g_N is positive for most nuclei but negative for some.

2.6 Hyperfine interaction

The hyperfine interaction is the interaction between the electrons and the nucleus beyond the Coulomb attraction, which we have already taken into account. The origin of the name is that this interaction leads to very small splittings in atomic spectra.

The first obvious contribution to the hyperfine interaction is the magnetic dipole-dipole interaction between electrons and nucleus. Naively, we would write

$$V_{\mathrm{dip}} = \frac{\mu_0}{4\pi} \frac{3(\mathbf{m}_e \cdot \mathbf{r})(\mathbf{m}_N \cdot \mathbf{r}) - r^2 \mathbf{m}_e \cdot \mathbf{m}_N}{r^5}. \qquad (2.27)$$

This leads to a divergence if the elctron can by at the position of the nucleus, which is the case for s-orbitals.

We have to calculate the $\mathbf{B}$-field due to the nuclear moment more carefully. We have seen in Sec. 1.2 that the vector potential is

$$\mathbf{A} = \frac{\mu_0}{4\pi} \mathbf{m}_N \times \frac{\hat{\mathbf{r}}}{r^2} = \frac{\mu_0}{4\pi} \mathbf{m}_N \times \frac{\mathbf{r}}{r^3}. \qquad (2.28)$$

Thus

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} = \frac{\mu_0}{4\pi} \boldsymbol{\nabla} \times \left(\mathbf{m}_N \times \frac{\mathbf{r}}{r^3} \right) = -\frac{\mu_0}{4\pi} \boldsymbol{\nabla} \times \left(\mathbf{m}_N \times \boldsymbol{\nabla} \frac{1}{r} \right)$$

$$= +\frac{\mu_0}{4\pi} \boldsymbol{\nabla} \times \left(\boldsymbol{\nabla} \frac{1}{r} \times \mathbf{m}_N \right) = \frac{\mu_0}{4\pi} \boldsymbol{\nabla} \times \left(\boldsymbol{\nabla} \times \frac{\mathbf{m}_N}{r} \right). \qquad (2.29)$$

With $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{F}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{F}) - \nabla^2 \mathbf{F}$ we find

$$\mathbf{B} = \frac{\mu_0}{4\pi} \boldsymbol{\nabla} \left(\boldsymbol{\nabla} \cdot \frac{\mathbf{m}_N}{r} \right) - \frac{\mu_0}{4\pi} \nabla^2 \frac{\mathbf{m}_N}{r}. \tag{2.30}$$

Now we use a trick: we split the second term into two parts and apply $\nabla^2(1/r) = -4\pi\delta(\mathbf{r})$ to the second:

$$\mathbf{B} = \frac{\mu_0}{4\pi} \left(\underbrace{\boldsymbol{\nabla}\boldsymbol{\nabla}}_{\text{dyad}} - \frac{1}{3}\nabla^2 \right) \frac{\mathbf{m}_N}{r} + \frac{2}{3}\mu_0 \mathbf{m}_N \delta(\mathbf{r}). \tag{2.31}$$

Why did we do this? The first term is well defined for $\mathbf{r} \neq 0$ but what happens at $\mathbf{r} = 0$? We can show that the first term does not have a singularity there and can thus be analytically continued to $\mathbf{r} = 0$ by choosing an irrelevant finite value.

Proof: consider a sphere $\mathcal{S}$ centered at the origin. We integrate

$$\left(\boldsymbol{\nabla}\boldsymbol{\nabla} - \frac{1}{3}\nabla^2 \right) \frac{\mathbf{m}_N}{r} \tag{2.32}$$

over $\mathcal{S}$:

$$\int_{\mathcal{S}} \left(\boldsymbol{\nabla}\boldsymbol{\nabla} - \frac{1}{3}\nabla^2 \right) \frac{\mathbf{m}_N}{r} = \underbrace{\left[\int_{\mathcal{S}} d^3r \left(\boldsymbol{\nabla}\boldsymbol{\nabla} - \frac{1}{3}\nabla^2 \right) \frac{1}{r} \right]}_{=:Q} \mathbf{m}_N. \tag{2.33}$$

The first factor Q is a matrix acting on $\mathbf{m}_N$. But Q does not distinguish any direction in space and thus has to be of the form $Q = q\mathbf{1}$ with $q \in \mathbb{R}$. On the other hand, the trace of Q is

$$\begin{aligned}
3q &= \operatorname{Tr} Q = Q_{xx} + Q_{yy} + Q_{zz} \\
&= \int_{\mathcal{S}} d^3r \operatorname{Tr} \begin{pmatrix} \frac{2}{3}\frac{\partial^2}{\partial x^2} - \frac{1}{3}\frac{\partial^2}{\partial y^2} - \frac{1}{3}\frac{\partial^2}{\partial z^2} & \frac{\partial^2}{\partial x \partial y} & \frac{\partial^2}{\partial x \partial z} \\ \frac{\partial^2}{\partial x \partial y} & -\frac{1}{3}\frac{\partial^2}{\partial x^2} + \frac{2}{3}\frac{\partial^2}{\partial y^2} - \frac{1}{3}\frac{\partial^2}{\partial z^2} & \frac{\partial^2}{\partial y \partial z} \\ \frac{\partial^2}{\partial x \partial z} & \frac{\partial^2}{\partial y \partial z} & -\frac{1}{3}\frac{\partial^2}{\partial x^2} - \frac{1}{3}\frac{\partial^2}{\partial y^2} + \frac{2}{3}\frac{\partial^2}{\partial z^2} \end{pmatrix} \frac{1}{r} \\
&= \int_{\mathcal{S}} d^3r \left(\frac{2}{3}\frac{\partial^2}{\partial x^2} - \frac{1}{3}\frac{\partial^2}{\partial y^2} - \frac{1}{3}\frac{\partial^2}{\partial z^2} + \frac{2}{3}\frac{\partial^2}{\partial y^2} - \frac{1}{3}\frac{\partial^2}{\partial x^2} - \frac{1}{3}\frac{\partial^2}{\partial z^2}\frac{2}{3}\frac{\partial^2}{\partial z^2} - \frac{1}{3}\frac{\partial^2}{\partial x^2} - \frac{1}{3}\frac{\partial^2}{\partial y^2} \right) \frac{1}{r} \\
&= 0. \tag{2.34}
\end{aligned}$$

We have thus found that $Q = 0$. We now make the radius of $\mathcal{S}$ arbitrarily small and find that there is no singularity at $\mathbf{r} = 0$.

For $\mathbf{r} \neq 0$ we can evaluate $\mathbf{B}$ as in Sec. 1.2 and find

$$\mathbf{B} = \frac{\mu_0}{4\pi} \frac{3(\mathbf{m}_N \cdot \mathbf{r})\mathbf{r} - r^2 \mathbf{m}_N}{r^5} \bigg|_{\mathbf{r} \neq 0} + \frac{2}{3}\mu_0 \mathbf{m}_N \delta(\mathbf{r}). \tag{2.35}$$

What we have achieved so far is to make the singularity explicit as the second term. The resulting interaction energy between the nuclear moment and the *spin* moment of an electron is

$$\begin{aligned}
E_1 &= -\mathbf{m}_{e,\text{spin}} \cdot \mathbf{B} \\
&= \frac{\mu_0}{4\pi} \frac{\mathbf{m}_{e,\text{spin}} \cdot \mathbf{m}_N}{r^3} - \frac{\mu_0}{4\pi} 3 \frac{(\mathbf{r} \cdot \mathbf{m}_{e,\text{spin}})(\mathbf{r} \cdot \mathbf{m}_N)}{r^5} - \frac{2}{3}\mu_0 \mathbf{m}_{e,\text{spin}} \cdot \mathbf{m}_N \delta\mathbf{r} \\
&= -\frac{\mu_0}{4\pi} g\mu_B \frac{\mathbf{s} \cdot \mathbf{m}_N}{r^3} + \frac{\mu_0}{4\pi} 3 g\mu_B \frac{(\mathbf{r} \cdot \mathbf{s})(\mathbf{r} \cdot \mathbf{m}_N)}{r^5} + \frac{2}{3}\mu_0 g\mu_B \mathbf{s} \cdot \mathbf{m}_N \delta(\mathbf{r}). \tag{2.36}
\end{aligned}$$

The last, δ-function term is called the *Fermi contact interaction*. There are two cases:

- For s-orbitals the probability density is spherically symmetric. The expectation value of the normal dipole interaction then vanishes by the same argument as in the proof above. Only the contact term remains:

$$E_1 = \frac{2}{3} g\mu_0 \mu_B \mathbf{s} \cdot \mathbf{m}_N |\psi_s(0)|^2, \tag{2.37}$$

  where we have averaged over space but kept the spin degrees of freedom. Thus

$$E_1 = \frac{2}{3} g g_{\text{N}} \mu_0 \mu_B \, \mathbf{s} \cdot \mathbf{I} \, |\psi_s(0)|^2 =: -J_{\text{hyper}} \, \mathbf{s} \cdot \mathbf{I}. \tag{2.38}$$

  This term clearly leads to a splitting of ionic energy spectra.

- For p-, d-, and f-orbitals the contact term vanishes due to $\psi(0) = 0$ and only the dipole interaction survives. It is typically smaller.

  However, 3d transition metals like Fe show a rather large hyperfine splitting of the form $-J_{\text{hyper}}\, \mathbf{S}\cdot\mathbf{I}$ of a contact term. But the spin magnetic moment results from the partially filled 3d-shell with $|\psi_d(0)|^2 = 0$, while the s-shells are all filled or empty. The explanation is that the filled s-shells are polarized by the exchange interaction (see below) with the d-electrons. There is no resulting s-spin or s-moment but the probability to find an s-electron at the position of the nucleus ($\mathbf{r} = 0$) is different for $\uparrow$ and $\downarrow$.

spin down

spin up

nucleus

2s–shell
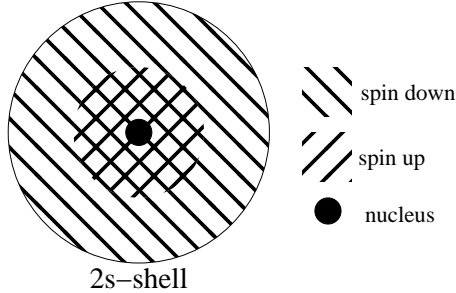
The above arguments do not apply to the orbital angular momentum of the electron since we cannot claim that it leads to a magnetic moment localized at the position $\mathbf{r}$ of the electron. However, we can calculate the corresponding hyperfine interaction from magnetostatics, since the electron is much faster than the typical timescale of the nuclear motion. Thus we can treat the orbital motion as a stationary current and use the Biot-Savart law. This is similar to Sec. 1.2 but we now need the $\mathbf{B}$ field at the nuclear position $\mathbf{r} = 0$ and not for $r \gg R$. We start from

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} I \int \frac{d\mathbf{l}' \times \Delta\mathbf{r}}{\Delta r^3}, \tag{2.39}$$

which implies

$$\mathbf{B}(0) = \frac{\mu_0}{4\pi} I \int \frac{d\mathbf{l}' \times (-\mathbf{r}')}{(\mathbf{r}')^3} = -\frac{\mu_0}{4\pi} I \int_0^{2\pi} d\phi' \frac{R\hat{\phi}' \times R\hat{\mathbf{r}}'}{R^3} = -\frac{\mu_0}{4\pi} \frac{I}{R} \int_0^{2\pi} d\phi' (-\hat{\mathbf{z}})$$
$$= \frac{\mu_0 I}{2R} \hat{\mathbf{z}} = -\frac{\mu_0 e}{4\pi R} \boldsymbol{\omega} = -\frac{\mu_0 e}{4\pi} \frac{\hbar\mathbf{l}}{m_e R^3} = -\frac{\mu_0}{2\pi} \frac{\mu_B}{R^3} \mathbf{l}. \tag{2.40}$$

This leads to the interaction energy

$$E_2 = -\mathbf{m}_N \cdot \mathbf{B} = \frac{\mu_0}{4\pi} \frac{2\mu_B}{r^3} \mathbf{m}_N \cdot \mathbf{l}, \tag{2.41}$$

where we have written r for the electron-nucleus distance. Setting $g = 2$, we can rewrite this as

$$E_2 = \frac{\mu_0}{4\pi} g\mu_B \frac{\mathbf{l} \cdot \mathbf{m}_N}{r^3}. \tag{2.42}$$

Compare this with Eq. (2.36): The first term has the same form but opposite sign.

We mention in passing two further contributions to the hyperfine interaction:

- the interaction between the *electric quadrupole moment* of the nucleus with the electric field generated by the electrons (nuclei do not have electric dipole moments),

- the so-called *isomer shift* due to the non-zero size of the nucleus appearing in the Coulomb interaction.

Chapter 3

Magnetic ions in crystals

In this chapter, we study magnetic ions in crystal lattices. The crystal breaks the isotropy of space, which mainly affects the spatial motion of the electrons, i.e., the orbital angular momentum and its contribution to the magnetic moment. We assume that all electrons remain bound to their ions. In particular, we do not yet consider metals—these will be discussed in chapters 9 and 10 below.

3.1 Crystal field effects: general considerations

Crystal field effects are, as the name implies, the effects of the crystal on an ion. We consider the most important cases of 3d (4d, 5d) and 4f (5f) ions. In stable states, these ions typically lack the s-electrons from the outermost shell and sometimes some of the d- and/or f-electrons. d- and f- ions behave quite differently in crystals. We consider the examples for Fe^{2+} and Gd^{3+}:

3d	4f
e.g., Fe^{2+}	e.g. Gd^{3+}

$3d^6$
3s, 3p
2s, 2p
1s

$4f^7$
5s,5p
4s, 4f, 4d
3s, 3p, 3d
2s, 2p
1s

partially filled d-shell on the outside of ion

partially filled f-shell on the inside of ion

$\Downarrow$

$\Downarrow$

The d-shell overlaps strongly with surrounding ions, thus crystal field effects are strong due to hybridization. We have to treat the crystal filed first, then LS-coupling as a perturbation.

The f-shell hardly overlaps, thus crystal-field effects are due to electro-static potential and therefore weak. We have to treat LS-coupling first, then the crystal field as a perturbation.

3.2 Rare-earth ions and the electrostatic potential

We first consider the electrostatic potential $\phi_{\text{cryst}}(\mathbf{r})$ due to the other ions acting on the electrons. As noted, this is most relevant for f-ions. The electrostatic potential leads to a potential energy $V_{\text{cryst}}(\mathbf{r}) = -e\phi_{\text{cryst}}(\mathbf{r})$,

which we can expand into multipoles:

$$V_{\text{cryst}}(\mathbf{r}) = -\frac{1}{4\pi\epsilon_0} \int d^3r' \frac{e\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

$$= -\frac{e}{4\pi\epsilon_0} \int d^3r' \, \rho(\mathbf{r}') \, 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{1}{2l+1} \frac{r^l}{(r')^{l+1}} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi), \tag{3.1}$$

where we have assumed that r (inside the ion) is smaller than r' (outside of the ion). With

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) \, e^{im\phi}, \tag{3.2}$$

where the $P_l^m(x)$ are the associated Legendre functions, we can write

$$V_{\text{cryst}}(\mathbf{r}) = \sum_{lm} K_{lm} \, r^l \, P_l^m(\cos\theta) \, e^{im\phi} \tag{3.3}$$

with the coefficients given by comparison with the previous expression,

$$K_{lm} = -\frac{e}{4\pi\epsilon_0} \frac{(l-m)!}{(l+m)!} \int d^3r' \frac{\rho(\mathbf{r}')}{(r')^{l+1}} P_l^m(\cos\theta') \, e^{-im\phi'}. \tag{3.4}$$

We note that for $m \neq 0$

$$K_{lm} = -\frac{e}{4\pi\epsilon_0} \frac{(l+m)!}{(l-m)!} \int d^3r' \frac{\rho(\mathbf{r}')}{(r')^{l+1}} P_l^{-m}(\cos\theta') e^{im\phi'}$$

$$= -\frac{e}{4\pi\epsilon_0} \frac{\cancel{(l+m)!}}{\cancel{(l-m)!}} (-1)^m \frac{\cancel{(l-m)!}}{\cancel{(l+m)!}} \int d^3r' \frac{\rho(\mathbf{r}')}{(r')^{l+1}} P_l^m(\cos\theta') e^{im\phi'} = (-1)^m \frac{(l+m)!}{(l-m)!} K_{lm}^* \tag{3.5}$$

so that

$$K_{lm} P_l^m(\cos\theta) e^{im\phi} + K_{l,-m} P_l^{-m}(\cos\theta) e^{-im\phi}$$

$$= K_{lm} P_l^m(\cos\theta) e^{im\phi} + (-1)^m \frac{(l+m)!}{(l-m)!} K_{lm}^* (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(\cos\theta) e^{-im\phi}$$

$$= \left(K_{lm} e^{im\phi} + K_{lm}^* e^{-im\phi} \right) P_l^m(\cos\theta). \tag{3.6}$$

This shows that the terms combine to make the potential energy real.

Which coefficients K_{lm} are non-zero depends on the symmetry of the crystal, i.e., of $\rho(\mathbf{r}')$, under (proper) rotations and rotation inversions about the ion position. For example, for a cubic lattice, one can easily check that $K_{2m} = 0$, $K_{4,\pm1} = K_{4,\pm2} = K_{4,\pm3} = 0$, and $K_{44} = K_{4,-4} = K_{40}/336$ so that

$$V_{\text{cryst}}(\mathbf{r}) \cong \underbrace{K_{00}}_{\text{irrelevant constant}} + K_{40} \, r^4 \left[P_4^0(\cos\theta) + \frac{e^{4i\phi} + e^{-4i\phi}}{336} P_4^4(\cos\theta) \right] + \dots$$

$$= \text{const} + K_{40} \, r^4 \left[\frac{1}{8}(35\cos^4\theta - 30\cos^2\theta + 3) + \frac{\cos 4\phi}{168} 105(\cos^4\theta - 2\cos^2\theta + 1) \right] + \dots$$

$$= \text{const} + \frac{5}{2} K_{40} \left(x^4 + y^4 + z^4 - \frac{3}{5} r^4 \right) + \dots \tag{3.7}$$

after some algebra.

We consider an ion with groundstate multiplet characterized by the total angular momentum J. We only operate within this $(2J+1)$-dimensional subspace. The ion is subjected to the potential V_{cryst} as a weak perturbation. Within this subspace, V_{cryst} has the matrix elements $\langle m_J | V_{\text{cryst}} | m_J' \rangle$, $m_J, m_J' = -J, \dots, J$. The main result is that x^p, y^p, z^p have the same matrix elements within this subspace as J_x^p, J_y^p, J_z^p up to a scalar factor. This factor contains the average $\langle r^p \rangle$ for the relevant shell—this is already dictated by symmetry and dimensional analysis—and a number that depends on the power p and on the shell, i.e., on the quantum numbers n, l, S, L, J. This rule is ambiguous if we have products like xy, since J_x and J_y do not commute so that and $J_x J_y$ and $J_y J_x$ are both plausible but not identical. The rule here is to symmetrize such products: $(J_x J_y + J_y J_x)/2$.

In particular, we get

$$r^2 = x^2 + y^2 + z^2 \quad \rightarrow \quad J_x^2 + J_y^2 + J_z^2 = \mathbf{J}^2 = J(J+1) \tag{3.8}$$

and

$$r^4 = \left(x^2 + y^2 + z^2\right)^2 = x^4 + y^4 + z^4 + 2x^2y^2 + 2y^2z^2 + 2z^2x^2$$

$$\rightarrow J_x^4 + J_y^4 + J_z^4 + \frac{1}{3}\left(J_x^2 J_y^2 + J_x J_y J_x J_y + J_x J_y^2 J_x + J_y J_x^2 J_y + J_y J_x J_y J_x + J_y^2 J_x^2 + \right.$$
$$\left. + \text{(12 more terms)}\right)$$

$$= \left(J_x^2 + J_y^2 + J_z^2\right)^2 + \frac{1}{3}\left(-2J_x^2 J_y^2 + J_x J_y J_x J_y + J_x J_y^2 J_x + J_y J_x^2 J_y + J_y J_x J_y J_x - 2J_y^2 J_x^2 + \right.$$
$$\left. + \text{(12 more terms)}\right)$$

$$= [J(J+1)]^2 + \frac{1}{3}\left(-2\cancel{J_x^2 J_y^2} + \cancel{J_x^2 J_y^2} - iJ_x J_z J_y + \cancel{J_x^2 J_y^2} - iJ_x J_y J_z - iJ_x J_z J_y + \cancel{J_x^2 J_y^2} - \right.$$
$$- iJ_x J_z J_y - iJ_z J_x J_y + \underbrace{J_x J_y^2 J_x}_{=\cancel{J_x^2 J_y^2} - iJ_x J_y J_z - iJ_x J_z J_y} - iJ_z J_y J_x - 2\cancel{J_x^2 J_y^2} + 2iJ_y J_x J_z +$$
$$\left. + 2iJ_y J_z J_x + 2iJ_x J_z J_y + 2iJ_z J_x J_y + \text{(12 more)}\right)$$

$$= J^2(J+1)^2 + \frac{1}{3}\left(-\cancel{iJ_x J_y J_z} - \cancel{J_x^2} - \cancel{iJ_x J_y J_z} - \cancel{J_x^2} - \cancel{iJ_x J_y J_z} - \cancel{J_x^2} - \cancel{iJ_x J_y J_z} - \right.$$
$$- \cancel{J_x^2} + \cancel{J_y^2} - \cancel{iJ_x J_y J_z} - \cancel{iJ_x J_y J_z} - J_x^2 - \cancel{iJ_x J_y J_z} - J_z^2 + \cancel{J_y^2} - J_x^2 + \underline{2iJ_x J_y J_z} + 2J_z^2 +$$
$$+ \underline{2iJ_x J_y J_z} + 2J_z^2 - 2\cancel{J_y^2} + \underline{2iJ_x J_y J_z} + 2\cancel{J_x^2} + \underline{2iJ_x J_y J_z} + 2\cancel{J_x^2} - 2J_y^2 + \text{(12 more)}\Big)$$

$$= J^2(J+1)^2 + \frac{1}{3}\left(-2J_x^2 - 2J_y^2 + 3J_z^2 - 2J_y^2 - 2J_z^2 + 3J_x^2 - 2J_z^2 - 2J_x^2 + 3J_y^2\right)$$

$$= J^2(J+1)^2 - \frac{1}{3}\left(J_x^2 + J_y^2 + J_z^2\right) = J^2(J+1)^2 - \frac{1}{3}J(J+1). \tag{3.9}$$

Thus in the above example of a *cubic* lattice, the leading crystal-field term in the Hamiltonian can be written as

$$\beta\left\langle r^4\right\rangle\left(J_x^4 + J_y^4 + J_z^4 - \frac{3}{5}\mathbf{J}^4 + \frac{1}{5}\mathbf{J}^2\right) = \beta\left\langle r^4\right\rangle\left[J_x^4 + J_y^4 + J_z^4 - \frac{3}{5}J^2(J+1)^2 + \frac{1}{5}J(J+1)\right]. \tag{3.10}$$

Here, β is a number. How does this *single-ion-anisotropy* term split the $2J+1$ states of a multiplet? the last two terms are proportional to the idendity operator and thus irrelevant for the splitting. The first term we rewrite as

$$\beta\left\langle r^4\right\rangle\left(J_x^4 + J_y^4 + J_z^4\right) = \beta\left\langle r^4\right\rangle\left[\frac{1}{16}(J_+ + J_-)^4 + \frac{1}{16}(J_+ - J_-)^4 + J_z^4\right]$$

$$= \beta\left\langle r^4\right\rangle\left[\frac{1}{8}J_+^4 + \frac{1}{8}J_+^2 J_-^2 + \frac{1}{8}J_+ J_- J_+ J_- + \frac{1}{8}J_+ J_-^2 J_+ + \frac{1}{8}J_- J_+^2 J_- + \right.$$
$$\left. + \frac{1}{8}J_- J_+ J_- J_+ + \frac{1}{8}J_-^2 J_+^2 + \frac{1}{8}J_-^4 + J_z^4\right]. \tag{3.11}$$

Working in the eigenbasis $\{|m_J\rangle\}$ of J_z, we see that $J_\pm^4$ only connect states that differ in m_J by 4. Furthermore, the states $|m_J\rangle$ are eigenstates of the mixed terms $J_+^2 J_-^2$, $J_+ J_- J_+ J_-$ etc. and of J_z^4. This means that we can express all the mixed terms as polynomials of J_z. Since spin inversion commutes with the anisotropy term, all terms in these polynomials have to be even. The resulting matrix $\langle m_J | J_x^4 + J_y^4 + J_z^4 | m_J'\rangle$ has the general structure

$$\begin{pmatrix} \star & 0 & 0 & 0 & \star & 0 & 0 \cdots \\ 0 & \star & 0 & 0 & 0 & \star & 0 \cdots \\ 0 & 0 & \star & 0 & 0 & 0 & \star \cdots \\ 0 & 0 & 0 & \star & 0 & 0 & 0 \cdots \\ \star & 0 & 0 & 0 & \star & 0 & 0 \cdots \\ 0 & \star & 0 & 0 & 0 & \star & 0 \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \ddots \end{pmatrix}, \tag{3.12}$$

where "$\star$" denotes a non-zero component. For given J one can easily diagonalize the matrix explictly and thereby find the eigenvalues and degeneracies. The general result is surprisingly complex, as the plots below show. The theory of irreducible representations in group theory allows to find the splitting but does not give any information on how the states are ordered in energy. For $J < 2$ there is no splitting. For $J = 2$ we obtain a triplet below a doublet. The first few spectra are sketched here, where lines close together represent a degenerate multiplet:

$$J = \quad 2 \qquad \tfrac{5}{2} \qquad 3 \qquad \tfrac{7}{2} \qquad 4$$
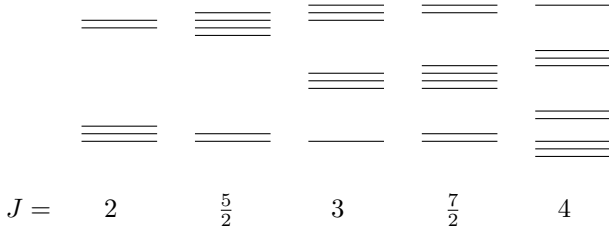
Over a larger range of J we find the following spectra, where each dot now represents a multiplet without indication of its multiplicity:
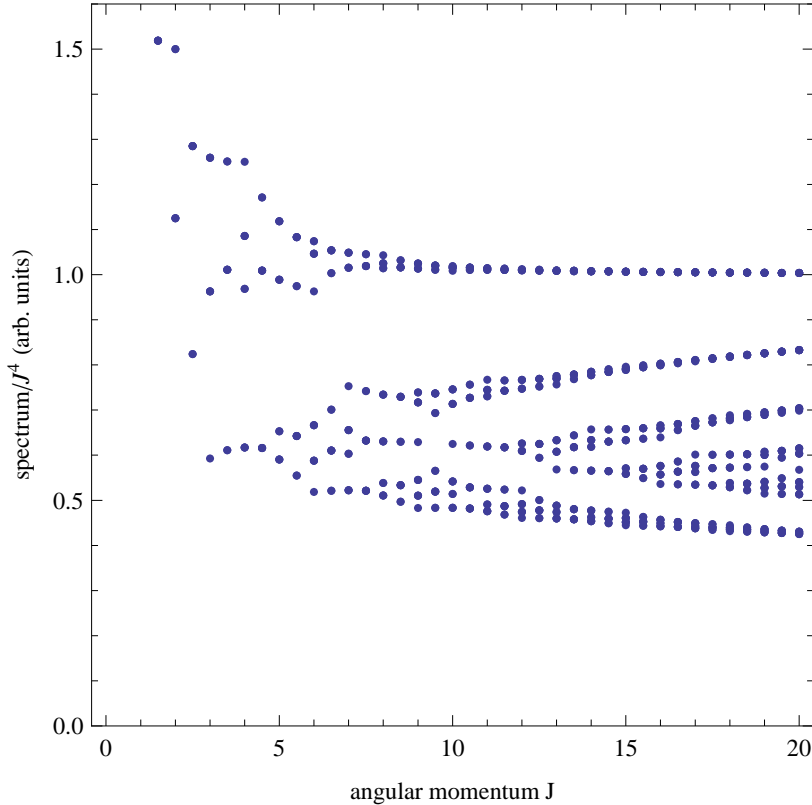
If the symmetry is not cubic, anisotropy terms tend to appear already at second order in $\mathbf{r}$ and, consequently, in $\mathbf{J}$. For an *orthorhombic* lattice the second-order terms are

$$3z^2 - r^2 \rightarrow \alpha \left\langle r^2 \right\rangle \left[3J_z^2 - J(J+1) \right], \tag{3.13}$$

$$x^2 - y^2 \rightarrow \alpha \left\langle r^2 \right\rangle \left[J_x^2 - J - y^2) \right]. \tag{3.14}$$

Only the first term exists for *tetragonal* symmetry. A term proportional to J_z^2 is the simplest and the most inportant anisotropy that can occur. For negative prefactor it favors large $|m_J|$ and is an *easy-axis anisotropy*, whereas for positive prefactor it favors small $|m_J|$ and is a *hard-axis* or *easy-plane anisotropy*. We note that $J_x^2 + J_y^2 + J_z^2 = J(J+1)$ is a constant and thus does not introduce any anisotropy. Therefore $-J_z^2$ is equivalent to $+J_x^2 + J_y^2$.

Where are the rules $x^p \rightarrow \langle r^p \rangle J_x^p$ etc. coming from? The formal proof requires group theory and can be given, for example, in terms of irreducible tensor operators, see Stevens, Proc. Phys. Soc. A **65**, 209 (1952) for a discussion. Also using group theory, it can be shown that terms of order higher than $2l$ ($l = 2$ for d-shells, $l = 3$ for f-shells) in V_{cryst} do not contribute.

3.3 Transition-metal ions

As noted, in transition-metal ions the hybridization between the d-orbitals and orbitals of neighboring ions dominates the crystal-field effects. However, although the mechanism is thus different from the electrostatic potential discussed previously, the splitting of multiplets is only a consequence of the reduced symmetry. We

can therefore discuss hybridization in terms of an *effective* potential having the correct symmetry. We must keep in mind, though, that for transition-metal ions crystal-field effects ars stronger than the LS coupling so that we should apply crystal-field theory to multiplets ignoring the LS coupling. LS coupling can later be treated as a weak perturbation.

For example, for an ion in a cubic environement we obtain the term

$$\beta \left\langle r^4 \right\rangle \left[L_x^4 + L_y^4 + L_z^4 - \frac{3}{5} L^2(L+1)^2 + \frac{1}{5} L(L+1) \right] \tag{3.15}$$

in the single-ion Hamiltonian. To calculate the coefficient β, we reiterate that the matrix elements of $(5/2) K_{40}(x^4 + y^4 + z^4 - 3/5 \, r^4)$ are proportional to those of the above operator (3.15). We can thus calculate *one* non-zero matrix element of each operator to obtain β. The result is a lengthy expression, see Yosida's book. In particular, we find alternations in the sign of β with the number $n_{n,2}$ of electrons in the d-shell.

If Hund's 1st and 2nd rules win over the crystal field, the only possible values for L are 0, 2, 3. Together with the schemes on p. 18 we obtain the following splittings:

configuration	d^1, d^6	d^2, d^7	d^3, d^8	d^4, d^9	d^5
L	2	3	3	2	0
β	> 0	< 0	> 0	< 0	

3.3.1 The Jahn-Teller effect

If the symmetry is lower than cubic, the levels will be split further. For sufficiently low symmetry, all degeneracies (except for spin degeneracy) are lifted. In cases where the ground state would still be degenerate for a putative crystal structure, the crystal will often distort to break the degeneracy and lower the symmetry. This is the *Jahn-Teller effect*. For example, in the $3d^9$ configuration of Cu^{2+} or in the $3d^4$ configuration of Mn^{3+}, the ground state would be twofold degenerate in a cubic environement. A distortion of the crystal causes two energy contributions:

  (a) an elastic energy, which is increasing quadratically with the distortion for small distortions if the cubic crystal was stable neglecting the Jahn-Teller effect,

  (b) a linear splitting of the angular momentum doublet as given by first-order perturbation theory.

Thus the total ground-state energy generically has a minimum for non-zero distortion:
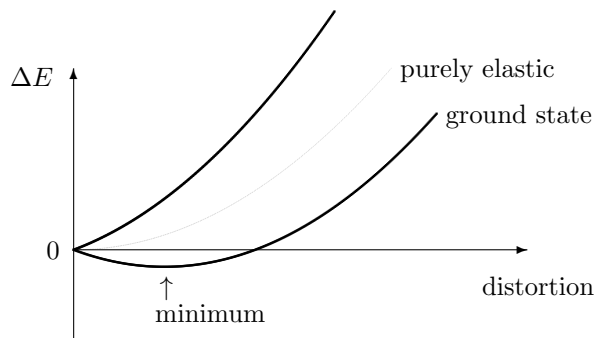
3.3.2 Quenching of the orbital angular momentum

If the many-particle ground state is not degenerate with respect to the orbital angular momentum $\mathbf{L}$ we find the phenomenon of *quenching*. (We neglect spin-orbit coupling for the moment and only talk about the

spatial state; it may be twofold spin degenerate.) To understand it, we calculate $\langle 0|\mathbf{L}|0\rangle$, where $|0\rangle$ is the non-degenerate many-electron ground state. $\langle 0|\mathbf{L}|0\rangle$ must be real since $\mathbf{L}$ is a hermitian operator. Since the Hamiltonian H including the crystal field V_{cryst} is real in the Schrödinger (real-space) representation, we can choose the many-electron wave function $\Psi_0(\mathbf{r}_1,\dots)$ to be real. However, $\mathbf{L}$ is purely imaginary in this representation:

$$\mathbf{L} = \sum_j \mathbf{l}_j = \sum_j \frac{\hbar}{i}\,\mathbf{r}_j \times \boldsymbol{\nabla}_j. \tag{3.16}$$

Thus

$$\langle 0|\mathbf{L}|0\rangle = \frac{\hbar}{i}\underbrace{\int d^3 r_1 \dots \Psi_0(\mathbf{r}_1,\dots)\sum_j \mathbf{r}_j \times \boldsymbol{\nabla}_j \Psi_0(\mathbf{r}_1,\dots)}_{\text{real}} \tag{3.17}$$

is imaginary. Since we have already seen that it must be real, we conclude

$$\langle 0|\mathbf{L}|0\rangle = 0. \tag{3.18}$$

(This argument fails if the ground state is degenerate since then the ion can be in a superposition of the degenerate states with complex coefficients.) We can draw the important conclusion that in transition-metal salts the magnetization is mainly carried by the electron spins, not the orbital angular momentum.

A small orbital contribution is restored if we take spin-orbit coupling into account as a weak perturbation: Let us treat

$$H_1 = \lambda \mathbf{L}\cdot\mathbf{S} + \mu_B \mathbf{B}\cdot(2\mathbf{S}+\mathbf{L}) \tag{3.19}$$

(where $g = 2$ has been assumed) as a small perturbation. At first order we get simply

$$\Delta E^{(1)} = 2\mu_B \mathbf{B}\cdot\mathbf{S} \tag{3.20}$$

but at second order we have

$$\Delta E^{(2)} = \sum_{\mu\nu}\left(\lambda^2 \Lambda_{\mu\nu} S_\mu S_\nu + 2\lambda\mu_B \Lambda B_\mu S_\nu + \mu_B^2 \Lambda_{\mu\nu} B_\mu B_\nu\right) \tag{3.21}$$

with

$$\Lambda_{\mu\nu} := \sum_{|\Psi\rangle \neq |0\rangle}\frac{\langle 0|L_\mu|\Psi\rangle\,\langle\Psi|L_\nu|0\rangle}{E_\Psi - E_0}. \tag{3.22}$$

Thus we obtain the effective spin Hamiltonian

$$H_s = \sum_{\mu\nu}\left[2\mu_B B_\mu(\delta_{\mu\nu} - \lambda\Lambda_{\mu\nu})S_\nu - \lambda^2 S_\mu \Lambda_{\mu\nu} S_\nu - \mu_B^2 B_\mu \Lambda_{\mu\nu} B_\nu\right]. \tag{3.23}$$

The magnetic moment is

$$M_\mu = -\frac{\partial H_s}{\partial B_\mu} = -\sum_\nu\left[2\mu_B(\delta_{\mu\nu} - \lambda\Lambda_{\mu\mu})S_\nu - 2\mu_B^2 \Lambda_{\mu\nu} B_\nu\right]$$
$$= -2\mu_B S_\mu + 2\mu_B \sum_\nu \Lambda_{\mu\nu}(\lambda S_\nu + \mu_B B_\nu), \tag{3.24}$$

where the first term is the usual spin magnetic moment, whereas the second is the *induced orbital magnetic moment*. Or, equivalently,

$$L_\mu^{\text{ind}} = -2\sum_{nu}\Lambda_{\mu\nu}(\lambda S_\nu + \mu_B B_\nu) \tag{3.25}$$

is the induced angular momentum. The term $2\mu_B^2 \sum_\nu \Lambda_{\mu\nu} B_\nu$ gives rise to the *Van Vleck orbital paramagnetism*. Note that it is independent of temperature.

The remaining term $-\lambda^2 \sum_{\mu\nu} S_\mu \Lambda_{\mu\nu} S_\nu$ in H_s describes the anisotropy energy of the *spin* as opposed to the orbital angular momentum discussed earlier. Higher-order anisotropy terms are obtained in higher orders of perturbation theory. Not surprisingly, the allowed anisotropy terms are dictated by crystal symmetry so that our earlier discussion carries over.

3.3.3 The Kramers theorem

It is natural to ask how far the splitting of multiplets by weak perturbations can go. Can all degeneracies be removed for sufficiently low crystal symmetry? For the answer, the Kramers theorem is important. It states that for a system invariant under time reversal and containing an odd number of fermions, the degeneracy cannot be completely lifted for any state. Thus there always remains at least a two-fold degeneracy ("Kramers doublet").

We sketch the proof: Note that in the absence of a magnetic field the Hamiltonian H commutes with the anti-unitary time-reveral operator K: $[H, K] = 0$. This is what being invariant under time reversal means. Let $|\Psi\rangle$ be an eigenstate to H with eigenenergy E_Ψ. For an odd number of fermions, the total spin expectation value $\langle\Psi|\mathbf{S}\cdot\mathbf{S}|\Psi\rangle$ is non-zero. But the spin is flipped by K, thus $K|\Psi\rangle$ is distinct from $|\Psi\rangle$. Since $HK|\Psi\rangle = KH|\Psi\rangle$, $K|\Psi\rangle$ has the same eigenenergy as $|\Psi\rangle$, i.e., the two states are degenerate.

3.3.4 Low-spin ions and spin crossover

We have so far assumed that the crystal field is weaker than the first (and second) Hund couplings. In this case we obtain a *high-spin state* with maximum S.

What happens if the crystal field is the strongest effect? This can happen for certain ions, Fe^{2+} ($3d^6$ configuration) in complex salts is an important example. Clearly, we now have to treat V_{cryst} first and the Hund rules as weaker perturbation. This means that we can consider the effect of V_{cryst} on *non-interacting* electrons since interactions first enter with the Hund rules. Thus we can start from a single-electron picture. We consider the example of $3d^6$ in a cubic crystal field. The splitting of the *single-electron orbitals* is identical to what we have obtained for the d^1 configuration on p. 19 :
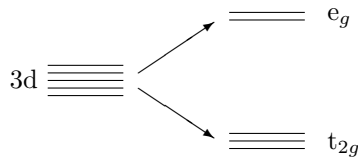
These orbitals are filled by 6 electrons. If the crystal-field splitting is strong, all 6 go into the lower-energy triplet (called "t_{2g}", the labels denote certain representations in group theory):
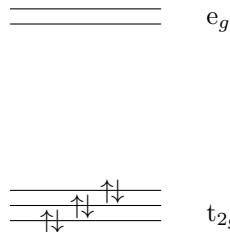
In principle, we should now consider the Hund rules, but they are all simple in this case, since

  (a) $S = 0$ is obvious.

  (b) $L = 0$: for this we have to consider the orbital content of the three t_{2g} orbitals. By diagonalizing the cubic single-ion anisotropy operator of Eq. (3.15) we find the eigenstates

$$\frac{i}{\sqrt{2}}\left(|2,-1\rangle + |2,1\rangle\right),$$

$$\frac{i}{\sqrt{2}}\left(|2,-1\rangle - |2,1\rangle\right),$$

$$\frac{i}{\sqrt{2}}\left(|2,-2\rangle - |2,2\rangle\right),$$

    using the notation $|l, m\rangle$. This implies $\langle L_z\rangle = 0$ and by symmetry $\langle\mathbf{L}\rangle = 0$.

  (c) $J = 0$ follows from (a) and (b).

The ion thus is in a *low-spin state*, which for the present example has $S = 0$. (Remember that $S = 2$ in the high-spin state.) If now the energy difference ΔE between the low-spin (LS) and the high-spin (HS) state is

not too large we find *spin crossover*: The average spin should be

$$\bar{S} = \frac{S_{\mathrm{LS}}G_{\mathrm{LS}} + S_{\mathrm{HS}}G_{\mathrm{HS}}\,e^{-\Delta E/k_B T}}{G_{\mathrm{LS}} + G_{\mathrm{HS}}\,e^{-\Delta E/k_B T}}, \tag{3.26}$$

where G_{LS} and G_{HS} are the degeneracies of the LS and HS state, respectively. If these are only due to the spin we have $G = 2S + 1$. In our example with $S_{\mathrm{LS}} = 0$ and $S_{\mathrm{HS}} = 2$ we obtain

$$\bar{S} = \frac{10\,e^{-\Delta E/k_B T}}{1 + 5\,e^{-\Delta E/k_B T}} \tag{3.27}$$
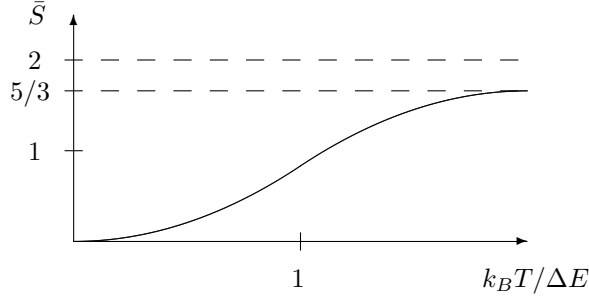
We thus find a crossover from LS behavior for $k_B T \ll \Delta E$ to predominantly HS behavior for $k_B T \gg \Delta E$, driven by the higher degeneracy of the HS state, i.e., by entropy.

Chapter 4

Exchange interactions between local spins

We have seen in Sec. 1.2 that the dipole-dipole interaction between the magnetic moments of electrons is much too weak to explain magnetic order at high temperatures. Therefore, we have to find a *strong* interaction between electrons to explain the observations. At first glance, it seems that we need an interaction that depends explicitly on the spins (or magnetic moments) of the electrons. However, no strong interaction of this type is known. Heisenberg realized in 1928 that the responsible interaction is the *Coulomb repulsion* between electrons, which is strong but does not explicitly depend on the spin. The spin selectivity is coming from quantum mechanics, specifically from the Pauli principle: Two electrons with parallel or antiparallel spins behave differently, even though the fundamental interaction is the same , because the (spatial) wave function $\psi(\mathbf{r}_1, \mathbf{r}_2)$ has to be antisymmetric and symmetric in these cases, respectively. This means for example that two electrons with parallel spins cannot be at the same place.

4.1 Direct ferromagnetic exchange interaction

The Coulomb interaction reads

$$H_{\text{Coulomb}} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \tag{4.1}$$

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

$$\rho(\mathbf{r}) = -e \sum_\sigma \Psi_\sigma^\dagger(\mathbf{r})\Psi_\sigma(\mathbf{r}), \tag{4.2}$$

where $\sigma = \uparrow, \downarrow$ is the spin orientation and $\Psi_\sigma(\mathbf{r})$ is the *field operator*. The field operator satisfies the *anti-commutation relations*

$$\left\{\Psi_{\sigma_1}(\mathbf{r}_1), \Psi_{\sigma_2}^\dagger(\mathbf{r}_2)\right\} \equiv \Psi_{\sigma_1}(\mathbf{r}_1)\Psi_{\sigma_2}^\dagger(\mathbf{r}_2) + \Psi_{\sigma_2}^\dagger(\mathbf{r}_2)\Psi_{\sigma_1}(\mathbf{r}_1) = \delta_{\sigma_1,\sigma_2}\delta\left(\mathbf{r}_1 - \mathbf{r}_2\right), \tag{4.3}$$

$$\left\{\Psi_{\sigma_1}(\mathbf{r}_1), \Psi_{\sigma_2}(\mathbf{r}_2)\right\} = 0, \tag{4.4}$$

$$\left\{\Psi_{\sigma_1}^\dagger(\mathbf{r}_1), \Psi_{\sigma_2}^\dagger(\mathbf{r}_2)\right\} = 0 \tag{4.5}$$

since it is a *fermionic* field. We obtain

$$
\begin{aligned}
H_{\text{Coulomb}} &= \frac{1}{2} \frac{1}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 \sum_{\sigma_1,\sigma_2} \Psi_{\sigma_1}^\dagger(\mathbf{r}_1)\Psi_{\sigma_1}(\mathbf{r}_1) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \Psi_{\sigma_2}^\dagger(\mathbf{r}_2)\Psi_{\sigma_2}(\mathbf{r}_2) \\
&= \frac{1}{2} \frac{1}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 \sum_{\sigma_1,\sigma_2} \Psi_{\sigma_1}^\dagger(\mathbf{r}_1)\Psi_{\sigma_2}^\dagger(\mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \Psi_{\sigma_2}(\mathbf{r}_2)\Psi_{\sigma_1}(\mathbf{r}_1) \\
&\quad + \frac{1}{2} \frac{1}{4\pi\epsilon_0} \int d^3r_1 \sum_{\sigma_1} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_1|} \Psi_{\sigma_1}^\dagger(\mathbf{r}_1)\Psi_{\sigma_1}(\mathbf{r}_1).
\end{aligned} \tag{4.6}
$$

The last, clearly singular, term is unphysical. Indeed, the derivation of the second-quantization formalism shows that the field operators should be written in normal order ($\Psi^\dagger\Psi^\dagger\Psi\Psi$) from the start.

We can expand Ψ into any orthonormal set of single-electron wave functions. For an ionic crystal, a set of orthonormal functions $\phi_{\mathbf{R}m}(\mathbf{r})$ localized at the ionic positions $\mathbf{R}$ (Wannier functions) is advantageous. Here, m includes all orbital quantum numbers but not the spin. We also introduce spinors

$$\chi_\uparrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_\downarrow = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.7}$$

which are eigenvectors of

$$s^z = \frac{1}{2}\sigma^z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.8}$$

to eigenvalues $\pm 1/2$. Then

$$\Psi_\sigma(\mathbf{r}) = \sum_{\mathbf{R}m} a_{\mathbf{R}m\sigma}\phi_{\mathbf{R}m}(\mathbf{r})\chi_\sigma, \tag{4.9}$$

where $a_{\mathbf{R}m\sigma}$ is a fermionic annihilation operator satisfying $\{a_{\mathbf{R}m\sigma}, a^\dagger_{\mathbf{R}'m'\sigma'}\} = \delta_{\mathbf{R}\mathbf{R}'}\delta_{mm'}\delta_{\sigma\sigma'}$ etc. The Coulomb interaction then reads

$$H_{\text{Coulomb}} = \frac{1}{2}\sum_{\mathbf{R}_1 m_1} \cdots \sum_{\mathbf{R}_4 m_4} \int d^3r_1 d^3r_2 \, \phi^*_{\mathbf{R}_1 m_1}(\mathbf{r}_1)\phi^*_{\mathbf{R}_2 m_2}(\mathbf{r}_2)\frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|}\phi_{\mathbf{R}_3 m_3}(\mathbf{r}_2)\phi_{\mathbf{R}_4 m_4}(\mathbf{r}_1)$$
$$\times \sum_{\sigma_1 \sigma_2}\chi^\dagger_{\sigma_1}\chi^\dagger_{\sigma_2}\chi_{\sigma_2}\chi_{\sigma_1}\, a^\dagger_{\mathbf{R}_1 m_1 \sigma_1}a^\dagger_{\mathbf{R}_2 m_2 \sigma_2}a_{\mathbf{R}_3 m_3 \sigma_2}a_{\mathbf{R}_4 m_4 \sigma_1}. \tag{4.10}$$

The scalar products of the spinors are simple: $\chi^\dagger_{\sigma_1}\chi_{\sigma_1} = \chi^\dagger_{\sigma_2}\chi_{\sigma_2} = 1$. We further define the integral

$$\left\langle \mathbf{R}_1 m_1, \mathbf{R}_2 m_2 \left| \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} \right| \mathbf{R}_3 m_3, \mathbf{R}_4 m_4 \right\rangle$$
$$:= \int d^3r_1 d^3r_2 \, \phi^*_{\mathbf{R}_1 m_1}(\mathbf{r}_1)\phi^*_{\mathbf{R}_2 m_2}(\mathbf{r}_2)\frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|}\phi_{\mathbf{R}_3 m_3}(\mathbf{r}_2)\phi_{\mathbf{R}_4 m_4}(\mathbf{r}_1) \tag{4.11}$$

and thus obtain

$$H_{\text{Coulomb}} = \frac{1}{2}\sum_{\mathbf{R}_1 m_1} \cdots \sum_{\mathbf{R}_4 m_4} \left\langle \mathbf{R}_1 m_1, \mathbf{R}_2 m_2 \left| \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} \right| \mathbf{R}_3 m_3, \mathbf{R}_4 m_4 \right\rangle$$
$$\times \sum_{\sigma_1 \sigma_2} a^\dagger_{\mathbf{R}_1 m_1 \sigma_1}a^\dagger_{\mathbf{R}_2 m_2 \sigma_2}a_{\mathbf{R}_3 m_3 \sigma_2}a_{\mathbf{R}_4 m_4 \sigma_1}. \tag{4.12}$$

4.1.1 On-site Coulomb interaction

We first consider the contribution of $\mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3 = \mathbf{R}_4$. We now drop the $\mathbf{R}_i$ where this is safe to do without causing confusion. In general, the quantum numbers $m_1, \ldots, m_4$ in $\langle m_1, m_2 | e^2/4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2| | m_3, m_4 \rangle$ can be all different and still lead to a non-zero integral. However, if we treat H_{Coulomb} as a perturbation, we only obtain a non-zero first-order contribution if the creation and annihilation operators $a^\dagger$, a are paired for each orbital. This requires $m_1 = m_4$ and $m_2 = m_3$ or $m_1 = m_3$ and $m_2 = m_4$. We thus have to consider the *direct Coulomb integrals*

$$K_{m_1 m_2} := \left\langle m_1, m_2 \left| \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} \right| m_2, m_1 \right\rangle$$
$$= \int d^3r_1 d^3r_2 \, |\phi_{m_1}(\mathbf{r}_1)|^2 \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|}|\phi_{m_2}(\mathbf{r}_2)|^2 \tag{4.13}$$

and the *exchange integrals*

$$J_{m_1 m_2} := \left\langle m_1, m_2 \left| \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} \right| m_1, m_2 \right\rangle$$
$$= \int d^3r_1 d^3r_2 \, \phi^*_{m_1}(\mathbf{r}_1)\phi^*_{m_2}(\mathbf{r}_2)\frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|}\phi_{m_1}(\mathbf{r}_2)\phi_{m_2}(\mathbf{r}_1), \tag{4.14}$$

so called because m_1 and m_2 are exchanged in the last factor compared to the direct integrals.

We obtain, to first order,

$$H_{\text{Coulomb}} \cong \frac{1}{2}\sum_{\mathbf{R}_1}\sum_{m_1 m_2}\sum_{\sigma_1 \sigma_2}\left\{ K_{m_1 m_2}a^\dagger_{\mathbf{R}m_1 \sigma_1}a^\dagger_{\mathbf{R}m_2 \sigma_2}a_{\mathbf{R}m_2 \sigma_2}a_{\mathbf{R}m_1 \sigma_1} \right.$$
$$\left. + J_{m_1 m_2}a^\dagger_{\mathbf{R}m_1 \sigma_1}a^\dagger_{\mathbf{R}m_2 \sigma_2}a_{\mathbf{R}m_1 \sigma_2}a_{\mathbf{R}m_2 \sigma_1} \right\} \tag{4.15}$$

24

(there is no double counting of the contribution from $m_1 = m_2$ and $\sigma_1 = \sigma_2$ since the corresponding terms contain $a_{\mathbf{R}m_1\sigma_1} a_{\mathbf{R}m_1\sigma_1} = 0$). Thus

$$
H_{\text{Coulomb}} \cong \frac{1}{2} \sum_{\mathbf{R}_1} \sum_{m_1 m_2} \sum_{\sigma_1 \sigma_2} \Big\{ K_{m_1 m_2} a^\dagger_{\mathbf{R}m_1\sigma_1} a_{\mathbf{R}m_1\sigma_1} a^\dagger_{\mathbf{R}m_2\sigma_2} a_{\mathbf{R}m_2\sigma_2} - J_{m_1 m_2} a^\dagger_{\mathbf{R}m_1\sigma_1} a_{\mathbf{R}m_1\sigma_2}
$$
$$
\times\, a^\dagger_{\mathbf{R}m_2\sigma_2} a_{\mathbf{R}m_2\sigma_1} \Big\} + (\text{irrelevant potential term}). \tag{4.16}
$$

We define the number operators $n_{\mathbf{R}m} := \sum_\sigma a^\dagger_{\mathbf{R}m\sigma} a_{\mathbf{R}m\sigma}$ and the spin operators $s^\alpha_{\mathbf{R}m} := \sum_{\sigma\sigma'} a^\dagger_{\mathbf{R}m\sigma} (\sigma^\alpha_{\sigma\sigma'}/2) a_{\mathbf{R}m\sigma'}$, where σ^α, $\alpha = x,\, y\, z$, are the Pauli matrices. With some algebra, one shows that

$$
\sum_{\sigma_1 \sigma_2} a^\dagger_{\mathbf{R}m_1\sigma_1} a_{\mathbf{R}m_1\sigma_2} a^\dagger_{\mathbf{R}m_2\sigma_2} a_{\mathbf{R}m_2\sigma_1} = \frac{1}{2}\, n_{\mathbf{R}m_1} n_{\mathbf{R}m_2} + 2 s^z_{\mathbf{R}m_1} s^z_{\mathbf{R}m_2} + s^+_{\mathbf{R}m_1} s^-_{\mathbf{R}m_2} + s^-_{\mathbf{R}m_1} s^+_{\mathbf{R}m_2}
$$
$$
= \frac{1}{2}\, n_{\mathbf{R}m_1} n_{\mathbf{R}m_2} + 2\, \mathbf{s}_{\mathbf{R}m_1} \cdot \mathbf{s}_{\mathbf{R}m_2}. \tag{4.17}
$$

Thus we obtain

$$
H_{\text{Coulomb}} \cong \sum_{\mathbf{R}} \frac{1}{2} \sum_{m_1 m_2} \left\{ \left(K_{m_1 m_2} - \frac{1}{2} J_{m_1 m_2} \right) n_{\mathbf{R}m_1} n_{\mathbf{R}m_2} - 2\, J_{m_1 m_2}\, \mathbf{s}_{\mathbf{R}m_1} \cdot \mathbf{s}_{\mathbf{R}m_2} \right\}. \tag{4.18}
$$

The first term is the on-site Coulomb interaction. The definition of $K_{m_1 m_2}$ shows immediately that $K_{m_1 m_2} > 0$. This contribution would also be there in a classical theory.

We now show that $J_{m_1 m_2} \geq 0$: Since

$$
\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \int \frac{d^3 k}{(2\pi)^3}\, e^{i\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2)} \frac{4\pi}{k^2} \tag{4.19}
$$

we have

$$
J_{m_1 m_2} = \int \frac{d^3 k}{(2\pi)^3} \frac{e^2}{\epsilon_0 k^2} \underbrace{\int d^3 r_1 \phi^*_{m_1}(\mathbf{r}_1) \phi_{m_2}(\mathbf{r}_1) e^{i\mathbf{k}\cdot\mathbf{r}_1}}_{=:\, I(\mathbf{k})} \underbrace{\int d^3 r_2 \phi_{m_1}(\mathbf{r}_2) \phi^*_{m_2}(\mathbf{r}_2) e^{-i\mathbf{k}\cdot\mathbf{r}_2}}_{=:\, I^*(\mathbf{k})}
$$
$$
= \int \frac{d^3 k}{(2\pi)^3} \frac{e^2}{\epsilon_0 k^2}\, |I(\mathbf{k})|^2 \geq 0. \tag{4.20}
$$

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

$$
K_{m_1 m_2} - J_{m_1 m_2} = \frac{1}{2} (K_{m_1 m_2} + K_{m_2 m_1} - J_{m_1 m_2} - J_{m_2 m_1})
$$
$$
= \frac{1}{2} \int d^3 r_1 d^3 r_2 \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} \underbrace{[\phi^*_{m_1}(\mathbf{r}_1) \phi^*_{m_2}(\mathbf{r}_2) - \phi^*_{m_2}(\mathbf{r}_1) \phi^*_{m_1}(\mathbf{r}_2)]}_{=:\, f(\mathbf{r}_1, \mathbf{r}_2)}
$$
$$
\times \underbrace{[\phi_{m_2}(\mathbf{r}_2) \phi_{m_1}(\mathbf{r}_1) - \phi_{m_1}(\mathbf{r}_2) \phi_{m_2}(\mathbf{r}_1)]}_{=:\, f^*(\mathbf{r}_1, \mathbf{r}_2)}
$$
$$
= \frac{1}{2} \int d^3 r_1 d^3 r_2 \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|}\, |f(\mathbf{r}_1, \mathbf{r}_2)|^2 \geq 0. \tag{4.21}
$$

This shows that $K_{m_1 m_2} - 1/2\, J_{m_1 m_2} > 0$ so that the corrected Coulomb term is reduced but still repulsive. Even more importantly, we obtain a spin-spin interaction of the form

$$
-J_{m_1 m_2} \mathbf{s}_{\mathbf{R}m_1} \cdot \mathbf{s}_{\mathbf{R}m_2} \tag{4.22}
$$

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

Note that all terms containig $J_{m_1 m_2}$ 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 orbital indices m even possible. There is no analogy in classical physics.

For a single relevant orbital $\phi(\mathbf{r})$, we of course just get

$$
\begin{aligned}
H_{\text{Coulomb}} &\cong \frac{1}{2} \sum_{\mathbf{R}} \int d^3r_1 d^3r_2 \, \phi^*(\mathbf{r}_1)\phi^*(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1-\mathbf{r}_2|} \phi(\mathbf{r}_2)\phi(\mathbf{r}_1) \sum_{\sigma_1\sigma_2} a^\dagger_{\mathbf{R}\sigma_1} a^\dagger_{\mathbf{R}\sigma_2} a_{\mathbf{R}\sigma_2} a_{\mathbf{R}\sigma_1} \\
&= \frac{1}{2} \sum_{\mathbf{R}} \int d^3r_1 d^3r_2 \, \phi^*(\mathbf{r}_1)\phi^*(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1-\mathbf{r}_2|} \phi(\mathbf{r}_2)\phi(\mathbf{r}_1) \left(a^\dagger_{\mathbf{R}\uparrow} a^\dagger_{\mathbf{R}\downarrow} a_{\mathbf{R}\downarrow} a_{\mathbf{R}\uparrow} + a^\dagger_{\mathbf{R}\downarrow} a^\dagger_{\mathbf{R}\uparrow} a_{\mathbf{R}\uparrow} a_{\mathbf{R}\downarrow} \right) \\
&= \sum_{\mathbf{R}} \int d^3r_1 d^3r_2 \, \phi^*(\mathbf{r}_1)\phi^*(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1-\mathbf{r}_2|} \phi(\mathbf{r}_2)\phi(\mathbf{r}_1) \, a^\dagger_{\mathbf{R}\uparrow} a^\dagger_{\mathbf{R}\downarrow} a_{\mathbf{R}\downarrow} a_{\mathbf{R}\uparrow} \\
&=: \sum_{\mathbf{R}} U \, a^\dagger_{\mathbf{R}\uparrow} a^\dagger_{\mathbf{R}\downarrow} a_{\mathbf{R}\downarrow} a_{\mathbf{R}\uparrow}, \tag{4.23}
\end{aligned}
$$

where $U > 0$ This is the famous "Hubbard U-term", which will become important later.

4.1.2   Inter-ion exchange interaction

Essentially everything from the previous discussion goes through if we allow the ionic sites $\mathbf{R}_1, \ldots, \mathbf{R}_4$ to be different. We just have to treat $\mathbf{R}_i$ as another quantum number besides m_i. We here restict ourselves to a model with a single, non-degenerate (except for spin) orbital per site. We can then drop the orbital quantum numbers m_i. As noted above, we assume the orbitals at different sites to have negligible overlap, i.e., they are orthogonal. We will have a first-order perturbation if $\mathbf{R}_1 = \mathbf{R}_4$ and $\mathbf{R}_2 = \mathbf{R}_3$ or if $\mathbf{R}_1 = \mathbf{R}_3$ and $\mathbf{R}_2 = \mathbf{R}_4$. In complete analogy to the previous subsection we obtain

$$
H_{\text{Coulomb}} \cong \frac{1}{2} \sum_{\mathbf{R}_1 \mathbf{R}_2} \left\{ \left(K_{12} - \frac{1}{2} J_{12} \right) n_1 n_2 - 2 J_{12} \, \mathbf{s}_1 \cdot \mathbf{s}_2 \right\}, \tag{4.24}
$$

where

$$
K_{12} \equiv K_{\mathbf{R}_1 \mathbf{R}_2} := \int d^3r_1 d^3r_2 \, |\phi_{\mathbf{R}_1}(\mathbf{r}_1)|^2 \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1-\mathbf{r}_2|} |\phi_{\mathbf{R}_2}(\mathbf{r}_2)|^2, \tag{4.25}
$$

$$
J_{12} \equiv J_{\mathbf{R}_1 \mathbf{R}_2} := \int d^3r_1 d^3r_2 \, \phi^*_{\mathbf{R}_1}(\mathbf{r}_1)\phi^*_{\mathbf{R}_2}(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1-\mathbf{r}_2|} \phi_{\mathbf{R}_1}(\mathbf{r}_2)\phi_{\mathbf{R}_2}(\mathbf{r}_1). \tag{4.26}
$$

In an ionic crystal, the charge $-en_i$ should not fluctuate much. In our simple model, an electron number of $n_i = 1$ is the only interesting case since otherwise there is no spin. The interaction then becomes

$$
H_{\text{exc}} = - \sum_{\mathbf{R}_1 \mathbf{R}_2} J_{12} \, \mathbf{s}_1 \cdot \mathbf{s}_2, \tag{4.27}
$$

disregarding a constant. By the same argument as above, $J_{12} \geq 0$. Thus Coulomb repulsion between electrons in *orthogonal* orbitals always leads to a *ferromagnetic* exchange interaction.

The physical interpretation is the following: electrons with parallel spins cannot occupy the same orbital and therefore avoid the strong intra-orbital Coulomb repulsion. Their energy is therefore lower than for antiparallel spins.

4.2   Kinetic antiferromagnetic exchange interaction

Above, we have neglected charge fluctuations. This is not usually a good approximation even for ionic crystals and we drop it here. In an independent-electron picture, charge fluctuations result from the hybridization between orbitals of different ions, which allows electrons to tunnel or hop from one ion to another. To study the effect of hybridization, we now neglect the non-local (inter-ionic) part of the Coulomb repulsion, which according to Sec. 4.1 leads to direct ferromagnetic exchange. In the case of a single relevant orbital per site, we thereby obtain the *Hubbard model*,

$$
\begin{aligned}
H &= - \sum_{\mathbf{R}\mathbf{R}'\sigma} t(\mathbf{R} - \mathbf{R}') \, a^\dagger_{\mathbf{R}'\sigma} a_{\mathbf{R}\sigma} + U \sum_{\mathbf{R}} a^\dagger_{\mathbf{R}\uparrow} a^\dagger_{\mathbf{R}\downarrow} a_{\mathbf{R}\downarrow} a_{\mathbf{R}\uparrow} \\
&\equiv - \sum_{\mathbf{R}\mathbf{R}'\sigma} t(\mathbf{R} - \mathbf{R}') \, a^\dagger_{\mathbf{R}'\sigma} a_{\mathbf{R}\sigma} + U \sum_{\mathbf{R}} a^\dagger_{\mathbf{R}\uparrow} a_{\mathbf{R}\uparrow} a^\dagger_{\mathbf{R}\downarrow} a_{\mathbf{R}\downarrow}. \tag{4.28}
\end{aligned}
$$

The first term describes the kinetic energy due to tunneling or hybridization (and a local chemical potential) and the second is the local Coulomb repulsion (note $U > 0$) known from Sec. 4.1.

We first consider the case of a *dimer* as a toy model:

$$H = -t \sum_\sigma \left(a_{1\sigma}^\dagger a_{2\sigma} + a_{2\sigma}^\dagger a_{1\sigma} \right) - \mu \sum_\sigma \left(a_{1\sigma}^\dagger a_{1\sigma} + a_{2\sigma}^\dagger a_{2\sigma} \right) + U \sum_{i=1,2} a_{i\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow}^\dagger a_{i\downarrow}. \tag{4.29}$$

Since each site can be in one of four states (empty, spin-up, spin-down, doubly occupied) the dimension of the Fock space is $4^2 = 16$. The Hamiltonian H conserves the total electron number. We consider the sector of two electrons, which corresponds to a Hilbert space of dimension six. Suitable basis vectors are, in an obvious notation, $\{| \uparrow\downarrow, 0\rangle, |0, \uparrow\downarrow\rangle, | \uparrow, \downarrow\rangle, | \downarrow, \uparrow\rangle, | \uparrow, \uparrow\rangle, | \downarrow, \downarrow\rangle\}$. In this space, the μ-term is of course an irrelevant constant. The remaining Hamiltonian is a 6×6 matrix in the above basis,

$$H' = \left(\begin{array}{cccc|cc} U & 0 & t & -t & 0 & 0 \\ 0 & U & t & -t & 0 & 0 \\ t & t & 0 & 0 & 0 & 0 \\ -t & -t & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right). \tag{4.30}$$

We can simplify H' further by transforming from $| \uparrow, \downarrow\rangle, | \downarrow, \uparrow\rangle$ onto $(| \uparrow, \downarrow\rangle - | \downarrow, \uparrow\rangle)/\sqrt{2}$, $(| \uparrow, \downarrow\rangle + | \downarrow, \uparrow\rangle)/\sqrt{2}$, which gives

$$H'' = \left(\begin{array}{ccc|ccc} U & 0 & \sqrt{2}t & 0 & 0 & 0 \\ 0 & U & \sqrt{2}t & 0 & 0 & 0 \\ \sqrt{2}t & \sqrt{2}t & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \tag{4.31}$$

with the eigenenergies

- U, $\dfrac{U \pm \sqrt{U^2 + 16t^2}}{2}$ in the first sector and

- 0, 0, 0 in the second sector, which corresponds to the states $(| \uparrow, \downarrow\rangle + | \downarrow, \uparrow\rangle)/\sqrt{2}, | \uparrow, \uparrow\rangle, | \downarrow, \downarrow\rangle$, i.e., to the spin *triplet*.
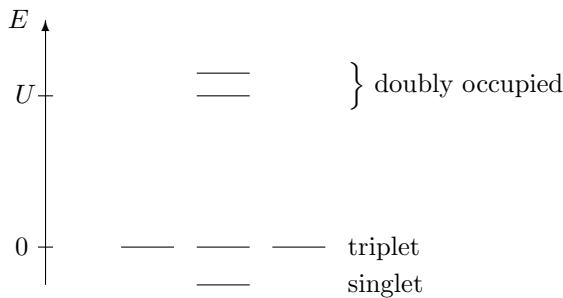
We are interested in ionic systems, for which t should be small, $t \ll U$. Then the first sector contains two very large energies

$$U \quad \text{and} \quad \frac{U + \sqrt{U^2 + 16t^2}}{2} \cong U + \frac{4t^2}{U} \tag{4.32}$$

and one small energy

$$\frac{U - \sqrt{U^2 + 16t^2}}{2} \cong -\frac{4t^2}{U} < 0. \tag{4.33}$$

For $U/|t| \to \infty$, 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. Thus 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 $s_1 = s_2 = 1/2$, we write

$$H_{\text{eff}} = -J\,\mathbf{s}_1 \cdot \mathbf{s}_2 = -\frac{J}{2}\left[\mathbf{S} \cdot \mathbf{S} - \mathbf{s}_1 \cdot \mathbf{s}_1 - \mathbf{s}_2 \cdot \mathbf{s}_2\right] \quad \text{with } \mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$$

$$= -\frac{J}{2}\Big[S(S+1) - \underbrace{\frac{3}{4} - \frac{3}{4}}_{\text{const}}\Big] = \text{const} - \frac{J}{2}S(S+1) = \text{const}\begin{cases} +0 & \text{for } S = 0, \\ -J & \text{for } S = 1. \end{cases} \tag{4.34}$$

Comparing this to Eq. (4.33), we read off

$$J = -\frac{4t^2}{U} \quad \text{for } U \gg t. \tag{4.35}$$

Here, we only state that an analogous result also holds for the Hubbard model on a lattice, not only for a dimer. The corresponding derivation will be given in Sec. 10.4 on the t-J model.

The result for a lattice is, at half filling and in the limit $U \gg t$,

$$H_{\text{eff}} = -J\sum_{\langle ij \rangle} \mathbf{s}_i \cdot \mathbf{s}_j \tag{4.36}$$

with $J = -4t^2/U$ and $\sum_{\langle ij \rangle}$ runs over all nearest-neighbor bonds, counting each bond once ($\langle ji \rangle$ and $\langle ij \rangle$ are the same and occur only once in the sum).

For the Hubbard model on a lattice, H_{eff} as written above is only the lowest-order term in an expansion in t/U. At order t^4/U^3 we obtain a *biquadratic exchange* term $\propto (\mathbf{s}_i \cdot \mathbf{s}_j)^2$ and for a square or cubic lattice a *ring-exchange* term $\propto (\mathbf{s}_i \cdot \mathbf{s}_j)(\mathbf{s}_k \cdot \mathbf{s}_l) + (\mathbf{s}_i \cdot \mathbf{s}_l)(\mathbf{s}_k \cdot \mathbf{s}_j) - (\mathbf{s}_i \cdot \mathbf{s}_k)(\mathbf{s}_j \cdot \mathbf{s}_l)$, where i, j, k, l form the corners of a square plaquette.
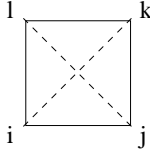

These are generally the most important higher-order terms but there are others.

4.3 Superexchange interaction

In ionic crystals, the magnetic ions are always separated by non-magnetic anions. Thus both the direct exchange interaction (note that the exchange integral contains $\phi^*_{\mathbf{R}_1}(\mathbf{r}_1)\phi_{\mathbf{R}_2}(\mathbf{r}_1)$) and the kinetic exchange interaction (note that J contains t^2) become very small. A larger exchange interaction results from electrons hopping from a magnetic cation to an anion and then to the next cation.

An important example is provided by the cuprates: Cu^{2+}-ions form a square lattice (maybe slightly deformed) with O^{2-}-ions centered on the bonds.
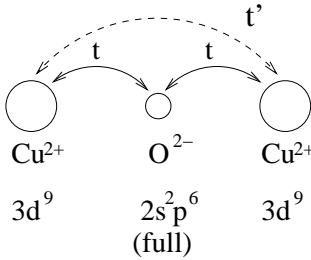
Hopping of a hole from Cu^{2+} via O^{2-} to the next Cu^{2+} (and back) leads to a strong antiferromagnetic exchange interaction proportional to t^4. The kinetic exchange from direct Cu-Cu hopping is proportional to $(t')^2$ but is nevertheless tiny in comparison because of $|t'| \ll |t|$.

This type of magnetic interaction is called *superexchange*. Note that this nomenclature is not consistently used. Goodenough calls this "semicovalent exchange". Since it is the most important exchange mechanism for ionic crystals, it has been studied extensively. Detailed calculations have led to a number of rules of thumb to estimate the strength and sign of the superexchange, which need not be antiferromagnetic in all cases. These are the *Goodenough-Kanamori rules*. We adapt them from the formulation given by Anderson (1963):

1. There is a *strong antiferromagnetic* exchange interaction if the half-filled orbitals of two cations overlap with the same empty or filled orbital of the intervening anion.

Examples:

(a)

$$d_{x^2-y^2} \qquad p_x \qquad d_{x^2-y^2}$$
$$(\text{of Cu}^{2+}) \qquad (\text{of O}^{2-}) \qquad (\text{of Cu}^{2+})$$

(b)

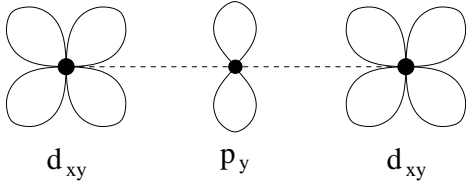$$d_{xy} \qquad p_y \qquad d_{xy}$$
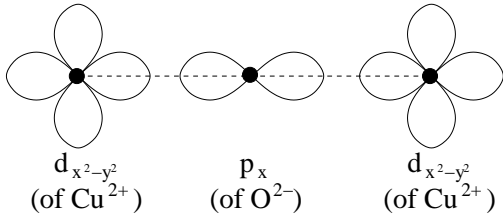
(c)
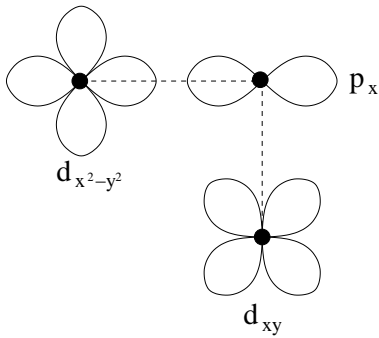
2. There is a *weaker ferromagnetic* exchange interaction if the half-filled orbitals of two cations overlap with orthogonal orbitals of the same intervening anion.

   Example:
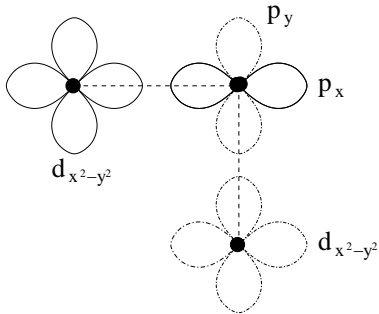
4.4 Dzyaloshinsky-Moriya interaction

There can be many other terms in the spin-spin interaction, which are of higher order or of more complicated form. They can be calculated similarly to the ones already discussed but the presence or absence of a certain term can be determined based solely on the crystal symmetry.

We only discuss the important example of the Dzyaloshinsky-Moriya interaction, which is of the form

$$H_{\mathrm{DM}} = \mathbf{D}_{12} \cdot (\mathbf{S}_1 \times \mathbf{S}_2) \tag{4.37}$$

with the time-independent but possibly spatially inhomogeneous vector $\mathbf{D}_{12}$ associated with the bond between $\mathbf{S}_1$ and $\mathbf{S}_2$. When is this term allowed by symmetry? We have to consider all symmetry operations of the crystal that leave the center point C on the bond between the two spins fixed.
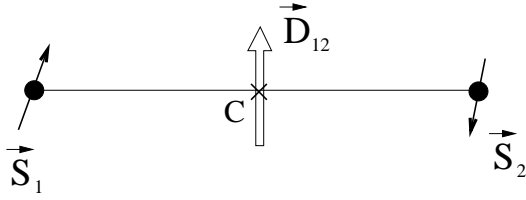
The whole Hamiltonian and in particular H_{DM} must remain the same if we apply any of these symmetry operations.

If C is an inversion center i of the crystal, i interchages the two spins but does not otherwise change them (spins are *pseudo*vectors). Thus

$$\mathbf{S}_1 \times \mathbf{S}_2 \overset{i}{\to} \mathbf{S}_2 \times \mathbf{S}_1 = -\mathbf{S}_1 \times \mathbf{S}_2. \tag{4.38}$$

Thus H is only invariant if $\mathbf{D}_{12} = 0$. In this case there cannot be a Dzyaloshinsky-Moriya term.

If the symmetry is lower, the term is possible. Moriya has given rules for the allowed directions of $\mathbf{D}_{12}$. We consider one example, a two-fold rotation axis C_2 through C perpendicular to the bond.
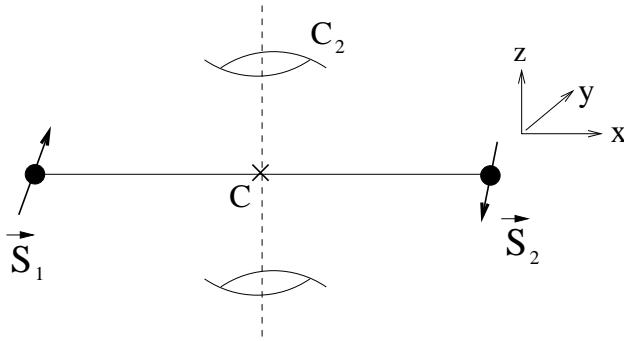
The mapping is

$$S_{1x} \overset{C_2}{\longleftrightarrow} -S_{2x},$$
$$S_{1y} \overset{C_2}{\longleftrightarrow} -S_{2y},$$
$$S_{1z} \overset{C_2}{\longleftrightarrow} +S_{2z}. \tag{4.39}$$

H_{DM} reads

$$H_{DM} = D_{12x}(S_{1y}S_{2z} - S_{1z}S_{2y}) + D_{12y}(S_{1z}S_{2x} - S_{1x}S_{2z}) + D_{12z}(\underline{S_{1x}S_{2y} - S_{1y}S_{2x}}). \tag{4.40}$$

Herein, the underlined term changes sign under C_2, the others do not. Thus H is invariant only if $D_{12z} = 0$. Therefore, $\mathbf{D}_{12}$ must be perpendicular to the C_2 symmetry axis, i.e., the z-axis. To find its orientation in the xy-plane we have to actually calculate it.

Chapter 5

The Heisenberg model

We have seen that the leading exchange interaction between spins in ionic crystals is usually of the form

$$H = -\frac{1}{2} \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j, \tag{5.1}$$

where in the sum i and j run over all sites, $J_{ij} = J_{ji}$ is symmetric and the factor $1/2$ corrects for double counting. Although we write $\mathbf{S}_i$, this model might also apply to the total angular momenta $\mathbf{J}_i$ in the case of rare-earth ions. This Hamiltonian represents the *Heisenberg model*. Of course, it omits a number of contributions:

- single-ion anisotropies such as $-\sum_i K_2 (S_i^z)^2$,

- anisotropic exchange interactions of the form $-\frac{1}{2} \sum_{ij} \mathbf{S}_i \mathsf{J}_{ij} \mathbf{S}_j$, where the J_{ij} are tensors, and the Dzyaloshinski-Moriya interaction, and

- higher-order terms such as ring exchange.

If some of these contributions are large, the Heisenberg model is clearly not a good starting point. One important limiting case is obtained when (a) the exchange interaction is anisotropic with the interaction between the z-components (or, equivalently, x or y) much stronger than all others and (b) the spins are $S = 1/2$:

$$H = -\frac{1}{2} \sum_{ij} J_{ij} S_i^z S_j^z \tag{5.2}$$

with $\mathbf{S} \cdot \mathbf{S} = (1/2)(1/2 + 1) = 3/4$. This constitutes the *Ising model*.

5.1   Ground state in the ferromagnetic case

A technical problem encountert for the Heisenberg model is that in

$$
\begin{aligned}
H &= -\frac{1}{2} \sum_{ij} J_{ij} \left(S_i^x S_j^x + S_i^y S_j^y + S_i^z S_j^z \right) \\
&= -\frac{1}{2} \sum_{ij} J_{ij} \left(\frac{S_i^+ + S_i^-}{2} \frac{S_j^+ + S_j^-}{2i} + \frac{S_i^+ - S_i^-}{2i} \frac{S_j^+ - S_j^-}{2} + S_i^z S_j^z \right) \\
&= -\frac{1}{2} \sum_{ij} J_{ij} \left(\frac{S_i^+ S_j^- + S_i^- S_j^+}{2} + S_i^z S_j^z \right)
\end{aligned} \tag{5.3}
$$

the terms in parentheses do not commute. The last term is obviously diagonal in the usual basis spanned by the product states $|S_1, m_1\rangle, |S_2, m_2\rangle \ldots |S_N, m_N\rangle$ but the other two terms change the quantum numbers m_i, m_j.

Defining the total spin $\mathbf{S} = \sum_i \mathbf{S}_i$, we find

$$[S^z, H] = -\frac{1}{2} \sum_{ijk} J_{ij} \left[S_k^z, S_i^x S_j^x + S_i^y S_j^y + S_i^z S_j^z \right]$$

$$= -\frac{1}{2} \sum_{ijk} J_{ij} \left(S_i^x [S_k^z, S_j^x] + [S_k^z, S_i^x] S_j^x + S_i^y [S_k^z, S_j^y] + [S_k^z, S_i^y] S_j^y + 0 \right)$$

$$= -\frac{1}{2} \sum_{ijk} J_{ij} \left(S_i^x \delta_{kj} i S_j^y + \delta_{ki} i S_i^y S_j^x + S_i^y \delta_{kj} (-i) S_j^x + \delta_{ki} (-i) S_i^x S_j^y \right) = 0 \qquad (5.4)$$

and analogously $[S^x, H] = [S^y, H] = 0$. Thus the total spin $\mathbf{S}$ is compatible with the Hamiltonian and we can find simultanious eigenstates of H, $\mathbf{S} \cdot \mathbf{S}$, and, for example, S^z.

Let us consider states with maximum total spin S_{tot}. Specifically, all spins are aligned in the z-direction, i.e., we consider the state

$$|\psi\rangle = \underbrace{|S, S\rangle_1}_{\text{spin 1}} \underbrace{|S, S\rangle_2}_{\text{spin 2}} \dots \underbrace{|S, S\rangle_N}_{\text{spin N}}. \qquad (5.5)$$

We have

$$H|\psi\rangle = -\frac{1}{2} \sum_{ij} J_{ij} \left(\frac{1}{2} \underbrace{S_i^+ S_j^- |\psi\rangle}_{=0} + \frac{1}{2} \underbrace{S_i^- S_j^+ |\psi\rangle}_{=0} + S_i^z S_j^z |\psi\rangle \right). \qquad (5.6)$$

The first two terms vanish since they contain $S_i^+ |S, S\rangle_i = 0$; the quantum number m_i cannot be further increased. Thus

$$H|\psi\rangle = -\frac{1}{2} \sum_{ij} J_{ij} S S |\psi\rangle = -\frac{1}{2} S^2 \sum_{ij} J_{ij} |\psi\rangle. \qquad (5.7)$$

We see that $|\psi\rangle$ is an eigenstate of H. Since H is invariant under rotations in spin space, this holds for *any* state with maximum S_{tot}. There are $2S_{\text{tot}} + 1$ such states. Note that we have not made any assumptions on the J_{ij}.

Now consider the expectation value $\langle \varphi | H | \varphi \rangle$ in an arbitrary product state $|\varphi\rangle = |S, m_1\rangle \dots |S, m_n\rangle$. $|\varphi\rangle$ is generally not an eigenstate of H. The expectation value is

$$\langle \varphi | H | \varphi \rangle = -\frac{1}{2} \sum_{ij} J_{ij} \left(\frac{1}{2} \underbrace{\langle \varphi | S_i^+ S_j^- | \varphi \rangle}_{=0} + \frac{1}{2} \underbrace{\langle \varphi | S_i^- S_j^+ | \varphi \rangle}_{=0} + \langle \varphi | S_i^z S_j^z | \varphi \rangle \right) = -\frac{1}{2} \sum_{ij} J_{ij} m_i m_j. \qquad (5.8)$$

If $J_{ij} \geq 0$ for all i, j then $J_{ij} m_i m_j \leq J_{ij} S^2$ and thus

$$\langle \varphi | H | \varphi \rangle \geq -\frac{1}{2} S^2 \sum_{ij} J_{ij}. \qquad (5.9)$$

Since the $|\varphi\rangle$ form a basis we conclude that all eigenenergies, in particular the ground-state energy, are larger than or equal to $-(1/2) S^2 \sum_{ij} J_{ij}$. Since $|\psi\rangle$ is an eigenstate to the eigenenergy $-(1/2) S^2 \sum_{ij} J_{ij}$, $|\psi\rangle$ must be a ground state if $J_{ij} \geq 0$ for all i, j.

In summary, if all exchange interactions are non-negative, the fully polarized states (i.e., with maximum S_{tot}) are ground states. One can further show that if every pair of sites i, j is connected by some string of bonds kl with $J_{kl} > 0$, these are the *only* ground states (the condition excludes cases where the system can by divided into non-interacting components). We have thus found the exact ground states of the fully *ferromagnetic* Heisenberg model.

5.1.1 Spontaneous symmetry breaking

We have found an example of spontaneous symmetry breaking: the ground-state manifold of the completely ferromagnetic Heisenberg model consists of fully polarized states (maximum S_{tot}). It is $(2S_{\text{tot}} + 1)$-fold degenerate, where $S_{\text{tot}} = NS$. Typical ground states thus have $\langle \mathbf{S} \rangle \neq 0$, i.e., they distinguish a certain direction in spin space. On the other hand, the Hamiltonian H does not. Thus the symmetry of a typical ground state is lower than that of the Hamiltonian. We say, the symmetry is "spontaneously broken". We conclude with two remarks:

- We have not assumed that our system is in the thermodynamic limit ($N \to \infty$). Spontaneous symmetry breaking in the ground state is not a statistical or thermodynamic concept and can also happen in finite systems.

- The complete ground-state manifold is invariant under spin rotations, only most individual elements are not.

5.2 Ground state in the antiferromagnetic case: Marshall's theorem

If some J_{ij} are negative, the fully polarized states are still eigenstates of H but generally no longer ground states. Practically nothing is known rigorously about the ground state of the general Heisenberg model. An important subclass for which rigorous results exist are antiferromagnets on *bipartite* systems.

In the context of magnetism, a system is bipartite if all sites can be divided into two disjoint subsystems A and B in such a way that $J_{ij} = 0$ if $i, j \in A$ or if $i, j \in B$. This means that the only interactions are between sites from different subsystems. The most important case of course involves spins on a crystal lattice, in which case a bipartite *lattice* has the obvious definition. Example: the square lattice with nearest-neighbor interactions is bipartite:

$\bullet \in A$

$\circ \in B$
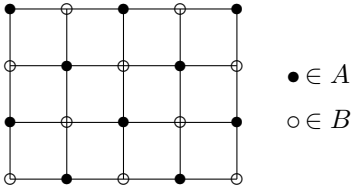
An obvious guess would be that, for $J_{ij} \leq 0$ on a bipartite lattice, all spins are fully polarized but in opposite directions for the two sublattices. This is called a *Néel state*. We consider the state

$$|\psi\rangle = \prod_{i \in A} |S, S\rangle_i \prod_{j \in B} |S, -S\rangle_j. \tag{5.10}$$

Then

$$H|\psi\rangle = -\frac{1}{2} \sum_{ij} J_{ij}(\dots)|\psi\rangle = -\sum_{i \in A} \sum_{j \in B} J_{ij}\left(\frac{1}{2} \underbrace{S_i^+ S_j^- |\psi\rangle}_{=0} + \frac{1}{2} S_i^- S_j^+ |\psi\rangle + S_i^z S_j^z |\psi\rangle \right). \tag{5.11}$$

$|\psi\rangle$ is an eigenstate of the first (trivially) and the third term but not of the second, which does not vanish in this case:

$$H|\psi\rangle = -\sum_{i \in A} \sum_{j \in B} J_{ij}\left(\frac{1}{2} \underbrace{[S(S+1) - S(S-1)]}_{2S} \times \dots |S, S-1\rangle_i \dots |S, -S+1\rangle_j \dots - S^2|\psi\rangle \right), \tag{5.12}$$

which is not of the form $E|\psi\rangle$, i.e., $|\psi\rangle$ is not even an eigenstate of H and thus certainly not a ground state. There is a rigorous statement, though:

> *Marshall's theorem* (extended by Lieb and Mattis): for the Heisenberg model on a bipartite lattice with sublattices of equal size and $J_{ij} \leq 0$ for all $i \in A$ and $j \in B$ or $i \in B$ and $j \in A$ and every pair of sites i, j is connected by a string of bonds with $J_{kl} \neq 0$, the ground state $|\psi_0\rangle$ is *non-degenerate* and is a *singlet of total spin*:
>
> $$\mathbf{S}|\psi_0\rangle = 0, \tag{5.13}$$
>
> where $\mathbf{S} = \sum_i \mathbf{S}_i$.

The proof uses similar ideas to the one given for the ferromagnetic ground state but also some additional concepts, see Auerbach's book. Recall that because of $[\mathbf{S}, H] = 0$ we can choose the ground state to be a simultaneous eigenstate of $\mathbf{S}$. For a fully antiferromagnetic model it is plausible that the corresponding eigenvalue is zero. That the ground state is also non-degenerate is not obvious, though.

Note that Marshall's theorem does not uniquely determine the ground state. There are many other total-spin singlets that are not the ground state.

5.3 Helical ground states of the classical Heisenberg model

We have seen that it is difficult to find the exact quantum ground state of the Heisenberg model, unless all $J_{ij} \geq 0$. However, many properties are well described by a classical approximation. In this approximation, we replace the spin operators $\mathbf{S}_i$ by real three-component vectors of fixed magnitude $|\mathbf{S}_i| = S$. This leads to the *classical* Heisenberg model with Hamilton *function*

$$H = -\frac{1}{2} \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j. \tag{5.14}$$

This turns out to be a good approximation if

- S is large ("classical limit") and

- the dimensionality d of space is high (it works better in 3D than in 2D).

The ground state is clearly obtained by minimizing H under the conditions $|\mathbf{S}_i| = S$.

We consider the classical Heisenberg model an a (Bravais) lattice $\{\mathbf{R}_i\}$ with J_{ij} only depending on the separation vector, $J_{ij} = J_{ji} = J(\mathbf{R}_i - \mathbf{R}_j)$. We define the Fourier transform

$$\mathbf{S_q} := \frac{1}{\sqrt{N}} \sum_i e^{-i\mathbf{q}\cdot\mathbf{R}_i} \mathbf{S}_i \tag{5.15}$$

(N is the number of sites) and thus

$$\mathbf{S}_i = \frac{1}{\sqrt{N}} \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{R}_i} \mathbf{S_q}, \tag{5.16}$$

where the sum is over the first Brillouin zone. Inserting this into H we get

$$
\begin{aligned}
H &= -\frac{1}{2N} \sum_{\mathbf{qq'}} \sum_{ij} J(\underbrace{\mathbf{R}_i - \mathbf{R}_j}_{\Delta\mathbf{R}}) e^{i\mathbf{q}\cdot\mathbf{R}_i} e^{i\mathbf{q'}\cdot\mathbf{R}_j} \mathbf{S_q} \cdot \mathbf{S_{q'}} \\
&= -\frac{1}{2N} \sum_{\mathbf{qq'}} \sum_{\Delta\mathbf{R}} J(\Delta\mathbf{R}) \underbrace{\sum_{\mathbf{R}_i} e^{i(\mathbf{q}+\mathbf{q'})\cdot\mathbf{R}_i}}_{N\delta_{\mathbf{q}+\mathbf{q'},0}} e^{i\mathbf{q}\cdot\Delta\mathbf{R}} \mathbf{S_q} \cdot \mathbf{S_{q'}} \\
&= -\frac{1}{2} \sum_{\mathbf{q}} \underbrace{\sum_{\Delta\mathbf{R}} e^{-i\mathbf{q}\cdot\Delta\mathbf{R}} J(\Delta\mathbf{R})}_{=: J(\mathbf{q})} \mathbf{S_q} \cdot \mathbf{S_{-q}} \\
&= -\frac{1}{2} \sum_{\mathbf{q}} J(\mathbf{q}) \mathbf{S_q} \cdot \mathbf{S_{-q}}.
\end{aligned}
\tag{5.17}
$$

Here, $J(\mathbf{q})$ is real and even since $J(\Delta\mathbf{R})$ is real and even. $\mathbf{S_q}$ has to respect the normalization condition on $\mathbf{S}_i$:

$$S^2 = \mathbf{S}_i \cdot \mathbf{S}_i = \frac{1}{N} \sum_{\mathbf{qq'}} e^{i(\mathbf{q}+\mathbf{q'})\cdot\mathbf{R}_i} \mathbf{S_q} \cdot \mathbf{S_{q'}} = \frac{1}{N} \sum_{\mathbf{qq'}} e^{i(\mathbf{q}-\mathbf{q'})\cdot\mathbf{R}_i} \mathbf{S_q} \cdot \mathbf{S_{-q'}}. \tag{5.18}$$

Let $\mathbf{Q}$ be a wavevector for which $J(\mathbf{q})$ assumes a global maximum. We expect but do not show rigorously that H is minimized if $\mathbf{S_Q} \neq 0$ and $\mathbf{S_{-Q}} \neq 0$ and all other $S_{\mathbf{q}} = 0$ for $\mathbf{q} \neq \pm\mathbf{Q}$.

Case 1: $\mathbf{Q} = 0$, then $\mathbf{S}_i = N^{-1/2} \mathbf{S_{Q=0}}$ for all i. We obtain a homogeneous spin polarization, i.e., a *ferromagnet.*

Case 2: $\mathbf{Q} \neq 0$. Condition (5.18) reads

$$S^2 = \frac{1}{N} \left(2\mathbf{S_Q} \cdot \mathbf{S_{-Q}} + e^{2i\mathbf{Q}\cdot\mathbf{R}_i} \mathbf{S_Q} \cdot \mathbf{S_Q} + e^{-2i\mathbf{Q}\cdot\mathbf{R}_i} \mathbf{S_{-Q}} \cdot \mathbf{S_{-Q}} \right) \tag{5.19}$$

for all $\mathbf{R}_i$. The right-hand side must be independent of $\mathbf{R}_i$. We thus get $\mathbf{S_Q} \cdot \mathbf{S_Q} = \mathbf{S_{-Q}} \cdot \mathbf{S_{-Q}} = 0$. This looks like we get a trivial zero result. However, we have to take into account that $\mathbf{S_Q}$ is a Fourier transform and is thus generally complex. We therefore write

$$\mathbf{S_Q} = \mathbf{R_Q} + i\mathbf{I_Q} \quad (\mathbf{R_Q}, \mathbf{I_Q} \in \mathbb{R}^3). \tag{5.20}$$

Since $\mathbf{S}_i$ is real, we require

$$\mathbf{S_{-Q}} = \mathbf{S_Q^*} = \mathbf{R_Q} - i\mathbf{I_Q} \tag{5.21}$$

and thus

$$\mathbf{S_Q} \cdot \mathbf{S_{-Q}} = \mathbf{R_Q^2} + \mathbf{I_Q^2} \tag{5.22}$$

and

$$\mathbf{S_Q} \cdot \mathbf{S_Q} = \mathbf{R_Q^2} - i\mathbf{I_Q^2} + 2i\mathbf{R_Q} \cdot \mathbf{I_Q} \overset{!}{=} 0. \tag{5.23}$$

Thus

$$\mathbf{R_Q^2} = \mathbf{I_Q^2} \quad \text{and} \quad \mathbf{R_Q} \cdot \mathbf{I_Q} = 0, \tag{5.24}$$

showing that the real and imaginary parts of $\mathbf{S_Q}$ are orthogonal vectors of equal magnitude. Furthermore,

$$S^2 = \frac{1}{N} 2\mathbf{S_Q} \cdot \mathbf{S_{-Q}} = \frac{2}{N} \left(\mathbf{R_Q^2} + \mathbf{I_Q^2} \right) \tag{5.25}$$

so that

$$\mathbf{R}_{\mathbf{Q}}^2 = \mathbf{I}_{\mathbf{Q}}^2 = \frac{NS^2}{4}. \tag{5.26}$$

Then the ground-state energy is

$$H_0 = -\frac{1}{2}J(\mathbf{Q})\mathbf{S}_{\mathbf{Q}} \cdot \mathbf{S}_{-\mathbf{Q}} - \frac{1}{2}J(-\mathbf{Q})\mathbf{S}_{-\mathbf{Q}} \cdot \mathbf{S}_{\mathbf{Q}} = -J(\mathbf{Q})\left(\mathbf{R}_{\mathbf{Q}}^2 + \mathbf{I}_{\mathbf{Q}}^2\right) = -J(\mathbf{Q})\frac{NS^2}{2}. \tag{5.27}$$

Reverse Fourier transformation yields

$$
\begin{aligned}
\mathbf{S}_i &= \frac{1}{\sqrt{N}}\left(e^{i\mathbf{Q}\cdot\mathbf{R}_i}\mathbf{S}_{\mathbf{Q}} + e^{-i\mathbf{Q}\cdot\mathbf{R}_i}\mathbf{S}_{-\mathbf{Q}}\right) \\
&= \frac{1}{\sqrt{N}}\left[e^{i\mathbf{Q}\cdot\mathbf{R}_i}(\mathbf{R}_{\mathbf{Q}} + i\mathbf{I}_{\mathbf{Q}}) + e^{-i\mathbf{Q}\cdot\mathbf{R}_i}(\mathbf{R}_{\mathbf{Q}} - i\mathbf{I}_{\mathbf{Q}})\right] \\
&= \frac{2}{\sqrt{N}}\left(\mathbf{R}_{\mathbf{Q}}\cos\mathbf{Q}\cdot\mathbf{R}_i - \mathbf{I}_{\mathbf{Q}}\sin\mathbf{Q}\cdot\mathbf{R}_i\right).
\end{aligned} \tag{5.28}
$$

Defining the unit vectors along $\mathbf{R}_{\mathbf{Q}}$ and $\mathbf{I}_{\mathbf{Q}}$, $\hat{\mathbf{R}} := \mathbf{R}_{\mathbf{Q}}/|\mathbf{R}_{\mathbf{Q}}|$, $\hat{\mathbf{I}} := \mathbf{I}_{\mathbf{Q}}/|\mathbf{I}_{\mathbf{Q}}|$ with $\hat{\mathbf{R}} \cdot \hat{\mathbf{I}} = 0$, we get

$$\mathbf{S}_i = S\left(\hat{\mathbf{R}}\cos\mathbf{Q}\cdot\mathbf{R}_i - \hat{\mathbf{I}}\sin\mathbf{Q}\cdot\mathbf{R}_i\right). \tag{5.29}$$

Apart from the condition $\hat{\mathbf{R}} \perp \hat{\mathbf{I}}$, the directions of $\hat{\mathbf{R}}$ and $\hat{\mathbf{I}}$ are not fixed by minimizing H. Any choice minimizes the energy. Anisotropic terms in the Hamilton function would break this degeneracy, though.

For illustration, we choose coordinates in spin space such that $\hat{\mathbf{x}} = \hat{\mathbf{R}}$, $\hat{\mathbf{y}} = -\hat{\mathbf{I}}$. Then

$$
\begin{aligned}
S_i^x &= S\cos\mathbf{Q}\cdot\mathbf{R}_i \\
S_i^y &= S\sin\mathbf{Q}\cdot\mathbf{R}_i \\
S_i^z &= 0.
\end{aligned} \tag{5.30}
$$

We see that we have found a *helical* state, which is coplanar but not collinear.

Note that $\hat{\mathbf{R}}, \hat{\mathbf{I}}$ can be arbitrarily oriented relative to $\mathbf{Q}$. ($\mathbf{Q}$ itself is fixed by $J(\mathbf{q})$ having a maximum at $\mathbf{q} = \mathbf{Q}$.) For example:
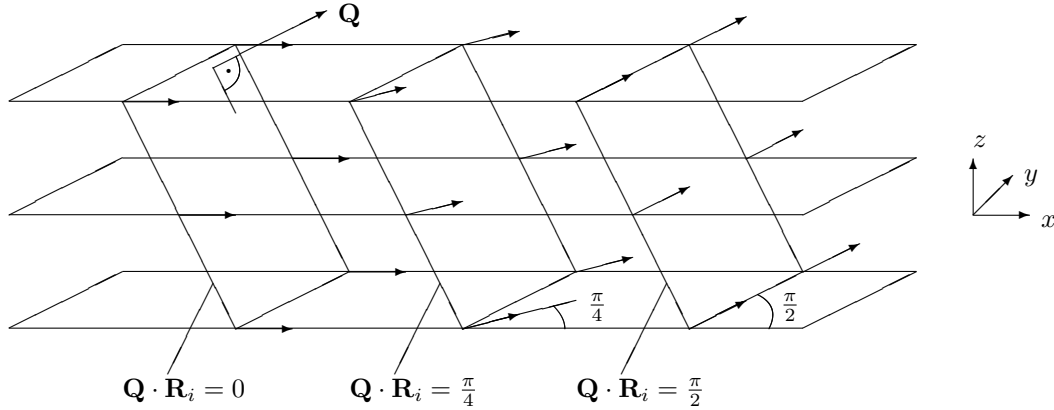
Note that $\mathbf{S}_i = \text{const}$ on any plane orthogonal to $\mathbf{Q}$.

Since $\mathbf{Q}$ is where $J(\mathbf{q})$ assumes a maximum, there is no *general* reason why $\mathbf{Q}$ should represent a special (high symmetry) point in the first Brillouin zone. Thus the helical order is generally *incommensurate*, i.e., $n\mathbf{Q}$ is not a reciprocal lattice vector for any $n \neq 0$. Then the lattice with helical spin order is not invariant under any translation in real space. Examples for compounds showing incommensurate order are $LiCu_2O_2$ and $NaCu_2O_2$. Some rare-earth metals such as Ho also show incommensurate helical order but are, of course, not ionic crystals.

On the other hand, quite often $\mathbf{Q}$ *is* a special point. For $\mathbf{Q} = 0$ we find ferromagnetism. For $\mathbf{Q}$ at a high-symmetry point at the edge of the Brillouin zone we find simple antiferromagnetic orderings. We consider a few examples for a simple cubic lattice:

$$\vec{Q}=(\tfrac{\pi}{a},\tfrac{\pi}{a},\tfrac{\pi}{a}) \qquad\qquad \vec{Q}=(\tfrac{\pi}{a},\tfrac{\pi}{a},0) \qquad\qquad \vec{Q}=(\tfrac{\pi}{a},0,0)$$
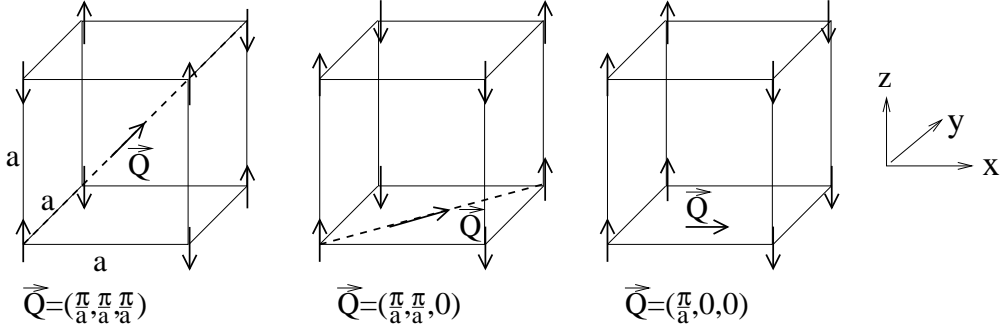
5.4 Schwinger bosons

It is often useful to map a certain model onto a different but equivalent one. In many cases in many-body theory this procedure has led to new insights. Several such mappings take a spin (e.g., Heisenberg) model into a *bosonic* one. Bosonic models are often easier to study since boson creation and annihilation operators have simpler commutation relations than spin operators. Also, a bosonic model has a straightforward physical interpretation as a system of, usually coupled, harmonic oscillators.

The *Schwinger bosons* are introduced as follows:

$$S_i^+ = a_i^\dagger b_i, \tag{5.31}$$

$$S_i^- = b_i^\dagger a_i, \tag{5.32}$$

$$S_i^z = \frac{a_i^\dagger a_i - b_i^\dagger b_i}{2}, \tag{5.33}$$

where i is the site index. One easily checks that these operators have the correct commutation relations, for example

$$[S_i^+, S_i^-] = [a_i^\dagger b_i, b_i^\dagger a_i] = a_i^\dagger [b_i, b_i^\dagger] a_i + b_i^\dagger [a_i^\dagger, a_i] b_i = a_i^\dagger a_i - b_i^\dagger b_i = 2S_i^z. \tag{5.34}$$

We also see that

$$\begin{aligned}
\mathbf{S}_i \cdot \mathbf{S}_i &= \frac{a_i^\dagger b_i^\dagger b_i a_i + b_i^\dagger a_i a_i^\dagger b_i}{2} + \frac{(a_i^\dagger a_i - b_i^\dagger b_i)^2}{4} \\
&= \frac{1}{4}\left(2a_i^\dagger b_i b_i^\dagger a_i + 2a_i^\dagger a_i + 2b_i^\dagger a_i^\dagger a_i b_i + 2b_i^\dagger b_i + a_i^\dagger a_i a_i^\dagger a_i - a_i^\dagger a_i b_i^\dagger b_i - b_i^\dagger b_i a_i^\dagger a_i + b_i^\dagger b_i b_i^\dagger b_i \right) \\
&= \frac{1}{4}\left(2a_i^\dagger a_i b_i^\dagger b_i + 2a_i^\dagger a_i + 2b_i^\dagger b_i + a_i^\dagger a_i a_i^\dagger a_i + b_i^\dagger b_i b_i^\dagger b_i \right) \\
&= \frac{a_i^\dagger a_i + b_i^\dagger b_i}{2}\left(\frac{a_i^\dagger a_i + b_i^\dagger b_i}{2} + 1 \right). \tag{5.35}
\end{aligned}$$

The spins should have a fixed spin quantum number S so that $\mathbf{S}_i \cdot \mathbf{S}_i = S(S+1)$. This requires

$$a_i^\dagger a_i + b_i^\dagger b_i = 2S. \tag{5.36}$$

This is a *local constraint* on the number of a- and b-bosons. It restricts the Fock space of the bosons to the physically meaningful subspace for given S. The constraint is what makes the bosonized model difficult. It is the prize to pay for the simple bosonic commutations relations.

The nearest-neighbor Heisenberg model on an arbitrary lattice then maps onto the following bosonic model, where every bond $\langle ij \rangle$ is counted once:

$$\begin{aligned}
H &= -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J \sum_{\langle ij \rangle} \left(\frac{S_i^+ S_j^-}{2} + S_i^z S_j^z \right) \\
&= -\frac{J}{2} \sum_{\langle ij \rangle} \left(a_i^\dagger b_i b_j^\dagger a_j + b_i^\dagger a_i a_j^\dagger b_j + \frac{1}{2} a_i^\dagger a_i a_j^\dagger a_j - \frac{1}{2} a_i^\dagger a_i \underbrace{b_j^\dagger b_j}_{=2S - a_j^\dagger a_j} - \frac{1}{2} b_i^\dagger b_i \underbrace{a_j^\dagger a_j}_{=2S - b_j^\dagger b_j} + \frac{1}{2} b_i^\dagger b_i b_j^\dagger b_j \right) \\
&= -\frac{J}{2} \sum_{\langle ij \rangle} \left(a_i^\dagger b_i b_j^\dagger a_j + b_i^\dagger a_i a_j^\dagger b_j + a_i^\dagger a_i a_j^\dagger a_j - S a_i^\dagger a_i + b_i^\dagger b_i b_j^\dagger b_j - S b_i^\dagger b_i \right) \\
&= -\frac{J}{2} \sum_{\langle ij \rangle} \left(a_i^\dagger a_i a_j^\dagger a_j + a_i^\dagger b_i b_j^\dagger a_j + b_i^\dagger a_i a_j^\dagger b_j + b_i^\dagger b_i b_j^\dagger b_j \right) + \underbrace{\text{const}}_{\text{irrelevant}}. \tag{5.37}
\end{aligned}$$

36

Thus the bosonic Hamiltonian is biquadratic, i.e., explicitly interacting. Specifically, H describes a density-density interaction and correlated hopping of an a- or b-boson from i to j accompanied by the hopping of another a- or b-boson from j to i. One easily sees that H conserves the total a- and b-boson numbers $N_a = \sum_i a_i^\dagger a_i$ and $N_b = \sum_i b_i^\dagger b_i$ seperately and also conserves the local boson number $a_i^\dagger a_i + b_i^\dagger b_i (= 2S)$. We will use the Schwinger-boson representation in the following section.

5.5 Valence-bond states

We have seen that we cannot find the exact quantum ground state of the antiferromagnetic Heisenberg model. It is sometimes possible to find a good approximation from a suitable *variational* ansatz. The *valence bond states* are of this type. They are defined as follows:

1. Let Λ_α be a *configuration of bonds* (ij) on the lattice such that exactly $2S$ bonds end at each lattice site.

   Examples: $S = \dfrac{1}{2}$

$S = 1$
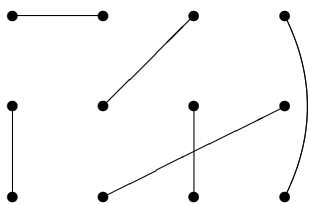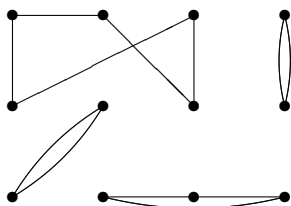
2. For any Λ_α, we define a state

$$|\alpha\rangle := \prod_{(ij)\in\Lambda_\alpha} (a_i^\dagger b_j^\dagger - b_i^\dagger a_j^\dagger)\,|0\rangle, \tag{5.38}$$

   where $a_i^\dagger$ and $b_i^\dagger$ create Schwinger bosons and $|0\rangle$ is the vacuum state without any bosons. One sees that every term in $|\alpha\rangle$ contains exactly $2S$ creation operators at each site, i.e., $|\alpha\rangle$ statisfies the constraint (5.36) for local spin quantum number S and is thus a (generally not normalized) admissable state of the Heisenberg model with spin S.

   One can show that $|\alpha\rangle$ is a spin singlet: Rotation of all spins around the x-axis by π leads to

$$\left.\begin{array}{l} S_i^x \to \ \ S_i^x \\ S_i^y \to -S_i^y \\ S_i^z \to -S_i^z \end{array}\right\} \Rightarrow \left\{\begin{array}{l} S_i^+ \to S_i^- \\ S_i^- \to S_i^+ \end{array}\right. \tag{5.39}$$

   and thus to

$$a_i \leftrightarrow b_i. \tag{5.40}$$

   This takes $|\alpha\rangle := \prod_{(ij)\in\Lambda_\alpha}(a_i^\dagger b_j^\dagger - b_i^\dagger a_j^\dagger)|0\rangle$ into

$$|\alpha\rangle' := \prod_{(ij)\in\Lambda_\alpha} (b_i^\dagger a_j^\dagger - a_i^\dagger b_j^\dagger)|0\rangle = (-1)^{NS} \prod_{(ij)\in\Lambda_\alpha} (a_i^\dagger b_j^\dagger - b_i^\dagger a_j^\dagger)|0\rangle$$
$$= (-1)^{NS}|\alpha\rangle. \tag{5.41}$$

   Here, $(-1)^{NS}$ is an irrelevant total phase factor. Thus $|\alpha\rangle$ is invariant under this rotation. Analogously one can show that it is also invariant under any other spin rotation. Thus $|\alpha\rangle$ is a spin singlet. This is of course desirable in the case of antiferromagnetic interactions because of Marshall's theorem.

3. A general valence-bond state has the form

$$|\{C_\alpha\}, S\rangle := \sum_\alpha C_\alpha |\alpha\rangle, \quad C_\alpha \in \mathbb{C}, \tag{5.42}$$

i.e., it is a superposition of states $|\alpha\rangle$ defined above and thus also an admissable state of the Heisenberg model. Like any $|\alpha\rangle$, the superposition is a spin singlet.

For the case that macroscopically many (i.e., a number growing sufficiently rapidly with the lattice size N) states $|\alpha\rangle$ contibute to $|\{C_\alpha\}, \alpha\rangle$, Anderson has introduced the term *resonating valence bond* (RVB) states. RVB states have been used by Anderson to describe the ground state of (doped) two-dimensional Mott antiferromagnets. These are relevant for the cuprates. Note, however, that the two-dimensional Heisenberg antiferromagnet on the square lattice very probably has an ordered ground state, not an RVB ground state.

5.5.1 The case $S = 1/2$

We consider the case of spin $S = 1/2$. In this case a basis for the Hilbert space of a single spin is

$$|\uparrow_i\rangle := a_i^\dagger |0\rangle, \tag{5.43}$$

$$|\downarrow_i\rangle := b_i^\dagger |0\rangle. \tag{5.44}$$

We consider only *bipartite* valence-bond configurations Λ_α, i.e., $(ij) \in \Lambda_\alpha$ only if $i \in A$ and $j \in B$ or $i \in B$ and $j \in A$ for sublattices A and B. This is natural in the case of bipartite models, see Sec. 5.2. We can assume $i \in A$ and $j \in B$ without loss of generality. Then the states $|\alpha\rangle$ take the form

$$|\alpha\rangle = \prod_{(ij) \in \Lambda_\alpha} \frac{|\uparrow_i\rangle |\downarrow_i\rangle - |\downarrow_i\rangle |\uparrow_i\rangle}{\sqrt{2}}, \tag{5.45}$$

where we have introduced a normalization factor.

One can calculate the spin correlation function in this state, which reads

$$\langle\alpha|\mathbf{S}_k \cdot \mathbf{S}_l|\alpha\rangle - \langle\alpha|\mathbf{S}_k|\alpha\rangle \cdot \langle\alpha|\mathbf{S}_l|\alpha\rangle. \tag{5.46}$$

The case $k = l$ trivially gives $\langle\alpha|\mathbf{S}_k \cdot \mathbf{S}_l|\alpha\rangle = S(S+1) = 3/4$. We thus assume $k \neq l$ in the following. Now note that for $S = 1/2$, any site only belongs to a single bond in Λ_α. If k and l do not belong to the same bond, $|\alpha\rangle$ does not contain any correlation between $\mathbf{S}_k$ and $\mathbf{S}_l$ so that $\langle\alpha|\mathbf{S}_k \cdot \mathbf{S}_l|\alpha\rangle - \langle\alpha|\mathbf{S}_k|\alpha\rangle \cdot \langle\alpha|\mathbf{S}_l|\alpha\rangle = 0$. If k and l belong to the same bond, we have, for $k \in A$ and $l \in B$ without loss of generality,

$$
\begin{aligned}
\langle\alpha|\mathbf{S}_k \cdot \mathbf{S}_l|\alpha\rangle &= \frac{\langle\downarrow_l|\langle\uparrow_k| - \langle\uparrow_l|\langle\downarrow_k|}{\sqrt{2}} \left(\frac{S_k^+ S_l^- + S_k^- S_l^+}{2} + S_k^z S_l^z \right) \frac{|\uparrow_k\rangle|\downarrow_l\rangle - |\downarrow_k\rangle|\uparrow_l\rangle}{\sqrt{2}} \\
&= \frac{1}{2}\langle\downarrow_l|\langle\uparrow_k|S_k^z S_l^z|\uparrow_k\rangle|\downarrow_l\rangle - \frac{1}{4}\langle\downarrow_l|\langle\uparrow_k|S_k^+ S_l^-|\downarrow_k\rangle|\uparrow_l\rangle \\
&\quad - \frac{1}{4}\langle\uparrow_l|\langle\downarrow_k|S_k^- S_l^+|\uparrow_k\rangle|\downarrow_l\rangle - \frac{1}{2}\langle\uparrow_l|\langle\downarrow_k|S_k^z S_l^z|\downarrow_k\rangle|\uparrow_l\rangle \\
&= -\frac{1}{8} - \frac{1}{4} - \frac{1}{4} - \frac{1}{8} = -\frac{3}{4}.
\end{aligned} \tag{5.47}
$$

In summary,

$$\langle\alpha|\mathbf{S}_k \cdot \mathbf{S}_l|\alpha\rangle - \langle\alpha|\mathbf{S}_k|\alpha\rangle \cdot \langle\alpha|\mathbf{S}_l|\alpha\rangle = \begin{cases} \dfrac{3}{4} & \text{if } k = l \\[2mm] -\dfrac{3}{4} & \text{if } (kl) \in \Lambda_\alpha \text{ or } (lk) \in \Lambda_\alpha \\[2mm] 0 & \text{otherwise.} \end{cases} \tag{5.48}$$

If all bonds in Λ_α are of short range, i.e., if there is a length Ξ so that Λ_α does not contain bonds (ij) with $|\mathbf{R}_i - \mathbf{R}_j| \geq \Xi$, the spin correlations in $|\alpha\rangle$ are of short range. (Note that it would be sufficient if the fraction of bonds of length $|\mathbf{R}_i - \mathbf{R}_j|$ decayed exponentially with $|\mathbf{R}_i - \mathbf{R}_j|$.) Then any superposition $|\{C_\alpha\}, 1/2\rangle = \sum_\alpha C_\alpha |\alpha\rangle$ of such states also has only short-range spin correlations. States of this type are called *spin-liquid* states.

We note that any $|\alpha\rangle$ for $S = 1/2$ necessarily breaks translational symmetry since Λ_α does (any site belongs only to a single bond). The valence-bond state $|\{C_\alpha\}, 1/2\rangle$ can restore translational symmetry, though.

5.5.2 The spin-1/2 chain

We consider a chain of spins wth $S = 1/2$ and antiferromagnetic interactions. If the interactions are only between neighbors, it is reasonable to consider only valence-bond states with nearest-neighbor bonds in Λ_α. (Recall that the valence-bond states represent a variational ansatz. We *choose* which states we want to include based on physical insight.) But then the valence-bond states become very simple: There are only two distinct bond confugrations:

$\Lambda_+ = $ •——• •——• •——• •——•

$\Lambda_- = $ —• •——• •——• •——• •—

Here we assume periodic boundary conditions and an even number N of sites. These two configurations lead to two *dimer states*

$$|\pm\rangle := \prod_{n=1}^{N/2} \frac{|\uparrow_{2n}\rangle|\downarrow_{2n\pm1}\rangle - |\downarrow_{2n}\rangle|\uparrow_{2n\pm1}\rangle}{\sqrt{2}}. \tag{5.49}$$

The Hamiltonian is

$$H = -J\sum_i \mathbf{S}_i \cdot \mathbf{S}_{i+1} \qquad \text{with } J < 0. \tag{5.50}$$

It is clear from symmetry that $\langle+|H|+\rangle = \langle-|H|-\rangle$. Thus the coefficients c_+, c_- are not useful variational parameters. But we can at least calculate this expectation value, which will give an upper bound for the true ground-state energy:

$$\langle+|H|+\rangle = -J\sum_i \langle+|\mathbf{S}_i \cdot \mathbf{S}_{i+1}|+\rangle = -J\sum_{n=1}^{N/2} \left(\langle+|\mathbf{S}_{2n} \cdot \mathbf{S}_{2n+1}|+\rangle + \langle+|\mathbf{S}_{2n} \cdot \mathbf{S}_{2n-1}|+\rangle\right). \tag{5.51}$$

We have calculated the expectation values of $\mathbf{S}_k \cdot \mathbf{S}_l$ above:

$$\langle+|H|+\rangle = -J\sum_{n=1}^{N/2} \left(\frac{3}{4} + 0\right) = +\frac{3}{4}\frac{N}{2}J = \frac{3}{8}NJ < 0. \tag{5.52}$$

Note that $|+\rangle$ cannot be the true ground state, since it is degenerate with $|-\rangle$, whereas Marshall's theorem shows that the ground state is non-degenerate.

It is instructive to compare $\langle\pm|H|\pm\rangle$ to the energy expectation value for the Néel state

$$|\text{Néel}\rangle := |\uparrow_1\rangle|\downarrow_2\rangle|\uparrow_3\rangle|\downarrow_4\rangle\ldots \tag{5.53}$$

We get

$$\langle\text{Néel}|H|\text{Néel}\rangle = -J\sum_i \langle\text{Néel}|\mathbf{S}_i \cdot \mathbf{S}_{i+1}|\text{Néel}\rangle = -J\sum_i \langle\text{Néel}|\mathbf{S}_i^z \cdot \mathbf{S}_{i+1}^z|\text{Néel}\rangle$$

$$= -J\sum_i \langle\uparrow|S^z|\uparrow\rangle\langle\downarrow|S^z|\downarrow\rangle$$

$$= -J\sum_i \frac{1}{2}\left(-\frac{1}{2}\right) = \frac{1}{4}NJ < 0. \tag{5.54}$$

We see that

$$\langle\pm|H|\pm\rangle < \langle\text{Néel}|H|\text{Néel}\rangle, \tag{5.55}$$

the valence-bond states thus provide a tighter bound on the ground-state energy. It is not in general true that variational states with lower energy are more similar to the true ground state. But in the present case numerical simulations do suggest that the ground state is indeed similar to the simple valence-bond states $|\pm\rangle$ considered here.

5.5.3 The Majumdar-Ghosh Hamiltonian

It is interesting that a Hamiltonian exists for which $|\pm\rangle$ are the true ground states. A Hamiltonian for which a given state is the ground state is called a *parent Hamiltonian* of this state. In our case it was found by Majumdar and Ghosh. It reads

$$H_{\text{MG}} := -K\sum_i \left(\frac{4}{3}\mathbf{S}_i \cdot \mathbf{S}_{i+1} + \frac{2}{3}\mathbf{S}_i \cdot \mathbf{S}_{i+2}\right) \quad \text{with } K < 0. \tag{5.56}$$

The exchange interaction is here denoted by the symbol K to avoid confusion later. Note that the second term introduces frustration since it is antiferromagnetic and not bipartite.

To show that $|\pm\rangle$ are the ground states of H_{MG} we first define the total spin of a triad of sites,

$$\mathbf{J}_i := \mathbf{S}_{i-1} + \mathbf{S}_i + \mathbf{S}_{i+1}. \tag{5.57}$$

Clearly, $\mathbf{J}_i \cdot \mathbf{J}_i$ has eigenvalues $J_i(J_i + 1)$ with $J_i = 1/2$ or $J_i = 3/2$. We now define

$$P_i := \frac{1}{3}\left(\mathbf{J}_i \cdot \mathbf{J}_i - \frac{3}{4}\right). \tag{5.58}$$

P_i has the eigenvalues

$$\frac{1}{3}\left[J(J+1) - \frac{3}{4}\right] = \begin{cases} 0 & \text{for } J_i = 1/2 \\ 1 & \text{for } J_i = 3/2. \end{cases} \tag{5.59}$$

Thus $P_i P_i = P_i$ so that P_i is a *projection operator*. It clearly projects onto the subspace of triad spin $3/2$. On the other hand,

$$\begin{aligned} P_i &= \frac{1}{3}\left[(\mathbf{S}_{i-1} + \mathbf{S}_i + \mathbf{S}_{i+1}) \cdot (\mathbf{S}_{i-1} + \mathbf{S}_i + \mathbf{S}_{i+1}) - \frac{3}{4}\right] \\ &= \frac{1}{3}\left[\frac{3}{4} + \frac{3}{4} + \frac{3}{4} + 2\mathbf{S}_{i-1} \cdot \mathbf{S}_i + 2\mathbf{S}_{i-1} \cdot \mathbf{S}_{i+1} + 2\mathbf{S}_i \cdot \mathbf{S}_{i+1} - \frac{3}{4}\right] \\ &= \frac{1}{2} + \frac{2}{3}\left(\mathbf{S}_{i-1} \cdot \mathbf{S}_i + \mathbf{S}_{i-1} \cdot \mathbf{S}_{i+1} + \mathbf{S}_i \cdot \mathbf{S}_{i+1}\right). \end{aligned} \tag{5.60}$$

Thus

$$-K\sum_i P_i = -\frac{1}{2}NK - K\sum_i \left(\frac{4}{3}\mathbf{S}_i \cdot \mathbf{S}_{i+1} + \frac{2}{3}\mathbf{S}_{i-1} \cdot \mathbf{S}_{i+1}\right) = H_{\mathrm{MG}} - \frac{1}{2}NK \tag{5.61}$$

and we can write the Hamiltonian as

$$H_{\mathrm{MG}} = -K\sum_i P_i + \frac{1}{2}NK. \tag{5.62}$$

In the dimer states $|\pm\rangle$, none of the triad spins can be $3/2$ since two of the three spins form a singlet. Thus $P_i|\pm\rangle = 0$ for all i. Therefore,

$$H_{\mathrm{MG}}|\pm\rangle = \frac{1}{2}NK|\pm\rangle \tag{5.63}$$

so that $|\pm\rangle$ are eigenstates of H_{MG} to the eigenenergy $(1/2)NK < 0$.

To see that they are also ground states, note that $-K\sum_i P_i$ has eigenvalues $0, -K, -2K, \ldots$, i.e., is non-negative ($K < 0$). Thus energy expectation values cannot be smaller than $(1/2)NK$. (Note that we have not shown that there are no other ground states besides $|\pm\rangle$.)

The usefulness of parent Hamiltonians such as the Majumdar-Ghosh model is that they allow us to derive exact results. While these do not directly apply to other models such as the nearest-neighbor Heisenberg model, they do give insight into the generic behavior of families of spin models.

Chapter 6

Mean-field theory for magnetic insulators

So far, we have considered the ground states of Heisenberg-type models. This turned out to be difficult, unless all interactions were ferromagnetic, and required approximations. We now turn to the *equilibrium state* of such models at non-zero temperatures. We should expect this to be an even harder problem and indeed it is. The equilibrium state is described by the canonical partition function

$$Z = \sum_{|\Psi\rangle} e^{-\beta E_\Psi}, \tag{6.1}$$

where $\beta = 1/k_B T$ and the sum is over all microstates with energies E_Ψ of the full systems. To find it exactly, we would need not only the exact ground state but *all* states. We evidently have to make approximations. The simplest—but quite powerful—one is the mean-field approximation.

6.1 Weiß' molecular-field theory

We return to the quantum-mechanical Heisenberg Hamiltonian

$$H = -\frac{1}{2} \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j \tag{6.2}$$

with $\mathbf{S}_i \cdot \mathbf{S}_i = S(S+1)$. We write

$$\mathbf{S}_i = \underbrace{\langle \mathbf{S}_i \rangle}_{\text{average}} + \underbrace{(\mathbf{S}_i - \langle \mathbf{S}_i \rangle)}_{\text{fluctuations}}, \tag{6.3}$$

where $\langle \mathbf{S}_i \rangle = \operatorname{Tr} \rho \, \mathbf{S}_i$ denotes the ensemble (thermal) average and ρ is the density operator. Equation (6.3) is trivially exact. We rewrite H in terms of averages and fluctuations,

$$H = -\frac{1}{2} \sum_{ij} J_{ij} \langle \mathbf{S}_i \rangle \cdot \langle \mathbf{S}_j \rangle - \frac{1}{2} \sum_{ij} J_{ij} \langle \mathbf{S}_i \rangle \cdot (\mathbf{S}_j - \langle \mathbf{S}_j \rangle)$$
$$- \frac{1}{2} \sum_{ij} J_{ij} (\mathbf{S}_i - \langle \mathbf{S}_i \rangle) \cdot \langle \mathbf{S}_j \rangle - \frac{1}{2} \sum_{ij} J_{ij} (\mathbf{S}_i - \langle \mathbf{S}_i \rangle) \cdot (\mathbf{S}_j - \langle \mathbf{S}_j \rangle). \tag{6.4}$$

We now assume the fluctuations $\mathbf{S}_i - \langle \mathbf{S}_i \rangle$ to be small in the sense that the contribution of the last term in H is negligible. Neglecting this term we obtain

$$H \cong +\frac{1}{2} \sum_{ij} J_{ij} \langle \mathbf{S}_i \rangle \cdot \langle \mathbf{S}_j \rangle - \frac{1}{2} \sum_j \mathbf{S}_j \cdot \sum_i J_{ij} \langle \mathbf{S}_i \rangle - \frac{1}{2} \sum_i \mathbf{S}_i \cdot \sum_j J_{ij} \langle \mathbf{S}_j \rangle$$
$$= -\sum_i \mathbf{S}_i \cdot \sum_j J_{ij} \langle \mathbf{S}_j \rangle + \frac{1}{2} \sum_{ij} J_{ij} \langle \mathbf{S}_i \rangle \cdot \langle \mathbf{S}_j \rangle$$
$$=: H_{\text{MF}}. \tag{6.5}$$

We have thus replaced the exchange interaction by a Zeeman term describing spins in an effective $\mathbf{B}$-field (or "*molecular field*")

$$\mathbf{B}_{\text{eff}}(\mathbf{R}_i) = \frac{1}{g \mu_B} \sum_j J_{ij} \langle \mathbf{S}_j \rangle \tag{6.6}$$

and a constant energy shift. In view of Sec. 5.3 we consider *helical* equilibrium states as a reasonably general class of solutions,

$$\langle \mathbf{S}_i \rangle = M \left(\cos(\mathbf{Q} \cdot \mathbf{R}_i + \alpha), \sin(\mathbf{Q} \cdot \mathbf{R}_i + \alpha), 0 \right), \tag{6.7}$$

where $0 \geq M \geq S$. Since H is isotropic in spin space, any uniform notation of $\langle \mathbf{S}_i \rangle$ is just as good. Then the effective field is

$$\mathbf{B}_{\text{eff}}(\mathbf{R}_i) = \frac{M}{g\mu_B} \sum_j J_{ij} \left(\cos(\mathbf{Q} \cdot \mathbf{R}_j + \alpha), \sin(\mathbf{Q} \cdot \mathbf{R}_j + \alpha), 0 \right). \tag{6.8}$$

In Sec. 5.3 we have already defined

$$
\begin{aligned}
J(\mathbf{q}) &= \sum_{\Delta\mathbf{R}} e^{-i\mathbf{q}\cdot\Delta\mathbf{R}} J(\Delta\mathbf{R}) = \sum_j e^{-i\mathbf{q}\cdot(\mathbf{R}_i - \mathbf{R}_j)} J_{ij} \\
&= e^{-i\mathbf{q}\cdot\mathbf{R}_i} \sum_j e^{i\mathbf{q}\cdot\mathbf{R}_i} J_{ij} = e^{-i\mathbf{q}\cdot\mathbf{R}_i - i\alpha} \sum_j e^{i\mathbf{q}\cdot\mathbf{R}_j + i\alpha} J_{ij}.
\end{aligned}
\tag{6.9}
$$

This implies

$$e^{i\mathbf{q}\cdot\mathbf{R}_i + i\alpha} J(\mathbf{q}) = \sum_j e^{i\mathbf{q}\cdot\mathbf{R}_j + i\alpha} J_{ij}. \tag{6.10}$$

Decomposing this equation into real and imaginary parts we obtain (note that $J(\mathbf{q})$ is real)

$$
\begin{aligned}
\sum_j J_{ij} \cos(\mathbf{q} \cdot \mathbf{R}_j + \alpha) &= J(\mathbf{q}) \cos(\mathbf{q} \cdot \mathbf{R}_i + \alpha), \\
\sum_j J_{ij} \sin(\mathbf{q} \cdot \mathbf{R}_j + \alpha) &= J(\mathbf{q}) \sin(\mathbf{q} \cdot \mathbf{R}_i + \alpha).
\end{aligned}
\tag{6.11}
$$

Inserting these equations into Eq. (6.8) gives

$$\mathbf{B}_{\text{eff}}(\mathbf{R}_i) = \frac{M}{g\mu_B} J(\mathbf{Q}) \left(\cos(\mathbf{Q} \cdot \mathbf{R}_i + \alpha), \sin(\mathbf{Q} \cdot \mathbf{R}_i + \alpha), 0 \right). \tag{6.12}$$

We find that $\mathbf{B}_{\text{eff}}$ is everywhere parallel to $\langle \mathbf{S}_i \rangle$. Since the $\langle \mathbf{S}_i \rangle$ are the spin averages in the effective field $\mathbf{B}_{\text{eff}}$ and H does not contain any anisotropy terms, this is required of a consistent approximation. Furthermore, $\mathbf{B}_{\text{eff}}(\mathbf{R}_i)$ and $\langle \mathbf{S}_i \rangle$ have to point in the same (and not opposite) direction. This is the case if $J(\mathbf{Q}) > 0$. This inequality is satisfied, as we will see.

We now calculate the absolute value of $\langle \mathbf{S}_i \rangle$ quantum-mechanically:

$$|\langle \mathbf{S}_i \rangle| \equiv M = \frac{\sum_{m=-S}^{S} m \exp\left(MJ(\mathbf{Q})m/k_BT\right)}{\sum_{m=-S}^{S} \exp\left(MJ(\mathbf{Q})m/k_BT\right)} =: SB_S \left(\frac{MJ(\mathbf{Q})m}{k_BT} \right), \tag{6.13}$$

which defines the *Brillouin function* $B_S(x)$ for spin S. One can show that $B_S(x)$ has the closed form

$$B_S(x) = \frac{2S+1}{2S} \coth\left(\frac{2S+1}{2S} x \right) - \frac{1}{2S} \coth \frac{x}{2S}. \tag{6.14}$$

The function $B_S(x)$ is plotted for several values of S in the following graph:
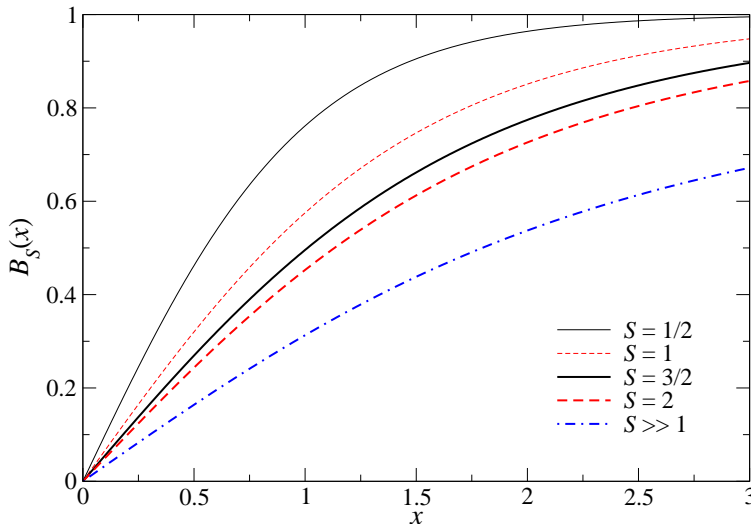
The mean-field approximation now amounts to solving the equation

$$M = SB_S\left(\frac{J(\mathbf{Q})SM}{k_BT}\right) \tag{6.15}$$

for the unknown $M \in [0, S]$. Since $B_S(0) = 0$, there is always a solution $M = 0$. Non-trivial solutions have to be found numerically or graphically.
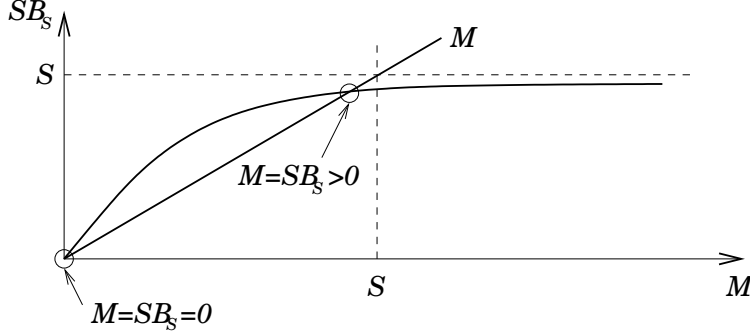
Solutions correspond to intersections between the straight line $M \to M$ (the identity function) and the curve $M \to SB_S(J(\mathbf{Q})M/k_BT)$. A non-trivial solution exists if SB_S has a larger slope at $M = 0$ than the identity function, i.e., if

$$\frac{d}{dM}SB_S\left(\frac{J(\mathbf{Q})M}{k_BT}\right)\bigg|_{M=0} > 1. \tag{6.16}$$

Since for small x, $B_S(x) \cong [(S+1)/3S]\,x$, we obtain the condition

$$\cancel{S}\frac{S+1}{3\cancel{S}}\frac{J(\mathbf{Q})S}{k_BT} = \frac{J(\mathbf{Q})S(S+1)}{3k_BT} > 1. \tag{6.17}$$

Thus in the mean-field approximation long-range order with the ordering vector $\mathbf{Q}$ can exist for temperatures $T < T_\mathbf{Q}$ with the *critical temperature*

$$T_\mathbf{Q} = \frac{J(\mathbf{Q})S(S+1)}{3k_B}. \tag{6.18}$$

Note that the prefactor depends on the definition of the Hamiltonian. Without the factor $1/2$ in H we would get $2J(\mathbf{Q})S(S+1)/3k_B$, like in Yosida's book.

We expect the system to order with a $\mathbf{Q}$-vector for which the interaction $J(\mathbf{Q})$ is strongest. This clearly corresponds to the maximum critical temperature

$$T_c = \frac{\max J(\mathbf{q})S(S+1)}{3k_B}. \tag{6.19}$$

If the maximum of $J(\mathbf{q})$ is at $\mathbf{Q} = 0$, we clearly get ferromagnetic order. In this case $T_c = T_C$ is called the *Curie temperature*. For antiferromagnetic order, $T_c = T_N$ is called the *Néel temperature*.

We now fill in a gap in the derivation by showing that $J(\mathbf{Q}) > 0$ at the maximum: Note that

$$\frac{1}{N}\sum_\mathbf{q}J(\mathbf{q}) = \frac{1}{N}\sum_\mathbf{q}\sum_{\Delta\mathbf{R}}e^{-i\mathbf{q}\cdot\Delta\mathbf{R}}J(\Delta\mathbf{R}) = \frac{1}{N}\sum_{\Delta\mathbf{R}}N\delta_{\Delta\mathbf{R},0}J(\Delta\mathbf{R}) = J(\Delta\mathbf{R}=0) = 0. \tag{6.20}$$

Thus the average of $J(\mathbf{q})$ is zero. If there is *any* exchange interaction, $J(\mathbf{q})$ is not identically zero. Consequently, $\max J(\mathbf{q}) > 0$.

To see that the non-trivial solution is the stable one, if it exists, one has to consider the free energy. The result within mean-field theory is that for $T < T_c$ the free energy has a minimum at $M > 0$ but a maximum at $M = 0$.
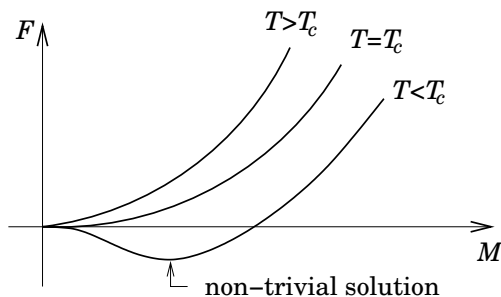
As noted above, any rotation gives another equilibrium solution. In addition, the phase α is also not constrained by the mean-field theory. Thus we obtain a hugely degenerate space of equilibrium solutions— another example of spontaneous symmetry breaking.

To actually find the magnetization M, we have to solve the mean-field equation numerically. The solution is of the form sketched here:

Analytical results exist in limiting cases:

1. $T \ll T_c$: Since for large x,

$$B_S(x) \cong 1 + \frac{2S+1}{S} \exp\left(-\frac{2S+1}{S}x\right) - \frac{1}{S}\exp\left(-\frac{x}{S}\right) = 1 + \frac{1}{S}\left[(2s+1)e^{-2x} - 1\right]e^{-x/S}$$

$$\cong 1 - \frac{1}{S}e^{-x/S}, \tag{6.21}$$

we have

$$M = SB_S \cong S - \exp\left(-\frac{\max J(\mathbf{q})M}{k_BT}\right) \cong S - \exp\left(-\frac{\max J(\mathbf{q})M}{k_BT}\right)$$
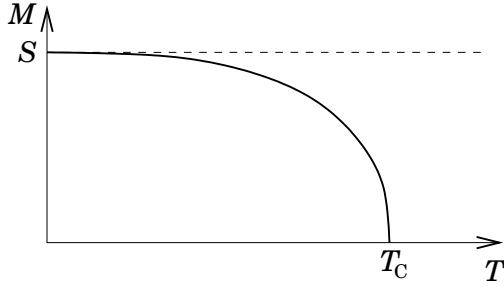
$$= S - \exp\left(-\frac{3}{S+1}\frac{\max J(\mathbf{q})S(S+1)}{3k_BT}\right) = S - \exp\left(-\frac{3}{S+1}\frac{T_c}{T}\right). \tag{6.22}$$

The correction is exponentially small. We will see that fluctuations change this behavior to a power law.

2. $T \to T_c^-$: We expand $B_S(x)$ for small x,

$$B_S(x) \cong \frac{S+1}{3S}x - \frac{(S+1)(2S^2+2S+1)}{90S^3}x^3. \tag{6.23}$$

This implies

$$M \cong \frac{S+1}{3}\frac{\max(\mathbf{q})SM}{k_BT} - \frac{(S+1)(2S^2+2S+1)}{90S^2}\frac{(\max J(\mathbf{q}))^3S^3M^3}{(k_BT)^3}$$

$$= \frac{T_c}{T}M - \frac{3(2S^2+2S+1)}{10S^2(S+1)^2}\left(\frac{T_c}{T}\right)^3M^3. \tag{6.24}$$

Solving for M^2, we get

$$M^2 \cong \frac{10S^2(S+1)^2}{3(2S^2+2S+1)}\left(\frac{T}{T_c}\right)^3\left(\frac{T_c}{T}-1\right) \cong \frac{10S^2(S+1)^2}{3(2S^2+2S+1)}\left(\frac{T_c}{T}-1\right) \tag{6.25}$$

and thus within the mean-field approximation M is proportional to

$$\sqrt{\frac{T_c}{T}-1} = \sqrt{\frac{T_c-T}{T}} \cong \sqrt{\frac{T_c-T}{T_c}} \quad \text{for } T \to T_c^-. \tag{6.26}$$

The critical exponent β is defined by $M \propto (T_c - T)^\beta$ for $T \to T_c^-$. In mean-field theory we thus find $\beta = 1/2$. Fluctuations also change this value.

6.2 Susceptibility: the Curie-Weiß Law

The uniform susceptibility is defined by $\chi = \partial M/\partial B$, where $\mathbf{M}$ is the magnetization and $\mathbf{B}$ is the uniform applied magnetic induction. Since the Heisenberg model is isotropic, χ is a scalar, more generally it would be a tensor with components

$$\chi_{ij} = \frac{\partial M_i}{\partial B_j}. \tag{6.27}$$

We are here interested in the case of a weak field $\mathbf{B}$, for which linear-response theory is valid. In the *paramagnetic phase* for $T > T_c$, the applied field induces a magnetization

$$M = S B_S \left(\frac{J(\mathbf{q} = 0) S M + g \mu_B S B}{k_B T} \right). \tag{6.28}$$

Since the magnetization is uniform, it depends on the interaction $J(\mathbf{q} = 0)$ at $\mathbf{q} = 0$ and not on $\max J(\mathbf{q})$. We define the *paramagnetic Curie temperature*

$$T_0 := \frac{J(0) S(S+1)}{3 k_B}. \tag{6.29}$$

Obviously $T_0 = T_c$ in the ferromagnetic case. Equation (6.28) now becomes

$$M = S B_S \left(\frac{3}{S+1} \frac{T_0}{T} M + \frac{g \mu_B S B}{k_B T} \right). \tag{6.30}$$

For small B and $M \ll S$ we can expand:

$$M \cong \cancel{S} \frac{S+1}{3 \cancel{S}} \left(\frac{3}{S+1} \frac{T_0}{T} M + \frac{g \mu_B S B}{k_B T} \right) = \frac{T_0}{T} M + \frac{S(S+1)}{3} \frac{g \mu_B B}{k_B T} \tag{6.31}$$

so that

$$\frac{T - T_0}{T} M \cong \frac{S(S+1)}{3} \frac{g \mu_B B}{k_B T} \tag{6.32}$$

$$\Rightarrow \qquad M \cong \frac{S(S+1)}{3} \frac{g \mu_B B}{k_B (T - T_0)}. \tag{6.33}$$

Recall that $M = |\langle \mathbf{S}_i \rangle|$. The proper magnetization is given by

$$\mathbf{M} = g \mu_B n \langle \mathbf{S}_i \rangle, \tag{6.34}$$

where n is the concentration of magnetic ions. With these additional factors we obtain

$$M \cong \frac{S(S+1)}{3} \frac{g^2 \mu_B^2 n}{k_B (T - T_0)} B \tag{6.35}$$

and thus

$$\chi \cong \frac{S(S+1)}{3} \frac{g^2 \mu_B^2 n}{k_B (T - T_0)}. \tag{6.36}$$

This is the *Curie-Weiß Law*, which holds for $T > T_0$, but not too close to T_0.

For a ferromagnet, we have $T_0 = T_c$ and χ diverges for $T \to T_c^+$. Consequently, $1/\chi$ approaches zero. Experimental results are therefore often plotted as $1/\chi$ vs. T:
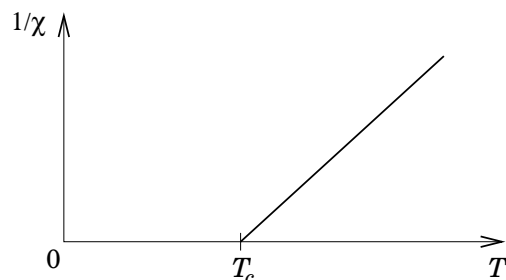
In general, one expect a divergence of the form

$$\chi \propto \frac{1}{(T - T_c)^\gamma} \tag{6.37}$$

as T approaches the critical point at T_c. Thus in mean-field theory, we find $\gamma = 1$. This exponent is also changed by fluctuations ignored in mean-field theory. See also Sec. 7.2.

For general helical order, $T_0 < T_c$ and χ grows for $T \to T_c^+$ but does not diverge. (The linear response to a field modulated with wavevector $\mathbf{Q}$ does diverge at T_c.) The following sketch shows $1/\chi$ for three different materials:
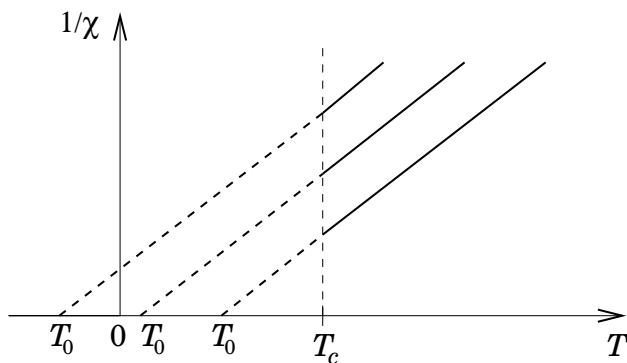
Note that $T_0 < 0$ is possible since $J(\mathbf{q} = 0)$ can be negative. This is often the case for antiferromagnets.

6.3 Validity of the mean-field approximation

The mean-field approximation replaces the fluctuating effective field acting on a spin by its time average. This is clearly good if the fluctuations in the effective field are small. Fluctuations are supressed if many random quantities are added together (this is expressed by the *law of large numbers*). Thus the mean-field approximation is good if

- the exchange interaction J_{ij} is of long range or

- each spin has many nearest neighbors (high coordination number z, thus fcc, $z = 12$, is better than diamond lattice, $z = 4$).

We will see that fluctuations are stronger in lower dimensions beyond the effect of the usually lower co-ordination number. Furthermore, anisotropy suppresses fluctuations. Thus two additional criteria for the validity of the mean-field approximation are

- the dimension d is high (mean-field theory becomes exact for $d \to \infty$) or

- anisotropy is strong.

We will now consider cases where the mean-field approximation fails completely, i.e., makes even qualitatively incorrect predictions.

6.3.1 Weak bonds

The first example concerns systems with weak and strong interactions. Let us consider a lattice of *dimers* of spin S with strong ferromagnetic exchange interaction J, which are coupled to each other by a much weaker ferromagnetic exchange interaction $J' \ll J$.
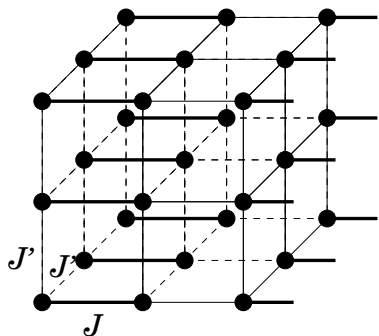
The mean-field Hamiltonian reads

$$H_{\mathrm{MF}} = -\sum_i g\mu_B \mathbf{B}_{\mathrm{eff}} \cdot \mathbf{S}_i + \mathrm{const}, \qquad (6.38)$$

where we have assumed a uniform effective field as appropriate for a ferromagnet. In our example,

$$\mathbf{B}_{\mathrm{eff}} = \frac{1}{g\mu_B}\left(J \langle \mathbf{S} \rangle + 5J' \langle \mathbf{S} \rangle\right) = \frac{1}{g\mu_B}(J + 5J') \langle \mathbf{S} \rangle. \qquad (6.39)$$

Choosing the z-axis along $\langle \mathbf{S} \rangle$ and $M := |\langle \mathbf{S} \rangle|$, we write

$$\mathbf{B}_{\text{eff}} = \frac{1}{g\mu_B}(J + 5J')\, M\hat{\mathbf{z}} \tag{6.40}$$

and

$$H_{\text{MF}} = -\sum_i (J + 5J')\, M S_i^z + \text{const.} \tag{6.41}$$

This leads to the mean-field equation

$$M = SB_S\left(\frac{[J + 5J']SM}{k_B T}\right). \tag{6.42}$$

There is a non-zero solution for M for all

$$T < T_c = \frac{(J + 5J')S(S + 1)}{3k_B} \cong \frac{JS(S + 1)}{3k_B} \quad (\text{since } J \gg J'). \tag{6.43}$$

However, this is completely wrong: In the limit $J'/J \to 0$ the system consists of non-interacting dimers and can by solved exactly: In the presence of an applied B-field, the dimer Hamiltonian reads

$$H_{\text{dimer}} = -J\mathbf{S}_1 \cdot \mathbf{S}_2 - g\mu_B B(S_1^z + S_2^z) = -\frac{J}{2}\left[S_{\text{tot}}(S_{\text{tot}} + 1) - 2S(S + 1)\right] - g\mu_B B\, m_{\text{tot}},$$

where $S_{\text{tot}} = 0, 1, \ldots, 2S$ is the total spin and $m_{\text{tot}} = -S_{\text{tot}}, \ldots, S_{\text{tot}}$. For the example of $S = 1/2$, the partition function reads

$$Z_{\text{dimer}} = \underbrace{\exp\left(\frac{\beta J}{4} - \beta g\mu_B B\right) + \exp\left(\frac{\beta J}{4}\right) + \exp\left(\frac{\beta J}{4} + \beta g\mu_B B\right)}_{\text{triplet}} + \underbrace{\exp\left(-\frac{3}{4}\beta J\right)}_{\text{singlet}} \tag{6.44}$$

so that the magnetization for $S = 1/2$ is

$$\begin{aligned}
M_{\text{dimer}} &= -\frac{\partial F_{\text{dimer}}}{\partial B} = k_B T \frac{\partial}{\partial B} \ln Z_{\text{dimer}} \\
&= g\mu_B\left[-\exp\left(\frac{\beta J}{4} - \beta g\mu_B B\right) + \exp\left(\frac{\beta J}{4} + \beta g\mu_B B\right)\right],
\end{aligned} \tag{6.45}$$

which goes to zero for $B \to 0$, as expected.

More generally, in mean-field approximations one finds that T_c is proportional to the *strongest* interaction in the system, whereas the true T_c is governed by the *weakest* interactions that have to be taken into account to obtain a percolating model (i.e., all spins are directly or indirectly coupled).

6.3.2 The Hohenberg-Mermin-Wagner theorem

There is a strong rigorous result regarding the absence of long-range order in low-dimensional systems. A somewhat sloppy statement is that continuous symmetries are not spontaneously broken at nonzero temperature in one and two dimensions, provided the interactions are of short range. Considering special cases of this, Hohenberg showed that superfluidity does not exist at $T > 0$ in one and two dimensions and Mermin and Wagner proved the absence of long-range order in one- and two-dimensional spin systems. The proofs of both results use the same ideas and rely on an inequality found by Bogoliubov, which we shall derive first.

Let A, B be linear operators on a Hilbert space and H a hermitian operator on that space. We think of H as a Hamiltonian. Let $|n\rangle$ be the eigenstates of H to eigenvalues E_n. We assume the set $\{|n\rangle\}$ to be orthonormalized.

Define

$$(A, B) := \frac{1}{Z}\sum_{\substack{mn \\ E_m \neq E_n}} \langle n|A^\dagger|m\rangle \langle m|B|n\rangle \frac{e^{-\beta E_m} - e^{-\beta E_n}}{E_n - E_m} \tag{6.46}$$

with $Z := \sum_m e^{-\beta E_m}$. Note that $(e^{-\beta E_m} - e^{-\beta E_n})/(E_n - E_m)$ is positive for all $E_m \neq E_n$. Thus (A, B) is a *scalar product* on the space of linear operators. Furthermore,

$$
\begin{aligned}
\frac{e^{-\beta E_m} - e^{-\beta E_n}}{E_n - E_m} &= \frac{\dfrac{e^{-\beta E_m} - e^{-\beta E_n}}{e^{-\beta E_m} + e^{-\beta E_n}}}{\frac{\beta}{2} E_n - E_m} \frac{\beta}{2} \left(e^{-\beta E_m} + e^{-\beta E_n} \right) \\
&= \frac{\dfrac{e^{\beta/2(E_n - E_m)} - e^{-\beta/2(E_n - E_m)}}{e^{\beta/2(E_n - E_m)} + e^{-\beta/2(E_n - E_m)}}}{\frac{\beta}{2} E_n - E_m} \frac{\beta}{2} \left(e^{-\beta E_m} + e^{-\beta E_n} \right) \\
&= \underbrace{\frac{\tanh \frac{\beta}{2}(E_n - E_m)}{\frac{\beta}{2}(E_n - E_m)}}_{<1 \text{ for } E_n \neq E_m} \frac{\beta}{2} \left(e^{-\beta E_m} + e^{-\beta E_n} \right) < \frac{\beta}{2} \left(e^{-\beta E_m} + e^{-\beta E_n} \right).
\end{aligned} \tag{6.47}
$$

Thus

$$
\begin{aligned}
(A, A) &\leq \frac{1}{Z} \sum_{\substack{mn \\ E_m \neq E_n}} \langle n|A^\dagger|m \rangle \langle m|A|n \rangle \frac{\beta}{2} \left(e^{-\beta E_m} + e^{-\beta E_n} \right) \\
&\leq \frac{\beta}{2} \frac{1}{Z} \underbrace{\sum_{mn}}_{\text{unrestricted}} \langle n|A^\dagger|m \rangle \langle m|A|n \rangle \left(e^{-\beta E_m} + e^{-\beta E_n} \right) \\
&= \frac{\beta}{2} \left(\frac{1}{Z} \sum_m e^{-\beta E_m} \langle m|AA^\dagger|m \rangle + \frac{1}{Z} \sum_n e^{-\beta E_n} \langle n|AA^\dagger|n \rangle \right) \\
&= \frac{\beta}{2} \left\langle AA^\dagger + A^\dagger A \right\rangle.
\end{aligned} \tag{6.48}
$$

Since (A, B) is a scalar product, it satifies the Cauchy-Schwartz inequality

$$
|(A, B)|^2 \leq (A, A)(B, B) \tag{6.49}
$$

and thus

$$
|(A, B)|^2 \leq \frac{\beta}{2} \left\langle AA^\dagger + A^\dagger A \right\rangle (B, B). \tag{6.50}
$$

We choose a special form for B:

$$
B := [C^\dagger, H] \tag{6.51}
$$

where C is another linear operator. Then

$$
\begin{aligned}
(A, B) &= \left(A, [C^\dagger, H] \right) \\
&= \frac{1}{Z} \sum_{\substack{mn \\ E_m \neq E_n}} \langle n|A^\dagger|m \rangle \langle m|(C^\dagger H - H C^\dagger)|n \rangle \frac{e^{-\beta E_m} - e^{-\beta E_n}}{E_n - E_m} \\
&= \frac{1}{Z} \sum_{\substack{mn \\ E_m \neq E_n}} \langle n|A^\dagger|m \rangle \cancel{(E_n - E_m)} \langle m|C^\dagger|n \rangle \frac{e^{-\beta E_m} - e^{-\beta E_n}}{\cancel{E_n - E_m}} \\
&= \frac{1}{Z} \underbrace{\sum_{mn}}_{\text{unrestricted}} \left(e^{-\beta E_m} - e^{-\beta E_n} \right) \langle n|A^\dagger|m \rangle \langle m|C^\dagger|n \rangle \\
&= \left\langle C^\dagger A^\dagger - A^\dagger C^\dagger \right\rangle = \left\langle [C^\dagger, A^\dagger] \right\rangle
\end{aligned} \tag{6.52}
$$

and in particular

$$
(B, B) = \left\langle [C^\dagger, B^\dagger] \right\rangle = \left\langle [C^\dagger, [H, C]] \right\rangle. \tag{6.53}
$$

With Eq. (6.50) we obtain *Bogoliubov's inequality*

$$
\left| \left\langle [C^\dagger, A^\dagger] \right\rangle \right|^2 \leq \frac{\beta}{2} \left\langle AA^\dagger + A^\dagger A \right\rangle \left\langle [C^\dagger, [H, C]] \right\rangle \tag{6.54}
$$

for linear operators A, C, and a hermitian operator H.

We now formulate the Mermin-Wagner theorem cleanly: For the quantum Heisenberg model in one and two dimensions with Hamiltonian

$$H = -\frac{1}{2} \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j - B S_{\mathbf{q}}^z, \tag{6.55}$$

where

$$S_{\mathbf{q}}^z := \sum_i e^{i\mathbf{q}\cdot\mathbf{R}_i} S_i^z, \tag{6.56}$$

and with interactions that obey

$$\bar{J} := \frac{1}{2N} \sum_{ij} |J_{ij}||\mathbf{R}_i - \mathbf{R}_j|^2 < \infty, \tag{6.57}$$

there is no spontaneously broken spin symmetry at $T > 0$, i.e.,

$$\lim_{B \to 0^+} \lim_{N \to \infty} \frac{1}{N} \left\langle S_{\mathbf{q}}^z \right\rangle = 0. \tag{6.58}$$

The order of limits in this expression matters. (Note that the forms of spin order included here are not quite as general as the helical states studied earlier. The generalization is straightforward, though.)

Proof: We use Bogoliubov's inequality with

$$C = S_{\mathbf{k}}^x \quad \text{and} \quad A = S_{-\mathbf{k}-\mathbf{q}}^y. \tag{6.59}$$

Then

$$C^\dagger = S_{-\mathbf{k}}^x \quad \text{and} \quad A = S_{\mathbf{k}+\mathbf{q}}^y \tag{6.60}$$

and

$$\frac{1}{2} \left\langle AA^\dagger + A^\dagger A \right\rangle = \left\langle S_{\mathbf{k}+\mathbf{q}}^y S_{-\mathbf{k}-\mathbf{q}}^y \right\rangle =: N \, \mathcal{C}_{S^y S^y}(\mathbf{k}+\mathbf{q}), \tag{6.61}$$

where we have defined the Fourier-transformed spin-spin correlation function.

Furthermore,

$$\left\langle [C^\dagger, A^\dagger] \right\rangle = \left\langle [S_{-\mathbf{k}}^x, S_{\mathbf{k}+\mathbf{q}}^y] \right\rangle = i \left\langle S_{\mathbf{q}}^z \right\rangle. \tag{6.62}$$

The Bogoliubov inequality then reads

$$\left| \left\langle S_{\mathbf{q}}^z \right\rangle \right|^2 \leq \beta N \, \mathcal{C}_{S^y S^y}(\mathbf{k}+\mathbf{q}) \left\langle [S_{-\mathbf{k}}^x, [H, S_{\mathbf{k}}^x]] \right\rangle. \tag{6.63}$$

Herein,

$$\begin{aligned}
[H, S_{\mathbf{k}}^x] &= \sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i}[H, S_i^x] = -\frac{1}{2} \sum_{ijl} e^{i\mathbf{k}\cdot\mathbf{R}_i} J_{jl}[\mathbf{S}_j \cdot \mathbf{S}_l, S_i^x] - iBS_{\mathbf{k}+\mathbf{q}}^y \\
&= -\frac{1}{2} \sum_{jl} J_{jl} \left(e^{i\mathbf{k}\cdot\mathbf{R}_l} S_j^y(-i)S_l^z + e^{i\mathbf{k}\cdot\mathbf{R}_l} S_j^z iS_l^y + e^{i\mathbf{k}\cdot\mathbf{R}_j}(-i)S_j^z S_l^y + e^{i\mathbf{k}\cdot\mathbf{R}_j} iS_j^y S_l^z \right) - iBS_{\mathbf{k}+\mathbf{q}}^y \\
&= -\frac{i}{2} \sum_{jl} J_{jl} \left(e^{i\mathbf{k}\cdot\mathbf{R}_l} - e^{i\mathbf{k}\cdot\mathbf{R}_j} \right)(S_j^z S_l^y - S_j^y S_l^z) - iBS_{\mathbf{k}+\mathbf{q}}^y.
\end{aligned} \tag{6.64}$$

This implies

$$\begin{aligned}
\left[S_{-mk}^x, [H, S_{\mathbf{k}}^x] \right] &= \sum_i e^{-i\mathbf{k}\cdot\mathbf{R}_i} \left[S_i^x, [H, S_{\mathbf{k}}^x] \right] \\
&= -\frac{i}{2} \sum_{ijl} e^{-i\mathbf{k}\cdot\mathbf{R}_i} J_{jl} \left(e^{i\mathbf{k}\cdot\mathbf{R}_l} - e^{i\mathbf{k}\cdot\mathbf{R}_j} \right) \left[S_i^x, S_j^z S_l^y - S_j^y S_l^z \right] + BS_{\mathbf{q}}^z \\
&= -\frac{i}{2} \sum_{jl} J_{jl} \left(e^{i\mathbf{k}\cdot\mathbf{R}_l} - e^{i\mathbf{k}\cdot\mathbf{R}_j} \right) \left(e^{-i\mathbf{k}\cdot\mathbf{R}_l} S_j^z iS_l^z + e^{-i\mathbf{k}\cdot\mathbf{R}_j}(-i)S_j^y S_l^y + \right. \\
&\qquad \left. e^{-i\mathbf{k}\cdot\mathbf{R}_l} S_j^y(-i)S_l^y + e^{-i\mathbf{k}\cdot\mathbf{R}_j} iS_j^z S_l^z \right) + BS_{\mathbf{q}}^z \\
&= -\frac{1}{2} \sum_{jl} J_{jl} \left(e^{i\mathbf{k}\cdot(\mathbf{R}_l-\mathbf{R}_j)} - 1 - 1 + e^{i\mathbf{k}\cdot(\mathbf{R}_j-\mathbf{R}_l)} \right)(S_j^y S_l^y + S_j^z S_l^z) + BS_{\mathbf{q}}^z \\
&= \sum_{jl} J_{jl} \left(1 - \cos\mathbf{k}\cdot(\mathbf{R}_j - \mathbf{R}_l) \right)(S_j^y S_l^y + S_j^z S_l^z) + BS_{\mathbf{q}}^z.
\end{aligned} \tag{6.65}$$

49

We now note that $1 - \cos x \le x^2/2$ and, trivially, $\cos x - 1 \le 0 \le x^2/2$ so that

$$\langle [S^x_{-mk}, [H, S^x_{\mathbf{k}}]] \rangle = \sum_{jl} J_{jl} \left(1 - \cos \mathbf{k} \cdot (\mathbf{R}_j - \mathbf{R}_l) \right) \langle S^y_j S^y_l + S^z_j S^z_l \rangle + B \langle S^z_{\mathbf{q}} \rangle$$

$$\le \frac{k^2}{2} \sum_{jl} |J_{jl}| |\mathbf{R}_j - \mathbf{R}_l|^2 |\langle S^y_j S^y_l + S^z_j S^z_l \rangle| + B \langle S^z_{\mathbf{q}} \rangle. \tag{6.66}$$

Moreover, $|\langle S^y_j S^y_l + S^z_j S^z_l \rangle| \le |\langle \mathbf{S}_j \cdot \mathbf{S}_l \rangle| \le S^2$ (since $j \ne l$). Thus with Eq. (6.57)

$$\langle [S^x_{-mk}, [H, S^x_{\mathbf{k}}]] \rangle \le N k^2 \bar{J} S^2 + B \langle S^z_{\mathbf{q}} \rangle \tag{6.67}$$

and, putting everything together,

$$|\langle S^z_{\mathbf{q}} \rangle|^2 \le \beta N \, \mathcal{C}_{S^y S^y}(\mathbf{k} + \mathbf{q}) \left(N k^2 \bar{J} S^2 + B \langle S^z_{\mathbf{q}} \rangle \right). \tag{6.68}$$

Thus for the correlation function we find the lower bound

$$\mathcal{C}_{S^y S^y}(\mathbf{k} + \mathbf{q}) \ge \frac{1}{\beta N} \frac{|\langle S^z_{\mathbf{q}} \rangle|^2}{B \langle S^z_{\mathbf{q}} \rangle + N k^2 \bar{J} S^2}. \tag{6.69}$$

We sum this inequality over all $\mathbf{k}$ and note that

$$\sum_{\mathbf{k}} \mathcal{C}_{S^y S^y}(\mathbf{k} + \mathbf{q}) = \sum_{\mathbf{k}} \mathcal{C}_{S^y S^y}(\mathbf{k}) = \frac{1}{N} \sum_{\mathbf{k}} \langle S^y_{\mathbf{k}} S^y_{-\mathbf{k}} \rangle = \frac{1}{N} \sum_{ij} \underbrace{\sum_{\mathbf{k}} e^{i\mathbf{k}(\mathbf{R}_i - \mathbf{R}_j)}}_{N \delta_{ij}} \langle S^y_i S^y_j \rangle$$

$$= \sum_i \langle (S^y_i)^2 \rangle \le N S^2. \tag{6.70}$$

Thus we get

$$S^2 \ge \frac{1}{N} \sum_{\mathbf{k}} \mathcal{C}_{S^y S^y}(\mathbf{k} + \mathbf{q}) \ge \frac{1}{\beta N^2} \sum_{\mathbf{k}} \frac{|\langle S^z_{\mathbf{q}} \rangle|^2}{B \langle S^z_{\mathbf{q}} \rangle + N k^2 \bar{J} S^2}. \tag{6.71}$$

We now replace the sum over $\mathbf{k}$ by an integral. We will see that the theorem relies on the small-$\mathbf{k}$ contribution and it is therefore rather uncritical what we do at large $\mathbf{k}$. However, we have to restrict the integral at large $\mathbf{k}$ since $\mathbf{k}$ is from the first Brillouin zone. We introduce a cutoff Λ of the order of the diameter of the Brillouin zone.

Thus

$$S^2 \ge \frac{1}{\beta N^2} \int_{|\mathbf{k}| < \Lambda} \frac{d^d k}{\frac{(2\pi)^d}{V}} \frac{|\langle S^z_{\mathbf{q}} \rangle|^2}{B \langle S^z_{\mathbf{q}} \rangle + N \bar{J} S^2 k^2} = k_B T \frac{V}{N} \int_{|\mathbf{k}| < \Lambda} \frac{d^d k}{(2\pi)^d} \frac{|\frac{1}{N} \langle S^z_{\mathbf{q}} \rangle|^2}{B/N \langle S^z_{\mathbf{q}} \rangle + \bar{J} S^2 k^2}. \tag{6.72}$$

Now we consider the interesting cases of $d = 1, 2, 3$ separately.

(a) $d = 1$:

$$S^2 \ge k_B T \frac{V}{N} \frac{1}{2\pi} 2 \frac{\left(\frac{1}{N} \langle S^z_{\mathbf{q}} \rangle \right)^{3/2}}{\sqrt{B \bar{J} S^2}} \arctan \frac{\sqrt{\bar{J} S^2} \Lambda}{\sqrt{B/N \langle S^z_{\mathbf{q}} \rangle}}. \tag{6.73}$$

For small applied field, $B \to 0^+$, the arc tangent approaches $\pi/2$ so that

$$S^2 \ge k_B T \frac{V}{N} \frac{1}{2} \frac{\left(\frac{1}{N} \langle S^z_{\mathbf{q}} \rangle \right)^{3/2}}{\sqrt{B \bar{J} S^2}} \tag{6.74}$$

and

$$\frac{1}{N} \langle S^z_{\mathbf{q}} \rangle \le \left(\frac{2N}{k_B T V} \right)^{2/3} \bar{J}^3 S^2 B^{1/3}. \tag{6.75}$$

Since $1/T$ and $\bar{J}$ are finite by assumption, the magnetization indeed approaches zero for $B \to 0^+$, at least as fast as $B^{1/3}$.

(b) $d = 2$:

$$S^2 \geq k_B T \frac{V}{N} \frac{1}{2\pi} \int_0^\Lambda dk\, k \frac{\left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2}{B/N\langle S_{\mathbf{q}}^z\rangle + \bar{J}S^2 k^2}$$

$$= k_B T \frac{V}{N} \frac{1}{2\pi} \frac{1}{2\bar{J}S^2} \left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2 \ln\left(1 + \frac{\bar{J}S^2\Lambda^2}{B/N\langle S_{\mathbf{q}}^z\rangle}\right)$$

$$\geq k_B T \frac{V}{N} \frac{1}{4\pi\bar{J}S^2} \left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2 \ln\left(\frac{\bar{J}S^2\Lambda^2}{B/N\langle S_{\mathbf{q}}^z\rangle}\right)$$

$$= k_B T \frac{V}{N} \frac{1}{4\pi\bar{J}S^2} \left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2 \ln\left(\frac{\bar{J}S^2\Lambda^2}{\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle} \underbrace{-\ln B}_{\substack{=+|\ln B| \\ \text{since } B \text{ is assumed} \\ \text{to be small}}}\right)$$

$$\geq k_B T \frac{V}{N} \frac{1}{4\pi\bar{J}S^2} \left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2 |\ln B| \tag{6.76}$$

so that

$$\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle \leq \frac{\sqrt{4\pi\bar{J}S^2}}{\sqrt{k_B T}} \sqrt{\frac{N}{V}} \frac{1}{\sqrt{|\ln B|}}. \tag{6.77}$$

Again, for finite $1/T$ and $\bar{J}$ the magnetization vanishes for $B \to 0^+$. However, in 2D the bound is much weaker than in 1D.

(c) $d = 3$:

$$S^2 \geq k_B T \frac{V}{N} \frac{1}{2\pi^2} \int_0^\Lambda dk\, k^2 \frac{\left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2}{B/N\langle S_{\mathbf{q}}^z\rangle + \bar{J}S^2 k^2}$$

$$= k_B T \frac{V}{N} \frac{1}{2\pi^2} \frac{\left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2}{\bar{J}^2 S^2} \left(\bar{J}S\Lambda - \sqrt{B\bar{J}\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle} \arctan \frac{\sqrt{\bar{J}}S\Lambda}{\sqrt{B/N\langle S_{\mathbf{q}}^z\rangle}}\right), \tag{6.78}$$

which for $B \to 0^+$ converges to

$$S^2 \geq k_B T \frac{V}{N} \frac{1}{2\pi^2} \frac{\left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2}{2\pi^2 \bar{J}^2 S^2} \bar{J}S\Lambda = k_B T \frac{V}{N} \frac{1}{2\pi^2} \frac{\left|\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle\right|^2}{2\pi^2 \bar{J}S} \Lambda, \tag{6.79}$$

from which

$$\frac{1}{N}\langle S_{\mathbf{q}}^z\rangle \leq \frac{\pi\sqrt{2\bar{J}}S^{3/2}}{\sqrt{k_B T}} \sqrt{\frac{V}{N}} \frac{1}{\sqrt{\Lambda}}. \tag{6.80}$$

Thus the magnetization in 3D need not vanish for $B \to 0^+$. The Mermin-Wagner theorem does not make any statement on the existence of long-range order in 3D. Note that it is really the dimension that enters, not the coordination number z. Thus the theorem forbids long-range order for the triangular lattice with $z = 6$ but not for the simple cubic lattice also with $z = 6$.

Chapter 7

The paramagnetic phase of magnetic insulators

Beyond the cases discussed in the previous chapter for which the mean-field approximation fails, there is a regime where it does not make any meaningful prediction. This is the paramagnetic phase in the absence of an applied magnetic field. As we have seen, the mean-field Hamiltonian for the Heisenberg model reads

$$H_{\mathrm{MF}} = -\sum_i \mathbf{S}_i \cdot \sum_j J_{ij} \langle \mathbf{S}_j \rangle + \frac{1}{2} \sum_{ij} J_{ij} \langle \mathbf{S}_i \rangle \cdot \langle \mathbf{S}_j \rangle, \tag{7.1}$$

which for $T > T_c$ of course gives

$$H_{\mathrm{MF}} = 0. \tag{7.2}$$

Thus we have lost all information on the interaction. On the other hand, the thermal average of the exact Hamiltonian is

$$\langle H \rangle = -\frac{1}{2} \sum_{ij} J_{ij} \langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle, \tag{7.3}$$

which is genarally non-zero if there are non-vanishing spin correlations $\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle$. In the equilibrium state, we expect the correlations to be such that they reduce the energy, i.e., in the paramagnetic phase,

$$\langle H \rangle < H_{\mathrm{MF}} = 0. \tag{7.4}$$

In this chapter we discuss properties, in particular spin correlations, in the paramagnetic phase.

7.1   Spin correlations and susceptibility

We consider a Heisenberg model in an applied magnetic field,

$$H = \underbrace{-\frac{1}{2} \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j}_{H_0} - g\mu_B \sum_i \mathbf{B}_i \cdot \mathbf{S}_i. \tag{7.5}$$

In the present section, $\mathbf{B}_i$ may be non-uniform and noncollinear.

The generalized spin susceptibility is defined by

$$
\begin{aligned}
\chi_{ij}^{\alpha\beta} &:= \left. \frac{\partial M_i^\alpha}{\partial B_j^\beta} \right|_{\mathbf{B}=0} = - \left. \frac{\partial^2 F}{\partial B_i^\alpha \partial B_j^\beta} \right|_{\mathbf{B}=0} \\
&= k_B T \left. \frac{\partial^2}{\partial B_i^\alpha \partial B_j^\beta} \ln \mathrm{Tr}\, e^{-\beta H} \right|_{\mathbf{B}=0} = k_B T \left. \frac{\partial}{\partial B_i^\alpha} \frac{\mathrm{Tr}\,(-\beta g \mu_B S_j^\beta) e^{-\beta H}}{\mathrm{Tr}\, e^{-\beta H}} \right|_{\mathbf{B}=0} \\
&= k_B T \left(\frac{\mathrm{Tr}\, \beta^2 g^2 \mu_B^2 S_i^\alpha S_j^\beta e^{-\beta H}}{\mathrm{Tr}\, e^{-\beta H}} - \frac{\mathrm{Tr}\, \beta g \mu_B S_i^\alpha e^{-\beta H}}{\mathrm{Tr}\, e^{-\beta H}} \frac{\mathrm{Tr}\, \beta g \mu_B S_j^\beta e^{-\beta H}}{\mathrm{Tr}\, e^{-\beta H}} \right) \Bigg|_{\mathbf{B}=0} \\
&= \frac{g^2 \mu_B^2}{k_B T} \left(\langle S_i^\alpha S_j^\beta \rangle_0 - \langle S_i^\alpha \rangle_0 \langle S_j^\beta \rangle_0 \right),
\end{aligned} \tag{7.6}
$$

where the subscript "0" indicates that the thermal averages are to be taken for $\mathbf{B} = 0$, i.e., with H_0 only:

$$\langle A \rangle_0 := \frac{\text{Tr}\, A\, e^{-\beta H_0}}{\text{Tr}\, e^{-\beta H_0}}. \tag{7.7}$$

We call

$$\mathcal{C}_{ij}^{\alpha\beta} := \left\langle S_i^\alpha S_j^\beta \right\rangle_0 - \left\langle S_i^\alpha \right\rangle_0 \left\langle S_j^\beta \right\rangle_0 \tag{7.8}$$

the *spin correlation function*. We have found a relation between the susceptibility and the spin correlation function:

$$\chi_{ij}^{\alpha\beta} = \frac{g^2 \mu_B^2}{k_B T} \mathcal{C}_{ij}^{\alpha\beta}. \tag{7.9}$$

It is useful to generalize this result to time-dependent quantities, which leads to the fluctuation-dissipation theorem. We here only state the result.

For a small $\mathbf{B}$-field, which can now be time-dependent, applied to a system without spontaneous magnetic order, the magnetization is linear in the field,

$$M_\alpha(\mathbf{r}, t) = \int d^3 r'\, dt' \sum_\beta \chi_{\alpha\beta}(\mathbf{r}, \mathbf{r}', t, t')\, B_\beta(\mathbf{r}', t'). \tag{7.10}$$

This defines the time-dependent spin susceptibility. Note that causality implies $\chi_{\alpha\beta}(\mathbf{r}, \mathbf{r}', t, t') = 0$ for $t < t'$. Due to translational invariance in time, we can write

$$\chi_{\alpha\beta}(\mathbf{r}, \mathbf{r}', t, t') = 2i\, \Theta(t - t')\, \chi_{\alpha\beta}(\mathbf{r}, \mathbf{r}', t - t'), \tag{7.11}$$

where the factor $2i$ is conventional.

The *fluctuation-dissipation theorem* relates the imaginary part of the Fourier transform $\chi_{\alpha\beta}(\mathbf{r}, \mathbf{r}', \omega) = \chi_{\alpha\beta}'(\mathbf{r}, \mathbf{r}', \omega) + i\chi_{\alpha\beta}''(\mathbf{r}, \mathbf{r}', \omega)$ of $\chi_{\alpha\beta}(\mathbf{r}, \mathbf{r}', t)$ to the time-dependent spin correlation function,

$$\chi_{\alpha\beta}''(\mathbf{R}_i, \mathbf{R}_j, \omega) = \frac{g^2 \mu_B^2}{4} \frac{1 - e^{-\hbar\omega/k_B T}}{\hbar} \int_{-\infty}^{\infty} dt\, e^{-i\omega t} \left[\left\langle S_i^\alpha(t) S_j^\beta(0) \right\rangle_0 + \left\langle S_j^\beta(t) S_i^\alpha(0) \right\rangle_0 - 2\left\langle S_i^\alpha \right\rangle_0 \left\langle S_j^\beta \right\rangle_0 \right]. \tag{7.12}$$

For $\omega \to 0$ we recover the static result. This relates the full, dynamical spin correlation function to the observable dynamical susceptibility.

7.2   High-temperature expansion

The paramagnetic phase is generally realized at high temperatures. It is thus natural to consider $J/k_B T$ as a small parameter, where J is a measure of the exchange interaction. $J = \max J(\mathbf{q})$ would be a possible choice. We show in this section that the susceptibility and other observable quantities, such as the specific heat, can be obtained from an expansion in $\beta J = J/k_B T$. The corresponding method is called *high-temperature expansion* or, more specifically, *moment expansion*.

For simplicity, we consider a Heisenberg model in a uniform applied magnetic field,

$$H = -\frac{1}{2} \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j - g\mu_B B \sum_i S_i^z. \tag{7.13}$$

The partition function $Z = \text{Tr}\, e^{-\beta H}$ can be expanded in powers of β,

$$Z = \text{Tr} \sum_{n=0}^{\infty} \frac{1}{n!} (-\beta)^n H^n = \sum_{n=0}^{\infty} \frac{1}{n!} (-\beta)^n \,\text{Tr}\, H^n. \tag{7.14}$$

To understand the physical meaning of the expressions $\text{Tr}\, H^n$, we write the equilibrium average of H^n in the limit $T \to \infty$ or $\beta \to 0$ as

$$\langle H^n \rangle_\infty = \text{Tr}\, \rho_\infty H^n = \frac{\text{Tr}\, e^{-0H} H^n}{\text{Tr}\, e^{-0H}} = \frac{\text{Tr}\, H^n}{\text{Tr}\, \mathbf{1}}. \tag{7.15}$$

For N spins, the dimension of the Hilbert space is $(2S+1)^N$ so that $\text{Tr}\, \mathbf{1} = (2S+1)^N$ and

$$\text{Tr}\, H^n = (2S+1)^N \langle H^n \rangle_\infty. \tag{7.16}$$

Thus

$$Z = (2S+1)^N \sum_{n=0}^{\infty} \frac{1}{n!} (-\beta)^n \langle H^n \rangle_\infty . \tag{7.17}$$

The free energy is then

$$F = -\frac{1}{\beta} \ln Z = -\frac{N \ln(2S+1)}{\beta} - \frac{1}{\beta} \ln \sum_{n=0}^{\infty} \frac{1}{n!} (-\beta)^n \langle H^n \rangle_\infty$$

$$= -\frac{N \ln(2S+1)}{\beta} - \frac{1}{\beta} \ln \left[1 + \sum_{n=2}^{\infty} \frac{1}{n!} (-\beta)^n \langle H^n \rangle_\infty \right], \tag{7.18}$$

where we have used $\langle H \rangle_\infty = 0$. This follows from

$$\mathrm{Tr}\, H = -\frac{1}{2} \sum_{ij} J_{ij} \mathrm{Tr}\, \mathbf{S}_i \cdot \mathbf{S}_j - g\mu_B B \sum_i \mathrm{Tr}\, S_i^z$$

$$= -\frac{1}{2} \sum_{ij} J_{ij} (\mathrm{Tr}\, \mathbf{S}_i) \cdot (\mathrm{Tr}\, \mathbf{S}_j)(2S+1)^{(N-2)} - g\mu_B B \sum_i \mathrm{Tr}\, S_i^z (2S+1)^{(N-1)} = 0. \tag{7.19}$$

One can perform a second expansion of the logarithm and group terms of the same order in β together. This leads to

$$F = -\frac{N \ln(2S+1)}{\beta} - \frac{\beta}{2} \left[\langle H^2 \rangle_\infty - \frac{\beta}{3} \langle H^3 \rangle_\infty - \frac{\beta^2}{12} \left(\langle H^4 \rangle_\infty - 3 \langle H^2 \rangle_\infty^2 \right) + \dots \right]. \tag{7.20}$$

One can show that this is an expansion in terms of *cumulants*. Their significance is that terms of higher than first order in N cancel in the cumulants. For example, $\langle H^4 \rangle_\infty$ and $\langle H^2 \rangle_\infty^2$ are themselves proportional to N^2 to leading order, but the N^2 terms cancel in $\langle H^4 \rangle_\infty - 3 \langle H^2 \rangle_\infty^2$. The cancelation of higher-order terms in N is reasonable since the free energy should be proportional to N.

The uniform susceptibility can now be obtained from

$$\chi = -\left. \frac{\partial^2 F}{\partial B^2} \right|_{B=0}. \tag{7.21}$$

For the example of the nearest-neighbor ferromagnetic Heisenberg model one obtains

$$\chi = \frac{N g^2 \mu_B^2 S(S+1)}{3 k_B T} \left\{ 1 + \frac{zJS(S+1)}{3 k_B T} + \frac{zJ^2 S(S+1)}{6 k_B^2 T^2} \left[\frac{2}{3}(z-1)S(S+1) - \frac{1}{2} \right] + \dots \right\}, \tag{7.22}$$

where z is the coordination number of the lattice. (Note that higher-order terms depend on details of the lattice structure that cannot be expressed in terms of z alone.) There is a slight difference in the definition compared to the one we used for the Curie-Weiß Law, essentially a factor of volume.

We can analogously find the expansion for $1/\chi$,

$$\frac{1}{\chi} = \frac{3 k_B T}{N g^2 \mu_B^2 S(S+1)} \left\{ 1 - \frac{zJS(S+1)}{3 k_B T} + \frac{zJ^2 S(S+1)}{12 k_B^2 T^2} \left[1 + \frac{4}{3} S(S+1) \right] + \dots \right\}. \tag{7.23}$$

The leading term in $1/\chi$ or χ is the exact susceptibility for non-interacting spins. If we take the first two terms in $1/\chi$, we see that $1/\chi \to 0$ or $\chi \to \infty$ for

$$T \to \frac{zJS(S+1)}{3 k_B T}. \tag{7.24}$$

This equals the mean-field Curie temperature for this model:

$$J(\mathbf{q} = 0) \sum_{\Delta \mathbf{R}} e^0 J = zJ \tag{7.25}$$

and thus

$$T_c = \frac{J(0)S(S+1)}{3 k_B} = \frac{zJS(S+1)}{3 k_B}. \tag{7.26}$$

We thus recover the mean-field critical temperature from the expansion up to first order in βJ. Using this expression for T_c, we reobtain the Curie-Weiß Law for χ from the first two terms in Eq. (7.23).

Going to order $(\beta J)^2$, i.e., including the first three terms, we find that the divergence occurs for $T = T_c$ with

$$1 - \frac{zJS(S+1)}{3k_B T_c} + \frac{zJ^2 S(S+1)}{12k_B^2 T_c^2}\left[1 + \frac{4}{3}S(S+1)\right] \cong 0. \tag{7.27}$$

This is a quadratic equation for $k_B T_c$. The physical solution is the larger one since we expand about $k_B T = \infty$. It is

$$\frac{J}{k_B T_c} \cong \frac{\frac{1}{2}}{1 + \frac{4}{3}S(S+1)}\left[1 - \sqrt{1 - 3\frac{1 + \frac{4}{3}S(S+1)}{zS(S+1)}}\right]. \tag{7.28}$$

For spin $S = 1/2$ on the fcc lattice ($z = 12$) this gives

$$k_B T_c \cong 2.37\,J. \tag{7.29}$$

Compare the mean-field result from Eq. (7.26),

$$k_B T_c^{MF} = \frac{12 \times 3/4}{3}\,J = 3J. \tag{7.30}$$

The correction thus reduces T_c. If more and more orders in βJ are taken into account, the result for T_c approaches the (not analytically known) exact value.

High-temperature series for the susceptibility, the specific heat, etc. have been calculated to high orders. We will now discuss how one might extract high-precision values of T_c and of the critical exponents—for the example of the susceptibility exponent γ—from such an expansion. The expansion of χ in units of the Curie susceptibility of non-interacting spins takes the form

$$\frac{3k_B T}{Ng^2\mu_B^2 S(S+1)}\chi = \sum_{n=0}^{\infty} a_n(S)\,K^n, \tag{7.31}$$

where the $a_n(S)$ are functions of the spin quantum number S only and

$$K := \frac{JS^2}{k_B T}. \tag{7.32}$$

For high orders n, thr ratio of coefficients a_n/a_{n-1} assumes the form

$$\frac{a_n}{a_{n-1}} = \frac{1}{K_c}\left[1 + \frac{\gamma - 1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right]. \tag{7.33}$$

Thus $\lim_{n\to\infty}(a_n/a_{n-1}) = 1/K_c$, where K_c is the radius of convergence of the series. $\gamma - 1$ is a constant, which we have given a peculiar name for later convenience.

For large n, the higher-order terms in a_n/a_{n-1} become irrelevant. The large-n form

$$\frac{a_n}{a_{n-1}} \cong \frac{1}{K_c}\left(1 + \frac{\gamma - 1}{n}\right) \tag{7.34}$$

is in fact satisfied by the expansion coefficients of the function

$$f(K) = \frac{A}{(1 - K/K_c)^\gamma}. \tag{7.35}$$

The proof is straightforward: we expand $f(K)$ in a Taylor series,

$$f(K) = \sum_{n=0}^{\infty}\frac{1}{n!}f^{(n)}(0)\,K^n \overset{!}{=} \sum_{n=0}^{\infty} a_n K^n \tag{7.36}$$

and obtain for the ratio

$$\begin{aligned}
\frac{a_n}{a_{n-1}} &= \frac{1}{n}\frac{f^{(n)}(0)}{f^{(n-1)}(0)} \\
&= \frac{1}{n}\left.\frac{A(1/K_c)^n\,\gamma(\gamma+1)\ldots(\gamma+n-1)\dfrac{1}{(1-K/K_c)^{\gamma+n}}}{A(1/K_c)^{n-1}\,\gamma(\gamma+1)\ldots(\gamma+n-2)\dfrac{1}{(1-K/K_c)^{\gamma+n-1}}}\right|_{K=0} \\
&= \frac{1}{nK_c}(\gamma+n-1) = \frac{1}{K_c}\left(1 + \frac{\gamma-1}{n}\right).
\end{aligned} \tag{7.37}$$

Therefore, we find

$$\frac{3k_BT}{Ng^2\mu_B^2S(S+1)}\chi = \frac{A}{(1-K/K_c)^\gamma} + \Delta\chi = \frac{A}{(1-T_c/T)^\gamma} + \Delta\chi = \frac{AT^\gamma}{(T-T_c)^\gamma} + \Delta\chi$$

$$= \frac{AT_c^\gamma}{(T-T_c)^\gamma} + \Delta\widetilde{\chi}, \tag{7.38}$$

where the functions $\Delta\chi$ and $\Delta\widetilde{\chi}$ do not diverge for $T \to T_c^+$ or diverge less strongly. We have made the leading divergence explicit. By calculating a_n for a few large n, we can obtain $K_c = JS^2/k_BT_c$ and the critical exponent γ.

High-temperature expansion of this type, albeit with some added improvements, leads to $k_BT_c \approx 2.01\,J$, for our case of $S = 1/2$ and $z = 12$. See Baker *et al.*, Phys. Rev. B **2**, 706 (1970). This result can again be compared the mean-field prediction $k_BT_c^{\mathrm{MF}} = 3J$.

Chapter 8

Excitations in the ordered state: magnons and spinons

In the magnetically ordered phase, the high-temperature expansion method of Ch. 7 does not work. On the other hand, the mean-field approximation is too crude to describe many phenomena correctly. In the present chapter, we will consider low-energy excitations over an ordered ground state. These are expected to dominate the physical properties at low temperatures.

8.1 Ferromagnetic spin waves and magnons

The ground state of a Heisenberg ferromagnet with only non-negative exchange interactions, $J_{ij} \geq 0$, is completely aligned. We are now interested in the low-energy excitations. An obvious candidate excited state is a state with one spin flipped or, more generally, reduced by one:

$$|\psi_1\rangle = |S, \ldots, S, S-1, S, \ldots, S\rangle. \tag{8.1}$$

This is not an eigenstate since the $S_i^+ S_j^- + S_i^- S_j^+$ terms in the Hamiltonian shift the flipped spin to neighboring sites. This is not very critical in itself—the state could still be a good approximation.

However, we can see that $|\psi_1\rangle$ has a rather high energy relative to the ground state,

$$\Delta E_1 := \langle\psi_1|H|\psi_1\rangle - \langle\psi_0|H|\psi_0\rangle$$

$$= -\sum_{j\neq 0} J_{0j} \langle\psi_1| \left(\frac{S_0^+ S_j^- + S_0^- S_j^+}{2} + S_0^z S_j^z \right) |\psi_1\rangle + \sum_{j\neq 0} J_{0j} \langle\psi_0| \left(\frac{S_0^+ S_j^- + S_0^- S_j^+}{2} + S_0^z S_j^z \right) |\psi_0\rangle, \tag{8.2}$$

where we have assumed that the reduced spin is at site $i = 0$. This gives

$$\Delta E_1 := -\sum_{j\neq 0} J(\mathbf{R}) \langle\psi_1|S_0^z S_j^z|\psi_1\rangle + \sum_{j\neq 0} J(\mathbf{R}) \langle\psi_0|S_0^z S_j^z|\psi_0\rangle$$

$$= -\sum_{j\neq 0} J(\mathbf{R})(S-1)S + \sum_{j\neq 0} J(\mathbf{R})S^2 = S \sum_{\mathbf{R}=0} = S J(\mathbf{q}=0). \tag{8.3}$$

For the nearest-neighbor model, this is $\Delta E_1 = zJS$. This is a large energy of the order of $k_B T_C$. We will see that lower-energy excitations exist.

8.1.1 Bloch spin-wave theory

To find lower-energy excitations, we write down the equation for the spin operator $\mathbf{S}_i$ in the Heisenberg picture,

$$\frac{d\mathbf{S}_i}{dt} = \frac{i}{\hbar}[H, \mathbf{S}_i], \tag{8.4}$$

where we take H to be the Heisenberg Hamiltonian in a uniform field,

$$H = -\frac{1}{2}\sum_{ij} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j - g\mu_B B \sum_i S_i^z. \tag{8.5}$$

This gives

$$\frac{d\mathbf{S}_i}{dt} = -\frac{1}{\hbar} \mathbf{H}(\mathbf{R}_i) \times \mathbf{S}_i \tag{8.6}$$

with

$$\mathbf{H}(\mathbf{R}_i) := \sum_j J_{ij}\mathbf{S}_j + g\mu_B B\hat{\mathbf{z}}. \tag{8.7}$$

This is still exact since we have not replaced $\mathbf{H}$ by its average, which would be $g\mu_B\mathbf{B}_{\text{eff}}$, see Sec. 6.1. Writing Eq. (8.6) in components, we obtain

$$\hbar\frac{dS_i^x}{dt} = -\sum_j J_{ij}(S_j^y S_i^z - S_j^z S_i^y) + g\mu_B B S_i^y, \tag{8.8}$$

$$\hbar\frac{dS_i^y}{dt} = -\sum_j J_{ij}(S_j^z S_i^x - S_j^x S_i^z) - g\mu_B B S_i^x, \tag{8.9}$$

$$\hbar\frac{dS_i^z}{dt} = -\sum_j J_{ij}(S_j^x S_i^y - S_j^y S_i^x). \tag{8.10}$$

Note that these equations also make sense for a purely classical model (by restoring the units of the spin and absorbing $\hbar$ into J_{ij} and μ_B we obtain equations without the quantum parameter $\hbar$).

Since we are interested in low-energy excitations, we expect the state of the system to be close to the fully polarized ground state. What this actually means is defined by the approximation we are making: we replace S_i^z in products of two spin operators by S. This gives for the first two equations

$$\hbar\frac{dS_i^x}{dt} = -S\sum_j J_{ij}(S_j^y - S_i^y) + g\mu_B B S_i^y, \tag{8.11}$$

$$\hbar\frac{dS_i^y}{dt} = -S\sum_j J_{ij}(S_i^x - S_j^x) - g\mu_B B S_i^x. \tag{8.12}$$

To decouple these equations, we introduce $S_i^{\pm} := S_i^x \pm S_i^y$ and obtain

$$\hbar\frac{dS_i^{\pm}}{dt} = \mp i\left[S\sum_j J_{ij}(S_i^{\pm} - S_j^{\pm}) + g\mu_B B S_i^{\pm}\right]. \tag{8.13}$$

We assume for simplicity that the sites $\mathbf{R}_i$ form a Bravais lattice. With the Fourier transform

$$S_{\mathbf{q}}^- = \frac{1}{\sqrt{N}}\sum_i e^{-i\mathbf{q}\cdot\mathbf{R}_i}S_i^-, \tag{8.14}$$

$$S_i^- = \frac{1}{\sqrt{N}}\sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{R}_i}S_{\mathbf{q}}^- \tag{8.15}$$

we find

$$\hbar\frac{dS_{\mathbf{q}}^-}{dt} = iS\frac{1}{N}\sum_{ij} e^{-i\mathbf{q}\cdot\mathbf{R}_i}J_{ij}\sum_{\mathbf{q}'}\left(e^{i\mathbf{q}'\cdot\mathbf{R}_i} - e^{i\mathbf{q}'\cdot\mathbf{R}_j}\right)S_{\mathbf{q}'}^- + ig\mu_B B S_{\mathbf{q}}^-$$

$$= iS\frac{1}{N}\sum_{i,\Delta\mathbf{R}} J_{\Delta\mathbf{R}}\sum_{\mathbf{q}'} e^{-i(\mathbf{q}-\mathbf{q}')\cdot\mathbf{R}_i}\left(1 - e^{-i\mathbf{q}'\cdot\Delta\mathbf{R}}\right)S_{\mathbf{q}'}^- + ig\mu_B B S_{\mathbf{q}}^-, \tag{8.16}$$

where $\Delta\mathbf{R} := \mathbf{R}_i - \mathbf{R}_j$ and $J_{\Delta\mathbf{R}} := J_{ij}$ (we assume, as usual, that J_{ij} only depends on the separation vector). Then we get

$$\hbar\frac{dS_{\mathbf{q}}^-}{dt} = iS\frac{1}{N}\sum_{\Delta\mathbf{R}} J_{\Delta\mathbf{R}}\sum_{\mathbf{q}'} N\delta_{\mathbf{q}\mathbf{q}'}\left(1 - e^{-i\mathbf{q}\cdot\Delta\mathbf{R}}\right)S_{\mathbf{q}}^- + ig\mu_B B S_{\mathbf{q}}^-$$

$$= iS\sum_{\Delta\mathbf{R}} J_{\Delta\mathbf{R}}\left(1 - e^{-i\mathbf{q}\cdot\Delta\mathbf{R}}\right) + ig\mu_B B S_{\mathbf{q}}^-. \tag{8.17}$$

Defining

$$J(\mathbf{q}) := \sum_{\Delta\mathbf{R}} J_{\Delta\mathbf{R}}\, e^{-i\mathbf{q}\cdot\Delta\mathbf{R}}, \tag{8.18}$$

we obtain

$$\hbar\frac{dS_{\mathbf{q}}^-}{dt} = i\left[J(0) - J(\mathbf{q})\right]S S_{\mathbf{q}}^- + ig\mu_B B S_{\mathbf{q}}^- = i\left(\left[J(0) - J(\mathbf{q})\right]S + g\mu_B B\right)S_{\mathbf{q}}^-. \tag{8.19}$$

This has the obvious solutions

$$S_{\mathbf{q}}^-(t) = M_{\mathbf{q}} \, e^{i\omega_{\mathbf{q}}t + i\alpha_{\mathbf{q}}} \tag{8.20}$$

with

$$\hbar\omega_{\mathbf{q}} = [J(0) - J(\mathbf{q})]S + g\mu_B B \tag{8.21}$$

and an arbitrary phase $\alpha_{\mathbf{q}}$. If only one mode $S_{\mathbf{q}}^-$ is excited, the excitation in real space is

$$S_i^-(t) = \frac{1}{\sqrt{N}} e^{i\mathbf{q}\cdot\mathbf{R}_i} S_{\mathbf{q}}^- = \frac{M_{\mathbf{q}}}{\sqrt{N}} e^{i(\mathbf{q}\cdot\mathbf{R}_i + \omega_{\mathbf{q}}t + \alpha_{\mathbf{q}})}. \tag{8.22}$$
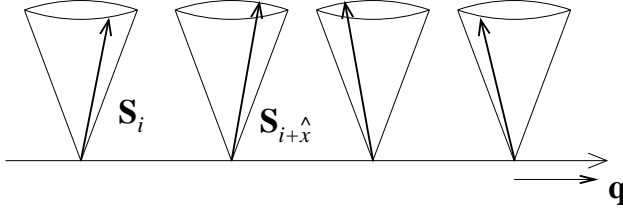
The physical components are then

$$S_i^x(t) = \frac{M_{\mathbf{q}}}{\sqrt{N}} \cos(\mathbf{q}\cdot\mathbf{R}_i + \omega_{\mathbf{q}}t + \alpha_{\mathbf{q}}), \tag{8.23}$$

$$S_i^y(t) = -\frac{M_{\mathbf{q}}}{\sqrt{N}} \sin(\mathbf{q}\cdot\mathbf{R}_i + \omega_{\mathbf{q}}t + \alpha_{\mathbf{q}}), \tag{8.24}$$

$$S_i^z(t) \cong S. \tag{8.25}$$

$S_i^{x,y}$ have the form of plane waves. Note that the x- and y-components are out of phase by $\pi/2$ or $90°$. Classically, $\mathbf{S}_i(t)$ can be said to precess on a cone.

In the classical limit, this kind of excitations is called a *spin wave*. Its frequency is evidently

$$\omega_{\mathbf{q}} = \frac{J(0) - J(\mathbf{q})}{\hbar} S + \frac{g\mu_B}{\hbar} B. \tag{8.26}$$

For example, in the absense of an applied field and for nearest-neighbor exchange on the simple cubic lattice, we have

$$J(\mathbf{q}) = \sum_{\text{NN } \Delta\mathbf{R}} J e^{-i\mathbf{q}\cdot\Delta\mathbf{R}} = 2J \left(\cos q_x a + \cos q_y a + \cos q_z a\right), \tag{8.27}$$

where a is the lattice constant. Thus

$$\omega_{\mathbf{q}} = \frac{2JS}{\hbar} \left(3 - \cos q_x a - \cos q_y a - \cos q_z a\right). \tag{8.28}$$

For small $\mathbf{q}$, $\omega_{\mathbf{q}}$ approaches zero like

$$\omega_{\mathbf{q}} \cong \frac{2JS}{\hbar} \left(3 - 1 + \frac{1}{2}q_x^2 a^2 - 1 + \frac{1}{2}q_y^2 a^2 - 1 + \frac{1}{2}q_z^2 a^2\right) = \frac{JS}{\hbar}q^2 a^2. \tag{8.29}$$

This behavior, $\omega_{\mathbf{q}} \propto q^2$ is generic for ferromagnets. It is different from lattice vibrations in crystals, which have $\omega_{\mathbf{q}} \propto q$. We see that the group velocity approaches zero for $\mathbf{q} \to 0$ and is thus not very useful for the characterization of ferromagnetic spin waves. Instead one uses the *spin-wave stiffness* ρ defined by

$$\omega_{\mathbf{q}} = \rho q^2 \qquad \text{for small } \mathbf{q}, \tag{8.30}$$

i.e., here $\rho = JSa^2/\hbar$. Note that there are other definitions that differ by q-independent factors.

Recall that S_i^- and $S_{\mathbf{q}}^-$ are really operators in the Heisenberg picture. From

$$\frac{dS_i^-}{dt} = \frac{i}{\hbar}[H, S_i^-] \tag{8.31}$$

we immediately obtain

$$\frac{dS_{\mathbf{q}}^-}{dt} = \frac{i}{\hbar}[H, S_{\mathbf{q}}^-]. \tag{8.32}$$

Applying this operator to the fully polarized ground state $|\psi_0\rangle$, we obtain

$$\hbar\omega_{\mathbf{q}}S_{\mathbf{q}}^-|\psi_0\rangle = [H, S_{\mathbf{q}}^-]|\psi_0\rangle = HS_{\mathbf{q}}^-|\psi_0\rangle - S_{\mathbf{q}}E_0|\psi_0\rangle, \tag{8.33}$$

which implies

$$HS_{\mathbf{q}}^-|\psi_0\rangle = (E_0 + \hbar\omega_{\mathbf{q}})S_{\mathbf{q}}^-|\psi_0\rangle. \tag{8.34}$$

Thus $S_{\mathbf{q}}^-|\psi_0\rangle$ is an eigenstate to H with eigenenergy $E_0 + \hbar\omega_{\mathbf{q}}$, at least in this approximation. Hence, $S_{\mathbf{q}}^-$ creates one quantum of the spin wave, called a *magnon*, with the energy $\hbar\omega_{\mathbf{q}}$.

Now note that $\hbar\omega_{\mathbf{q}=0} = 0$. The excitation of the $\mathbf{q} = 0$ magnon mode thus does not cost any energy. This has to be so since it corresponds to a uniform spin rotation, under which the Hamiltonian is invariant. This result is a special case of a general statement: the spontaneous breaking of a continuous symmetry, here spin rotation, is always accompanied by the appearance of a zero-energy mode. This statement is the *Goldstone theorem* and the zero-energy mode is called a *Goldstone mode*.

One can show that $S_{\mathbf{q}}^-|\psi_0\rangle$ is in fact rigorously an eigenstate of H with energy $\hbar\omega_{\mathbf{q}}$. If more than one magnon is excited, states like $S_{\mathbf{q}}^- S_{\mathbf{q}'}^-|\psi_0\rangle$ are no longer rigorously eigenstates, but they are eigenstates in the linear approximation introduced by setting $S_i^z \cong S$ above. Thus this approximation is equivalent to assuming non-interacting magnons. We come back to this point below.

8.1.2 Equilibrium properties at low temperatures in three dimensions

At not too high temperatures, the assumption of non-interacting magnons is well justified. The average number of magnons with wave-vector $\mathbf{q}$ is then

$$\begin{aligned}
\langle n_{\mathbf{q}}\rangle &= \frac{\sum_{n=0}^\infty n\exp(-\beta n\hbar\omega_{\mathbf{q}})}{\sum_{n=0}^\infty \exp(-\beta n\hbar\omega_{\mathbf{q}})} = -\frac{1}{\hbar\omega_{\mathbf{q}}}\frac{\partial}{\partial\beta}\ln\sum_n \exp(-\beta n\hbar\omega_{\mathbf{q}}) \\
&= -\frac{1}{\hbar\omega_{\mathbf{q}}}\frac{\partial}{\partial\beta}\ln\frac{1}{1-\exp(-\beta\hbar\omega_{\mathbf{q}})} = -\frac{1}{\hbar\omega_{\mathbf{q}}}\frac{\partial}{\partial\beta}\ln[1-\exp(-\beta\hbar\omega_{\mathbf{q}})] \\
&= \frac{\exp(-\beta\hbar\omega_{\mathbf{q}})}{1-\exp(-\beta\hbar\omega_{\mathbf{q}})} = \frac{1}{\exp(\beta\hbar\omega_{\mathbf{q}})-1} \equiv n_B(\hbar\omega_{\mathbf{q}}).
\end{aligned} \tag{8.35}$$

This is the Bose-Einstein distribution function showing that magnons behave statistically like bosons.

The total number of excited magnons is then

$$\langle n\rangle = \sum_{\mathbf{q}}\langle n_{\mathbf{q}}\rangle = V\int\frac{d^3q}{(2\pi)^3}\frac{1}{\exp(\beta\hbar\omega_{\mathbf{q}})-1} \tag{8.36}$$

in the thermodynamic limit, where V is the volume of the system. Since each magnon reduces the total spin by one, the magnetization is

$$\langle M\rangle = \frac{g\mu_B(NS - \langle n\rangle)}{V} = \underbrace{\frac{g\mu_B NS}{V}}_{=M_{\text{sat}}} - g\mu_B\int\frac{d^3q}{(2\pi)^3}\frac{1}{\exp(\beta\hbar\omega_{\mathbf{q}})-1}. \tag{8.37}$$

M_{sat} is the saturation magnetization.

We now consider the example of a nearest-neighbor model on the simple cubic lattice. The integral is expected to be dominated by small $\mathbf{q}$, since the integrand diverges for $\mathbf{q} \to 0$. It is then justified to extend the integral, which is really over the Brillouin zone, to infinite $\mathbf{q}$-space and to replace $\omega_{\mathbf{q}}$ by its small-$\mathbf{q}$ limit:

$$\begin{aligned}
\langle M\rangle &\cong M_{\text{sat}} - g\mu_B\int_{\mathbb{R}^3}\frac{d^3q}{(2\pi)^3}\frac{1}{\exp(\beta JSa^2q^2)-1} = M_{\text{sat}} - \frac{g\mu_B}{2\pi^2}\int_0^\infty dq\,\frac{q^2}{\exp(\beta JSa^2q^2)-1} \\
&= M_{\text{sat}} - g\mu_B\frac{\zeta(3/2)}{8}\frac{1}{(\pi\beta JSa^2)^{3/2}},
\end{aligned} \tag{8.38}$$

where $\zeta(x)$ is the zeta function and $\zeta(3/2) \approx 2.612$. We find that the magnetization at low T starts to deviate from its maximum value like $T^{3/2}$. This result is called the *Bloch law*. Note that mean-field theory had incorrectly predicted an exponentially small deviation, see Sec. 6.1. A similar derivation leads to a $T^{3/2}$ behavior of the specific heat at low temperatures, also in contradiction to mean-field theory.

8.1.3 The infrared catastrophy in one and two dimensions

Under the assumption of non-interacting magnons, we can perform the derivation for $\langle n \rangle$ in any dimension d. This results in

$$\langle n \rangle = \sum_{\mathbf{q}} \langle n_{\mathbf{q}} \rangle = V_d \int \frac{d^d q}{(2\pi)^d} \frac{1}{\exp(\beta \hbar \omega_{\mathbf{q}}) - 1}, \tag{8.39}$$

where V_d is the generalized d-dimensional volume. Arguing as above, we find

$$\langle M \rangle \cong M_{\text{sat}} - g\mu_B \int \frac{d^d q}{(2\pi)^d} \frac{1}{\exp(\beta J S a^2 q^2) - 1} = M_{\text{sat}} - g\mu_B \frac{\Omega_d}{(2\pi)^d} \int_0^\infty dq \frac{q^{d-1}}{\exp(\beta J S a^2 q^2) - 1}, \tag{8.40}$$

where Ω_d is the surface of the d-dimensional unit sphere ($\Omega_1 = 2, \Omega_2 = 2\pi, \Omega_3 = 4\pi, \dots$). At the lower limit $q \to 0$, the integrand behaves like

$$\frac{q^{d-1}}{\exp(\beta J S a^2 q^2) - 1} \cong \frac{1}{\beta J S a^2} \frac{q^{d-1}}{q^2} = \frac{1}{\beta J S a^2} q^{d-3}. \tag{8.41}$$

Introducing a lower cutoff Λ into the integral, we find

$$\int_\Lambda dq\, q^{d-3} \sim \begin{cases} \Lambda^{d-2} & \text{for } d \neq 2 \\ \ln \Lambda & \text{for } d = 2. \end{cases} \tag{8.42}$$

Thus for $\Lambda \to 0$ we obtain a logarithmic divergence for $d = 2$ and a stronger $1/\Lambda$ divergence for $d = 1$.

For non-interacting magnons, infinitely many of them are thermally excited in one and two dimensions for any $T > 0$. The result $\langle M \rangle \to -\infty$ is of course unphysical, since the true magnetization is bounded. The approximation of non-interacting magnons fails, but the results give the correct idea: thermal fluctuations destroy the magnetic order. This is the physics behind the Mermin-Wagner theorem.

8.2 Magnon-magnon interaction

The previously discussed approach is not the most convenient for the description of magnon interactions. These are naturally included in the *Holstein-Primakoff bosonization* scheme. We introduce a single bosonic mode for every spin, unlike in the Schwinger scheme, where we needed two. Let

$$S_i^- = \sqrt{2S}\, a_i^\dagger \sqrt{1 - \frac{a_i^\dagger a_i}{2S}}, \tag{8.43}$$

$$S_i^+ = \sqrt{2S} \sqrt{1 - \frac{a_i^\dagger a_i}{2S}}\, a_i, \tag{8.44}$$

$$S_i^z = S - a_i^\dagger a_i. \tag{8.45}$$

One can show that the spin commutation relations for $[S_i^\alpha, S_i^\beta]$ are satisfied in this representation and that $\mathbf{S}_i \cdot \mathbf{S}_i = S(S+1)\,\mathbf{1}$. The number operators $a_i^\dagger a_i$ of course have the eigenvalues $0, 1, 2, \dots$, whereas S_i^z must only have the eigenvalues $-S, -S+1, \dots, S$. Thus we have to impose a *constraint*

$$a_i^\dagger a_i \leq 2S. \tag{8.46}$$

This is an inequality, i.e., a non-holonomic constraint, which makes it much more difficult to handle than an equality (a holonomic constraint) would be. Recall the Schwinger-boson scheme: there we had a holonomic constraint but needed two boson species.

Note that we have, for any spin $\mathbf{S}_i$,

$$S_i^- \,|m = -S\rangle_i = \sqrt{2S}\, a_i^\dagger \sqrt{1 - \frac{a_i^\dagger a_i}{2S}} \,|n = 2S\rangle_i = \sqrt{2S}\, a_i^\dagger \sqrt{1 - \frac{2S}{2S}} \,|n = 2S\rangle_i = 0 \tag{8.47}$$

and

$$S_i^+ \,|m = -S-1\rangle_i = \sqrt{2S}\sqrt{1 - \frac{a_i^\dagger a_i}{2S}}\, a_i \,|n = 2S+1\rangle_i = \sqrt{2S}\sqrt{1 - \frac{a_i^\dagger a_i}{2S}} \sqrt{2S+1} \,|n = 2S\rangle_i$$

$$= \sqrt{2S}\sqrt{1 - \frac{2S}{2S}} \sqrt{2S+1} \,|n = 2S\rangle_i = 0. \tag{8.48}$$

Thus $S_i^{\pm}$ do not connect the physical subspace ($n \le 2S$) and the unphasical one ($n \ge 2S+1$).

We see that the vacuum of the Hostein-Primakoff bosons satisfies

$$a_i^{\dagger} a_i |0\rangle_i = 0 \quad \Rightarrow \quad S_i^z |0\rangle_i = S|0\rangle_i \tag{8.49}$$

and is thus the fully polarized ferromagnetic ground state. Holstein-Primakoff bosons are therefore better suited for expansions around this polarized state, while Schwinger bosons are better suited in the paramagnetic phase. Both schemes are exact, though.

The roots of operators are rather inconvenient and the practical usefulness of the scheme lies in the expansion of these roots in orders of $a_i^{\dagger} a_i / 2S$. For example, the nearest-neighbor Hamiltonian

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J \sum_{\langle ij \rangle} \left(\frac{S_i^+ S_j^- + S_i^- S_j^+}{2} + S_i^z S_j^z \right)$$

$$= -J \sum_{\langle ij \rangle} \left[S \sqrt{1 - \frac{a_i^{\dagger} a_i}{2S}} a_i a_j^{\dagger} \sqrt{1 - \frac{a_j^{\dagger} a_j}{2S}} + S a_i^{\dagger} \sqrt{1 - \frac{a_i^{\dagger} a_i}{2S}} \sqrt{1 - \frac{a_j^{\dagger} a_j}{2S}} a_j + (S - a_i^{\dagger} a_i)(S - a_j^{\dagger} a_j) \right] \tag{8.50}$$

is expanded as

$$H = -Nz \frac{JS^2}{2} + JS \sum_{\langle ij \rangle} \left(a_i^{\dagger} a_i + a_j^{\dagger} a_j - a_i^{\dagger} a_j - a_j^{\dagger} a_i \right)$$

$$- J \sum_{\langle ij \rangle} \left[a_i^{\dagger} a_i a_j^{\dagger} a_j - \frac{1}{4} \left(a_i^{\dagger} a_i^{\dagger} a_i a_j + a_i^{\dagger} a_j^{\dagger} a_j a_j + a_j^{\dagger} a_i^{\dagger} a_i a_i + a_j^{\dagger} a_j^{\dagger} a_j a_i \right) \right] + \mathcal{O}\left(\frac{1}{S} \right). \tag{8.51}$$

This is an expansion in $1/S$, starting with the order $(1/S)^{-2}$. This suggests that in the classical limit, $S \to \infty$, we can get away with keeping only the first two terms (keeping only the first is too crude since it is a constant energy).

Keeping only terms up to the order $(1/S)^{-1} = S^1$, we can diagonalize H by a Fourier transformation,

$$a_i = \frac{1}{\sqrt{N}} \sum_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{R}_i} a_{\mathbf{q}}. \tag{8.52}$$

This gives

$$H = -Nz \frac{JS^2}{2} + \frac{JS}{N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{\langle ij \rangle} \left(e^{-i\mathbf{q}\mathbf{R}_i + i\mathbf{q}'\mathbf{R}_i} + e^{-i\mathbf{q}\mathbf{R}_j + i\mathbf{q}'\mathbf{R}_j} - e^{-i\mathbf{q}\mathbf{R}_i + i\mathbf{q}'\mathbf{R}_j} - e^{-i\mathbf{q}\mathbf{R}_j + i\mathbf{q}'\mathbf{R}_i} \right) a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}'}$$

$$= -Nz \frac{JS^2}{2} + \frac{JS}{2N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{i} \sum_{\Delta \mathbf{R}} e^{-i(\mathbf{q}-\mathbf{q}') \cdot \mathbf{R}_i} \left(1 + e^{-i\mathbf{q} \cdot \Delta\mathbf{R} + i\mathbf{q}' \cdot \Delta\mathbf{R}} - e^{i\mathbf{q}' \cdot \Delta\mathbf{R}} - e^{-i\mathbf{q} \cdot \Delta\mathbf{R}} \right) a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}'}$$

$$= -Nz \frac{JS^2}{2} + \frac{JS}{2} \sum_{\mathbf{q}} \sum_{\Delta \mathbf{R}} (2 - 2\cos \mathbf{q} \cdot \Delta\mathbf{R}) a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}}$$

$$= -Nz \frac{JS^2}{2} + \sum_{\mathbf{q}} JS \sum_{\Delta \mathbf{R}} (1 - \cos \mathbf{q} \cdot \Delta\mathbf{R}) a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}}, \tag{8.53}$$

where $\sum_{\Delta \mathbf{R}}$ is a sum over all nearest-neighbor vectors. For the simple cubic lattice we get

$$H = \frac{NJS^2}{3} + \sum_{\mathbf{q}} 2JS \left(3 - \cos q_x a - \cos q_y a - \cos q_z a \right) a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}} = \text{const} + \sum_{\mathbf{q}} \hbar \omega_{\mathbf{q}} a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}}. \tag{8.54}$$

This in fact holds in general: to order S^1 we recover the magnon dispersion of Sec. 8.1.

The terms of order S^0 in H contain four bosonic operators and thus describe magnon-magnon two-particle interactions. An analytical solution is no longer possible. There are various approaches for including this term approximately, going back to Anderson, Tyablikov, and others. We here consider a mean-field decoupling, which is essentially equivalent to the *random phase approximation* (RPA) employed by Anderson.

With Eq. (8.52) we write, up to a constant,

$$H \cong \underbrace{\sum_{\mathbf{q}} \hbar \omega_{\mathbf{q}} a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}}}_{=H_0} - \frac{J}{N^2} \sum_{\mathbf{q}\mathbf{q}'\mathbf{q}''\mathbf{q}'''} \sum_{\langle ij \rangle} \left[e^{-i\mathbf{q}\cdot\mathbf{R}_i - i\mathbf{q}'\cdot\mathbf{R}_j + i\mathbf{q}''\cdot\mathbf{R}_j + i\mathbf{q}'''\cdot\mathbf{R}_i} - \frac{1}{4} \left(e^{-i\mathbf{q}\cdot\mathbf{R}_i - i\mathbf{q}'\cdot\mathbf{R}_i + i\mathbf{q}''\cdot\mathbf{R}_i + i\mathbf{q}'''\cdot\mathbf{R}_j} \right. \right.$$

$$\left. \left. + e^{-i\mathbf{q}\cdot\mathbf{R}_i - i\mathbf{q}'\cdot\mathbf{R}_j + i\mathbf{q}''\cdot\mathbf{R}_j + i\mathbf{q}'''\cdot\mathbf{R}_j} + e^{-i\mathbf{q}\cdot\mathbf{R}_j - i\mathbf{q}'\cdot\mathbf{R}_i + i\mathbf{q}''\cdot\mathbf{R}_i + i\mathbf{q}'''\cdot\mathbf{R}_i} + e^{-i\mathbf{q}\cdot\mathbf{R}_j - i\mathbf{q}'\cdot\mathbf{R}_j + i\mathbf{q}''\cdot\mathbf{R}_j + i\mathbf{q}'''\cdot\mathbf{R}_i} \right) \right]$$

$$\times a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}'}^{\dagger} a_{\mathbf{q}''} a_{\mathbf{q}'''}. \tag{8.55}$$

Writing again $\sum_{\langle ij \rangle} = (1/2) \sum_i \sum_{\Delta\mathbf{R}}$, we obtain

$$
H \cong H_0 - \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}'\mathbf{q}''} \sum_{\Delta\mathbf{R}} \left[e^{-i\mathbf{q}'\cdot\Delta\mathbf{R} + i\mathbf{q}''\cdot\Delta\mathbf{R}} - \frac{1}{4}\left(e^{i(\mathbf{q}+\mathbf{q}'-\mathbf{q}'')\cdot\Delta\mathbf{R}} + e^{-i\mathbf{q}'\cdot\Delta\mathbf{R}+i\mathbf{q}''\cdot\Delta\mathbf{R}+i(\mathbf{q}+\mathbf{q}'-\mathbf{q}'')\cdot\Delta\mathbf{R}} \right. \right.
$$
$$
\left. \left. + e^{-i\mathbf{q}\cdot\Delta\mathbf{R}} + e^{-i\mathbf{q}\cdot\Delta\mathbf{R}-i\mathbf{q}'\cdot\Delta\mathbf{R}+i\mathbf{q}''\cdot\Delta\mathbf{R}} \right) \right] a_{\mathbf{q}}^\dagger a_{\mathbf{q}'}^\dagger a_{\mathbf{q}''} a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''}
$$
$$
= H_0 - \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}''} \sum_{\Delta\mathbf{R}} \left[e^{-i(\mathbf{q}-\mathbf{q}')\cdot\Delta\mathbf{R}} - \frac{1}{2}\left(\cos\mathbf{q}\cdot\Delta\mathbf{R} + \cos(\mathbf{q}+\mathbf{q}'-\mathbf{q}'')\cdot\Delta\mathbf{R} \right) \right]
$$
$$
\times a_{\mathbf{q}}^\dagger a_{\mathbf{q}'}^\dagger a_{\mathbf{q}''} a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''}. \tag{8.56}
$$

Now we perform a Hartree-Fock decoupling, which for bosons reads

$$
a_{\mathbf{q}}^\dagger a_{\mathbf{q}'}^\dagger a_{\mathbf{q}''} a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''} \cong \left\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''} \right\rangle a_{\mathbf{q}'}^\dagger a_{\mathbf{q}''} + a_{\mathbf{q}}^\dagger a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''} \left\langle a_{\mathbf{q}'}^\dagger a_{\mathbf{q}''} \right\rangle - \left\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''} \right\rangle \left\langle a_{\mathbf{q}'}^\dagger a_{\mathbf{q}''} \right\rangle
$$
$$
+ \left\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}''} \right\rangle a_{\mathbf{q}'}^\dagger a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''} + a_{\mathbf{q}}^\dagger a_{\mathbf{q}''} \left\langle a_{\mathbf{q}'}^\dagger a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''} \right\rangle - \left\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}''} \right\rangle \left\langle a_{\mathbf{q}'}^\dagger a_{\mathbf{q}+\mathbf{q}'-\mathbf{q}''} \right\rangle, \tag{8.57}
$$

where the first three terms are the Hartree contribution and the last three terms the Fock contribution. Since we are considering a ferromagnet, we expect the averages only to be non-zero if the two wave vectors agree. Defining $n_{\mathbf{q}} := \left\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}} \right\rangle$ and dropping a constant in H we obtain

$$
H_{\text{HF}} = H_0 - \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{\Delta\mathbf{R}} [1 - \cos\mathbf{q}\cdot\Delta\mathbf{R}] n_{\mathbf{q}} a_{\mathbf{q}'}^\dagger a_{\mathbf{q}'} - \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{\Delta\mathbf{R}} [1 - \cos\mathbf{q}\cdot\Delta\mathbf{R}] a_{\mathbf{q}}^\dagger a_{\mathbf{q}} n_{\mathbf{q}'}
$$
$$
- \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{\Delta\mathbf{R}} \left[e^{-i(\mathbf{q}'-\mathbf{q})\cdot\Delta\mathbf{R}} - \frac{1}{2}(\cos\mathbf{q}\cdot\Delta\mathbf{R} + \cos\mathbf{q}'\cdot\Delta\mathbf{R}) \right] n_{\mathbf{q}} a_{\mathbf{q}'}^\dagger a_{\mathbf{q}'}
$$
$$
- \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{\Delta\mathbf{R}} \left[e^{-i(\mathbf{q}'-\mathbf{q})\cdot\Delta\mathbf{R}} - \frac{1}{2}(\cos\mathbf{q}\cdot\Delta\mathbf{R} + \cos\mathbf{q}'\cdot\Delta\mathbf{R}) \right] a_{\mathbf{q}}^\dagger a_{\mathbf{q}} n_{\mathbf{q}'}
$$
$$
= H_0 - \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{\Delta\mathbf{R}} \left[1 - \cos\mathbf{q}'\cdot\Delta\mathbf{R} + 1 - \cos\mathbf{q}\cdot\Delta\mathbf{R} + e^{-i(\mathbf{q}-\mathbf{q}')\cdot\mathbf{R}} - \frac{1}{2}(\cos\mathbf{q}'\cdot\Delta\mathbf{R} + \cos\mathbf{q}\cdot\Delta\mathbf{R}) \right.
$$
$$
\left. + e^{-i(\mathbf{q}'-\mathbf{q})\cdot\mathbf{R}} - \frac{1}{2}(\cos\mathbf{q}\cdot\Delta\mathbf{R} + \cos\mathbf{q}'\cdot\Delta\mathbf{R}) \right] \times n_{\mathbf{q}'} a_{\mathbf{q}}^\dagger a_{\mathbf{q}}
$$
$$
= H_0 - \frac{J}{2N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{\Delta\mathbf{R}} [\not{2} - \not{2}\cos\mathbf{q}'\cdot\Delta\mathbf{R} - \not{2}\cos\mathbf{q}\cdot\Delta\mathbf{R} + \not{2}\cos(\mathbf{q}-\mathbf{q}')\cdot\Delta\mathbf{R}] n_{\mathbf{q}'} a_{\mathbf{q}}^\dagger a_{\mathbf{q}}
$$
$$
= \sum_{\mathbf{q}} \hbar\omega_{\mathbf{q}}^{\text{HF}} a_{\mathbf{q}}^\dagger a_{\mathbf{q}} \tag{8.58}
$$

with

$$
\hbar\omega_{\mathbf{q}}^{\text{HF}} := \underbrace{JS \sum_{\Delta\mathbf{R}}(1 - \cos\mathbf{q}\cdot\Delta\mathbf{R})}_{\hbar\omega_{\mathbf{q}}} - \frac{J}{N} \sum_{\mathbf{q}'} \sum_{\Delta\mathbf{R}} [1 - \cos\mathbf{q}\cdot\Delta\mathbf{R} - \cos\mathbf{q}'\cdot\Delta\mathbf{R} + \cos(\mathbf{q}-\mathbf{q}')\cdot\Delta\mathbf{R}] n_{\mathbf{q}'}
$$
$$
= \hbar\omega_{\mathbf{q}} - J \sum_{\Delta\mathbf{R}}(1 - \cos\mathbf{q}\cdot\Delta\mathbf{R}) \frac{1}{N} \sum_{\mathbf{q}'}(1 - \cos\mathbf{q}'\cdot\Delta\mathbf{R}) n_{\mathbf{q}'}
$$
$$
= JS \sum_{\Delta\mathbf{R}}(1 - \cos\mathbf{q}\cdot\Delta\mathbf{R}) \left[1 - \frac{1}{NS} \sum_{\mathbf{q}'}(1 - \cos\mathbf{q}'\cdot\Delta\mathbf{R}) n_{\mathbf{q}'} \right]. \tag{8.59}
$$

Herein, selfconsistency would require

$$
n_{\mathbf{q}} \equiv \left\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}} \right\rangle = n_B\left(\hbar\omega_{\mathbf{q}}^{\text{HF}}\right). \tag{8.60}
$$

However, the difference between $n_B\left(\hbar\omega_{\mathbf{q}}^{\text{HF}}\right)$ and $n_B(\hbar\omega_{\mathbf{q}})$ is of higher order in $1/S$ since the difference between $\hbar\omega_{\mathbf{q}}^{\text{HF}}$ and $\hbar\omega_{\mathbf{q}}$ is of higher order. Such terms are not correctly described by the Holstein-Primakoff Hamiltonian truncated after the S^0 term, anyway. We can thus set

$$
n_{\mathbf{q}} = n_B(\hbar\omega_{\mathbf{q}}) \tag{8.61}
$$

at the present order of approximation.

The boson number $n_{\mathbf{q}}$ occurs in

$$\frac{1}{N}\sum_{\mathbf{q}}(1-\cos\mathbf{q}\cdot\Delta\mathbf{R})n_{\mathbf{q}} = \frac{V}{N}\int\frac{d^3q}{(2\pi)^3}\frac{1-\cos\mathbf{q}\cdot\Delta\mathbf{R}}{\exp(\beta\hbar\omega_{\mathbf{q}})-1}. \tag{8.62}$$

We consider the low-T limit for the nearest-neighbor model on a simple cubic lattice. In analogy to Sec. 8.1 we obtain

$$\begin{aligned}
\ldots &\cong \frac{V}{N}\frac{1}{4\pi^2}\int_0^\infty dq\,q^2\int_{-1}^1 d(\cos\theta)\frac{1-\cos(qa\cos\theta)}{\exp(\beta JSa^2q^2)-1}\\
&= \frac{V}{N}\frac{1}{2\pi^2}\int_0^\infty dq\,q^2\frac{1-\sin qa/qa}{\exp(\beta JSa^2q^2)-1}\\
&\cong \frac{V}{N}\frac{1}{2\pi^2}\int_0^\infty dq\,q^2\frac{q^2a^2/6}{\exp(\beta JSa^2q^2)-1}\\
&= \underbrace{\frac{V}{N}}_{=a^3}\frac{\zeta(5/2)}{32}\frac{a^2}{\pi^{3/2}(\beta JSa^2)^{5/2}} = \frac{\zeta(5/2)}{32}\frac{1}{\pi^{3/2}(\beta JS)^{5/2}}.
\end{aligned} \tag{8.63}$$

Thus at low temperatures we find for small $\mathbf{q}$

$$\hbar\omega_{\mathbf{q}}^{\mathrm{HF}} \cong JSq^2a^2\left\{1-\frac{1}{S}\frac{\zeta(5/2)}{32}\frac{1}{\pi^{3/2}(\beta JS)^{5/2}}\right\}. \tag{8.64}$$

Note that this result can be understood as a temperature-dependent reduction of the spin-wave stiffness. In the calculation of the magnetization $\langle M\rangle$ in Sec. 8.1 we should just replace JS by $JS\{1-\ldots\}$. This gives

$$\langle M\rangle \cong M_{\mathrm{sat}} - g\mu_B\frac{\zeta(5/2)}{8}\frac{1}{(\pi\beta JSa^2)^{3/2}}\left\{1+\frac{3}{2}\frac{1}{S}\frac{\zeta(5/2)}{32}\frac{1}{\pi^{3/2}(\beta JSa^2q^2)^{5/2}}\right\}. \tag{8.65}$$

We thus find a correction of the form $-c\,T^{3/2}T^{5/2} = -c\,T^4$ due to magnon interactions, where c is a positive constant. The power of T is much higher than in the $T^{3/2}$ Bloch law, showing that the interactions play a minor role at low temperatures.

8.3 Antiferromagnetic spin waves and magnons

We now discuss excitations of antiferromagnets. We will here restrict ourselves to bipartite models without frustation and in fact for the most part to bipartite nearest-neighbor models. Then the mean-field approximation gives the Néel state with full but opposite spin polarization on the two sublattices as the ground state. We know that this is not the true ground state.

It seems dangerous to describe excitations of the system starting from an invalid approximate ground state. However, we will see that a Holstein-Primakoff bosonization scheme based on the Néel state does give good results. In fact it will naturally lead to an improved prediction for the ground state that is not the Néel state.

8.3.1 The ground state

We will reserve the index i for sublattice A and index j for sublattice B. On sublattice A we define, as above,

$$S_i^- = \sqrt{2S}\,a_i^\dagger\sqrt{1-\frac{a_i^\dagger a_i}{2S}}, \tag{8.66}$$

$$S_i^- = \sqrt{2S}\,\sqrt{1-\frac{a_i^\dagger a_i}{2S}}\,a_i, \tag{8.67}$$

$$S_i^z = S - a_i^\dagger a_i \tag{8.68}$$

and on sublattice B we define

$$S_j^- = \sqrt{2S}\,\sqrt{1-\frac{b_j^\dagger b_j}{2S}}\,b_j, \tag{8.69}$$

$$S_j^- = \sqrt{2S}\,b_j^\dagger\sqrt{1-\frac{b_j^\dagger b_j}{2S}}, \tag{8.70}$$

$$S_j^z = -S + b_j^\dagger b_j. \tag{8.71}$$

Evidently, the vacuum state with $a_i|0\rangle = b_j|0\rangle = 0$ is the Néel state

$$|0\rangle = \prod_{i \in A} |S\rangle_i \prod_{j \in B} |-S\rangle_j. \tag{8.72}$$

The Hamiltonian

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j \tag{8.73}$$

(with $J < 0$) then becomes

$$H = Nz\frac{JS^2}{2} - JS \sum_{\langle ij \rangle} (a_i^\dagger a_i + b_j^\dagger b_j + a_i b_j + b_j^\dagger a_i^\dagger) + \text{(interaction terms)}. \tag{8.74}$$

We now drop the interaction terms. The spatial dependence can be diagonalized by introducing the Fourier transformations

$$a_i = \sqrt{\frac{2}{N}} \sum_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{R}_i} a_{\mathbf{q}}, \tag{8.75}$$

$$b_j = \sqrt{\frac{2}{N}} \sum_{\mathbf{q}} e^{-i\mathbf{q} \cdot \mathbf{R}_j} b_{\mathbf{q}}. \tag{8.76}$$

Note the opposite signs in the exponentials and that the number of sites in each sublattice is $N/2$.

The Hamiltonian then becomes

$$H = Nz\frac{JS^2}{2} - \frac{2JS}{N} \sum_{\mathbf{q}\mathbf{q}'} \sum_{i \in A} \sum_{\Delta\mathbf{R}} \left(e^{-i(\mathbf{q}-\mathbf{q}')\cdot\mathbf{R}_i} a_{\mathbf{q}}^\dagger a_{\mathbf{q}'} + e^{i(\mathbf{q}-\mathbf{q}')\cdot(\mathbf{R}_i+\Delta\mathbf{R})} b_{\mathbf{q}}^\dagger b_{\mathbf{q}'} + \right.$$

$$\left. e^{i(\mathbf{q}-\mathbf{q}')\cdot\mathbf{R}_i - i\mathbf{q}'\cdot\Delta\mathbf{R}} a_{\mathbf{q}} b_{\mathbf{q}'} + e^{i(\mathbf{q}-\mathbf{q}')\cdot\mathbf{R}_i + i\mathbf{q}\cdot\Delta\mathbf{R}} b_{\mathbf{q}}^\dagger a_{\mathbf{q}'}^\dagger \right)$$

$$= Nz\frac{JS^2}{2} - JS \sum_{\mathbf{q}} \sum_{\Delta\mathbf{R}} \left(a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + b_{\mathbf{q}}^\dagger b_{\mathbf{q}} + e^{-i\mathbf{q}\cdot\Delta\mathbf{R}} a_{\mathbf{q}} b_{\mathbf{q}} + e^{i\mathbf{q}\cdot\Delta\mathbf{R}} b_{\mathbf{q}}^\dagger a_{\mathbf{q}}^\dagger \right). \tag{8.77}$$

For the example of the linear chain, the square lattice, and the simple cubic lattice we obtain

$$H = Nz\frac{JS^2}{2} - JS \sum_{\mathbf{q}} \left(z a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + z b_{\mathbf{q}}^\dagger b_{\mathbf{q}} + 2\sum_{\nu=1}^{d} \cos q_\nu a \, a_{\mathbf{q}} b_{\mathbf{q}} + 2\sum_{\nu=1}^{d} \cos q_\nu a \, b_{\mathbf{q}}^\dagger a_{\mathbf{q}}^\dagger \right), \tag{8.78}$$

where d is the spatial dimension. Noting that $z = 2d$ for these lattices and defining

$$\gamma_{\mathbf{q}} := \frac{1}{d} \sum_{\nu=1}^{d} \cos q_\nu a, \tag{8.79}$$

we get

$$H = Nz\frac{JS^2}{2} - zJS \sum_{\mathbf{q}} \left(a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + b_{\mathbf{q}}^\dagger b_{\mathbf{q}} + \gamma_{\mathbf{q}} a_{\mathbf{q}} b_{\mathbf{q}} + \gamma_{\mathbf{q}} a_{\mathbf{q}}^\dagger b_{\mathbf{q}}^\dagger \right). \tag{8.80}$$

Unlike the Hamiltonian for the ferromagnet, we here find terms that do not conserve the total number of bosons. Terms of this form, but for fermions, are known from the BCS theory of superconductivity. Like in BCS theory, we diagonalize H using a Bogoliubov-Valatin transformation,

$$a_{\mathbf{q}} = \cosh\theta_{\mathbf{q}} \alpha_{\mathbf{q}} - \sinh\theta_{\mathbf{q}} \beta_{\mathbf{q}}^\dagger, \tag{8.81}$$

$$b_{\mathbf{q}} = -\sinh\theta_{\mathbf{q}} \alpha_{\mathbf{q}}^\dagger + \cosh\theta_{\mathbf{q}} \beta_{\mathbf{q}}. \tag{8.82}$$

One can show that the mixed terms proportional to $\alpha_{\mathbf{q}} \beta_{\mathbf{q}}$ and $\beta_{\mathbf{q}}^\dagger \alpha_{\mathbf{q}}^\dagger$ vanish if

$$\tanh 2\theta_{\mathbf{q}} = \gamma_{\mathbf{q}}. \tag{8.83}$$

Insertion into H gives

$$H = Nz\frac{JS^2}{2} - zJS \sum_{\mathbf{q}} \left(\cosh^2\theta_{\mathbf{q}} + \sinh^2\theta_{\mathbf{q}} - 2\gamma_{\mathbf{q}} \cosh\theta_{\mathbf{q}} \sinh\theta_{\mathbf{q}} \right) (\alpha_{\mathbf{q}}^\dagger \alpha_{\mathbf{q}} + \beta_{\mathbf{q}}^\dagger \beta_{\mathbf{q}})$$

$$- zJS \sum_{\mathbf{q}} \left(2\sinh^2\theta_{\mathbf{q}} - 2\gamma_{\mathbf{q}} \cosh\theta_{\mathbf{q}} \sinh\theta_{\mathbf{q}} \right), \tag{8.84}$$

where the last term is coming from the commutation relations $[\alpha_{\mathbf{q}}, \alpha_{\mathbf{q}}^\dagger] = [\beta_{\mathbf{q}}, \beta_{\mathbf{q}}^\dagger] = 1$. Using identities for the hyperbolic functions, we obtain

$$
\begin{aligned}
H &= Nz\frac{JS^2}{2} - zJS\sum_{\mathbf{q}}\sqrt{1-\gamma_{\mathbf{q}}^2}(\alpha_{\mathbf{q}}^\dagger\alpha_{\mathbf{q}} + \beta_{\mathbf{q}}^\dagger\beta_{\mathbf{q}}) - zJS\sum_{\mathbf{q}}\left(\sqrt{1-\gamma_{\mathbf{q}}^2} - 1\right) \\
&= Nz\frac{JS(S+1)}{2} - zJS\sum_{\mathbf{q}}\sqrt{1-\gamma_{\mathbf{q}}^2}\left(\alpha_{\mathbf{q}}^\dagger\alpha_{\mathbf{q}} + \frac{1}{2} + \beta_{\mathbf{q}}^\dagger\beta_{\mathbf{q}} + \frac{1}{2}\right) \\
&=: Nz\frac{JS(S+1)}{2} + \sum_{\mathbf{q}}\hbar\omega_{\mathbf{q}}\left(\alpha_{\mathbf{q}}^\dagger\alpha_{\mathbf{q}} + \frac{1}{2} + \beta_{\mathbf{q}}^\dagger\beta_{\mathbf{q}} + \frac{1}{2}\right)
\end{aligned}
\tag{8.85}
$$

(note that the sums over $\mathbf{q}$ contain $N/2$ terms). We have made the zero-point energy explicit and kept the constant terms since they are important in this case.

The resulting effective Hamiltonian is approximate because we have neglected magnon interactions. This was the only approximation we have made. The ground state $|0\rangle_{\mathrm{NIM}}$ in this non-interacting-magnon (NIM) approximation satisfies

$$
\alpha_{\mathbf{q}}|0\rangle_{\mathrm{NIM}} = \beta_{\mathbf{q}}|0\rangle_{\mathrm{NIM}} = 0,
\tag{8.86}
$$

i.e., it is the vacuum of the new bosons. The ground-state energy is

$$
\begin{aligned}
E_0^{\mathrm{NIM}} &= Nz\frac{JS(S+1)}{2} + \sum_{\mathbf{q}}\hbar\omega_{\mathbf{q}} = Nz\frac{JS(S+1)}{2} - zJS\sum_{\mathbf{q}}\sqrt{1-\gamma_{\mathbf{q}}^2} \\
&= Nz\frac{JS(S+1)}{2} - zJS\frac{N_a^d}{2}\int\frac{d^dq}{(2\pi)^d}\sqrt{1-\left(\frac{1}{d}\sum_{\nu=1}^d\cos q_\nu a\right)^2}.
\end{aligned}
\tag{8.87}
$$

The integral can be evaluated numerically (and analytically for $d=1$). The results are

$$
E_0^{\mathrm{NIM}} = dNJS^2 \times
\begin{cases}
(1 + 0.363/S) & \text{for } d = 1 \\
(1 + 0.158/S) & \text{for } d = 2 \\
(1 + 0.097/S) & \text{for } d = 3\,.
\end{cases}
\tag{8.88}
$$

For comparison, the energy expectation value of the Néel state is simply

$$
E_{\mathrm{Néel}} = -J\sum_{\langle ij\rangle}\underbrace{\langle\mathrm{Néel}|\mathbf{S}_i\cdot\mathbf{S}_j|\mathrm{Néel}\rangle}_{=-S^2} = JS^2\frac{N}{2}z = zN\frac{JS^2}{2} = dNJS^2.
\tag{8.89}
$$

Noting that $J < 0$, E_0^{NIM} is clearly smaller. The correction is larger for lower dimensions d. For $d = 1$, the exact ground-state energy is known from the Bethe ansatz. It is very close to E_0^{NIM}.

It is also instructive to derive the sublattice polarization or staggered magnetization

$$
M := \langle S_i^z\rangle\big|_{i\in A} = -\langle S_j^z\rangle\big|_{j\in B}.
\tag{8.90}
$$

In the state $|0\rangle_{\mathrm{NIM}}$ we find

$$
\begin{aligned}
M_0^{\mathrm{NIM}} &= \langle 0|_{\mathrm{NIM}}S_i^z|0\rangle_{\mathrm{NIM}} = \langle 0|_{\mathrm{NIM}}(S - a_i^\dagger a_i)|0\rangle_{\mathrm{NIM}} = S - \frac{2}{N}\sum_{\mathbf{q}}\langle 0|_{\mathrm{NIM}}a_{\mathbf{q}}^\dagger a_{\mathbf{q}}|0\rangle_{\mathrm{NIM}} \\
&= S - \frac{2}{N}\sum_{\mathbf{q}}\langle 0|_{\mathrm{NIM}}\Big(\cosh^2\theta_{\mathbf{q}}\,\alpha_{\mathbf{q}}^\dagger\alpha_{\mathbf{q}} + \sinh^2\theta_{\mathbf{q}}\,\beta_{\mathbf{q}}^\dagger\beta_{\mathbf{q}} - \cosh\theta_{\mathbf{q}}\sinh\theta_{\mathbf{q}}\,\alpha_{\mathbf{q}}\beta_{\mathbf{q}} \\
&\qquad - \cosh\theta_{\mathbf{q}}\sinh\theta_{\mathbf{q}}\,\beta_{\mathbf{q}}^\dagger\alpha_{\mathbf{q}}^\dagger\Big)|0\rangle_{\mathrm{NIM}}.
\end{aligned}
\tag{8.91}
$$

Since $|0\rangle_{\mathrm{NIM}}$ is the vacuum state of α and β, we get

$$
M_0^{\mathrm{NIM}} = S - \frac{2}{N}\sum_{\mathbf{q}}\sinh^2\theta_{\mathbf{q}} = S - \frac{2}{N}\sum_{\mathbf{q}}\left(\frac{1}{2}\frac{1}{\sqrt{1-\gamma_{\mathbf{q}}^2}} - \frac{1}{2}\right) = S + \frac{1}{2} - \frac{1}{2}a^d\int\frac{d^dq}{(2\pi)^d}\frac{1}{\sqrt{1-\gamma_{\mathbf{q}}^2}}.
\tag{8.92}
$$

For $d = 1$, the integral is of the form

$$
\int dq\,\frac{1}{\sqrt{1-(1-q^2a^2/2)^2}} \cong \int dq\,\frac{1}{q^2a^2} = \frac{1}{a}\int\frac{dq}{q}
\tag{8.93}
$$

at small q and thus diverges logarithmically. This indicates that for the 1D Heisenberg antiferromagnet even the *ground state* does not show long-range order. For the 1D ferromagnet we know that the ground state does show long-range order but that the order is destroyed by thermal fluctuations for any $T > 0$. Since thermal fluctuations cannot play a role for the ground state, one says that the magnetic order in the 1D antiferromagnet is destroyed by *quantum fluctuations*.

For $d > 1$ the integral converges. For the models considered above we get

$$M_0^{\text{NIM}} = \begin{cases} S\left(1 - 0.197/S\right) \text{ for } d = 2 \\ S\left(1 - 0.078/S\right) \text{ for } d = 3 \,. \end{cases} \tag{8.94}$$

Note that for $d = 2$ and $S = 1/2$ we obtain a roughly 40% reduction compared to the Néel state due to quantum fluctuations.

8.3.2 Excited states

In the previous subsection, we found that the non-interacting-magnon approximation predicts two degenerate magnon species with dispersion

$$\hbar\omega_{\mathbf{q}} = -zJS\sqrt{1 - \gamma_{\mathbf{q}}^2} = \underbrace{-zJS}_{>0}\sqrt{1 - \left(\frac{1}{d}\sum_{\nu=1}^{d}\cos q_\nu a\right)^2} \tag{8.95}$$

for the chain/square/simple cubic nearest-neighbor models. Thus for small $\mathbf{q}$ we find

$$\hbar\omega_{\mathbf{q}} \cong -zJS\sqrt{1 - \left(1 - \frac{1}{2d}q^2a^2\right)^2} \cong -zJS\sqrt{\frac{1}{d}q^2a^2} = -\frac{zJS}{\sqrt{d}}\,qa = -2\sqrt{d}JS\,qa. \tag{8.96}$$

Unlike for the ferromagnet, the dispersion is *linear*. This is a general result for antiferromagnets, not restricted to our particular models. The spin-wave velocity is thus a meaningful quantity for small $\mathbf{q}$. For the models considered here it is $c_{\text{SW}} = -2\sqrt{d}JSa/\hbar$. We also find $\hbar\omega_{\mathbf{q}=0} = 0$, again satisfying the Goldstone theorem.
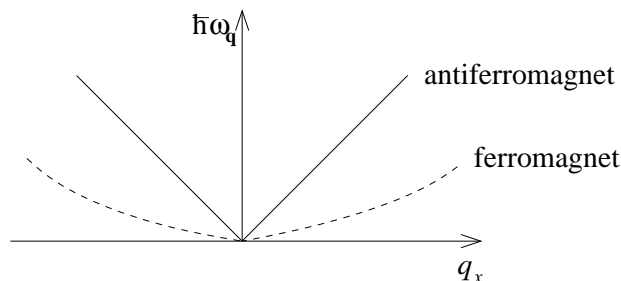
We can now calculate thermodynamic quantities in analogy to the ferromagnetic case. We only consider the staggered magnetization in 3D:

$$\begin{aligned} M = \langle S_i^z \rangle = \left\langle S - a_i^\dagger a_i \right\rangle &= S - \frac{2}{N}\sum_{\mathbf{q}}\left\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}}\right\rangle \\ &= S - \frac{2}{N}\sum_{\mathbf{q}}\left\langle\cosh^2\theta_{\mathbf{q}}\alpha_{\mathbf{q}}^\dagger\alpha_{\mathbf{q}} + \sinh^2\theta_{\mathbf{q}}\beta_{\mathbf{q}}^\dagger\beta_{\mathbf{q}} - \cosh\theta_{\mathbf{q}}\sinh\theta_{\mathbf{q}}\alpha_{\mathbf{q}}\beta_{\mathbf{q}} - \cosh\theta_{\mathbf{q}}\sinh\theta_{\mathbf{q}}\beta_{\mathbf{q}}^\dagger\alpha_{\mathbf{q}}^\dagger\right\rangle \\ &= S - \frac{2}{N}\sum_{\mathbf{q}}\left[\cosh^2\theta_{\mathbf{q}}n_B(\hbar\omega_{\mathbf{q}}) + \sinh^2\theta_{\mathbf{q}}(n_B(\hbar\omega_{\mathbf{q}}) + 1)\right] \\ &= M_0^{\text{NIM}} - \frac{2}{N}\sum_{\mathbf{q}}\left(\cosh^2\theta_{\mathbf{q}} + \sinh^2\theta_{\mathbf{q}}\right)n_B(\hbar\omega_{\mathbf{q}}) \\ &= M_0^{\text{NIM}} - \frac{2}{N}\sum_{\mathbf{q}}\frac{1}{\sqrt{1 - \gamma_{\mathbf{q}}^2}}\frac{1}{e^{\beta\hbar\omega_{\mathbf{q}}} - 1} \\ &= M_0^{\text{NIM}} - a^3\int\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{1 - \gamma_{\mathbf{q}}^2}}\frac{1}{e^{\beta\hbar\omega_{\mathbf{q}}} - 1}, \end{aligned} \tag{8.97}$$

where M_0^{NIM} is the ground-state staggered magnetization from the previous subsection. Since the integral is dominated by small $\mathbf{q}$, we insert the small-$\mathbf{q}$ limiting forms of $\sqrt{1 - \gamma_{\mathbf{q}}^2} \cong qa/\sqrt{3}$ and of $\hbar\omega_{\mathbf{q}}$,

$$
\begin{aligned}
M &\cong M_0^{\text{NIM}} - \frac{a^3}{2\pi^2} \int_0^\infty dq\, q^2 \frac{\sqrt{3}}{qa} \frac{1}{\exp(-2\sqrt{3}\beta J S q a) - 1} \\
&= M_0^{\text{NIM}} - \frac{\sqrt{3}}{2\pi^2} a^2 \int_0^\infty \frac{dq\, q}{\exp(-2\sqrt{3}\beta J S q a) - 1} \\
&= M_0^{\text{NIM}} - \frac{\sqrt{3}}{144} \frac{1}{(\beta J S)^3}.
\end{aligned}
\tag{8.98}
$$

Thus the staggered magnetization decreases like T^2 for low temperatures, not like $T^{3/2}$ as for the ferromagnet.

8.4 The antiferromagnetic chain

We have seen that the Heisenberg antiferromagnet in one dimension does not show long-range order even at $T = 0$. Marshall's theorem, which applies to the case of nearest-neighbor antiferromagnetic interactions, shows that the ground state is a non-degenerate singlet of the total spin. What can we say about the excited states and their energies? Perhaps surprisingly, the character of the excitations depends even qualitatively on the value of the single-spin quantum number S. For half-odd integer spins $S = 1/2, 3/2, 5/2, \ldots$, the excitation spectrum is gapless. This is shown by the Lieb-Schultz-Mattis theorem discussed below. For integer spins $S = 1, 2, 3, \ldots$, the excitation spectrum has a gap, this has been found by Haldane.

8.4.1 The Lieb-Schultz-Mattis theorem

We consider the Heisenberg model with nearest-neighbor antiferromagnetic interaction on a 1D chain with periodic boundary conditions,
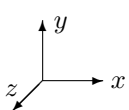
$$
H = -J \sum_{i=1}^{N} \mathbf{S}_i \cdot \mathbf{S}_{i+1},
\tag{8.99}
$$

where $J < 0$, $\mathbf{S}_{N+1} \equiv \mathbf{S}_1$, and N even. For half-odd integer spin $S = 1/2, 3/2, \ldots$, there is an excited state with eigenenergy E_1 that approaches the ground-state energy E_0 in the thermodynamic limit, $N \to \infty$.

Proof: Let $|0\rangle$ be the ground state of H. Define a *twist operator*

$$
U := \prod_{j=1}^{N} \exp\left(i \frac{2\pi j}{N} S_j^z \right) = \exp\left(i \sum_{j=1}^{N} \frac{2\pi j}{N} S_j^z \right).
\tag{8.100}
$$

U rotates spin j by an angle $2\pi j/N$ about the z-axis. The action of U is most easily pictured for a ferromagnetic state (which is otherwise irrelevant for the theorem):

We see that U introduces a twist. Now define $|1\rangle := U|0\rangle$.

(a) Next, define the unitary translation operator T_1 by

$$
T_1 \mathbf{S}_j T_1^\dagger = \mathbf{S}_{j+1}.
\tag{8.101}
$$

Since H is translationally invariant, we have $[H, T_1] = 0$. Thus eigenstates of H, in particular $|0\rangle$, can be chosen to be simultaneous eigenstates of T_1. Thus

$$
T_1 |0\rangle = e^{ik_0} |0\rangle, \quad k_0 \in \mathbb{R}.
\tag{8.102}
$$

The overlap between $|0\rangle$ and $|1\rangle$ is then

$$
\begin{aligned}
\langle 0|1\rangle &= \langle 0|U|0\rangle = \langle 0|e^{ik_0}Ue^{-ik_0}|0\rangle = \langle 0|T_1UT_1^\dagger|0\rangle \\
&= \langle 0|T_1\prod_{j=1}^{N}\exp\left(i\frac{2\pi j}{N}S_j^z\right)T_1^\dagger|0\rangle = \langle 0|\prod_{j=1}^{N}\exp\left(i\frac{2\pi j}{N}S_{j+1}^z\right)|0\rangle \\
&= \langle 0|\left[\prod_{j=1}^{N}\exp\left(i\frac{2\pi(j-1)}{N}S_j^z\right)\right]\exp\left(i\frac{2\pi}{N}S_1^z\right)|0\rangle \\
&= \langle 0|\left[\prod_{j=1}^{N}\exp\left(i\frac{2\pi j}{N}S_j^z\right)\exp\left(-i\frac{2\pi}{N}S_j^z\right)\right]\exp\left(i\frac{2\pi}{N}S_1^z\right)|0\rangle \\
&= \langle 0|U\exp\left(i\frac{2\pi}{N}S_1^z\right)\exp\left(-i\frac{2\pi}{N}S_{\text{tot}}^z\right)|0\rangle.
\end{aligned} \tag{8.103}
$$

Marshall's theorem shows that $|0\rangle$ is a singlet of $\mathbf{S}_{\text{tot}}$ so that $S_{\text{tot}}^z|0\rangle = 0$ and

$$
\exp\left(-i\frac{2\pi}{N}S_{\text{tot}}^z\right)|0\rangle = |0\rangle. \tag{8.104}
$$

Also,

$$
\exp\left(i\frac{2\pi}{N}S_1^z\right) = \begin{cases} 1 & \text{for } S=1,2,3,\dots \\ -1 & \text{for } S=1/2,3/2,5/2,\dots \end{cases} \tag{8.105}
$$

Thus for $S=1/2,3/2,5/2,\dots$ we find

$$
\langle 0|1\rangle = \langle 0|U(-1)|0\rangle = -\langle 0|1\rangle \quad \Rightarrow \quad \langle 0|1\rangle = 0. \tag{8.106}
$$

$|1\rangle$ is thus orthogonal to $|0\rangle$ for half-odd integer spins. (This is the point where the proof would fail for integer spins.)

(b) We now calculate the energy expectation value in the state $|1\rangle$. Since U is unitary, we get

$$
\begin{aligned}
\langle 1|H|1\rangle &= \langle 0|U^\dagger HU|0\rangle = -J\sum_j\langle 0|U^\dagger\left(S_j^xS_{j+1}^x + S_j^yS_{j+1}^y + S_j^zS_{j+1}^z\right)U|0\rangle \\
&= -J\sum_j\langle 0|\left[\left(\cos\frac{2\pi j}{N}S_j^x + \sin\frac{2\pi j}{N}S_j^y\right)\left(\cos\frac{2\pi(j+1)}{N}S_{j+1}^x + \sin\frac{2\pi(j+1)}{N}S_{j+1}^y\right) + \right. \\
&\quad \left.\left(\cos\frac{2\pi j}{N}S_j^y - \sin\frac{2\pi j}{N}S_j^x\right)\left(\cos\frac{2\pi(j+1)}{N}S_{j+1}^y - \sin\frac{2\pi(j+1)}{N}S_{j+1}^x\right) + S_j^zS_{j+1}^z\right]|0\rangle \\
&= -J\sum_j\langle 0|\left[\cos\frac{2\pi}{N}\left(S_j^xS_{j+1}^x + S_j^yS_{j+1}^y\right) + \sin\frac{2\pi}{N}\left(S_j^xS_{j+1}^y - S_j^yS_{j+1}^x\right) + S_j^zS_{j+1}^z\right]|0\rangle \\
&= E_0 - J\sum_j\langle 0|\left[\left(\cos\frac{2\pi}{N} - 1\right)\left(S_j^xS_{j+1}^x + S_j^yS_{j+1}^y\right) + \sin\frac{2\pi}{N}\left(S_j^xS_{j+1}^y - S_j^yS_{j+1}^x\right)\right]|0\rangle. \quad (8.107)
\end{aligned}
$$

Now the operator $S_j^xS_{j+1}^y - S_j^yS_{j+1}^x$ changes sign under rotation by π around $(1,1,0)$, whereas the singlet $|0\rangle$ is invariant under this (or any) rotation. Thus the expectation value of $S_j^xS_{j+1}^y - S_j^yS_{j+1}^x$ vanishes. We obtain

$$
\begin{aligned}
\langle 1|H|1\rangle - E_0 &= \underbrace{-J\left(\cos\frac{2\pi}{N} - 1\right)}_{<0}\sum_j\underbrace{\langle 0|\left(S_j^xS_{j+1}^x + S_j^yS_{j+1}^y\right)|0\rangle}_{\geq -S^2} \\
&\leq -J\left(\cos\frac{2\pi}{N} - 1\right)(-NS^2) = \underbrace{-JNS^2}_{>0}\left(1 - \cos\frac{2\pi}{N}\right) \\
&\cong -JNS^2\left(\frac{1}{2}\frac{4\pi^2}{N^2} + \mathcal{O}\left(\frac{1}{N^4}\right)\right) \\
&= -\frac{2\pi^2JS^2}{N} + \mathcal{O}\left(\frac{1}{N^3}\right).
\end{aligned} \tag{8.108}
$$

We see that the energy expectation value in the twisted state $|1\rangle$ is bounded from above by $E_0 - 2\pi^2JS^2/N \to E_0$ for $N \to \infty$ so that it approaches the ground-state energy for $N \to \infty$. Since $\langle 1|0\rangle = 0$, $|1\rangle$ can be written

as a superposition of eigenstates *not* including $|0\rangle$ so that at least one of these must have an eigenenergy approaching E_0 for $N \to \infty$. This completes the proof.

We have seen above that spontaneously broken symmetries lead to gapless excitations; this is Goldstone's theorem. The converse is evidently not true: the 1D antiferromagnetic Heisenberg chain has a gapless excitation spectrum but no long-range order.

8.4.2 The Jordan-Wigner transformation

This is a good point to introduce another mapping of a spin model onto a simpler model, in this case a fermionic one. The scheme discussed here is most suited to the antiferromagnetic spin chain. We start from a slightly more general spin Hamiltonian,

$$H = -J \sum_i \left(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta\, S_i^z S_{i+1}^z \right), \tag{8.109}$$

which includes an anisotropic exchange interaction. For $\Delta = 1$ we recover the Heisenberg model. The case $\Delta = 0$ is called *XY model* since the exchange interaction only involves the x- and y-components. We only consider the case $S = 1/2$.

The *Jordan-Wigner transformation* consists of the mapping

$$S_i^+ = a_i^\dagger \exp\left(i\pi \sum_{j=1}^{i-1} a_j^\dagger a_j \right), \tag{8.110}$$

$$S_i^- = \exp\left(-i\pi \sum_{j=1}^{i-1} a_j^\dagger a_j \right) a_i, \tag{8.111}$$

$$S_i^z = a_i^\dagger a_i - \frac{1}{2}. \tag{8.112}$$

Note that the factors $\exp(\pm i\pi \sum_{j=1}^{i-1} a_j^\dagger a_j)$ introduce a phase, which depends on the number of particles "to the left" of a given site. Due to these phase factors, a_i is a *fermionic* operator satisfying anticommutation relations

$$\{a_i, a_j^\dagger\} \equiv a_i a_j^\dagger + a_j^\dagger a_i = \delta_{ij}, \quad \{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0. \tag{8.113}$$

The Hamiltonian transforms into

$$
\begin{aligned}
H &= -J \sum_i \left(\frac{S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+}{2} + \Delta\, S_i^z S_{i+1}^z \right) \\
&= -J \sum_i \left[\frac{1}{2} a_i^\dagger e^{-i\pi a_i^\dagger a_i} a_{i+1} + \frac{1}{2} a_i a_{i+1}^\dagger e^{i\pi a_i^\dagger a_i} + \Delta \left(a_i^\dagger a_i - \frac{1}{2} \right) \left(a_{i+1}^\dagger a_{i+1} - \frac{1}{2} \right) \right] \\
&= -J \sum_i \left[\frac{1}{2} a_i^\dagger a_{i+1} + \frac{1}{2} a_i a_{i+1}^\dagger (-1) + \Delta \left(a_i^\dagger a_i - \frac{1}{2} \right) \left(a_{i+1}^\dagger a_{i+1} - \frac{1}{2} \right) \right] \\
&= -\frac{J}{2} \sum_i \left(a_i^\dagger a_{i+1} + a_{i+1}^\dagger a_i \right) - J\Delta \sum_i \left(a_i^\dagger a_i - \frac{1}{2} \right) \left(a_{i+1}^\dagger a_{i+1} - \frac{1}{2} \right). \tag{8.114}
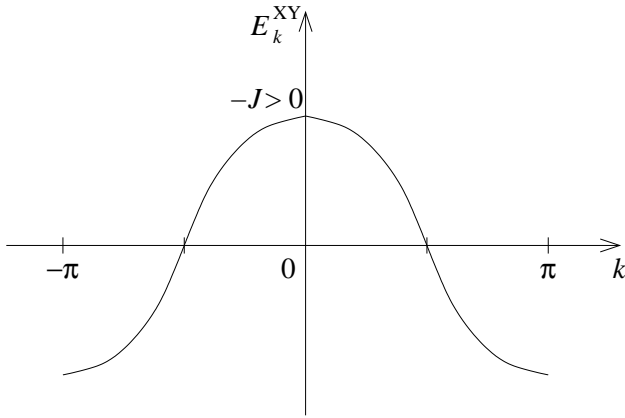\end{aligned}
$$

For the XY model ($\Delta = 0$), we clearly obtain a system of free fermions, which has a simple exact solution in terms of Slater determinants. With

$$a_j = \frac{1}{\sqrt{N}} \sum_k e^{ikj} a_k \tag{8.115}$$

we write

$$H_{\mathrm{XY}} = -J \sum_k \cos k\, a_k^\dagger a_k, \tag{8.116}$$

where k is from the first Brillouin zone, $k \in]-\pi, \pi]$.

For $\Delta \neq 0$ we find a nearest-neighbor density-density interaction. The model is still exactly solvable using the Bethe ansatz, but this solution is much more complicated than for $\Delta = 0$. We know, however, that the ground state $|0\rangle$ is a spin singlet. Thus

$$\sum_i S_i^z |0\rangle = \sum_i \left(a_i^\dagger a_i - \frac{1}{2} \right) |0\rangle = 0 \quad \Rightarrow \quad \sum_i a_i^\dagger a_i |0\rangle = \sum_i \frac{1}{2} |0\rangle = \frac{N}{2} |0\rangle \tag{8.117}$$

and $|0\rangle$ has sharp fermion number $N/2$. The ground state is thus half filled and we expect low-energy excitations to have fermion numbers close to $N/2$.
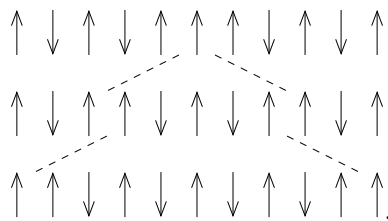
8.4.3 Spinons

We will now discuss the nature of the excitations of the antiferromagnetic spin-1/2 nearest-neighbor Heisenberg chain. We will see that while magnons are excitations of this system, as we have assumed in Sec. 8.3, they are not the most fundamental ones. We start with a qualitative discussion.

Recall how we arrived at the concept of magnons: the naive excitation with $|\Delta S^z| = 1$ is the flipping of a single spin $1/2$. However, this costs a high energy on the order of $|J|$. Forming instead superpositions of such spin flips over the whole system, we obtain magnons with energy approaching zero in the small-q limit. However, in 1D there is another possibility. A single spin flipped in the Néel state,

$$\uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow \, \uparrow \, \uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow,$$

can be split into two *kinks* (domain walls) in the Néel order,

$$\uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow \, \uparrow \, \uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow$$

$$\uparrow \, \downarrow \, \uparrow \, \uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow \, \uparrow \, \downarrow \, \uparrow$$

$$\uparrow \, \uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow \, \downarrow \, \uparrow \, \uparrow,$$

without an additional change of the energy or the total spin S^z. Note that the kinks correspond to π phase jumps of the Néel order. A single kink costs an energy of the order of $|J|/2$. Now one considers superpositions of such kinks over the whole chain and finds again that this leads to excitations with energy going to zero for $q \to 0$. These excitations are called *spinons*. We see that a spin-flip, which is the physical excitation for example in neutron scattering, generates a *pair* of spinons, which are then free to propagate independently from one another. One says that they are *deconfined*. A spinon is evidently a spin-1/2 excitation since it is "half a magnon". Also, in the above sketch the spinons do indeed carry an excess spin of $1/2$.

The above discussion does not carry over to higher dimensions: the separation of a spin-flip into a pair of kinks, or of a magnon into a pair of spinons, is associated with an energy increase in higher dimensions. Thus spinons are *confined* for $d \geq 2$. The reason is that the two spinons are connected by a string of flipped spins with frustated bonds to their unflipped neighbors:

string

 frustrated bond
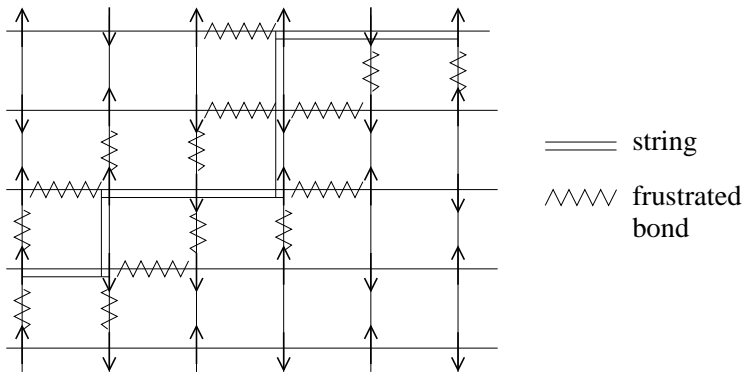
The energy is therefore linear in the length of the string.

In 1D it is possible to obtain the spinon dispersion from the Bethe ansatz. We quote the result without derivation:

$$\hbar\omega_{\mathbf{q}}^{\mathrm{spinon}} = \underbrace{-\frac{\pi}{2}J}_{>0} |\sin q|. \tag{8.118}$$

However, since physical excitations generate spinon pairs, it is useful to consider the energy vs. momentum q of such pairs. Since the momentum can be distributed between the two spinons in different ways, there is a *continuum* of excitations for given q. The exact result is sketched here:
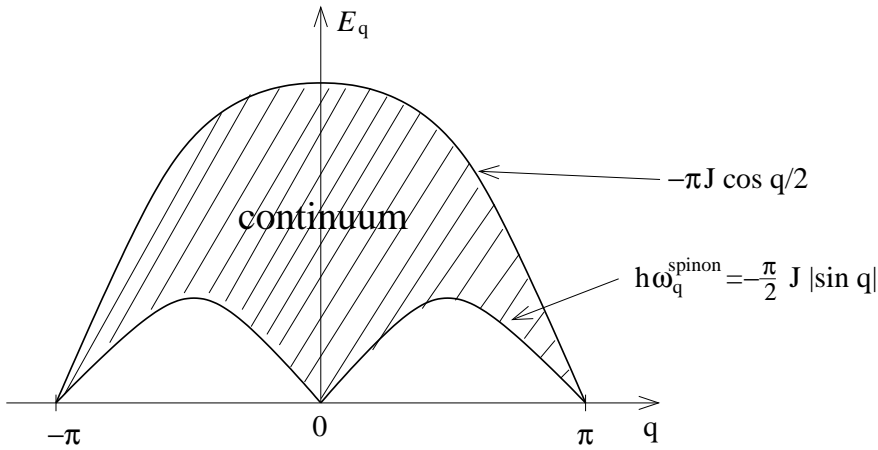
This dispersion has been observed with neutron scattering for KCuF$_3$ by Tennant *et al.* in 1993.

Chapter 9

Paramagnetism and diamagnetism of metals

Many magnetic materials are metallic, not insulating. Iron is the best known example. We thus have to understand how magnetic ordering arises in metals. As a prerequisite, we first study the magnetic properties of metals in the absense of interactions, i.e., the magnetism of the free electron gas.

9.1 Paramagnetism of the electron gas

We have seen in Sec. 1.2 that electrons possess a spin $s = 1/2$ and a spin magnetic moment of $m_s = g\mu_B/2 \approx \mu_B$ oriented oppositely to the spin. Thus the energy of a free electron in a uniform magnetic field is

$$\epsilon_{\mathbf{k}\sigma} = \frac{\hbar^2 k^2}{2m} + \sigma \frac{g\mu_B B}{2} \quad \text{with } \sigma = \uparrow, \downarrow \equiv \pm 1. \tag{9.1}$$

The total energy of the electron gas is then

$$E = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}\sigma} n_F(\epsilon_{\mathbf{k}\sigma} - \mu), \tag{9.2}$$

where μ is the chemical potential. The magnetization is given by

$$M = -\frac{g\mu_B}{2V} \sum_{\mathbf{k}\sigma} \sigma \, n_F(\epsilon_{\mathbf{k}\sigma} - \mu) \tag{9.3}$$

and thus the susceptibility is

$$\chi = \left.\frac{\partial M}{\partial B}\right|_{B=0} = -\frac{g\mu_B}{2V} \sum_{\mathbf{k}\sigma} \sigma \, n_F'\left(\frac{\hbar^2 k^2}{2m} - \mu\right) \sigma \frac{g\mu_B}{2} = -\frac{g^2 \mu_B^2}{4V} \underbrace{2}_{\text{from } \sum_\sigma} \sum_{\mathbf{k}} n_F'\left(\frac{\hbar^2 k^2}{2m} - \mu\right). \tag{9.4}$$

With the density of states (for one spin direction) $D(\epsilon) = 1/V \sum_{\mathbf{k}} \delta(\epsilon - \hbar^2 k^2/2m)$ we obtain

$$\chi = -\frac{g^2 \mu_B^2}{2} \int d\epsilon \, D(\epsilon) \, n_F'(\epsilon - \mu). \tag{9.5}$$

As long as $k_B T \ll \mu$, which is typically the case for metals (recall that a thermal energy of $1\,\text{eV}$ requires $T \approx 10^4\,\text{K}$), we can approximate n_F by a step function and thus n_F' by a δ-function: $n_F'(\epsilon) \cong -\delta(\epsilon)$. Thus

$$\chi = \frac{g\mu_B D(\mu)}{2} =: \chi_{\text{Pauli}}. \tag{9.6}$$

This is called the *Pauli susceptibility*, which describes *Pauli paramagnetism*. The result is valid for any dispersion, not just for free electrons, if the appropriate density of states is inserted. The result is essentially temperature-independent as long as $k_B T \ll \mu$. This is quite different from the Curie law for local magnetic moments, where $\chi \propto 1/T$.

Next, we want to find the corresponding susceptibility in a non-uniform field. We decompose $\mathbf{B}(\mathbf{r})$ into Fourier components. Since we are interested in the linear response, our equation for $\mathbf{M}(\mathbf{r})$ is linear in $\mathbf{B}(\mathbf{r})$

and we can just consider a single mode $\mathbf{B_q} \cos \mathbf{q} \cdot \mathbf{r}$. In linear response, we can also assume $|\mathbf{B_q}|$ to be small and therefore treat the Zeeman energy

$$E_Z = \sigma \, \frac{g\mu_B B_{\mathbf{q}}}{2} \, \cos \mathbf{q} \cdot \mathbf{r} \tag{9.7}$$

as a weak pertubation.

The unperturbed states of free electrons can be written as products of plane waves $\psi_{\mathbf{k}}(\mathbf{r}) = V^{-1/2} \, e^{i\mathbf{k}\cdot\mathbf{r}}$ and spinors $|\uparrow\rangle$, $|\downarrow\rangle$. The first-order correction to an eigenstate $|\mathbf{k}\sigma\rangle$ is

$$
\begin{aligned}
|\mathbf{k}\sigma\rangle^{(1)} &= \sum_{\mathbf{k}'\sigma'} \frac{|\mathbf{k}'\sigma'\rangle\langle\mathbf{k}'\sigma'|}{\Sigma_{\mathbf{k}\sigma} - \Sigma_{\mathbf{k}'\sigma'}} \, \sigma \, \frac{g\mu_B B_{\mathbf{q}}}{2} \, \cos \mathbf{q} \cdot \mathbf{r} \, |\mathbf{k}\sigma\rangle \\
&= \sigma \, \frac{g\mu_B B_{\mathbf{q}}}{2} \sum_{\mathbf{k}'} \frac{|\mathbf{k}'\sigma\rangle\langle\mathbf{k}'\sigma|}{\Sigma_{\mathbf{k}\sigma} - \Sigma_{\mathbf{k}'\sigma}} \, \cos \mathbf{q} \cdot \mathbf{r} \, |\mathbf{k}\sigma\rangle \\
&= \sigma \, \frac{g\mu_B B_{\mathbf{q}}}{2} \sum_{\mathbf{k}'} \frac{|\mathbf{k}'\sigma\rangle}{\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 (k')^2}{2m}} \, \frac{1}{V} \int d^3 r \, e^{-i\mathbf{k}'\cdot\mathbf{r}} \, \cos(\mathbf{q}\cdot\mathbf{r}) \, e^{i\mathbf{k}\cdot\mathbf{r}} \\
&= \sigma \, \frac{g\mu_B B_{\mathbf{q}}}{2} \, \frac{m}{\hbar^2} \sum_{\mathbf{k}'} \frac{|\mathbf{k}'\sigma\rangle}{k^2 - (k')^2} \, \frac{1}{V} \int d^3 r \, e^{-i\mathbf{k}'\cdot\mathbf{r}} \left(e^{i\mathbf{q}\cdot\mathbf{r}} + e^{-i\mathbf{q}\cdot\mathbf{r}} \right) e^{i\mathbf{k}\cdot\mathbf{r}} \\
&= \sigma \, \frac{g\mu_B B_{\mathbf{q}}}{2} \, \frac{m}{\hbar^2} \sum_{\mathbf{k}'} \frac{|\mathbf{k}'\sigma\rangle}{k^2 - (k')^2} \left(\delta_{\mathbf{k}',\mathbf{k}+\mathbf{q}} + \delta_{\mathbf{k}',\mathbf{k}-\mathbf{q}} \right) \\
&= \sigma \, \frac{g\mu_B B_{\mathbf{q}}}{2} \, \frac{m}{\hbar^2} \left(\frac{|\mathbf{k}+\mathbf{q},\sigma\rangle}{-2\mathbf{k}\cdot\mathbf{q} - q^2} + \frac{|\mathbf{k}-\mathbf{q},\sigma\rangle}{2\mathbf{k}\cdot\mathbf{q} - q^2} \right).
\end{aligned}
\tag{9.8}
$$

The magnetization is then, to first order,

$$
\begin{aligned}
M_{\mathbf{q}} &\cong -\frac{g\mu_B}{2V} \sum_{\mathbf{k}\sigma} \left(\langle\mathbf{k}\sigma| + {}^{(1)}\langle\mathbf{k}\sigma| \right) \sigma \cos \mathbf{q}\cdot\mathbf{r} \left(|\mathbf{k}\sigma\rangle + |\mathbf{k}\sigma\rangle^{(1)} \right) n_F(\epsilon_{\mathbf{k}} - \mu) \\
&\cong -\frac{g\mu_B}{2V} \sum_{\mathbf{k}\sigma} \sigma \left(\langle\mathbf{k}\sigma| \cos \mathbf{q}\cdot\mathbf{r} |\mathbf{k}\sigma\rangle^{(1)} + {}^{(1)}\langle\mathbf{k}\sigma| \cos \mathbf{q}\cdot\mathbf{r} |\mathbf{k}\sigma\rangle \right) n_F(\epsilon_{\mathbf{k}} - \mu)
\end{aligned}
\tag{9.9}
$$

with $\epsilon_{\mathbf{k}} := \hbar^2 k^2 / 2m$. Note that the magnetization vanishes in the unperturbed state and that the first-order correction to the energy also vanishes if $\mathbf{q} \neq 0$. Inserting $|\mathbf{k}\sigma\rangle^{(1)}$, we obtain

$$
\begin{aligned}
M_{\mathbf{q}} &\cong -\frac{g^2\mu_B^2 B_{\mathbf{q}}}{2V} \, \frac{m}{\hbar^2} \sum_{\mathbf{k}\sigma} \frac{1}{2} \left(\frac{\cancel{\langle\mathbf{k}\sigma|e^{i\mathbf{q}\cdot\mathbf{r}}|\mathbf{k}+\mathbf{q},\sigma\rangle}}{-2\mathbf{k}\cdot\mathbf{q} - q^2} + \frac{\langle\mathbf{k}\sigma|e^{-i\mathbf{q}\cdot\mathbf{r}}|\mathbf{k}+\mathbf{q},\sigma\rangle}{-2\mathbf{k}\cdot\mathbf{q} - q^2} + \frac{\langle\mathbf{k}\sigma|e^{i\mathbf{q}\cdot\mathbf{r}}|\mathbf{k}-\mathbf{q},\sigma\rangle}{2\mathbf{k}\cdot\mathbf{q} - q^2} \right. \\
&\qquad \left. + \frac{\cancel{\langle\mathbf{k}\sigma|e^{-i\mathbf{q}\cdot\mathbf{r}}|\mathbf{k}-\mathbf{q},\sigma\rangle}}{2\mathbf{k}\cdot\mathbf{q} - q^2} + \text{complex conjugate} \right) n_F(\epsilon_{\mathbf{k}} - \mu) \\
&= g^2\mu_B^2 B_{\mathbf{q}} \frac{m}{\hbar^2} \int \frac{d^3 k}{(2\pi)^3} \left(\frac{1}{q^2 + 2\mathbf{k}\cdot\mathbf{q}} + \frac{1}{q^2 - 2\mathbf{k}\cdot\mathbf{q}} \right) n_F(\epsilon_{\mathbf{k}} - \mu).
\end{aligned}
\tag{9.10}
$$

Assuming $k_B T \ll \mu$, the integral becomes

$$\int_{k \leq k_F} \frac{d^3 k}{(2\pi)^3} \left(\frac{1}{q^2 + 2\mathbf{k}\cdot\mathbf{q}} + \frac{1}{q^2 - 2\mathbf{k}\cdot\mathbf{q}} \right), \tag{9.11}$$

where $\hbar^2 k_F^2 / 2m = E_F = \mu$. This can be evaluated as

$$
\begin{aligned}
\ldots &= \frac{1}{(2\pi)^2} \int_0^{k_F} dk \, k^2 \int_{-1}^{1} d(\cos\theta) \left(\frac{1}{q^2 + 2kq\cos\theta} + \frac{1}{q^2 - 2kq\cos\theta} \right) \\
&= \frac{1}{(2\pi)^2} \int_0^{k_F} dk \, k^2 \int_{-1}^{1} d(\cos\theta) \, \frac{2}{q^2 - 4k^2\cos^2\theta}.
\end{aligned}
\tag{9.12}
$$

Note that this integral is singular if $q < 2k_F$. In this case we should treat it as a principal-value integral. After some calculation we find

$$\ldots = \frac{k_f}{8\pi^2} \left(1 + \frac{4k_F^2 - q^2}{4k_F q} \ln\left| \frac{2k_F + q}{2k_F - q} \right| \right). \tag{9.13}$$

We thus obtain

$$\chi_{\mathbf{q}} = \frac{M_{\mathbf{q}}}{B_{\mathbf{q}}} = g^2\mu_B^2 \frac{m}{\hbar^2} \frac{k_F}{8\pi^2} f\left(\frac{q}{2k_F} \right) \tag{9.14}$$

with

$$f(x) := 1 + \frac{1 - x^2}{2x} \ln \left| \frac{1 + x}{1 - x} \right|. \tag{9.15}$$

For a parabolic dispersion, the density of states is

$$\begin{aligned}
D(\epsilon) &= \frac{1}{V} \sum_{\mathbf{k}} \delta \left(\epsilon - \frac{\hbar^2 k^2}{2m} \right) = \int \frac{d^3 k}{(2\pi)^3} \delta \left(\epsilon - \frac{\hbar^2 k^2}{2m} \right) = \frac{2m}{\hbar^2} \int \frac{d^3 k}{(2\pi)^3} \delta \left(\frac{2m\epsilon}{\hbar^2} - k^2 \right) \\
&= \frac{1}{\pi^2} \frac{m}{\hbar^2} \int_0^\infty dk \, k^2 \, \delta \left[\left(k + \frac{\sqrt{2m\epsilon}}{\hbar} \right) \left(k - \frac{\sqrt{2m\epsilon}}{\hbar} \right) \right] \\
&= \frac{1}{\pi^2} \frac{m}{\hbar^2} \int_0^\infty dk \, \frac{k^2}{k + \frac{\sqrt{2m\epsilon}}{\hbar}} \delta \left(k - \frac{\sqrt{2m\epsilon}}{\hbar} \right) \\
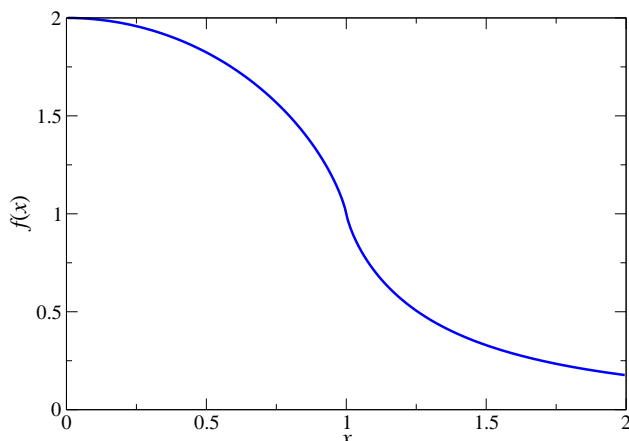&= \frac{1}{2\pi^2} \frac{m}{\hbar^2} \frac{\sqrt{2m\epsilon}}{\hbar} \tag{9.16}
\end{aligned}$$

and thus at the Fermi energy,

$$D(\mu) \cong D(E_F) = \frac{1}{2\pi^2} \frac{m}{\hbar^2} k_F. \tag{9.17}$$

Thus

$$\chi_{\mathbf{q}} = \frac{g^2 \mu_B^2}{4} D(\mu) f \left(\frac{q}{2k_F} \right) = \frac{1}{2} \chi_{\text{Pauli}} f \left(\frac{q}{2k_F} \right). \tag{9.18}$$

The following plot shows the function $f(x)$, Eq. (9.15):

For $q \to 0$ we have $f \to 2$ so that we recover the Pauli susceptibility. Note that the $\mathbf{q}$-dependent susceptibility $\chi_{\mathbf{q}}$ has a singularity at $q = 2k_F$; the derivate $d\chi_{\mathbf{q}}/dq$ diverges there. Since $2k_F$ is the diameter of the Fermi sea, a magnetic field modulated with wave vector $\mathbf{q}$ of magnitude $q = 2k_F$ mixes electrons at opposite points on the Fermi surface.
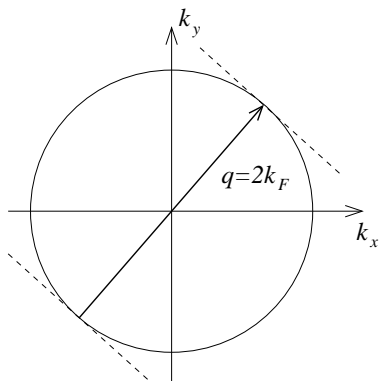
The singularity in $\chi_{\mathbf{q}}$ is actually a "nesting" feature due to the Fermi surface portions at $\mathbf{k}$ and $\mathbf{k} + \mathbf{q}$ being locally parallel.

In real materials the Fermi surface is not a sphere. The qualitative $\mathbf{q}$ dependence of $\chi_{\mathbf{q}}$ remains the same, though. $\chi_{\mathbf{q}}$ decreases as a function of $|\mathbf{q}|$ and shows singularities for $\mathbf{q}$ vectors that connect parallel portions of the Fermi surface.

9.2 Diamagnetism of the electron gas

There is also a diamagnetic contribution to the susceptibility of free electrons, which is due to the magnetic moments generated by charge currents. It would still be present if the electrons had no spin (or rather no spin magnetic moment). We know from the Bohr-van Leeuwen theorem that this diamagnetic response has to be a quantum-mechanical phenomenon.

9.2.1 The two-dimensional electron gas

It is easier to first consider a two-dimensional electron gas in a uniform magnetic field with the single-electron Hamiltonian

$$H = \frac{1}{2m} \left[\mathbf{p} + e\mathbf{A}(\mathbf{r})\right]^2, \tag{9.19}$$

where the charge is $-e$ and we have ignored the Zeeman term, which we know leads to Pauli paramagnetism. Without loss of generality, the uniform field is assumed to be $\mathbf{B} = B\,\hat{\mathbf{z}}$. We choose the so-called Landau gauge

$$\mathbf{A}(\mathbf{r}) = (-By, 0, 0), \tag{9.20}$$

which indeed gives

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} = \begin{pmatrix} 0 - 0 \\ 0 - 0 \\ 0 - (-B)\partial y/\partial y \end{pmatrix} = B\,\hat{\mathbf{z}}. \tag{9.21}$$

Then

$$H = \frac{1}{2m}\left(p_x - eBy\right)^2 + \frac{1}{2m}\,p_y^2. \tag{9.22}$$

We find $[H, p_x] = 0$ since H does not contain x. Thus p_x is a constant of motion and can be replaced by its eigenvalue $\hbar k_x$. Defining

$$y_0 := \frac{\hbar}{eB}k_x \quad \text{and} \quad \omega_c := \frac{eB}{m}, \tag{9.23}$$

we obtain

$$H = \frac{1}{2m}\,p_y^2 + \frac{1}{2}m\omega_c^2\,(y - y_0)^2. \tag{9.24}$$

Here, ω_c is the *cyclotron frequency*. The resulting Hamiltonian describes a harmonic oscillator with the potential minimum shifted to y_0. It of course has the eigenvalues

$$E_{n,k_x} = \hbar\omega_c\left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, \ldots \tag{9.25}$$

The energies do not depend on k_x and are thus hugely degenerate. Note that the apparent asymmetry between the x- and y-direction in H is gauge-dependent and therefore without physical consequence. We could have chosen the vector potential $\mathbf{A}$ to point in any direction within the xy-plane. The isotropy of two-dimensional space is not broken by the choice of a special gauge.

The magnetic field transforms the continuous spectrum of the two-dimensional electron gas, $E_{\mathbf{k}} = \hbar^2(k_x^2 + k_y^2)/2m$, into a discrete spectrum of *Landau levels* enumerated by n. For $B = 0$, the density of states is

$$\begin{aligned} D(\epsilon) &= \int \frac{d^2k}{(2\pi)^2}\delta\left(\epsilon - \frac{\hbar^2 k^2}{2m}\right) = \frac{1}{2\pi}\int_0^\infty dk\, k\, \delta\left(\epsilon - \frac{\hbar^2 k^2}{2m}\right) \\ &= \frac{1}{4\pi}\int_0^\infty dk^2\, \delta\left(\epsilon - \frac{\hbar^2 k^2}{2m}\right) = \frac{1}{4\pi}\int_0^\infty dE\, \frac{2m}{\hbar}\,\delta(\epsilon - E) \\ &= \frac{m}{2\pi\hbar^2} \quad \text{for } \epsilon \geq 0. \end{aligned} \tag{9.26}$$

The density of states is thus constant. For $B > 0$, it is replaces by equidistant δ-function peaks:
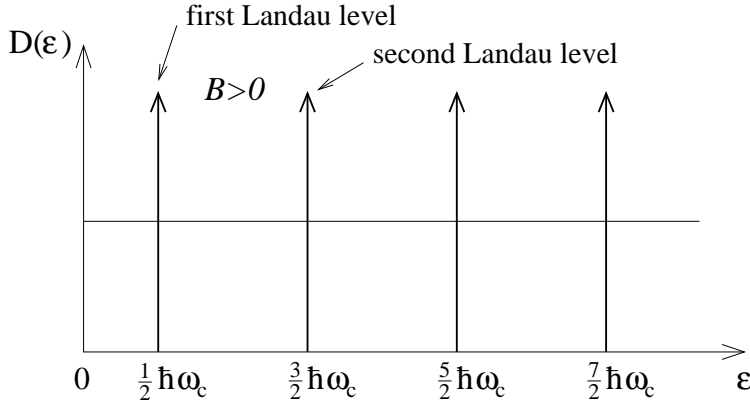
We can obtain the degeneracy of the Landau levels as follows: since the total number of states does not change, the Landau levels must accomodate these states, thus the degeneracy of the first one (and all the others) is

$$N_L = \underbrace{L^2}_{\text{area}} \int_0^{\hbar\omega_c} d\epsilon \, D(\epsilon) = \frac{m}{2\pi\hbar^2}\hbar\omega_c L^2 = \frac{m}{2\pi\hbar}\frac{eB}{m}L^2 = \frac{eB}{h}L^2. \tag{9.27}$$

At low temperatures, the low-energy states are filled successively until all N electrons are accomodated. If

$$N = 2nN_L \quad \text{with} \quad n = 1, 2, \ldots, \tag{9.28}$$

the lowest n Landau leves are completely filled and the others empty. The factor 2 is due to the two spin directions. In the generic case that $N/2N_L$ is not an integer, the highest Landau level is partially filled. (Landau-level quantization is of course one of the key ingredients of the integer quantum Hall effect.) The total energy of N electrons is

$$E = \sum_{n=0}^{\underline{n}-1} 2N_L\hbar\omega_c\left(n + \frac{1}{2}\right) + (N - \underline{n}\,2N_L)\,\hbar\omega_c\left(\underline{n} - \frac{1}{2}\right) \tag{9.29}$$

with

$$\underline{n} := \left\lfloor \frac{N}{2N_L} \right\rfloor. \tag{9.30}$$

The function $\lfloor x \rfloor$ is the largest integer smaller or equal to x. The first term in E describes the filled Landau levels and the second the partially filled one.
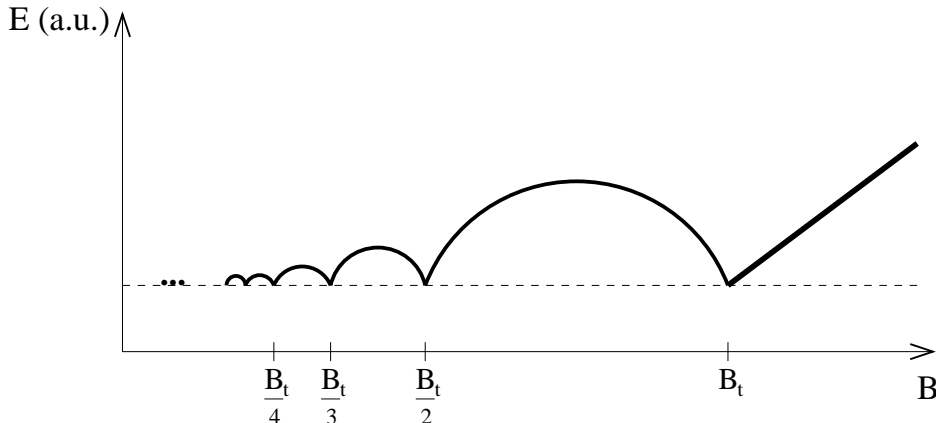
With the *filling factor* $\nu := N/N_L$, we obtain the energy per electron as

$$\frac{E}{N} = \sum_{n=0}^{\underline{n}-1} \frac{2}{\nu}\,\hbar\omega_c\left(n + \frac{1}{2}\right) + \left(1 - \frac{2\underline{n}}{\nu}\right)\hbar\omega_c\left(\underline{n} + \frac{1}{2}\right) \tag{9.31}$$

with

$$\underline{n} = \left\lfloor \frac{\nu}{2} \right\rfloor. \tag{9.32}$$

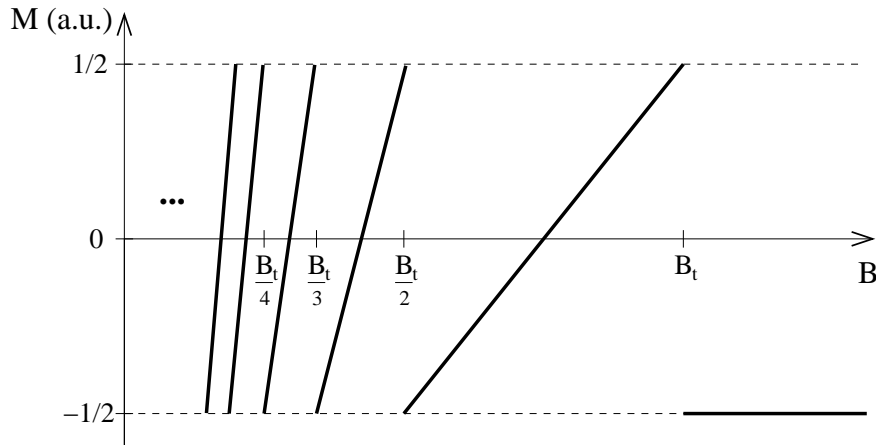This function is continous but not everywhere differentiable:

Here, B_t is the field for which $\nu = 2$, i.e., for which the lowest Landau level is completely filled,

$$\frac{N}{N_L} = \frac{N}{eB_t L^2/h} = 2 \quad \Rightarrow \quad B_t = \frac{h}{2e}\frac{N}{L^2}. \tag{9.33}$$

The areal magnetization

$$M = -\frac{1}{L^2}\frac{\partial E}{\partial B} \tag{9.34}$$

shows oscillations that are periodic in $1/B$. These are the de Haas-van Alpen oscillations.

The limit $\lim_{B\to 0} M$ does not exsist and neither does the limit $\lim_{B\to 0} \chi$ for the susceptibility $\chi = \partial M/\partial B$.

This unphysical result is due to our assumption of $T = 0$. 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. In this regime, we can neglect the discreteness of the Landau levels and write

$$E \overset{N \gg 2N_L}{\cong} \int_0^{\frac{N}{2N_L}} dn\, 2N_L \hbar\omega_c\, n = \frac{2eB}{h} L^2 \frac{\hbar eB}{m}\left[\frac{n^2}{2}\right]_0^{Nh/2eBL^2}$$

$$= \frac{e^2 B^2}{\pi m} L^2 \frac{1}{2}\frac{N^2 h^2}{4e^2 B^2 L^4} = \frac{N^2 h^2}{8\pi m}\frac{1}{L^2}. \tag{9.35}$$

The areal magnetization thus vanishes,

$$M = -\frac{1}{L^2}\frac{\partial E}{\partial B} = 0. \tag{9.36}$$

We could have guessed this result from the plot of $M(B)$ for $T = 0$. If we smear out the rapid oscilltions seen at small B, it is plausible that we obtain $M = 0$.

It is important that powers of B have cancelled in $E(B)$ above because this shows that the expressions for E and M are correct at least to order B^2 and B^1, respectively, not just to order B^0, which would be trivial (of course the magnetization vanishes if we ignore B!).

Thus we find $\chi = \partial M/\partial B = 0$. The diamagnetic susceptibility of the 2D electron gas vanishes.

9.2.2 The three-dimensional electron gas

The Hamiltonian of free electrons in three dimensions in the presence of a uniform magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$ reads
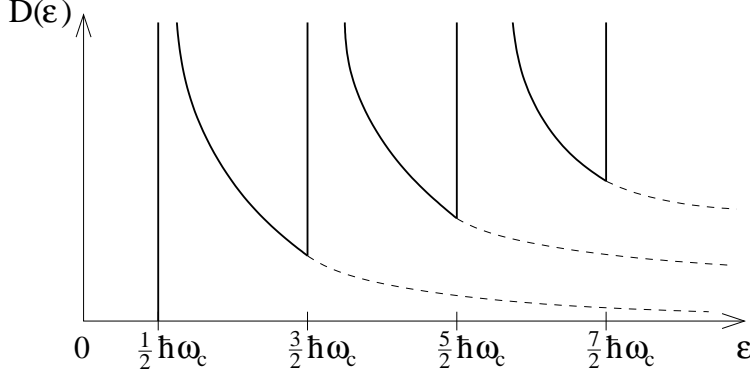
$$H = \frac{1}{2m}[\mathbf{p} + e\mathbf{A}(\mathbf{r})]^2 = \frac{1}{2m}p_y^2 + \frac{1}{2}m\omega_c^2(y - y_0)^2 + \frac{1}{2m}p_z^2, \tag{9.37}$$

where we have again chosen the Landau gauge $\mathbf{A} = (-By, 0, 0)$. Evidently we find free motion in the z-direction in addition to shifted harmonic oscillators in the xy-plane. The eigenenergies are

$$E_{n,k_x,k_y,k_z} = \hbar\omega_c\left(n + \frac{1}{2}\right) + \frac{\hbar^2 k_z^2}{2m}. \tag{9.38}$$

The density of states is thus the sum of the densities of states of the one-dimensional electron gas, shifted to the minimum energies $\hbar\omega_c(n + 1/2)$, $n = 0, 1, 2, \ldots$ This gives, for one spin direction,

$$D(\epsilon) = \frac{N_L}{L^2}\sum_{n=0}^{\infty}\frac{1}{\pi\hbar}\sqrt{\frac{m}{2(\epsilon - [n + 1/2]\hbar\omega_c)}}\,\Theta\left(\epsilon - \left[n + \frac{1}{2}\right]\hbar\omega_c\right)$$

$$= \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\frac{N_L}{L^2}\sum_{n=0}^{\infty}\frac{\Theta(\epsilon - [n + 1/2]\,\hbar\omega_c)}{\sqrt{\epsilon - [n + 1/2]\hbar\omega_c}}. \tag{9.39}$$

For $k_B T \ll \hbar\omega_c$, the low-energy states are again filled up until all electrons are accomodated. We again expect to find special features whenever the chemical potential μ reaches $\hbar\omega_c(n + 1/2)$ with $n = 0, 1, 2\ldots$ Since μ is roughly constant, while $\hbar\omega_c \propto B$, we find in the total energy $E(B)$ and thus in the magnetization $M(B) = -(1/V)\partial E/\partial B$ features periodic in $1/B$. These are again de Haas-van Alphen oscillations. They are visible if B is large.

We are interested in the susceptibility for small fields. Thus we have $\hbar\omega_c \ll k_B T$ and we of course still assume $k_B T \ll \mu$ for a typical metal. The rigorous calculation is rather involved. It is presented for example in "Solid State Physics" by Grosso and Pastori Parravicini. We here give a simplified and less rigorous version.

It is useful to define the iterated integrals of the density of states

$$P_1(x) := \int_0^x d\epsilon\, D(\epsilon), \tag{9.40}$$

$$P_2(x) := \int_0^x d\epsilon\, P_1(\epsilon). \tag{9.41}$$

Obviously the total electron number is

$$N = 2V P_1(\mu), \tag{9.42}$$

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

$$E = 2V \int_0^\mu d\epsilon\, \epsilon\, D(\epsilon). \tag{9.43}$$

Integration by parts gives

$$E = 2V \left[\mu P_1(\mu) - \int_0^\mu d\epsilon\, P_1(\epsilon) \right] = \mu N - 2V P_2(\mu). \tag{9.44}$$

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

$$P_2(\mu) = \frac{1}{\pi\hbar} \sqrt{\frac{m}{2}} \frac{N_L}{L^2} \sum_{n=0}^\infty \frac{4}{3} \left(\mu - \left[n + \frac{1}{2} \right] \hbar\omega_c \right)^{3/2} \Theta\left(mu - \left[n + \frac{1}{2} \right] \hbar\omega_c \right). \tag{9.45}$$

Since $\hbar\omega_c \ll \mu$, we can replace the sum over n by an integral. The *Poisson summation formula*

$$\sum_{n=0}^\infty f\left(n + \frac{1}{2} \right) = \int_0^\infty dx\, f(x) + 2 \sum_{s=0}^\infty (-1)^s \int_0^\infty dx\, f(x) \cos 2\pi s x \tag{9.46}$$

states how to do that correctly. We obtain

$$P_2(\mu) = \frac{1}{\pi\hbar} \sqrt{\frac{m}{2}} \frac{N_L}{L^2} \frac{4}{3} \left\{ \int_0^{\frac{\mu}{\hbar\omega_c} - \frac{1}{2}} dx\, (\mu - x\hbar\omega_c)^{3/2} + 2 \sum_{s=1}^\infty (-1)^s \int_0^{\frac{\mu}{\hbar\omega_c} - \frac{1}{2}} dx\, (\mu - x\hbar\omega_c)^{3/2} \cos 2\pi s x \right\}$$

$$= \frac{1}{\pi\hbar} \sqrt{\frac{m}{2}} \frac{N_L}{L^2} \frac{4}{3} \left\{ \frac{2}{5} \frac{\mu^{5/2}}{\hbar\omega_c} - \frac{1}{10\sqrt{2}} (\hbar\omega_c)^{3/2} + 2 \sum_{s=1}^\infty (-1)^s \frac{3}{8\pi^2} \frac{\hbar\omega_c \sqrt{\mu}}{s^2} + \text{(oscillating terms)} \right\}. \tag{9.47}$$

Here, the oscillating terms contain $\cos(2\pi s\mu/\hbar\omega_c)$ or $\sin(2\pi s\mu/\hbar\omega_c)$, become rapidly oscillating in the limit $\mu/\hbar \to \infty$ (corresponding to $B \to 0$), and are neglected. Then with

$$\sum_{s=1}^\infty \frac{(-1)^s}{s^2} = -\frac{\pi^2}{12} \tag{9.48}$$

we obtain

$$P_2 \cong \frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\,\frac{N_L}{L^2}\,\frac{4}{3}\left\{\frac{2}{5}\frac{\mu^{5/2}}{\hbar\omega_c} - \frac{1}{10\sqrt{2}}(\hbar\omega_c)^{3/2} - \frac{1}{16}\hbar\omega_c\sqrt{\mu}\right\}. \tag{9.49}$$

Since $N_L \propto B$ and $\hbar\omega_c \propto B$, the second term is of order $B^{5/2}$ and is thus irrelevant for the susceptibility at $B \to 0$. The first term is of order B^0 and thus determines the energy in the absence of a magnetic field. Thus we can write

$$E = E(B=0) + 2V\frac{1}{\pi\hbar}\sqrt{\frac{m}{2}}\,\frac{N_L}{L^2}\,\frac{4}{3}\,\frac{1}{16}\,\hbar\omega_c\sqrt{\mu} = E(B=0) + \frac{1}{12\pi^2}V\frac{e^2\sqrt{\mu}}{\sqrt{2m\hbar}}B^2 \tag{9.50}$$

and

$$M = -\frac{1}{V}\frac{\partial E}{\partial B} = -\frac{1}{6\pi^2}\frac{e^2\sqrt{\mu}}{\sqrt{2m\hbar}}B \tag{9.51}$$

and finally

$$\chi = \frac{\partial M}{\partial B} = -\frac{1}{6\pi^2}\frac{e^2\sqrt{\mu}}{\sqrt{2m\hbar}}. \tag{9.52}$$

With the zero-field density of states $D(\mu) = m/(2\pi^2\hbar^3)\sqrt{2m\mu}$ we get

$$\chi = -\frac{e^2\hbar^2}{6m^2}\,D(\mu) \stackrel{g\approx 2}{\approx} -\frac{g^2\mu_B^2}{6}\,D(\mu). \tag{9.53}$$

Thus if we take $g = 2$ we obtain a very simple result for the diamagnetic susceptibility,

$$\chi = -\frac{1}{3}\,\chi_{\text{Pauli}} =: \chi_{\text{Landau}}. \tag{9.54}$$

This is called the *Landau susceptibility* describing *Landau diamagnetism*. It is essentially temperature-independent as long as $k_B T \ll \mu$. For the free electron gas in three dimensions we thus find the total susceptibility

$$\chi = \chi_{\text{Pauli}} + \chi_{\text{Landau}} = \frac{2}{3}\,\chi_{\text{Pauli}}. \tag{9.55}$$

In real metals, the band structure deviates from the parabolic form and the ratio $\chi_{\text{Landau}}/\chi_{\text{Pauli}}$ differs from $-1/3$. In many metals Landau diamagnetism even dominates over Pauli paramagnetism.

We note without derivation that the **q**-dependence of the non-uniform diamagnetic susceptibility is different from the one of the paramagnetic susceptibility:

$$\chi_{\mathbf{q}}^{\text{dia}} = \chi_{\text{Landau}}\,L\left(\frac{q}{2k_F}\right) \tag{9.56}$$

with

$$L(x) = \frac{3}{8x^2}\left[1 + x^2 - \frac{(1-x^2)^2}{2x}\ln\left|\frac{1+x}{1-x}\right|\right] \tag{9.57}$$

for the free electron gas.

Chapter 10

Magnetic order in metals

In the previous chapter, we have neglected the Coulomb interaction between electrons. It is surprising that the omission of the strong Coulomb interaction leads to reasonable results for many metals. The reason for this is that the interaction is screened by the mobile electrons themselves. The formal justification was given by Landau in his Fermi-liquid theory. However, in the case of magnetic order, this picture breaks down and we have to include the Coulomb interaction explicitly.

10.1 Bloch theory

The Coulomb interaction between electrons can be written as a Fourier transform,

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{1}{r} = \frac{1}{V} \sum_{\mathbf{q}\neq 0} e^{i\mathbf{q}\cdot\mathbf{r}} \frac{e^2}{\epsilon_0 q^2}. \tag{10.1}$$

The $\mathbf{q} = 0$ term has been omitted since it is canceled by the interaction of the electrons with the average potential of the nuclei.

We can thus write the Hamiltonian of the interacting electron gas as

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} + \frac{1}{2V} \sum_{\mathbf{q}\neq 0} \frac{e^2}{\epsilon_0 q^2} \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma_1 \sigma_2} a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_2\sigma_2} a_{\mathbf{k}_1\mathbf{q}\sigma_1} =: H_0 + H_c, \tag{10.2}$$

compare chapter 4.

We now treat this Hamiltonian in a variational approach, where the set of variational states comprises the Slater determinants of single-particle states $|\psi\rangle$ of the unperturbed Hamiltonian H_0. This is equivalent to the Hartree-Fock decoupling. The expectation value of the Coulomb term is

$$\langle\psi|H_c|\psi\rangle = \frac{1}{2V} \sum_{\mathbf{q}\neq 0} \frac{e^2}{\epsilon_0 q^2} \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma_1 \sigma_2} \langle\psi|a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_2\sigma_2} a_{\mathbf{k}_1\mathbf{q}\sigma_1}|\psi\rangle. \tag{10.3}$$

The expectation value under the sum is non-zero only if creation and annihilation operators are paired, i.e., if (a) $\mathbf{k}_1 + \mathbf{q} = \mathbf{k}_1$ and $\mathbf{k}_2 - \mathbf{q} = \mathbf{k}_2$ or (b) $\mathbf{k}_1 + \mathbf{q} = \mathbf{k}_2$ and $\sigma_1 = \sigma_2$. Case (a) is excluded by $\mathbf{q} \neq 0$. Thus

$$\langle\psi|H_c|\psi\rangle = \frac{1}{2V} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \\ \mathbf{k}_1 \neq \mathbf{k}_2}} \frac{e^2}{\epsilon_0 |\mathbf{k}_1 - \mathbf{k}_2|^2} \sum_{\sigma_1} \langle\psi|a_{\mathbf{k}_2\sigma_1}^\dagger a_{\mathbf{k}_1\sigma_1}^\dagger a_{\mathbf{k}_2\sigma_1} a_{\mathbf{k}_1\mathbf{q}\sigma_1}|\psi\rangle$$

$$= -\frac{1}{2V} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \\ \mathbf{k}_1 \neq \mathbf{k}_2}} \frac{e^2}{\epsilon_0 |\mathbf{k}_1 - \mathbf{k}_2|^2} \sum_{\sigma_1} \langle\psi|a_{\mathbf{k}_1\sigma_1}^\dagger a_{\mathbf{k}_1\sigma_1} a_{\mathbf{k}_2\sigma_1}^\dagger a_{\mathbf{k}_2\sigma_1}|\psi\rangle. \tag{10.4}$$

Since the variational states $|\psi\rangle$ are eigenstates of H_0, they have sharp and uncorrelated particle numbers

$$n_{\mathbf{k}\sigma} := \langle\psi|a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma}|\psi\rangle \in \{0, 1\}. \tag{10.5}$$

Thus

$$\langle\psi|H|\psi\rangle = \langle\psi|H_0|\psi\rangle + \langle\psi|H_c|\psi\rangle$$
$$= \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} n_{\mathbf{k}\sigma} - \frac{1}{2V} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \\ \mathbf{k}_1 \neq \mathbf{k}_2}} \frac{e^2}{\epsilon_0 |\mathbf{k}_1 - \mathbf{k}_2|^2} \sum_{\sigma} n_{\mathbf{k}_1\sigma} n_{\mathbf{k}_2\sigma}. \tag{10.6}$$

We see that the energy is *reduced* by each pair of electrons with *parallel* spins. There is thus a *ferromagnetic* exchange interaction between the electrons in a metal.

Following Bloch, we now consider special variational states for which the spin-up and spin-down electrons each fill a Fermi sea but with generally different volume. We assume a free-electron dispersion for simplicitiy. The Fermi seas are then characterized by the two Fermi wave numbers $k_{F\uparrow}$ and $k_{F\downarrow}$. The total density of spin-σ electrons is $n_\sigma = k_{F\sigma}^3 / 6\pi^2$ so that $k_{F\sigma} = (6\pi^2 n_\sigma)^{1/3}$.

The unperturbed energy is

$$\langle\psi|H_0|\psi\rangle = \sum_\sigma V \int_{k \leq k_{F\sigma}} \frac{d^3k}{(2\pi)^3} \frac{\hbar k^2}{2m} = \frac{V}{2\pi^2} \frac{\hbar^2}{2m} \sum_\sigma \int_0^{k_{F\sigma}} dk \, k^4$$
$$= \frac{V}{10\pi^2} \frac{\hbar^2}{2m} \sum_\sigma k_{F\sigma}^5 = V \frac{\hbar^2}{2m} \frac{6\pi}{5} \left(\frac{9\pi}{2}\right)^{1/3} \sum_\sigma n_\sigma^{5/3} \tag{10.7}$$

and the Coulomb energy is

$$\langle\psi|H_c|\psi\rangle = -\sum_\sigma \frac{V}{2} \int_{\substack{k_1 \leq k_{F\sigma} \\ k_2 \leq k_{F\sigma}}} \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{e^2}{\epsilon_0 |\mathbf{k}_1 - \mathbf{k}_2|^2}$$
$$= -\frac{V}{2} \frac{1}{8\pi^4} \sum_\sigma \int_0^{k_{F\sigma}} dk_1 \, k_1^2 \int_0^{k_{F\sigma}} dk_2 \, k_2^2 \int_0^\pi d\theta \sin\theta \frac{e^2}{\epsilon_0 |\mathbf{k}_1 - \mathbf{k}_2|^2}$$
$$= -\frac{Ve^2}{16\pi^4 \epsilon_0} \sum_\sigma \int_0^{k_{F\sigma}} dk_1 \, k_1^2 \int_0^{k_{F\sigma}} dk_2 \, k_2^2 \int_0^\pi d\theta \sin\theta \frac{1}{k_1^2 + k_2^2 - 2k_1 k_2 \cos\theta}, \tag{10.8}$$

where θ is the angle between $\mathbf{k}_1$ and $\mathbf{k}_2$. The integrals can be evaluated:

$$\langle\psi|H_c|\psi\rangle = -\frac{Ve^2}{16\pi^4 \epsilon_0} \sum_\sigma k_{F\sigma}^4 \underbrace{\int_0^1 dx_1 \, x_1^2 \int_0^1 dx_2 \, x_2^2 \int_{-1}^1 \frac{du}{x_1^2 + x_2^2 - 2x_1 x_2 u}}_{=1/2}$$
$$= -\frac{Ve^2}{32\pi^4 \epsilon_0} \sum_\sigma k_{F\sigma}^4 = -V \frac{1}{4\pi\epsilon_0} \frac{9}{4} \left(\frac{2}{9\pi}\right)^{1/3} \sum_\sigma n_\sigma^{4/3}. \tag{10.9}$$
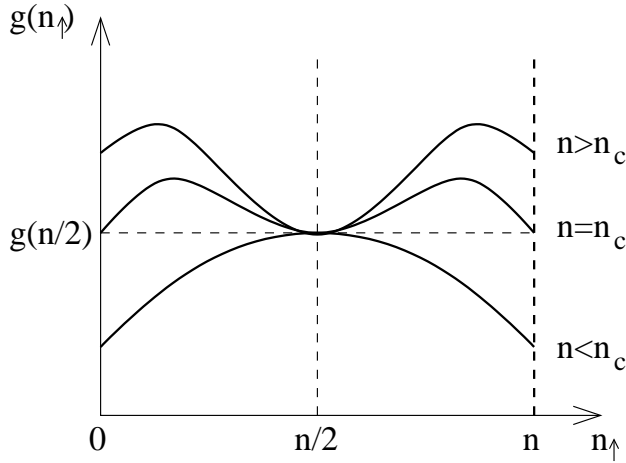
The total energy density is thus

$$\frac{1}{V} \langle\psi|H|\psi\rangle = \frac{\hbar^2}{2m} \frac{6\pi}{5} \left(\frac{9\pi}{2}\right)^{1/3} \left[\left(n_\uparrow^{5/3} + n_\downarrow^{5/3}\right) - \alpha \left(n_\uparrow^{4/3} + n_\downarrow^{4/3}\right) \right]$$
$$=: \frac{\hbar^2}{2m} \frac{6\pi}{5} \left(\frac{9\pi}{2}\right)^{1/3} g(n_\uparrow) \tag{10.10}$$

with

$$\alpha := \frac{5}{12\pi^2} \left(\frac{9\pi}{2}\right)^{1/3} \frac{1}{4\pi\epsilon_0} \frac{2m}{\hbar^2} \tag{10.11}$$

and in writing $g(n_\uparrow)$ as a function of a single argument, we imply $n_\downarrow = n - n_\uparrow$ with the total electron concetration n. To find the variational ground state, we have to minimize $g(n_\uparrow)$.

There is a special concentration

$$n_c := \left(\frac{\alpha}{1 + 2^{-1/3}} \right)^3, \tag{10.12}$$

which separates two phases with different properties:

- for $n > n_c$, $g(n_\uparrow)$ has its global minimum at $n_\uparrow = n/2$, the electron gas is unpolarized,

- for $n < n_c$, $g(n_\uparrow)$ has its global minimum at $n_\uparrow = 0$ and $n_\uparrow = n$, the electron gas is *completely spin polarized*. Note that theses minima occur at the edges of the allowed range; $g'(n_\uparrow)$ does not vanish there. A material with complete spin polarization of the valence band is also called a *half-metallic ferromagnet*, since it only has electrons of spin direction at the Fermi energy.

One can find a simpler form of the condition $n < n_c$ for ferromagnetism by introducing the parameter

$$r_s := \frac{r_0}{a_B}, \tag{10.13}$$

where $4\pi/3 r_0^3 := 1/n$ is the volume per electron and $a_B := \hbar^2/m\,(4\pi\epsilon_0)/e^2$ is the Bohr radius. Then ferromagnetic order occurs for

$$r_s > \frac{2\pi}{5}(1 + 2^{1/3}) \left(\frac{9\pi}{4} \right)^{1/3} \approx 5.4531. \tag{10.14}$$

Note that ferromagnetism occurs for *low* electron concentrations.

Compared to experiment, the Bloch theory suffers from problems:

- r_s for ferromagnetic metals is not particulary large in reality.

- The theory can only predict vanishing or complete spin polarization in the ground state, in contradiction to experiments; iron, cobalt, and nickel are all not completely spin polarized, for example.

10.2 Stoner mean-field theory of the Hubbard model

One problem of the Bloch theory is the insufficient treatment of the screening of the Coulomb interaction. The model that incorporates the screened Coulomb interaction in the simplest possible way is the *Hubbard model* introduced in Sec. 4.2. It is described by the Hamiltonian

$$H = -\sum_{ij\sigma} t_{ij} a_{i\sigma}^\dagger a_{j\sigma} + U \sum_i a_{i\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow}^\dagger a_{i\downarrow}. \tag{10.15}$$

It keeps only the local part of the Coulomb interaction , which is a reasonable approximation if the screening length r_{scr} in the screened Coulomb (Yukawa) potential

$$V(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0} \frac{e^{-r/r_{\mathrm{scr}}}}{r} \tag{10.16}$$

is of the order of the lattice constant. This is the case in good metals, i.e., metals with large density of states at the Fermi energy.

The exact ground state of the Hubbard model is not known, except in one dimension, where it can be obtained by the Bethe ansatz. We thus have to rely on approximations. The most straightforward one is the

mean-field decoupling of the interaction, which we will now consider for $T = 0$. The same results can also be obtained from a variational ansatz in terms of Slater determinants, distinct, however, from Bloch theory in that a short-range, screened Coulomb interaction is considered.

We first rewrite the Hamiltonian as

$$H = -\sum_{ij\sigma} t_{ij} a_{i\sigma}^\dagger a_{j\sigma} + U\sum_i a_{i\uparrow}^\dagger a_{i\downarrow}^\dagger a_{i\downarrow} a_{i\uparrow}$$

$$= -\sum_{ij\sigma} t_{ij} a_{i\sigma}^\dagger a_{j\sigma} + \frac{U}{2} \sum_{i\sigma_1\sigma_2} a_{i\sigma_1}^\dagger a_{i\sigma_2}^\dagger a_{i\sigma_2} a_{i\sigma_1}, \tag{10.17}$$

where we have used that $a_{i\sigma} a_{i\sigma} = 0$. This apparently more complicated form of H will be convenient later. In $\mathbf{k}$-space we have

$$H = \sum_{\mathbf{k}\sigma} \epsilon_\mathbf{k} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} + \frac{U}{2N} \sum_{\mathbf{k}_1\mathbf{k}_2} \sum_{\mathbf{q}\neq 0} \sum_{\sigma_1\sigma_2} a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_2\sigma_2} a_{\mathbf{k}_1\mathbf{q}\sigma_1}. \tag{10.18}$$

See Sec. 10.1 for why $\mathbf{q} = 0$ is excluded. The interaction is decoupled in the Hartree-Fock approximation,

$$a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_2\sigma_2} a_{\mathbf{k}_1\sigma_1} \cong \left\langle a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_1\sigma_1} \right\rangle a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_2\sigma_2} + a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_1\sigma_1} \left\langle a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_2\sigma_2} \right\rangle$$

$$- \langle\ldots\rangle\langle\ldots\rangle - \left\langle a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_2\sigma_2} \right\rangle a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_1\sigma_1}$$

$$- a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_2\sigma_2} \left\langle a_{\mathbf{k}_2-\mathbf{q},\sigma_2}^\dagger a_{\mathbf{k}_1\sigma_1} \right\rangle + \langle\ldots\rangle\langle\ldots\rangle. \tag{10.19}$$

Since we are interested in ferromagnetic order, we only consider homogeneous solutions. Then, with $\mathbf{q} \neq 0$, we have $\left\langle a_{\mathbf{k}_1+\mathbf{q},\sigma_1}^\dagger a_{\mathbf{k}_1\sigma_1} \right\rangle = 0$ etc. and the whole first (Hartree) term vanishes. The second (Fock) term is non-zero if $\mathbf{k}_1 + \mathbf{q} = \mathbf{k}_2$ and $\sigma_1 = \sigma_2$. Thus

$$H \cong H_{\text{Stoner}} := \sum_{\mathbf{k}\sigma} \epsilon_\mathbf{k} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} - \frac{U}{2N} \sum_{\substack{\mathbf{k}_1\mathbf{k}_2 \\ \mathbf{k}_1\neq\mathbf{k}_2}} \sum_\sigma \left(n_{\mathbf{k}_2\sigma} a_{\mathbf{k}_1\sigma}^\dagger a_{\mathbf{k}_1\sigma} + n_{\mathbf{k}_1\sigma} a_{\mathbf{k}_2\sigma}^\dagger a_{\mathbf{k}_2\sigma} - n_{\mathbf{k}_1} n_{\mathbf{k}_2} \right) \tag{10.20}$$

with

$$n_{\mathbf{k}\sigma} := \left\langle a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} \right\rangle. \tag{10.21}$$

The restriction $\mathbf{k}_1 \neq \mathbf{k}_2$ becomes irrelevant for large N. Denoting the average total number of spin-σ electrons by $N_\sigma := \sum_\mathbf{k} n_{\mathbf{k}\sigma}$, we get

$$H_{\text{Stoner}} = \sum_{\mathbf{k}\sigma} \epsilon_\mathbf{k} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} - \frac{U}{N} \sum_{\mathbf{k}\sigma} N_\sigma a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} + \frac{U}{2N} N_\sigma^2$$

$$= \sum_{\mathbf{k}\sigma} \epsilon_\mathbf{k} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} - \frac{U}{2N} \sum_\mathbf{k} \left[(N_\uparrow - N_\downarrow)(a_{\mathbf{k}\uparrow}^\dagger a_{\mathbf{k}\uparrow} - a_{\mathbf{k}\downarrow}^\dagger a_{\mathbf{k}\downarrow}) + (N_\uparrow + N_\downarrow)(a_{\mathbf{k}\uparrow}^\dagger a_{\mathbf{k}\uparrow} + a_{\mathbf{k}\downarrow}^\dagger a_{\mathbf{k}\downarrow}) \right]$$

$$+ \frac{U}{4N} \left[(N_\uparrow - N_\downarrow)^2 + (N_\uparrow + N_\downarrow)^2 \right]. \tag{10.22}$$

Now we assume a fixed total electron number $N_e := N_\uparrow + N_\downarrow$. We can thus replace the operator $\sum_{\mathbf{k}\sigma} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma}$ by N_e,

$$H_{\text{Stoner}} = \sum_{\mathbf{k}\sigma} \epsilon_\mathbf{k} a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} - \frac{U}{2N}(N_\uparrow - N_\downarrow) \sum_\mathbf{k} (a_{\mathbf{k}\uparrow}^\dagger a_{\mathbf{k}\uparrow} - a_{\mathbf{k}\downarrow}^\dagger a_{\mathbf{k}\downarrow}) + \frac{U}{4N}(N_\uparrow - N_\downarrow)^2 - \frac{U}{4N} N_e^2$$

$$= \sum_{\mathbf{k}\sigma} \left[\epsilon_\mathbf{k} - \sigma \frac{U}{2N}(N_\uparrow - N_\downarrow) \right] a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} + \frac{U}{4N}(N_\uparrow - N_\downarrow)^2 - \frac{U}{4N} N_e^2. \tag{10.23}$$

We define

$$k_B\theta := \frac{N_e}{2N} U. \tag{10.24}$$

This quantity is evidently the filling fraction multiplied by the Coulom interaction strength. Then

$$H_{\text{Stoner}} = \sum_{\mathbf{k}\sigma} \left(\epsilon_\mathbf{k} - \sigma k_B\theta \frac{N_\uparrow - N_\downarrow}{N_e} \right) a_{\mathbf{k}\sigma}^\dagger a_{\mathbf{k}\sigma} + \frac{k_B\theta}{2N_e}(N_\uparrow - N_\downarrow)^2 - \frac{k_B\theta}{2} N_e; \tag{10.25}$$

this Hamiltonain defines the *Stoner model*. H_{Stoner} describes electrons in an effective field $B_{\text{eff}} = 2/(g\mu_B)k_B\theta(N_\uparrow - N_\downarrow)/N_e$. This field or the polarization $(N_\uparrow - N_\downarrow)/N_e$ has to be determined selfconsistently. In the ground state this is done by minimizing the energy

$$E = \langle H_{\text{Stoner}}\rangle = N \int_{-\infty}^{E_F} d\epsilon\,\epsilon\, D\left(\epsilon + k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e}\right) + N \int_{-\infty}^{E_F} d\epsilon\,\epsilon\, D\left(\epsilon - k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e}\right)$$
$$+ \frac{k_B\theta}{2N_e}(N_\uparrow - N_\downarrow)^2 - \frac{k_B\theta}{2}N_e, \tag{10.26}$$

where $D(\epsilon) = m^{3/2}/(\sqrt{2}\pi^2\hbar^3)\sqrt{\epsilon}\,\Theta(\epsilon)$ is the density of states per spin direction and per site of the three-dimensional free electron gas. With

$$E_{F\sigma} := E_F + \sigma k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e} \tag{10.27}$$

we obtain

$$E = N \int_0^{E_{F\uparrow}} d\epsilon\left(\epsilon - k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e}\right)D(\epsilon) + N \int_0^{E_{F\downarrow}} d\epsilon\left(\epsilon + k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e}\right)D(\epsilon)$$
$$+ \frac{k_B\theta}{2N_e}(N_\uparrow - N_\downarrow)^2 - \frac{k_B\theta}{2}N_e$$
$$= N\,\frac{m^{3/2}}{\sqrt{2}\pi^2\hbar^3}\,\frac{2}{15}\left[E_{F\uparrow}^{3/2}\left(3E_{F\uparrow} - 5k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e}\right) + E_{F\downarrow}^{3/2}\left(3E_{F\downarrow} + 5k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e}\right)\right]$$
$$+ \frac{k_B\theta}{2N_e}(N_\uparrow - N_\downarrow)^2 - \frac{k_B\theta}{2}N_e. \tag{10.28}$$

Also, the electron numbers are

$$N_\sigma = N\int_{-\infty}^{E_F} d\epsilon\,\epsilon\, D\left(\epsilon + k_B\theta\,\frac{N_\uparrow - N_\downarrow}{N_e}\right) = N\int_0^{E_{F\sigma}} d\epsilon\,D(\epsilon) = N\,\frac{m^{3/2}}{\sqrt{2}\pi^2\hbar^3}\,\frac{2}{3}\,E_{F\sigma}^{3/2}. \tag{10.29}$$

Note that in the paramagnetic state we have

$$N_e = 2N\int_0^{E_F^{\text{para}}} d\epsilon\,D(\epsilon) = N\,\frac{m^{3/2}}{\sqrt{2}\pi^2\hbar^3}\,\frac{4}{3}\,(E_F^{\text{para}})^{3/2}. \tag{10.30}$$

so that we can write, now in the ferromagnetic state,

$$N_\sigma = \frac{N_e}{2}\,\frac{E_{F\sigma}^{3/2}}{(E_F^{\text{para}})^{3/2}}. \tag{10.31}$$

This implies

$$E_{F\sigma} = E_F^{\text{para}}\left(\frac{2N_\sigma}{N_e}\right)^{2/3} \tag{10.32}$$

and

$$E_F = \frac{E_{F\uparrow} + E_{F\downarrow}}{2} = \frac{E_F^{\text{para}}}{2^{1/3}}\left[\left(\frac{N_\uparrow}{N_e}\right)^{2/3} + \left(\frac{N_\downarrow}{N_e}\right)^{2/3}\right]. \tag{10.33}$$

Inserting this into E we obtain

$$E = \frac{3}{4}\,\frac{N_e}{(E_F^{\text{para}})^{3/2}}\left[\frac{2}{5}(E_F^{\text{para}})^{5/2}\,2^{5/3}\left(\left(\frac{N_\uparrow}{N_e}\right)^{5/3} + \left(\frac{N_\downarrow}{N_e}\right)^{5/3}\right) - \frac{2}{3}\,\frac{2}{N_e}(E_F^{\text{para}})^{3/2}k_B\theta\,\frac{(N_\uparrow - N_\downarrow)^2}{N_e}\right]$$
$$+ \frac{k_B\theta}{2N_e}(N_\uparrow - N_\downarrow)^2 - \frac{k_B\theta}{2}N_e. \tag{10.34}$$

Thus the energy per electron is

$$\frac{E}{N_e} = \frac{3}{5}\,2^{2/3}E_F^{\text{para}}\left[\left(\frac{N_\uparrow}{N_e}\right)^{5/3} + \left(\frac{N_\downarrow}{N_e}\right)^{5/3}\right] - k_B\theta\left(\frac{N_\uparrow}{N_e} - \frac{N_\downarrow}{N_e}\right)^2 + \frac{k_B\theta}{2}\left(\frac{N_\uparrow}{N_e} - \frac{N_\downarrow}{N_e}\right)^2 - \frac{k_B\theta}{2}. \tag{10.35}$$

Defining the polarization

$$\zeta := \frac{N_\uparrow - N_\downarrow}{N_e} \tag{10.36}$$

we find

$$\frac{N_\uparrow}{N_e} = \frac{1+\zeta}{2} \quad \text{and} \quad \frac{N_\downarrow}{N_e} = \frac{1-\zeta}{2}. \tag{10.37}$$

The energy per electron then reads

$$\frac{E}{N_e} = \frac{3}{5} 2^{2/3} E_F^{\text{para}} \left[\left(\frac{1+\zeta}{2}\right)^{5/3} + \left(\frac{1-\zeta}{2}\right)^{5/3} \right] - \frac{k_B\theta}{2}\zeta^2 - \frac{k_B\theta}{2}$$

$$= \frac{3}{5} \frac{E_F^{\text{para}}}{2} \left[(1+\zeta)^{5/3} + (1-\zeta)^{5/3} \right] - \frac{k_B\theta}{2}\zeta^2 - \frac{k_B\theta}{2}. \tag{10.38}$$
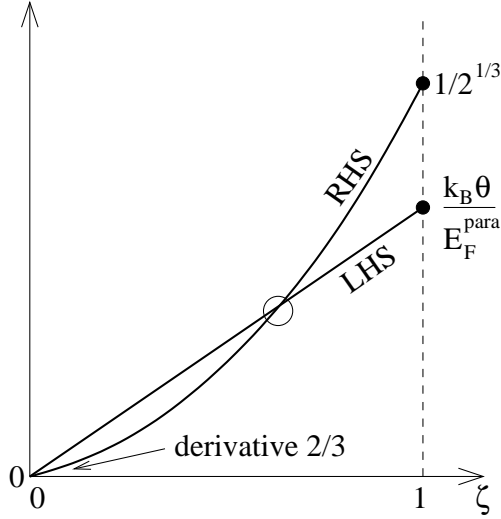
A local minimum not at the edges ($\zeta = \pm 1$) is found from the derivative,

$$0 = \frac{\partial}{\partial \zeta} \frac{E}{N_e} = \frac{E_F^{\text{para}}}{2} \left[(1+\zeta)^{2/3} + (1-\zeta)^{2/3} \right] - k_B\theta\zeta, \tag{10.39}$$

which leads to

$$\frac{k_B\theta}{E_F^{\text{para}}}\zeta = \frac{(1+\zeta)^{2/3} - (1-\zeta)^{2/3}}{2}. \tag{10.40}$$

Note that the right-hand expression has slope 2/3 at $\zeta = 0$. We plot both sides of this equation:

There are three cases:

1. for $k_B\theta/E_F^{\text{para}} < 2/3$: no intersection for $\zeta \neq 0$, the only solution is $\zeta = 0$, non-magnetic state,

2. $2/3 < k_B\theta/E_F^{\text{para}} < 1/2^{1/3}$: intersection for $0 < |\zeta| < 1$, partially polarized ferromagnetic state,

3. for $k_B\theta/E_F^{\text{para}} > 1/2^{1/3}$: One can see from E/N_e that the energy has a minimum at the edge $|\zeta| = 1$; the derivative does not vanish there. This is a completely polarized ferromagnetic state, i.e., only spin-$\uparrow$ electrons are present.

The condition for ferromagnetic order, be it partial or complete, reads

$$\frac{k_B\theta}{E_F^{\text{para}}} > \frac{2}{3}. \tag{10.41}$$

With $k_B\theta = N_e/(2N)U$, $D(E_F^{\text{para}}) = m^{3/2}/(\sqrt{2}\pi^2\hbar^3)\sqrt{E_F^{\text{para}}}$ and $N_e = Nm^{3/2}/(\sqrt{2}\pi^2\hbar^3)4/3(E_F^{\text{para}})^{3/2}$ we can rewrite this as

$$\left(\frac{N_e}{2N}U\right) \Big/ \left(\frac{3}{4N}\frac{N_e}{D(E_F^{\text{para}})}\right) = \frac{2}{3}D(E_F^{\text{para}})U > \frac{2}{3} \tag{10.42}$$

and thus as the very compact inequality

$$D(E_F^{\text{para}})U > 1. \tag{10.43}$$

This is the famous *Stoner criterion* for the occurance of ferromagnetic order in the Hubbard model. It is essentially independent of the dispersion realtion as long as the appropriate density of states is used. Note that ferromagnetism occurs for *strong* interactions.

There are examples for all three cases in real compounds: 1. Platinum just belongs to the first case—the local Coulomb interaction is rather strong but not strong enough to cause ferromagnetism. 2. The ferromagnetic transition metals iron, cobalt, and nickel belong to this case. One has to keep in mind that the underlying Hubbard model is a caricature of real materials, though. 3. The compounds CrO_2 and EuB_6 are completely polarized ferromagnetic metals.

10.3 Stoner excitations

After discussing the ground state in the previous section, we now turn to the low-energy excitations of metallic ferromagnets. We will give a partly qualitative presentation.

10.3.1 The particle-hole continuum

Let us consider excitations that decrease the z-component of the total spin by unity. (If the ground state is not completely polarized there are also excitations that increase it, as we shall see.) For a periodic lattice, the lattice momentum $\mathbf{q}$ in conserved and we can therefore choose the excited states to have sharp $\mathbf{q}$. The simplest excitations with spin reduced by one and with momentum $\mathbf{q}$ are particle-hole-excitations of the type

$$a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k},\uparrow}|\mathrm{GS}\rangle, \tag{10.44}$$

where $|\mathrm{GS}\rangle$ is the ground state. If this state exists, i.e., if $a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}|\mathrm{GS}\rangle \neq 0$, it is indeed an eigenstate of the Stoner Hamiltonian:

$$H_{\mathrm{Stoner}}\, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}|\mathrm{GS}\rangle = \left[H_{\mathrm{Stoner}}, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}\right] + a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow} H_{\mathrm{Stoner}}|\mathrm{GS}\rangle$$

$$= \left[H_{\mathrm{Stoner}}, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}\right] + E_{\mathrm{GS}}\, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}|\mathrm{GS}\rangle, \tag{10.45}$$

where

$$\left[H_{\mathrm{Stoner}}, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}\right] = \sum_{\mathbf{k}'} \left\{ (\epsilon_{\mathbf{k}'} - k_B\theta\zeta) \left[a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{k}'\uparrow}, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}\right] + (\epsilon_{\mathbf{k}'} + k_B\theta\zeta) \left[a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}'\downarrow}, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}\right] \right\}$$

$$= \sum_{\mathbf{k}'} \left\{ (\epsilon_{\mathbf{k}'} - k_B\theta\zeta) \left(-\delta_{\mathbf{k}'\mathbf{k}} a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}\right) + (\epsilon_{\mathbf{k}'} + k_B\theta\zeta)\, \delta_{\mathbf{k}',\mathbf{k}+\mathbf{q}} a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow} \right\}$$

$$= (-\epsilon_{\mathbf{k}} + k_B\theta\zeta + \epsilon_{\mathbf{k}+\mathbf{q}} + k_B\theta\zeta)\, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}. \tag{10.46}$$

We thus get

$$H_{\mathrm{Stoner}}\, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}|\mathrm{GS}\rangle = (E_{\mathrm{GS}} + \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} + 2k_B\theta\zeta)\, a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}|\mathrm{GS}\rangle. \tag{10.47}$$

Note that flipping a single spin does not appreciably change the polarization ζ in the thermodynamic limit so that we need not recalculate ζ selfconsistently. The state actually exists if the single-electron orbital $(\mathbf{k},\uparrow)$ is occupied in the ground state and the orbital $(\mathbf{k}+\mathbf{q},\downarrow)$ is empty. (Otherwise the above equation is simply 0=0.) Allowed values of $\mathbf{k}$ leading to excites states $a^\dagger_{\mathbf{k}+\mathbf{q},\downarrow} a_{\mathbf{k}\uparrow}|\mathrm{GS}\rangle \neq 0$ are best found graphically, as we show now.

(a) Partial polarization:
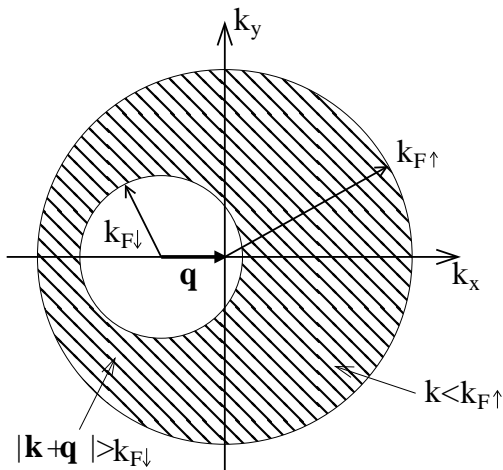
In the graph, $k_{F\sigma} = \sqrt{2mE_{F\sigma}}/\hbar$.

The excitation energy

$$\epsilon_{\mathbf{k+q}} - \epsilon_{\mathbf{k}} + 2k_B\theta\zeta = \frac{\hbar^2}{m}\mathbf{k}\cdot\mathbf{q} + \frac{\hbar^2 q^2}{2m} + 2k_B\theta\zeta \tag{10.48}$$

depends on $\mathbf{k}$ for given $\mathbf{q} \neq 0$. Therefore, for any $\mathbf{q} \neq 0$ we find a continuum of excitations, called the *particle-hole continuum*. Only for $\mathbf{q} = 0$, the excitation energy is sharp, $\Delta := 2k_B\theta\zeta > 0$. Excitations with vanishing energy exist when the two Fermi spheres above cross, since then we have at the crossing points

$$\epsilon_{\mathbf{k}} - k_B\theta\zeta = E_F \quad \text{and} \quad \epsilon_{\mathbf{k+q}} + k_B\theta\zeta = E_F \tag{10.49}$$

and thus

$$\epsilon_{\mathbf{k+q}} - \epsilon_{\mathbf{k}} + 2k_B\theta\zeta = \epsilon_{\mathbf{k+q}} + k_B\theta\zeta - (\epsilon_{\mathbf{k}} - k_B\theta\zeta) = E_F - E_F = 0. \tag{10.50}$$

Zero-energy excitations are dangerous, since they suggest an instability of the mean-field ground state. One has to go beyond Stoner theory to see that they do not destroy the ferromagnetic ground state in this case. A situation with crossing Fermi spheres is sketched here:
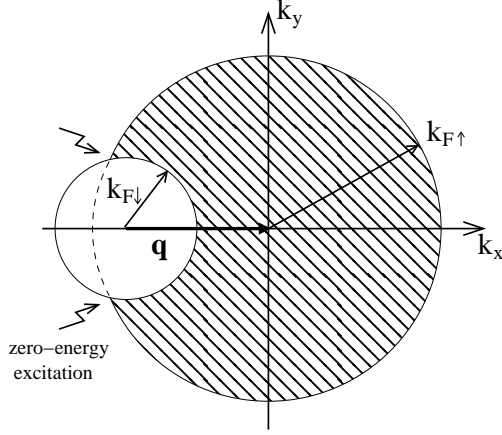

Note that in the crescend-shaped region to the left we find excitations that *increase* the total spin.

We can now sketch the dispersion of the particle-hole continuum for partial polarization:
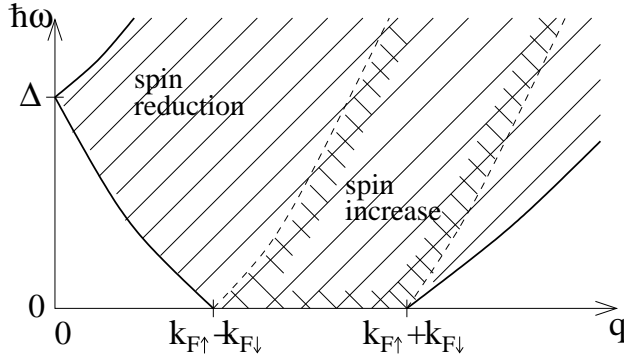
(b) Complete polarization: This case is in fact simpler since all spin-$\downarrow$ orbitals are empty in the ground state. Thus whenever $(\mathbf{k}, \uparrow)$ is occupied, i.e., $k \leq k_{F\uparrow}$, $a^\dagger_{\mathbf{k+q},\downarrow} a_{\mathbf{k}\uparrow}|\mathrm{GS}\rangle$ is an excited state. The excitation energy has a minimum when $a_{\mathbf{k}\uparrow}$ annihilates a spin-$\uparrow$ electron with the highest possible energy (E_F) and $a_{\mathbf{k+q},\downarrow}$ creates a spin-$\downarrow$ electron with the lowest possible energy ($\epsilon_{\mathbf{k}=0} + k_B\theta\zeta = k_B\theta\zeta$). The minumum excitation energy is thus $E_{\min} = k_B\theta\zeta - E_F$. Moreover, there are no excitations with increased spin since the spin is already maximal.
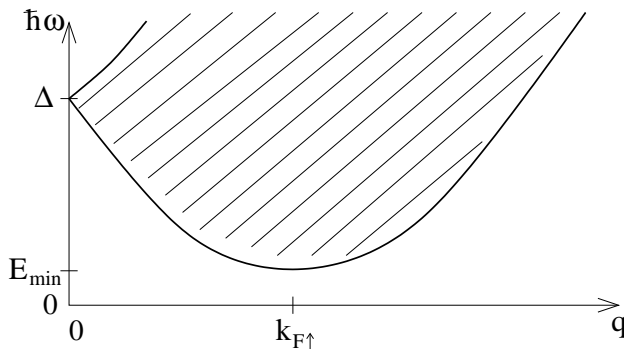
10.3.2 Spin waves and magnons

A uniform rotation of all spins should not cost any energy, regardless of whether the magnetic system is a metal or an insulator. Thus there should be a zero-energy Goldstone mode at $\mathbf{q} = 0$, which we evidently have not captured based on the Stoner Hamiltonian H_{Stoner}. We only discuss the physical ideas. One strategy is to calculate the dynamical susceptibility $\chi(\mathbf{q}, \omega)$, for example in the random-phase approximation (RPA) applied to the full Hubbard Hamiltonian. The linear response relation

$$\mathbf{M}(\mathbf{q}, \omega) = \chi(\mathbf{q}, \omega) B(\mathbf{q}, \omega) \tag{10.51}$$

shows that a pole in χ at a certain $(\mathbf{q}, \omega)$ permits the magnetization in this mode to be non-zero even in the absence of an applied field. Thus a wave-like excitation at this $(\mathbf{q}, \omega)$ can propagate. This is of course a spin wave and its quanta are magnons. Calculating $\chi(\mathbf{q}, \omega)$, one finds a single branch of $(\mathbf{q}, \omega = \omega_{\mathbf{q}})$ where χ diverges. $\hbar\omega_{\mathbf{q}}$ thus describes the spin-wave dispersion. It comes out as $\hbar\omega_{\mathbf{q}} \propto q^2$ for small $\mathbf{q}$, like for ferromagnetic insulators.

Note that the particle-hole continuum is also visible in $\chi(\mathbf{q}, \omega)$, specifically as a non-zero imaginary part $\operatorname{Im} \chi(\mathbf{q}, \omega) \neq 0$. Moreover, the spin waves do not survive as propagating modes when they are degenerate in energy with the particle-hole continuum. This is plausible: a magnon can here decay into many particle-hole excitations. This mechanism is called *Landau damping*. Hence, the full dispersion lools like this:
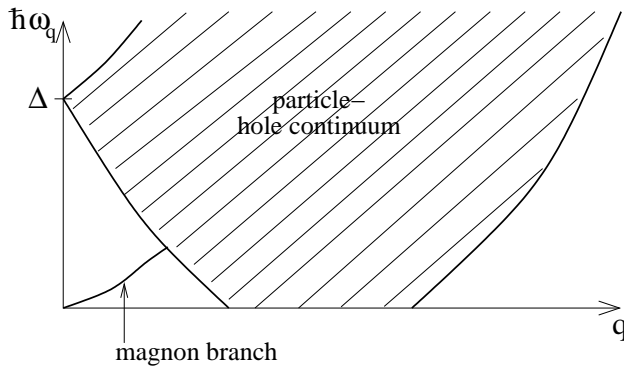
10.4 The *t-J* model

In Sec. 4.2 we have studied the half-filled Hubbard model for the case of weak hybridization, i.e., large U/t. The result was an *antiferromagentic* exchange interaction

$$J = -\frac{4t^2}{U} \tag{10.52}$$

for the nearest-neighbor model. We showed this explicitly for a dimer and stated that the result is in fact general. The result seems to contradict the prediction of Stoner theory of ferromagnetic order for $D(E_F^{\text{para}})U > 1$. For nearest-neighbor hopping, $D(E_F^{\text{para}})$ must be on the order of $1/t$ from dimensional analysis, thus the Stoner criterion becomes $U/t \gtrsim 1$. Our earlier result suggests that at half filling for $U/t \gg 1$, an antiferromagnetic phase will replace the Stoner ferromagnet. In the present section we consider this limit more carefully, at and away from half filling.

Since $t/U \ll 1$, we use pertubation theory for small t. Hence, we choose

$$V := U \sum_i n_{i\uparrow} n_{i\downarrow} \tag{10.53}$$

as the unperturbed Hamiltonian. The eigenstates of V are occupation-number states $|n_{1\uparrow}, n_{1\downarrow}, n_{2\uparrow}, n_{2\downarrow}, \dots\rangle$. We assume without loss of generality that the system is at most half filled,

$$N_e := \sum_{i\sigma} n_{i\sigma} \leq N. \tag{10.54}$$

If it is more than half filled, we can map the model onto the present case by means of a particle-hole transformation.

We divide the Fock space into two subspaces without and with doubly occupied sites:

$$S := \operatorname{span}\left\{ |n_{1\uparrow}, n_{1\downarrow}, \dots\rangle \,\middle|\, \forall i : n_{i\uparrow} + n_{i\downarrow} \leq 1 \right\}, \tag{10.55}$$

$$D := \operatorname{span}\left\{ |n_{1\uparrow}, n_{1\downarrow}, \dots\rangle \,\middle|\, \exists i : n_{i\uparrow} + n_{i\downarrow} = 2 \right\}. \tag{10.56}$$

Here, span B is the vector space spanned by the basis vectors in B, i.e., the space of all linear combinations. All states in D contain at least one doubly occupied site, the states in S do not. We further define projection operators P_S and P_D onto the subspaces S and D.

The hopping term

$$T := -\sum_{ij\sigma} t_{ij} a_{i\sigma}^\dagger a_{j\sigma} \tag{10.57}$$

is now treated as a pertubation. We write the full Hamiltonian $H = V + T$ in block form

$$H = \begin{pmatrix} P_S T P_S & P_S T P_D \\ P_D T P_S & P_D(V+T)P_D \end{pmatrix}. \tag{10.58}$$

We have used that $P_S V = V P_S = 0$ since V only contributes for doubly occupied sites.

The eigenstates of any Hamiltonian H are states for which the *resolvent*

$$G(E) := (E - H)^{-1} \tag{10.59}$$

has a singularity. (If $|\psi\rangle$ is an eigenstate to the eigenvalue E_ψ, then $G(E)|\psi\rangle = (E-H)^{-1}|\psi\rangle = (E-E_\psi)^{-1}|\psi\rangle$ and we find a pole at $E = E_\psi$.) The main idea is now to define an *effective Hamiltonian* H_{eff} for the low-energy degrees of freedom by the projection

$$P_S G(E) P_S =: P_S \left[E - H_{\text{eff}}(E)\right]^{-1} P_S. \tag{10.60}$$

The reasoning behind this is that doubly occupied sites cost the large energy U and are therefore unlikely at low energies. However, we cannot completely ignore them since states from D are mixed into states from S by the hopping term T.

The general formula for block matrices

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} \cdots \\ \cdots \quad\quad \cdots \end{pmatrix} \tag{10.61}$$

here gives

$$\begin{aligned} P_S G(E) P_S &= P_S \begin{pmatrix} E - P_S T P_S & -P_S T P_D \\ -P_D T P_S & E - P_D(V+T)P_D \end{pmatrix}^{-1} P_S \\ &= P_S\big(E - P_S T P_S - P_S T P_D[E - P_D(V+T)P_D]^{-1} P_D T P_S\big)^{-1} P_S. \end{aligned} \tag{10.62}$$

Thus we find

$$\begin{aligned} H_{\text{eff}}(E) &= P_S\big(T + T P_D[E - P_D T P_D - P_D V P_D]^{-1} P_D T\big) P_S \\ &\cong P_S\big(T + T P_D[E - P_D V P_D]^{-1} P_D T + T P_D[E - P_D V P_D]^{-1} P_D T P_D[E - P_D V P_D]^{-1} P_D T + \dots\big) P_S \\ &= P_S\big(T + T P_D[E - V]^{-1} P_D T + T P_D[E - V]^{-1} P_D T P_D[E - V]^{-1} P_D T + \dots\big) P_S. \end{aligned} \tag{10.63}$$

This is an expansion in T. We have used $[P_D, V] = 0$. We encounter the technical problem that $H_{\text{eff}}(E)$ depends on the energy. We now choose the zero of our energy scale as the true ground-state energy E_0. We do not know E_0 exactly but this is not required. E_0 itself is extensive, i.e., $E_0 \propto N$, but the excitation energies of low-lying excitations are not (compare particle-hole excitations in metals).

Now for any state $|\psi\rangle$, the state $P_D T P_S |\psi\rangle$ contains exactly one doubly occupied site. Thus we can replace V by U in the second-order term,

$$H_{\text{eff}} \cong P_S T P_S + P_S T P_D \frac{1}{E - U} P_D T P_S. \tag{10.64}$$

Furthermore, since E (relative to E_0) is small and intensive, we can neglect E compared to U and write

$$H_{\text{eff}} \cong P_S T P_S + P_S T P_D \frac{1}{U} P_D T P_S. \tag{10.65}$$

We write this so-called *t-J model* Hamiltonian more explicitly,

$$H_{t\text{-}J} = -P_S \sum_{ij\sigma} t_{ij} a_{i\sigma}^\dagger a_{j\sigma} P_S - P_S \frac{1}{U} \sum_{ijk} \sum_{\sigma\sigma'} t_{ij}\, a_{i\sigma}^\dagger a_{j\sigma}\, n_{j\uparrow} n_{j\downarrow}\, t_{jk}\, a_{j\sigma'}^\dagger a_{k\sigma'} P_S. \tag{10.66}$$

The factor of $n_{j\uparrow} n_{j\downarrow}$ comes from $P_D P_D \equiv P_D$ in Eq. (10.64); it is unity only if site j is doubly occupied and zero otherwise. The *t-J* model Hamiltonian can be rewritten as

$$H_{t\text{-}J} = P_S(T + H_{\text{Heisenberg}} + \Delta H) P_S \tag{10.67}$$

with

$$T = -\sum_{ij\sigma} t_{ij} a_{i\sigma}^\dagger a_{j\sigma}, \tag{10.68}$$

$$H_{\text{Heisenberg}} = -\frac{1}{2} \sum_{ij} J_{ij} \left(\mathbf{s}_i \cdot \mathbf{s}_j - \frac{n_i n_j}{4} \right), \tag{10.69}$$

$$\Delta H = -\frac{1}{2U} \sum_{ijk} t_{ij} t_{jk} \left[n_j \sum_\sigma a_{i\sigma}^\dagger a_{k\sigma} - 4\,\mathbf{s}_j \cdot \sum_{\sigma\sigma'} a_{i\sigma} \frac{\boldsymbol{\sigma}_{\sigma\sigma'}}{2} a_{k\sigma'} \right] \tag{10.70}$$

and

$$J_{ij} = -\frac{4t_{ij}^2}{U} \tag{10.71}$$

and the usual spin operators

$$\mathbf{s}_j = \sum_{\sigma\sigma'} a_{i\sigma} \frac{\boldsymbol{\sigma}_{\sigma\sigma'}}{2} a_{i\sigma'}. \tag{10.72}$$

We note that $H_{\text{Heisenberg}}$ describes a Heisenberg model for the *singly* occupied sites, except for the rather trivial term $-n_i n_j / 4$.

10.4.1 Half filling

At half filling, subspace S only contains states with $n_{i\uparrow} + n_{i\downarrow} \equiv n_i = 1$ for all sites i. Any hopping leads to the appearance of a doubly occupied site and an empty site so that $P_S T P_S = 0$ and also $P_S \Delta H P_S = 0$. Therefore, the low-energy physics is described by $H_{\text{Heisenberg}}$ alone, which now reads

$$H_{\text{Heisenberg}} = -\frac{1}{2} \sum_{ij} J_{ij} \left(\mathbf{s}_i \cdot \mathbf{s}_j - \frac{1}{4} \right), \tag{10.73}$$

and there is a spin $1/2$ at *every* site. The system is antiferromagnetic, though possibly frustrated if longer-range t_{ij} and J_{ij} contribute. Hopping is suppressed, i.e., the system is *insulating* at low energies. The insulating character results from the interaction—T alone describes a half-filled conduction band and thus a metal. This state of matter is called a *Mott* or *Mott-Hubbard insulator*.

10.4.2 Away from half filling

For $N_e < N$ ($N_e > N$ can be mapped onto $N_e < N$ as noted above) the *t-J* model contains on average $(N - N_e)/N$ holes per lattice site. The terms T and ΔH are important even for the low-energy properties. T is usually the more important term, since nearest-neighbor hopping frustrates the antiferromagnetic order in $d > 1$ dimensions:
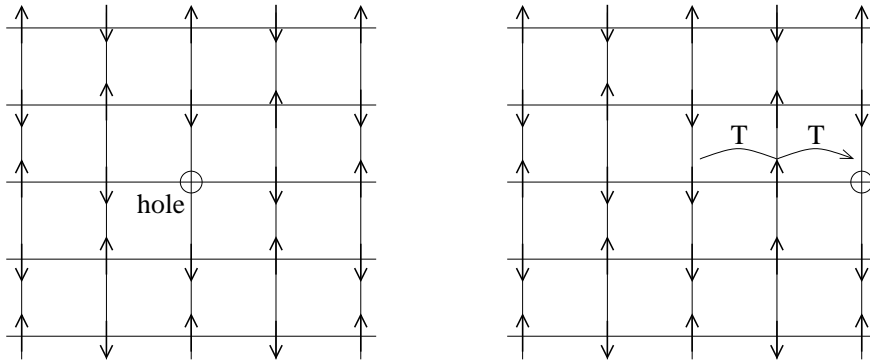
On the other hand, if t_{ij} only contains nearest-neighbor hopping, ΔH does *not* frustrate the order, since ΔH contains the product $t_{ij} t_{jk}$ describing next-nearest-neighbor hopping. ΔH is often neglected in practice. Note that it has been suggested and tested by numerical calculations but not rigorously shown that the ground state of the *t-J* model in a certain range of hole concentrations is a superconductor.

In one dimension, T does not frustrate the order. Compare our discussion of spinons in Sec. 8.4. In the one-dimensional *t-J* model with holes we therefore find *spin-charge-separation*: the charge of the hole and its spin $1/2$ can propagate independently—they are deconfined.

hole

↑ ↓ ↑ ↓ ○ ↓ ↑ ↓ ↑ ↓ ↑

↑ ↓ ↑ ↓ ↓ ○ ↑ ↓ ↑ ↓ ↑

↑ ↓ ↑ ↓ ↓ ↑ ○ ↓ ↑ ↓ ↑

...

↑ ↓ ↓ ↑ ↓ ↑ ↓ ↑ ↓ ○ ↑

kink, "spinon"   holon
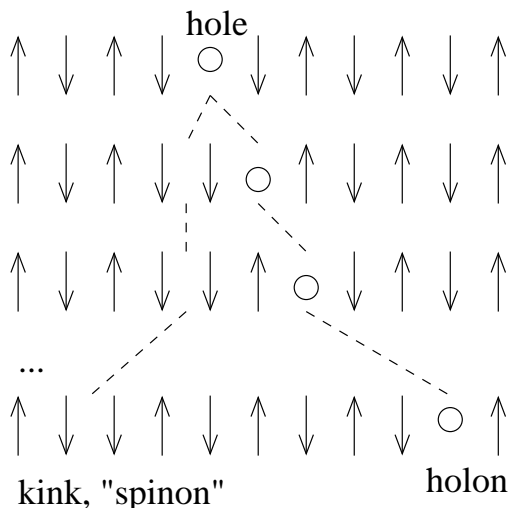
The spin excitation is clearly just the kink we have discussed in Sec. 8.4, the superositions of kinks make up the *spinons*. The charge excitation indicated by ○ does not carry a spin; the total spin in a symmetric interval centered at the charge is zero. This excitation is called a *holon*.

10.5   Nagaoka ferromagnetism

We have seen that mobile holes doped into a half-filled t-J model in two or three dimensions frustrate the antiferromagnetic order. We might ask how effectively the holes suppress antiferromagnetism. Experimentally, the CuO_2 planes in the cuprates, which are reasonably well described by a two-dimensional t-J model, loose antiferromagnetic order for only 2% of hole doping, i.e. 0.02 holes per lattice unit.

Recall that we have introduced the t-J model as an effective low-energy description of the Hubbard model. We now return to the original Hubbard model. There is an exact result for this case:

Nagaoka's theorem: The ground states of the Hubbard model on a bipartite lattice in the limit $U \to \infty$, described by

$$H_{\text{eff}} = -P_S \sum_{ij\sigma} t_{ij} a_{i\sigma}^\dagger a_{j\sigma} P_S, \tag{10.74}$$

with a single hole relative to half filling, $N_e = N - 1$, are completely polarized ferromagnets.

The proof, given by Y. Nagaoka, Phys. Rev. **147**, 392 (1966), is rather involved and is not presented here. We only make a few remarks.

The total spin is $S_{\text{tot}} = N_e/2 = (N-1)/2$. One of the ground states is

$$|S_{\text{tot}}, S_{\text{tot}}\rangle \propto \sum_i (-1)^i \prod_{j \neq i} a_{j\uparrow}^\dagger |0\rangle, \tag{10.75}$$

where $|0\rangle$ is the vacuum. All other ground states are obtained by spin rotation. Note that $|S_{\text{tot}}, S_{\text{tot}}\rangle$ is a superposition of states with a hole at site i and spin-up electrons at all other sites $j \neq i$.

Nagaoka's theorem is a rather peculiar result; the half-filled model for $U/t \gg 1$ is an antiferromagnet (in $d > 1$ dimensions) but introduction of a *single* hole into a macroscopic system changes the ground state completely. For large U, holes are thus extremely effective in destroying antiferromagnetism in the Hubbard model. Unfortunately it has so far not been possible to extend the theorem to finite U or to $N_e < N - 1$.

For this more general situation, exact numerical results only exist for very small systems. The reuslt of a mean-field theory in two dimensions is sketched here:
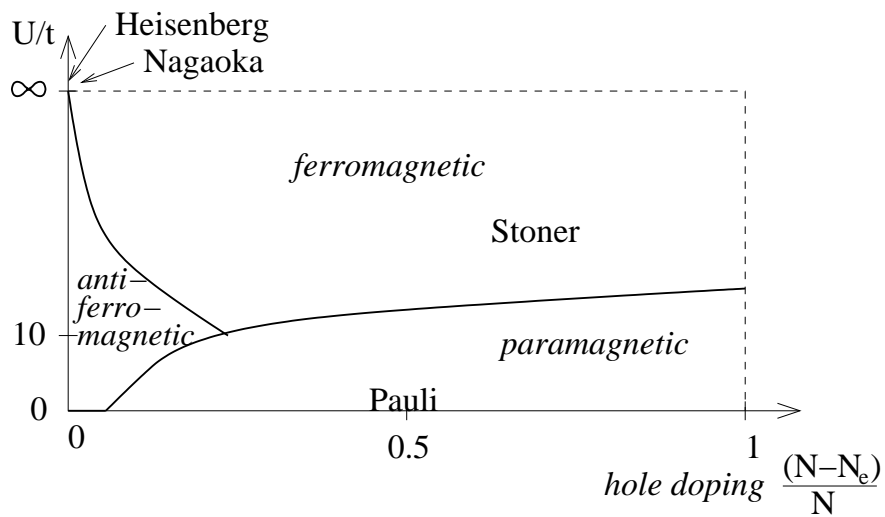
See Nielsen and Bhatt, arXiv:0907.3671. Although this phase diagram follows from mean-field theory, it appears to be qualitatively correct in that for finite U/t, the antiferromagnetic ground state survives over an extended hole-doping range. However, both experiments and more advanced numerical calculations indicate that it is then replaced by a state with short-range antiferromagnetic and possibly charge-density order.