Quantum Magnetism

Jonathan Keeling

http://www.tcm.phy.cam.ac.uk/~jmjk2/qm/

Contents

List of Figures

Introduction

These lectures aim to provide an overview on how a quantum description of electrons is responsible for magnetism in solid state systems. This subject is a vast subject, since to understand the correlations of electron spins, it is necessary to understand the correlations of electrons, and thus all possible phases of interacting electrons, structural phase transitions in crystals and so on. The aim of this course is not so ambitious, but merely to provide an overview of the main mechanisms leading to magnetic order (antiferromagnetism, ferromagnetism, ferrimagnetism); to indicate the relation between idealised theorists models and real experimental systems; and to connect such microscopic models of magnetism to the phenomenology of temperature dependent magnetisation and susceptibility relevant to continuum electrodynamics.

For the majority of the course, the aim will be to understand what magnetic interactions (interactions between magnetic spins) exist, and what ground states they lead to. In some cases, to understand the stability of various ground states, or to improve upon a first estimate, excitations will be considered, and at the end of the course, excitations will be relevant in relating the ground state structure to finite temperature susceptibilities, however it remains the case that the first task — that of understanding what interactions exist, and what ground state they lead to — is sufficiently involved to fill most of the course.

Questions exist both to illustrate the ideas, and to provide mathematical derivations of some statements not included in the lectures; you are strongly advised to attempt the questions marked with an asterisk as the results of these will be used in the next lecture.

Books and Review articles

> He could not conceive but that some one, somewhere, could tell him all about the magnet, if one could but find the book.
>
> *The education of Henry Adams,*
> *Twilight (1901)*
> HENRY ADAMS

There are a wide range of books on the subject, although most have quite overlapping content. Books generally divide into two classes; those which aim to study electron correlation to discover the magnetic properties of materials, and those which use magnetism as an example to study theoretical techniques in correlated electron systems:

Magnetic properties of materials

Fazekas [1] This is a very comprehensive discussion of almost all topics mentioned in these lectures, including a comprehensive bibliography of related works.

Yosida [2] This is a shorter, but still reasonably comprehensive book, but one which covers all the topics addressed in this course. It is often more focused on deriving the Hamiltonian, i.e. understanding the nature of the magnetic interactions, than on finding the resultant state, and requires some effort to extract the coherent picture of how different effects relate.

Mattis [3],[4] Mattis [3] is a two-volume book; the first volume deals mainly with models of magnetism, and the second with finite temperatures. It contains an illuminating and entertaining discussion of the history of magnetism, as well as a careful discussion of conditions under which ferromagnetism can be possible. Mattis [4] is an updated, single volume summary of the earlier textbook.

Anderson [5] This is a long review article discussing insulators, and in particular the idea of superexchange, but it does also include a general introduction to magnetic states of insulators.

Strongly correlated electron system

Auerbach [6] This is a rather formal, although very clear, discussion primarily of Hubbard and Heisenberg models and their extensions, and of how to produce approximation schemes for wavefunctions and correlation functions, with a focus on path integrals. It makes few connections to real materials, but instead uses magnetism as an example to discuss various fundamental properties of ordering in low dimensions, gapped and gapless spectra, and topological differences between integer and half integer spin.

Nagaosa [7] From this book, it is primarily chapter 4, discussing the Kondo effect that is relevant to this course, although other chapters, particular 3 and 5, in which hole doping of Mott insulators can be relevant. It also contains a detailed discussion of one dimensional systems, which these lectures do not discuss.

Lecture 1

Quantum magnetism and exchange

1.1 Need for Quantum Exchange

Why magnetism is quantum

The word “quantum” in the title of these lectures is almost redundant, as a simple argument — the Bohr – van Leuwen theorem — demonstrates. Neglecting classical spin (as spin is really a quantum concept itself), magnetism is associated with circulating currents. Thus, one may ask whether a magnetic moment:

$$\langle M \rangle = \langle \sum_i q_i \mathbf{v}_i \times \mathbf{r}_i \rangle \tag{1.1}$$

can exist, where expectations are with respect to classical statistical mechanical distributions:

$$\begin{aligned} \langle X \rangle &= \frac{1}{\mathcal{Z}} \iint d^d p d^d r X \exp\left[-\frac{\mathcal{H}(\{\mathbf{p}_i, \mathbf{r}_i\})}{k_B T} \right] & (1.2) \\ \mathcal{H}(\{\mathbf{p}_i, \mathbf{r}_i\}) &= \sum_i \left[\frac{(\mathbf{p}_i + q_i \mathbf{A}(\mathbf{r}_i))^2}{2m_i} + q_i \phi(\mathbf{r}_i) \right] + U(\{\mathbf{r}_1, \ldots, \mathbf{r}_N\}). & (1.3) \end{aligned}$$

Here, the only assumption made is that all interactions between particles depend only on their positions, and not velocities. By making a change of integrand from the canonical momentum $\mathbf{p}_i$ to the velocity, $\mathbf{v}_i = (\mathbf{p}_i + q_i \mathbf{A}(\mathbf{r}_i))/m_i$, one can straightforwardly see that the expectation $\langle \mathbf{v}_i \rangle$ vanishes by symmetry, and hence so does the magnetic moment. Hence, even in the presence of a magnetic field, a classical system does not possess a magnetic moment due to this symmetry. Quantum mechanically this problem is cured by replacing integration by summation.

Why magnetostatic interactions are not enough

Having discovered that something quantum mechanical is required, the next guess might be that one should consider atomic moments arising from electronic spins and electronic motion, and then consider the magnetostatic

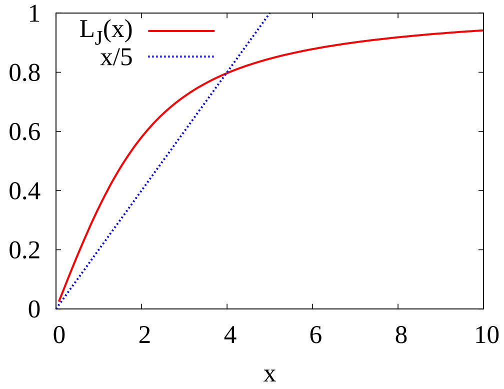


Figure 1.1: Brillouin function compare to a linear function of magnetisation

interaction between them. The first part of this idea is correct, but the second idea — magnetostatic interactions — is not sufficient to describe real magnetic interactions, and in fact magnetostatic interactions are generally irrelevant compared to the real interactions present.

Let us consider a magnetic ion, having spin J, and hence magnetic moment $J\mu_B$ in an effective magnetic field, $H_{\text{eff}} = H + \alpha M$, where αM describes the molecular field arising from the magnetisation of other ions. In such a field, one can use the Brillouin function, $L_J(x)$ [1] to write the magnetisation for a density N/V of such ions as:

$$M = \frac{N}{V} J\mu_B L_J \left(\frac{\mu_0 H_{\text{eff}} J \mu_B}{k_B T} \right) \tag{1.5}$$

This approach indeed will allow for a self-consistent magnetisation at zero field, as long as the right hand side has a steeper initial gradient than the left hand side, as seen in Fig. 1.1. As temperature increases, the gradient of the right hand side will decrease, and so above some temperature, ferromagnetism goes away, but a modified paramagnetic susceptibility survives. By comparing the form this paramagnetic susceptibility to experimental values, one may find the required value of α. For the linear paramagnetic susceptibility, we may assume that both M and H are small, and so expand L_J for small argument to give:

$$M \simeq \frac{N}{V} J\mu_B \frac{1}{3} \frac{J+1}{J} \frac{\mu_0 (H + \alpha M) J \mu_B}{k_B T} \tag{1.6}$$

writing $\langle \mu^2 \rangle = J(J+1)\mu_B/3$, we may rewrite this as:

$$M \left(1 - \alpha \frac{N}{V} \frac{\mu_0 \langle \mu^2 \rangle}{k_B T} \right) = \frac{N}{V} \frac{\mu_0 \langle \mu^2 \rangle}{k_B T} H. \tag{1.7}$$

---

[1] The Brillouin function $L_J(x)$ is defined by $L_J(x) = \sum_{m=-J}^{J} (m/J) e^{(m/J)x} / \mathcal{Z}$, and can be shown to be given by

$$L_J(x) = \frac{2J+1}{2J} \coth \left(\frac{2J+1}{2J} x \right) - \frac{1}{2J} \coth \left(\frac{x}{2J} \right) \tag{1.4}$$

By identifying the Curie-Weiss susceptibility $\chi = C/[T - T_c]$ via:

$$M = \chi H = \frac{C}{T - T_C} H \quad \Leftrightarrow \quad M\left(1 - \frac{T_C}{T}\right) = C\frac{H}{T} \tag{1.8}$$

we can identify the Curie temperature T_C in the Curie Weiss law as:

$$k_B T_c = \alpha \frac{N}{V}\langle \mu^2 \rangle \mu_0 \sim \alpha\mu_0 \left\langle \frac{\mu \cdot \mu}{r_{sep}^3} \right\rangle \tag{1.9}$$

where the last expression has identified the number density with average separation, showing that α acts as a numerical prefactor of some effective interaction between magnetic dipoles. For Iron, $T_C = 10^3 K$, and using the mass density of iron as 9000kg m^{-3}, one finds the required $\alpha \simeq 10^4$.

Let us briefly review what magnetostatic effective interaction one might expect to find — i.e. how large an α arises from Maxwell's equations alone. This is most simply done by the magnetic analogy of the Clausius-Mossotti cavity argument; since dipoles inside a spherical cavity can be shown to contribute no net field, the local field is given by the sum of the field from the sphere (which vanishes), and that from the bulk missing a sphere. Thus, using scalar magnetic potential:

$$\phi_M = \begin{cases} H_{\text{eff}}\; r\cos(\theta) & r < a \\ H \quad r\cos(\theta) + \frac{C}{r^2}\cos(\theta) & r > a \end{cases} \tag{1.10}$$

one finds that matching boundary conditions yields:

$$H_{\|}: \quad H_{\text{eff}} = H + \frac{C}{a^3}, \qquad B_{\perp}: \quad H_{\text{eff}} = H - 2\frac{C}{a^3} + M \tag{1.11}$$

produces $H_{\text{eff}} = H + M/3$, i.e. $\alpha = 1/3 \ll 10^4$. Thus, it is clear that the contribution of magnetostatic interactions to alignment between dipoles is not only insufficient, but is also negligible compared to the real source of interactions — exchange interactions, which will be the subject of the rest of this lecture.

There do exist some exceptions to the above statement; for magnetostatic dipolar interactions to be relevant, one requires a combination of several factors: large magnetic moments, so that the dipolar interaction is as large as possible, and frustration, so that exchange interactions cannot uniquely define a ground state.

1.2 Quantum Exchange

The interactions which provide such strong interactions between magnetic moments are nothing to do with the magnetic nature of the spin degree of freedom, but are to do with its symmetry properties. They exist because of the connection between symmetry of spin and spatial wavefunctions for electrons as fermions, and so the energy scales associated with these interactions are associated with the energy scales of the spatial degrees of freedom; i.e. electronic kinetic energy and Coulomb interaction. To summarise briefly; a given spin configuration restricts the range of possible spatial configurations, and because different spatial configurations can have significantly different energies, there is thus a large energy associated with different spin configurations.

Within this general framework, there exist two categories of types of magnetic interaction:

Direct (potential) exchange This is driven by minimising potential energy, by reducing wavefunction overlap. Wavefunction overlap is reduced by adding nodes to the wavefunction, producing antisymmetric spatial wavefunctions, and so favours symmetric spins, i.e. **ferromagnetic** interactions. This case arises when electrons occupy wavefunctions that overlap in space.

Kinetic exchange [2] This is driven by minimising kinetic energy, by reducing gradients of wavefunctions, i.e. allowing delocalisation of electrons. This corresponds to using symmetric superpositions of wavefunctions, and so favours antisymmetric spins, i.e. **antiferromagnetic** interactions. This case generally arises for localised electronic orbitals

There is a potential confusion in the above; these two effects compete, and it is not immediately clear whether in a given situation it is more important to delocalise electrons, or prevent their overlap with other electrons. We will address this point in Sec. 1.3, by showing cases where ferromagnetism is superficially expected, but does not occur, and thus finding that there are stricter criterion on when ferromagnetism may arise.

Minimising interaction energy: direct exchange

Let us consider the Coulomb interaction, written between fermions in some particular choice of second quantised basis:

$$H_{\text{Coulomb}} = \sum_{\sigma,\sigma'} \sum_{u,v,w,x} \frac{1}{2} \left\langle u, v \left| \frac{e^2}{4\pi\epsilon r_{12}} \right| w, x \right\rangle a^\dagger_{u\sigma} a^\dagger_{v\sigma'} a_{w\sigma'} a_{x\sigma}. \tag{1.12}$$

If we assume the spatial states are relatively localised, then for the expectation of the Coulomb term not to vanish, it is necessary that the left and right states have some overlap. This leaves two kinds of terms: direct, $u, v \rightarrow u, v$ and exchange $u, v \rightarrow v, u$, giving:

$$\begin{aligned} H_{\text{Coulomb}} = \frac{1}{2} \sum_{\sigma,\sigma'} & \left\langle u, v \left| \frac{e^2}{4\pi\epsilon r_{12}} \right| u, v \right\rangle a^\dagger_{u\sigma} a_{u\sigma} a^\dagger_{v\sigma'} a_{v\sigma'} \\ & - \left\langle u, v \left| \frac{e^2}{4\pi\epsilon r_{12}} \right| v, u \right\rangle a^\dagger_{u\sigma} a_{u\sigma'} a^\dagger_{v\sigma'} a_{v\sigma} \end{aligned} \tag{1.13}$$

---

[2] This effect is also sometimes referred to as superexchange; we will not use this terminology, as we will use superexchange to refer to a generalisation of this idea, involving hopping via intermediate atoms, to be discussed in Sec. 3.2.

The first term just sums over population of spin states. For the second, taking the two spin states, ↑, ↓ yields:

$$\begin{aligned} a^\dagger_{u\sigma} a_{u\sigma'} a^\dagger_{v\sigma'} a_{v\sigma} &= \frac{1}{2}\left(a^\dagger_{u\uparrow} a_{u\uparrow} + a^\dagger_{u\downarrow} a_{u\downarrow}\right)\left(a^\dagger_{v\uparrow} a_{v\uparrow} + a^\dagger_{v\downarrow} a_{v\downarrow}\right) \\ &+ \frac{1}{2}\left(a^\dagger_{u\uparrow} a_{u\uparrow} - a^\dagger_{u\downarrow} a_{u\downarrow}\right)\left(a^\dagger_{v\uparrow} a_{v\uparrow} - a^\dagger_{v\downarrow} a_{v\downarrow}\right) \\ &+ a^\dagger_{u\uparrow} a_{u\downarrow} a^\dagger_{v\downarrow} a_{v\uparrow} + a^\dagger_{u\downarrow} a_{u\uparrow} a^\dagger_{v\uparrow} a_{v\downarrow}. \end{aligned}$$

Writing spin operators as: $\mathbf{s}_u = a^\dagger_{u\sigma} \mathbf{S}_{\sigma,\sigma'} a_{u\sigma'}$ where $\mathbf{S}$ is the vector of spin 1/2 matrices:

$$S^x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S^y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S^x = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.14}$$

and the number operator as $n_u = a^\dagger_{u\sigma} a_{u\sigma}$, we thus have:

$$a^\dagger_{u\sigma} a_{u\sigma'} a^\dagger_{v\sigma'} a_{v\sigma} = \frac{n_u n_v}{2} + 2 s^z_u s^z_v + s^+_u s^-_v + s^-_u s^+_v = \frac{n_u n_v}{2} + 2\mathbf{s}_u \cdot \mathbf{s}_v. \tag{1.15}$$

i.e. the spin dependent part of the direct exchange term can be written as:

$$H_{\text{Coulomb}} = -J 2\mathbf{s}_u \cdot \mathbf{s}_v \tag{1.16}$$

By comparison to Eq. (1.9), we note that such an isotropic spin-spin interaction is exactly what we require in order to enhance the effective molecular field.

We can moreover show that the expression for J should be positive:

$$\begin{aligned} J &= \left\langle u, v \left| \frac{e^2}{4\pi\epsilon r_{12}} \right| v, u \right\rangle \\ &= \int d^3\mathbf{r}_1 \int d^3\mathbf{r}_2 \psi^*_u(\mathbf{r}_1)\psi^*_v(\mathbf{r}_2) \frac{e^2}{4\pi\epsilon|\mathbf{r}_1 - \mathbf{r}_2|} \psi_u(\mathbf{r}_2)\psi_v(\mathbf{r}_1) \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int d^d\mathbf{r}_1 \int d^d\mathbf{r}_2 \psi^*_u(\mathbf{r}_1)\psi^*_v(\mathbf{r}_2) \frac{e^2}{\epsilon k^2} e^{i\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \psi_u(\mathbf{r}_2)\psi_v(\mathbf{r}_1) \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{e^2}{\epsilon k^2} \left| \int d^d\mathbf{r}_1 \psi^*_u(\mathbf{r}_1) e^{i\mathbf{k}\cdot\mathbf{r}_1} \psi_v(\mathbf{r}_1) \right|^2 > 0, \end{aligned} \tag{1.17}$$

and we can estimate the characteristic energy scale of this ferromagnetic coupling to be $\simeq e^2/\epsilon_0 d \simeq \mathcal{R}y \times a_{Bohr}/d$ where d is the electron separation, and $\mathcal{R}y$ the atomic Rydberg, corresponding to a temperature of around 300K.

Minimising kinetic energy: kinetic exchange

In order to discuss the idea of kinetic exchange, we first introduce the Hubbard model[8], describing localised electronic orbitals, with on-site interactions (i.e. interactions between electrons on the same site). We will later, in Sec. 3.1 justify the validity of this simple model by considering the effect of the terms thus neglected. The model can be written as:

$$H = -t \sum_{\langle ij \rangle, \sigma} a^\dagger_{i\sigma} a_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \tag{1.18}$$

The form of the interaction makes use of the fact that Pauli exclusion prevents double occupancy of the same site with the same spin, so interactions only arise when there are two particles, one of each spin, on a given site.

Consider the case where there is one electron per site (i.e. where there is no doping, so the number of electrons and number of lattice sites will match), and consider the limit $U \gg t$. If interactions were absent, this model would be a conductor, since the band is half filled (one electron per site, but two possible spin states per site). However, with strong interactions, the model is insulating, as it costs an energy U to move an electron onto an already singly occupied site. One may then ask what the magnetic state of this insulating phase is.

Let us consider a pair of sites and compare the energies of the states:

$$|1,1\rangle = |\underline{\uparrow}\;\underline{\uparrow}\rangle, \quad |1,0\rangle = \frac{1}{\sqrt{2}}\left(|\underline{\uparrow}\;\underline{\downarrow}\rangle + |\underline{\downarrow}\;\underline{\uparrow}\rangle\right), \quad |0,0\rangle = \frac{1}{\sqrt{2}}\left(|\underline{\uparrow}\;\underline{\downarrow}\rangle - |\underline{\downarrow}\;\underline{\uparrow}\rangle\right).$$

(the state $|1,-1\rangle$ is omitted since its behaviour is clearly identical to $|1,1\rangle$). Since all of these states have the same occupation of each site, and since the hopping term will produce something orthogonal to each state, it is clear that at leading order:

$$\langle 1,1|H|1,1\rangle = \langle 1,0|H|1,0\rangle = \langle 0,0|H|0,0\rangle$$

However, since these states are eigenstates of the interaction part, let us now consider the hopping as a perturbation at second order in perturbation theory:

$$\delta E_n^{(2)} = \sum_m \frac{\langle n|\delta H|m\rangle\langle m|\delta H|n\rangle}{E_n - E_m}.$$

One can immediately see that $\delta H|1,1\rangle = 0$, since the hopping cannot move a spin up particle to a site already occupied by a spin up particle. However, for the other states, one has:

$$\delta H|\underline{\uparrow}\;\underline{\downarrow}\rangle = -t\left(+|\underline{\uparrow\downarrow}\;\underline{\quad}\rangle + |\underline{\quad}\;\underline{\uparrow\downarrow}\rangle\right) \tag{1.19}$$

$$\delta H|\underline{\downarrow}\;\underline{\uparrow}\rangle = -t\left(-|\underline{\uparrow\downarrow}\;\underline{\quad}\rangle - |\underline{\quad}\;\underline{\uparrow\downarrow}\rangle\right). \tag{1.20}$$

In order to relate the two terms above, it is necessary to choose a sign convention for how the diagramatic configurations relate to ordering of operators. The choice that has been used above is: $|\underline{\uparrow}\;\underline{\downarrow}\rangle = a^\dagger_{1\uparrow}a^\dagger_{2\downarrow}|\Omega\rangle$, i.e. operators appear in the same order as sites, and $|\underline{\uparrow\downarrow}\;\underline{\quad}\rangle = a^\dagger_{1\uparrow}a^\dagger_{1\downarrow}|\Omega\rangle$, i.e. spin up operators appear first. For example, Eq.(1.19) is given by:

$$\begin{aligned} -t\left(a^\dagger_{1\downarrow}a_{2\downarrow} + a^\dagger_{2\uparrow}a_{1\uparrow}\right)a^\dagger_{1\uparrow}a^\dagger_{2\downarrow} &= -t\left(-a^\dagger_{1\downarrow}a^\dagger_{1\uparrow} + a^\dagger_{2\uparrow}a^\dagger_{2\downarrow}\right) \\ &= -t\left(+a^\dagger_{1\uparrow}a^\dagger_{1\downarrow} + a^\dagger_{2\uparrow}a^\dagger_{2\downarrow}\right). \end{aligned}$$

It is thus clear that one has $\delta H|1,0\rangle = 0$ due to cancellation of the two terms, but $\delta H|0,0\rangle \neq 0$, and so:

$$\delta E^{(2)}_{0,0} = \frac{2t \cdot 2t}{0 - U} = -\frac{4t^2}{U}. \tag{1.21}$$

There is thus an antiferromagnetic interaction, strength $J = 4t^2/U$ favouring projection onto the singlet state. For a pair of spin one half particles, the operator:

$$\begin{aligned}\mathbf{S}_i \cdot \mathbf{S}_j = \frac{1}{2}\left[(\mathbf{S}_i + \mathbf{S}_j)^2 - \mathbf{S}_i^2 - \mathbf{S}_j^2\right] &= \begin{cases} \frac{1}{2}\left(0 - \frac{3}{2}\right) & \text{If singlet} \\ \frac{1}{2}\left(2 - \frac{3}{2}\right) & \text{If triplet} \end{cases} \\ &= \frac{1}{4} - \mathcal{P}_{\text{Singlet}} \end{aligned} \tag{1.22}$$

where $\mathcal{P}_{\text{Singlet}}$ projects onto the singlet state, thus this effective hopping term can again be written as $J\mathbf{S}_1 \cdot \mathbf{S}_2$ — another isotropic magnetic interaction.

1.3 Problems with ferromagnetism

As yet, the above discussion shows how ferromagnetic and antiferromagnetic exchange terms may arise, but leaves one with the feeling that which type of interaction one finds depends on what basis one chooses to work in. This evidently cannot be true, since the magnetic moment is physical and measurable. Thus, we must find that one point of view is more appropriate for some materials, and another for others. To show this, we show consider a case in which a superficial argument yields a ferromagnetic interaction, but the ground state must be antiferromagnetic, and use this to illustrate restrictions on problems in which ferromagnetism can physically occur.

A toy model with direct exchange

Our toy model is a pair of particles, in a harmonic trap, with some repulsive inter-particle interaction:

$$\begin{aligned} H &= \sum_{i=1,2}\left[\frac{m\omega^2}{2} r_i^2 - \frac{\nabla_i^2}{2m}\right] + U_{r_1 - r_2} \\ &= H_{COM} + \frac{m\omega^2}{4}\rho^2 - \frac{\nabla_\rho^2}{m} + U(\rho) \end{aligned} \tag{1.23}$$

where H_{COM} describes the centre of mass, and ρ is the separation of the particles. We will not choose any particular form for $U(\rho)$, but will assume it is symmetric, and decreases monotonically with increasing distance. For an antiferromagnetic spin configuration (i.e. spin singlet) the spatial wavefunction should satisfy $\phi(\rho) = \phi(-\rho)$, and for ferromagnetic, it should satisfy $\phi(\rho) = -\phi(-\rho)$. The question is then which boundary condition allows for a lower energy solution.

If one were to repeat the analysis of Sec. 1.2, and treat the interaction perturbatively, then the question concerns comparing the energies of the first few harmonic oscillator states, i.e.:

$$E_{AFM} = \frac{\hbar\omega}{2} + \int d\rho U(\rho)|\phi_0(\rho)|^2 \tag{1.24}$$

$$E_{FM} = \frac{3\hbar\omega}{2} + \int d\rho U(\rho)|\phi_1(\rho)|^2 \tag{1.25}$$

where $\phi_1(\rho = 0) = 0$, but $\phi_0(\rho = 0)$ has a maximum. If U becomes large enough compared to the harmonic oscillator energy (i.e. compared to the single particle energy scales), then it would appear that E_{FM} becomes lower in energy than E_{AFM}.

This is false. In a symmetric Sturm-Liouville problem such as Eq. (1.23), the ground state is always a spatially symmetric configuration, corresponding to antiferromagnetic configuration; the proof is straightforward by contradiction. Suppose that $\phi(r) = -\phi(-r)$ were the ground state. Then $|\phi(r)|$ is not the ground state (since it has a discontinuity of gradient, and so is not even an eigenstate. As such one has that $E[|\phi(r)|] > E_{g.s.}$. This is a strict inequality, since $|\phi(r)|$ is not an eigenstate. However, since $|\phi(r)|^2 = ||\phi(r)||^2$, and $|\nabla|\phi(r)||^2 = |\nabla\phi(r)|^2$ except at the single point $r = 0$, which has no weight in the integral, one finds that:

$$E_{g.s.} = E[\phi(r)] = E[|\phi(r)|] > E_{g.s.} \tag{1.26}$$

which is a contradiction,

Need for many-body physics

Physically, the error in Eqs. (1.24,1.25) is clear to see; if the interaction energy is sufficient to seem to favour the second excited state, it is also sufficient to favour admixtures of the third, fifth, seventh ... and all other even states) into the even ground state. This idea is rather general; for any apparently ferromagnetic configuration the interaction is strong enough to favour antisymmetric spatial configurations, and so is equally strong enough to cause significant changes to the spatial wavefunctions; i.e. admixtures of higher excited states, and so perturbative predictions of ferromagnetism are always suspicious.

In addition to the above argument, which shows that ferromagnetism cannot arise in a two particle ground state, there exists another strict theorem, the Lieb-Mattis theorem[9], which similarly shows the ground state is not ferromagnetic in one dimension — see also Question 1.1.

Under what conditions does ferromagnetism then exist? One clearly familiar case is Hund's rules for atomic configurations. These circumvent both of the above restrictions, since they require the presence of core electrons (the outer most orbitals must be orthogonal to the inner orbitals, thus one is automatically force to have a finite number of nodes), and they rely on orbital degeneracy (which requires angular momentum and hence $d > 1$), so that interactions need not be strong to overcome single particle energy costs. Thus, for many particles, since complete antisymmetry of the wavefunction ensures there must be at least some nodes, one cannot have a completely spatially symmetric ground state, and the argument of Sec. 1.3 ceases to hold.

Three particle exchange

As an illustration of a many-body effect leading to ferromagnetic exchange, let us consider the decomposition of three particle exchange into two particle exchange. The basic idea of n-particle exchange is that, given a set of orbitals, one may construct an effective Hamiltonian in that subspace, given by the overlap matrix elements between permutations of the wavefunctions:

$$H = H_0 + \sum_P J_{P_s} \left[P_s + (P_s)^{-1} \right] \tag{1.27}$$

where P_s is a permutation of spatial wavefunctions, and J_{P_s} is the matrix element between two such permutations. Note that since this matrix element is an overlap, i.e. $J_{P_s} = \langle \psi | H | P_s \psi \rangle = \langle \psi P_s | P_s^{-1} \psi P_s | \psi \rangle$ one has $J_{P_s} = J_{P_s^{-1}}$. The above made no restriction to overall antisymmetric wavefunctions, this is done by constructing total permutations $P = P_\sigma P_s$ where P_σ is a permutation of spins, and then requiring the wavefunction to be an eigenstate of P with eigenvalue $(-1)^{|P|}$ where $|P|$ is the number of particles exchanged. Restricting operators to this subspace, one may then write

$$(P_s)^{-1} P_\sigma P_s = P_\sigma = (-1)^{|P|} (P_S)^{-1} \tag{1.28}$$

so $(P_s)^{-1} = (-1)^{|P|} P_\sigma$, in terms of which:

$$H = H_0 + 2 \sum_P (-1)^{|P|} J_{P_S} P_\sigma \tag{1.29}$$

The spin permutation operator $P_\sigma = \pm 1$, for a triplet/singlet state respectively, and so:

$$P_\sigma^{ij} = 1 - 2\mathcal{P}^{ij}_{\text{singlet}} = \frac{1}{2} \left(1 + 4 \mathbf{S}_i \cdot \mathbf{S}_j \right) \tag{1.30}$$

making use of Eq. (1.22). By writing $\sigma_i = 2\mathbf{S}_i$, we can expand the three particle exchange term:

$$\begin{aligned} H_{(3)} &= J_{(3)} [P_{ijk} + P_{kji}] = J_{(3)} [P_{ki} P_{kj} + P_{ik} P_{ij}] \\ &= \frac{J_{(3)}}{4} \Big[1 + \sigma_k \cdot \sigma_j + \sigma_k \cdot \sigma_i + \sigma_i^\alpha \sigma_j^\beta \sigma_k^\alpha \sigma_k^\beta \\ &\qquad + 1 + \sigma_i \cdot \sigma_k + \sigma_i \cdot \sigma_j + \sigma_k^\alpha \sigma_j^\beta \sigma_i^\alpha \sigma_i^\beta \Big], \end{aligned} \tag{1.31}$$

and by using $\sigma_l^\alpha \sigma_l^\beta = \delta^{\alpha\beta} + i\epsilon^{\alpha\beta\gamma} \sigma_l^\gamma$ we get:

$$\begin{aligned} H_{(3)} &= \frac{J_{(3)}}{4} \left[2 + 2\sigma_k \cdot \sigma_j + 2\sigma_k \cdot \sigma_i + 2\sigma_i \cdot \sigma_j + i\epsilon^{\alpha\beta\gamma} \sigma_i^\alpha \sigma_j^\beta \sigma_k^\gamma + i\epsilon^{\alpha\beta\gamma} \sigma_k^\alpha \sigma_j^\beta \sigma_i^\gamma \right] \\ &= \frac{J_{(3)}}{2} \left[1 + \sigma_k \cdot \sigma_j + \sigma_k \cdot \sigma_i + \sigma_i \cdot \sigma_j \right]. \end{aligned} \tag{1.32}$$

Since the prefactor $J_{(3)}$ depended on spatial symmetry under exchange of three particles, it has the opposite sign $(-1)^{|P|}$ to that for two particle exchange, and can thus favour ferromagnetism.

Questions

* Question 1.1: Lieb-Mattis theorem: Absence of ferromagnetism in one dimension

1.1.(a) Consider an isotropic (i.e. no magnetic field) problem of interacting spin 1/2 particles. Show that, if the eigenstates are written in terms of

$|S,M\rangle$ with $(\sum_i S^i)^2|S,M\rangle = S(S+1)|S,M\rangle$ and $\sum_i S^i_z|S,M\rangle = M|S,M\rangle$, then the energy depends only on S, not on M.

The rest of this question has two sections; the first part is to find that the ground state spatial wavefunction for given M is nodeless in a particular sense. The second part is to show that this nodelessness implies $S = M$ for the ground state, and that is in turn implies the total ground state has maximum $S = 0$.

1.1.(b) For a given value of M, the spin is a superposition of parts each of which has $p = N/2 - M$ spin up particles, and $N - p = N/2 + M$ spin down. Considering one such part of this wavefunction, the spatial wavefunction must be antisymmetric under exchange of coordinates $\{r_1, \ldots, r_p\}$ and under exchange of $\{r_{p+1}, \ldots, r_N\}$. Such a wavefunction can be constructed by antisymmeterising a wavefunction for a particular spatial ordering of particles, defined as follows.

Consider a function defined in the domain $0 < r_1 < r_2 \ldots < r_p < R$ and $0 < r_{p+1} < r_{p+2} \ldots < r_N < R$, which vanishes on the boundary of this domain (i.e. vanishes if $r_1 = r_2$ etc.). Explain why (for a finite potential) the ground state wavefunction does not vanish anywhere else inside this domain, i.e. the ground state is always positive.

1.1.(c) Show that two wavefunctions with the same S_z component but different total spins are orthogonal.

1.1.(d) The state with $M = S$ must satisfy the relation $S_+|\Psi\rangle = 0$. This is in fact a necessary and sufficient condition. Show that an overall antisymmetric wavefunction built from the spatial function:

$$\Psi_S(\{r_1, \ldots, r_N\}) = \prod_{j,k=1;j>k}^{j,k<p} (r_j - r_k) \times \prod_{j,k=p+1;j>k}^{j,k<N} (r_j - r_k) \qquad (1.33)$$

must satisfy this condition, and show that it has a non-zero overlap with the ground state, as found in part 1.1.(b).

1.1.(e) Since the states with S_z component M can only belong to $S \geq M$, show that any state with $S > M$ has energy larger than $S = M$, hence that $S = 0$ is the overall ground state.

Question 1.2: Three electron atom and Hund's rules

1.2.(a) Consider an atom with three electrons, in spatial orbitals $\psi_{+1}, \psi_0, \psi_{-1}$. For the moment, take the electrons to be distinguishable, and write down the 6 distinct states of 3 electrons. The interaction Hamiltonian can be written as:

$$H_{int} = C\left[\frac{1}{r_{12}} + \frac{1}{r_{23}} + \frac{1}{r_{31}}\right] \qquad (1.34)$$

Show that the diagonal matrix elements are all identical, and that the off diagonal matrix elements either vanish, or can be written in terms of the integrals:

$$A = \iint d^3x d^3y \psi_1^*(x)\psi_{-1}^*(y)\frac{1}{|x-y|}\psi_1(y)\psi_{-1}(x)$$

$$B = \iint d^3x d^3y \psi_1^*(x)\psi_0^*(y)\frac{1}{|x-y|}\psi_1(y)\psi_0(x)$$

1.2.(b) Thus, writing $a = A/H_0, b = B/H_0$ where H_0 is the diagonal element, the Hamiltonian can be written as:

$$H \propto \begin{pmatrix} 1 & b & a & b & 0 & 0 \\ b & 1 & 0 & 0 & b & a \\ a & 0 & 1 & 0 & b & b \\ b & 0 & 0 & 1 & a & b \\ 0 & b & b & a & 1 & 0 \\ 0 & a & b & b & 0 & 1 \end{pmatrix} \tag{1.35}$$

Using the fact (see below) that $a = 2b$, find the eigenvalues of this matrix [Hint: guess the symmetry of wavefunctions to reduce the size of matrix you need to diagonalise], find the eigenvalues and eigenvectors. Find which of these eigenvectors can correspond to fermionic states (i.e. can be used to construct an overall antisymmetric wavefunction when combined with appropriate spin wavefunctions), thus find the ordering of spin states.

1.2.(c) Show $A = 2B$. Hint: one can write the integrals in terms of associated Legendre polynomials, and make use of two important results:

- The interaction can be written in terms of the generating functional of Legendre polynomials:

$$\frac{1}{\sqrt{1 - 2xz + z^2}} = \sum_{n=0}^{\infty} P_n(x) z^n. \tag{1.36}$$

- The spherical harmonic addition theorem:

$$P_n(\cos(\theta_1 - \theta_2)) = \frac{4\pi}{2n+1} \sum_{m=-n}^{n} (-1)^m Y_n^m(\theta_1, 0) Y_n^{-m}(\theta_2, 0) \tag{1.37}$$

Lecture 2

The Hubbard model and Heisenberg model

This lecture concerns the behaviour of magnetic insulators, described by the Hubbard model. Magnetic insulators provide yet another argument for the fact that interactions must be involved in magnetic behaviour: for a band insulator, one requires complete filling of a band, and thus equal numbers of spin up and down particles, thus to see insulating and magnetic behaviour together must require an insulator that is not a band insulator.

This lecture will therefore discuss the Hubbard model.

$$H = -t \sum_{\langle ij \rangle, \sigma} a^{\dagger}_{i\sigma} a_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow} \tag{2.1}$$

introduced in the previous lecture, and the resultant low energy Hamiltonians that correspond to this model; the $t - J$ model and its particular limit at half filling: the antiferromagnetic Heisenberg model. We will then discuss the nature of magnetic ground states in antiferromagnetic models. The next lecture will provide a firmer foundation for our starting with the Hubbard model, by describing the possible corrections to this model in real materials.

2.1 Away from half filling: the $t - J$ model

At half filling, we saw previously that antiferromagnetic coupling arises in the low-energy Hamiltonian due to virtual hopping processes reducing the energy for anti-parallel spin configurations. Away from half filling, there are in addition real hopping terms, allowing hopping of the excess electrons, or holes, according to the sign of deviation from half filling. Thus, the large U limit now contains two classes of interactions; those mediated by virtual hopping, and the real hopping of vacancies. We will find that these two kinds of interaction compete to set the sign of the magnetic interaction.

The problem can be seen as one of massively degenerate perturbation theory, since for N_s sites, and $n_{\uparrow,\downarrow}$ electrons of each spin, the number of possible single occupancy configurations is $N_s!/[n_\uparrow! n_\downarrow! (N_s - n_\uparrow - n_\downarrow)!]$. Our aim is to find how hopping acts within this degenerate subspace, including effects of virtual hops. To treat the low energy sector consistently away from half filling, one must treat separately hops that change the energy from those that do not, i.e. the amplitude associated with the two processes:

$$\text{(a)} \quad \underline{\uparrow}\ \underline{\ \ \ } \Rightarrow \underline{\ \ \ }\ \underline{\uparrow}, \qquad \text{(b)} \quad \underline{\uparrow}\ \underline{\downarrow} \Rightarrow \underline{\ \ \ }\ \underline{\uparrow\downarrow} \tag{2.2}$$

are different, since process (b) costs energy U, however both processes are at present described by the same term, $ta^{\dagger}_{2\uparrow}a_{1\uparrow}$.

Unitary transformation for degenerate perturbation theory

Bearing the above points in mind, we proceed by making a unitary transformation of the degenerate space, so as to eliminate the leading order effects of hopping terms such as those in Eq. (2.2 b). Formally, we divide the hopping term H_t into parts which are diagonal $H_{t,D}$ and off-diagonal $H_{t,OD}$ in terms of the interaction term H_U. Our aim is then to find a Hermitian operator S such that:

$$\tilde{H} = e^{iS}(H_U + H_{t,D} + H_{t,OD})e^{-iS} = H_U + H_{t,D} + \mathcal{O}(t^2/U) \tag{2.3}$$

which corresponds to working in a new basis, where all hopping terms $\propto t$ that exit the ground state manifold have been integrated out, leaving virtual hopping mediated contributions; and processes higher order in t which may exit the ground state manifold.

To make the division of H_t we write:

$$\begin{aligned} H_t = -t \sum_{\sigma,\langle ij\rangle} & \overbrace{n_{j-\sigma}a^{\dagger}_{j\sigma}a_{i\sigma}n_{i-\sigma}}^{H_{t,2}} + \overbrace{n_{j-\sigma}a^{\dagger}_{j\sigma}a_{i\sigma}(1-n_{i-\sigma})}^{H_{t,+}} \\ & + \underbrace{(1-n_{j-\sigma})a^{\dagger}_{j\sigma}a_{i\sigma}n_{i-\sigma}}_{H_{t,-}} + \underbrace{(1-n_{j-\sigma})a^{\dagger}_{j\sigma}a_{i\sigma}(1-n_{i-\sigma})}_{H_{t,0}} \end{aligned} \tag{2.4}$$

The factors introduced depend on the occupation of initial and final sites by the opposite spin, and are such that the sum of all four parts reproduces H_t. The occupation of the initial and final sites by the spin that hops need not be separately distinguished, since the hopping operator vanishes except when transferring from a filled to an empty site. Identifying $H_{t,0}$ as hopping of holes as in Eq. (2.2 a), and $H_{t,2}$ as hopping of doublets, i.e. $\underline{\uparrow\downarrow}\ \underline{\downarrow} \Rightarrow \underline{\downarrow}\ \underline{\uparrow\downarrow}$, these two parts together give the number diagonal part $H_{t,D} = H_{t,0} + H_{t,2}$, and so the remainder $H_{t,OD} = H_{t,+} + H_{t,-}$ correspond to creation of doublet-hole pair or annihilation of a doublet-hole pair.

To achieve the aim set out in Eq. (2.3), we will expand the unitary rotation, to give:

$$\tilde{H} = H_U + i[S, H_U] + H_{t,D} + H_{t,OD} + i[S, H_{t,D} + H_{t,OD}] - \frac{1}{2}[S,[S,H_U]]\ldots, \tag{2.5}$$

and then demand that $i[S, H_U] = -H_{t,OD}$ so as to remove the off-diagonal terms of order t. It is clear that this implies $S \sim t/U$, and so the last two terms are of order t^2/U, and the terms not written are at least of order t^3/U.

To find an explicit form for S, we note:

$$[H_{t,\pm}, H_U] = \mp U H_{t,\pm} \tag{2.6}$$

which can be easily seen by distinguishing whether double occupancy occurs in the initial state (for $H_{t,-}$) or final state (for $H_{t,+}$) and hence in which operator order H_U survives, since it annihilates anything without a doubly occupied site. It is thus clear that:

$$S = \frac{i}{U}\left(H_{t,-} - H_{t,+}\right), \quad i[S, H_U] = -\frac{1}{U}\left(H_{t,-} + H_{t,+}\right) \tag{2.7}$$

satisfies the requirement.

To find the terms in the second line, we first note that:

$$\begin{aligned}
[S,[S,H_U]] &= [S, iH_{t,OD}], \\
i[S, H_{t,OD}] &= -\frac{1}{U} 2[H_{t,-}, H_{t,+}]
\end{aligned} \tag{2.8}$$

and secondly note that the term $[S, H_{t,D}]$ describes terms that take one outside the ground state manifold, since they will involve a single operator $H_{t\pm}$, whereas the term in Eq. (2.8) can restore one to the ground state manifold by creating then destroying a doublet-hole pair.

Assuming one starts in a state without such excitations, one may further note that $H_{t,+}H_{t,-} = 0$, since the assumed state is annihilated by the excitation destruction operator, leaving:

$$\tilde{H} = H_U + H_{t,D} - \frac{1}{U} H_{t,-} H_{t,+}. \tag{2.9}$$

To be consistent, one should also drop H_U at this stage, since the number of doubly occupied states in the ground state manifold is a constant, so H_U contributes only a c-number to the Hamiltonian in this manifold.

It remains to work out what the term $H_{t,-}H_{t,+}$ does in this manifold. One may divide this into cases where the same pair of sites are acted on, $(ij)(ji)$, or terms where three sites are involved $(ij)(jk)$, with a doublet created then annihilated on site j.

The first process, involving two sites, is given by:

$$\begin{aligned}
H^{ij}_{t,-} H^{ji}_{t,+} &= t^2 \sum_\sigma \Big[(1 - n_{j-\sigma}) a^\dagger_{j\sigma} a_{i\sigma} n_{i-\sigma} + (1 - n_{j\sigma}) a^\dagger_{j-\sigma} a_{i-\sigma} n_{i\sigma}\Big] \\
&\qquad \times n_{i-\sigma} a^\dagger_{i\sigma} a_{j\sigma} (1 - n_{j-\sigma}) \\
&= t^2 \sum_\sigma (1 - n_{j-\sigma}) n_{i-\sigma} (1 - n_{i\sigma}) n_{j\sigma} - a^\dagger_{j-\sigma} a^\dagger_{i\sigma} a_{i-\sigma} a_{j\sigma}
\end{aligned} \tag{2.10}$$

where in the first line we have allowed the process $H_{t,-}$ to have either the same or opposite spin label to the process $H_{t,+}$, and in the second line we have: removed number operators that are redundant due to the effect of fermionic creation and annihilation operators; and we have replaced pairs of fermionic operators by number operators where possible. The two contributions both exist only if the sites i, j originally contain opposite spins, and so correspond to an exchange term, which either leaves the spins unchanged, or exchanges them and gives a minus sign. Using the $\mathbf{s}$ operators defined in Eq. (1.14) we may write:

$$\begin{aligned} H^{ij}_{t,-} H^{ji}_{t,+} &= t^2 \left[\left(\frac{n_i n_j}{2} - 2 s^z_i s^z_j \right) - \left(s^+_i s^-_j + s^-_i s^+_j \right) \right] \\ &= 2t^2 \left[\frac{n_i n_j}{4} - \mathbf{s}_i \cdot \mathbf{s}_j \right] \end{aligned} \tag{2.11}$$

After summing over configurations ij and ji, these class of terms then recover exactly the Heisenberg model we had described before, giving:

$$H_{t-J} = \text{const} - t \sum_{\sigma, \langle ij \rangle} a^\dagger_{i\sigma} a_{j\sigma} + \frac{4t^2}{U} \sum_{\langle ij \rangle} \left[\mathbf{s}_i \cdot \mathbf{s}_j - \frac{n_i n_j}{4} \right] + \mathcal{O}(t^2/U). \tag{2.12}$$

For completeness, we should also describe the effect of $H^{kj}_{t,-} H^{ji}_{t,+}$ involving three sites. This can only have an effect if site k was initially empty, so that a doublet created on site j by $H^{ji}_{t,+}$ is annihilated along with the hole on site k by $H^{kj}_{t,-}$. Schematically, the states involved might be written for sites i, j, k as:

$$\underline{\uparrow}\;\underline{\downarrow}\;\underline{\;\;\;} \Rightarrow \underline{\;\;\;}\;\underline{\uparrow\downarrow}\;\underline{\;\;\;} \Rightarrow \begin{cases} \underline{\;\;\;}\;\underline{\downarrow}\;\underline{\uparrow} \\ \underline{\;\;\;}\;\underline{\uparrow}\;\underline{\downarrow} \end{cases} \tag{2.13}$$

giving a contribution:

$$\delta H^{3\text{ site}}_{t-J} = -\frac{t^2}{U} \left[a^\dagger_{k\sigma} n_{j-\sigma} a_{i\sigma} + a^\dagger_{k-\sigma} (a^\dagger_{j\sigma} a_{j-\sigma}) a_{i\sigma} + H.c. \right]. \tag{2.14}$$

This term provides a mechanism for holes to hop two sites, at a cost t^2/U, and so in the large U limit (required for all derivations here), this is not particularly significant, as the order t hole hopping will dominate.

The approach in this section, of taking a massively degenerate problem, and eliminating the leading order effects of the perturbation by making an appropriate unitary transformation is known variously as a Schrieffer-Wolff transform, or as Löwdin partitioning, and will be used again several times in this course.

Nagaoka ferromagnetism

Considering a system just below half filling, the hopping of holes is believed to give a tendency toward ferromagnetism. While it is believed this exists in general for small densities of holes, it can be proved only for the case of exactly one hole in the limit of large U; this is Nagaoka's theorem[10]. We will discuss the idea of this effect and briefly outline the method of proof.

We already know this cannot be true in one dimension, so we will consider a two-dimensional case, and consider interference between different

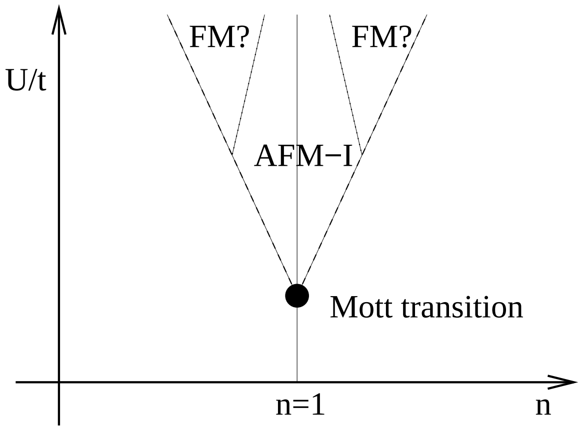


Figure 2.1: A possible phase diagram of the Hubbard model, inspired by mean-field theory results. Whether a finite region of ferromagnetism actually exists is not clear

hopping paths. If we consider four sites, hopping of a hole allows:

$$\begin{array}{cc} \underline{} & \underline{\uparrow} \\ \underline{\uparrow} & \underline{\downarrow} \end{array} \Rightarrow \left\{ \begin{array}{l} \begin{array}{cc} \underline{\uparrow} & \underline{} \\ \underline{\uparrow} & \underline{\downarrow} \end{array} \Rightarrow \begin{array}{cc} \underline{\uparrow} & \underline{\downarrow} \\ \underline{\uparrow} & \underline{} \end{array} \\ \\ \begin{array}{cc} \underline{\uparrow} & \underline{\uparrow} \\ \underline{} & \underline{\downarrow} \end{array} \Rightarrow \begin{array}{cc} \underline{\uparrow} & \underline{\uparrow} \\ \underline{\downarrow} & \underline{} \end{array} \end{array} \right. . \tag{2.15}$$

We find that for this initial antiferromagnetic situation, there is no overlap between the two processes of hole hopping — the trail of flipped spins records "which way" information on the path of the hole. Thus, in comparison, the ferromagnetic case has a larger hopping amplitude, since different paths will interfere constructively. This enhanced hopping means that there will be a wider bandwidth, and thus a greater energy gain from delocalising the single hole. Since this energy gain is of the order t, it will win against the Heisenberg antiferromagnetic term, of order t^2/U

The proof consists of rewriting the problem in a basis of configurations of spins, for a given number of spin up/spin down particles. It can be shown that the equation $E\Psi = H\Psi$ in this basis, with E equal to the ferromagnetic configuration energy, imposes the constraint that the weight of all configurations must be equal.

With finite numbers of holes, interactions between the holes become relevant, and as mentioned before, one has the generic problem that instabilities to ferromagnetism are susceptible to removal due to wavefunction renormalisation. The result of this is that while some simplistic variational calculations yield finite ranges of ferromagnetism in the $t - J$ model, the more sophisticated calculations yield only very narrow regions very close to half filling. The question of whether a finite volume of the phase diagram is really ferromagnetic remains unclear. One should however bear in mind that discussion of phases of the $t - J$ model at extreme parameter choices is somewhat philosophical: in such parameter regimes, it is almost always true that other terms, as we will discuss in lecture 3, become relevant.

2.2 Half filling: The Heisenberg model

Exactly at half filling, we may neglect all terms except the antiferromagnetic interaction, giving the isotropic Heisenberg model. We will now turn to discuss what states this model supports, but will generalise somewhat to discuss the case where $|S|$ is not necessarily 1/2, but can be larger — physically this corresponds to multiple electrons per atom; or, in rare-earth materials, to effects of spin-orbit coupling replacing spin by total angular momentum.

Inadequacy of Néel state; bounds on ground state energy

The Néel state is a classical spin antiferromagnet, i.e. a state of the form $\ldots \uparrow\downarrow\uparrow\downarrow\uparrow\downarrow \ldots$. This is clearly not the an eigenstate of the Heisenberg model since, writing:

$$H_H = J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = J \sum_{\langle ij \rangle} \left[S_i^z S_j^z + \frac{1}{2} \left(S_i^+ S_j^- + S_i^- S_j^+ \right) \right] \tag{2.16}$$

one sees that when this acts on the Néel state it will flip pairs of adjacent spins.

One may however consider the Néel state as a variational ansatz (albeit with no free parameters), and use it to provide an upper bound on the ground state energy. On a lattice with coordination number z, one clearly sees that:

$$\langle \Psi_{\text{Néel}} | H_H | \Psi_{\text{Néel}} \rangle = \langle \Psi_{\text{Néel}} | J \sum_{\langle ij \rangle} S_i^z S_j^z | \Psi_{\text{Néel}} \rangle = -\frac{JS^2 Nz}{2} \tag{2.17}$$

where S is the magnitude of the spin — i.e. the maximal S_z component, and N is the number of lattice sites, so that $Nz/2$ is the number of nearest neighbour pairs.

One may also consider a lower bound on the ground state energy, by considering the tendency to formation of singlets. For a single pair of atoms, the ground state is a singlet, i.e. $|\Psi_0\rangle = (\uparrow\downarrow - \downarrow\uparrow)/\sqrt{2}$. This can be shown by rewriting the Hamiltonian using:

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{1}{2} \left[(\mathbf{S}_1 + \mathbf{S}_2)^2 - \mathbf{S}_1^2 - \mathbf{s}_2^2 \right] \tag{2.18}$$

so eigenstates are of $J\mathbf{S}_1 \cdot \mathbf{S}_2$ are eigenstates of total spin S_{tot}, and have eigenvalue:

$$E = \frac{J}{2} \left[S_{tot}(S_{tot} + 1) - 2S(S+1) \right]. \tag{2.19}$$

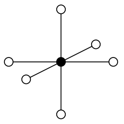
and so the ground state has energy $-JS(S+1)$ if $S_{tot} = 0$. This is lower than the Néel state which would have had $-JS^2$; it is lower because it respects the rotation symmetry of the Hamiltonian.

As an absolute lower bound, we could consider having every possible pair bond produce a singlet; this is not realisable in practice because the same spin cannot form a singlet bond with more than one other spin, but the energy of this configuration cannot be beaten, so we have:

$$-\frac{JS(S+1)Nz}{2} < E_{gs} < -\frac{JS^2Nz}{2} \tag{2.20}$$

This lower bound corresponds to dividing the Hamiltonian into pairs, and finding the ground state energy of each pair separately. A tighter lower bound comes from choosing larger subunits and finding the minimum energy of each subunit. Let us take our subunit to be one spin and all of its neighbours:

$$H = J\mathbf{S}_j \cdot \sum_{m\in\mathrm{NN}[j]} \mathbf{S}_m, \qquad \mathrm{NN}[j]:$$



$$= \frac{J}{2}\left[\left(\mathbf{S}_j + \sum_m \mathbf{S}_m\right)^2 - \left(\sum_m \mathbf{S}_m\right)^2 - \mathbf{S}_j^2\right]. \tag{2.21}$$

We take $\sum_m \mathbf{S}$ to be a spin of magnitude Σ, where Σ can range between $0 < \Sigma < zS$ for coordination z. It is clear that to minimise E we should take $\mathbf{S}_j + \sum_m \mathbf{S}_m = \Sigma - S$, and so the energy as a function of Σ is:

$$E = \frac{J}{2}\left[(\Sigma - S)(\Sigma - S + 1) - \Sigma(\Sigma+1) - S(S+1)\right] = -JS(\Sigma+1) \tag{2.22}$$

Hence the minimal energy of this cluster comes with $\Sigma = zS$ and is $-JS(zS+1)$. The global energy must be greater than the product of this value with $N/2$, since each bond appears in two clusters.

$$E_{gs} > -\frac{NJS(zS+1)}{2} = -\frac{JS(S+1/z)Nz}{2}. \tag{2.23}$$

One can see that if $z = 1$, this repeats the bound from singlet bonds, but generally this equation yields a tighter bound. As z increases, the difference between this bound and the Néel energy becomes smaller. Thus, in 3D, where $z = 6$ for SC lattice, and $z = 8$ for BCC, the Néel state may reasonably describe correlations of the Heisenberg Hamiltonian, but in 2D it is less clear. The FCC lattice in 3D, despite having $z = 12$, is also less clear, since a given site can be both the nearest neighbour of a site, and simultaneously the nearest neighbour of a nearest neighbour; i.e. the lattice is not bipartite. The Néel state relies on having a bipartite lattice, so that the alternating pattern is not frustrated.

Spin wave theory

In cases where the Néel state is a reasonable starting point, we may try to improve upon it by considering spin wave corrections, i.e. allowing "quantum fluctuations". To discuss this idea, we take the case of large

S (for which the fractional difference between Néel and singlet bonds is even smaller, $\sim 1/zS$). We also assume the lattice is bipartite, so that we can divide sites into class A and B such that the nearest neighbour of an A site is always of type B. We will then use a Holstein-Primakoff representation of spins to allow us to parameterise small fluctuations about the Néel state, i.e.:

$$S_i \to \begin{cases} S_i^z = S - a_i^\dagger a_i, \quad S_i^+ = \sqrt{2S - a_i^\dagger a_i} a_i, & A \text{ sites} \\ S_i^z = b_i^\dagger b_i - S, \quad S_i^+ = b_i^\dagger \sqrt{2S - b_i^\dagger b_i}, & B \text{ sites} \end{cases} \tag{2.24}$$

It can be shown that these definitions satisfy the spin commutation relations, and the Néel state corresponds to the vacuum state, i.e. no population of the modes $a^\dagger, b^\dagger$. Quantum fluctuations describe the occupations of these modes in the ground state; thermal fluctuations to the additional occupation at finite temperature. If we assume the occupation of fluctuations is small, $a^\dagger a, b^\dagger b \ll S$, then we may neglect the terms in the square root, and neglect any products of operators higher than second order, and thus write:

$$H = J \sum_{\langle ij \rangle} \left[-S^2 + S(a_i^\dagger a_i + b_j^\dagger b_j + \frac{1}{2} 2S \left(a_i^\dagger b_j^\dagger + a_i b_j \right) \right] \tag{2.25}$$

At quadratic level, it is straightforward to transform this expression into a ground state energy plus a cost of occupying excitations by making a Bogoliubov transform. We will outline this, and show how this procedure shows that the lowest energy state does not correspond to no bosons, but to a finite occupation, and thus the ground state energy is less than the Néel estimate. To perform this transformation there are two steps: Fourier transformation to see that the Hamiltonian is diagonal in momentum space (whereas it is off-diagonal in site index), and Bogoliubov transformation to find superpositions of particle and hole operators.

Fourier transformation Noting that each set of operators is defined only on $N/2$ sites, we write the Fourier transform as:

$$a_j^\dagger = \sum_{\mathbf{k}} \frac{1}{\sqrt{N/2}} A_k^\dagger e^{i\mathbf{k}\cdot\mathbf{r}_j} \tag{2.26}$$

where the number of $\mathbf{k}$ vectors in the sum is also $N/2$. The two types of term in the Hamiltonian can be written as:

$$\sum_{\langle ij \rangle} a_i^\dagger a_i = z \sum_i \frac{2}{N} \sum_{\mathbf{k},\mathbf{k}'} A_{\mathbf{k}}^\dagger A_{\mathbf{k}'} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}_i} = z \sum_{\mathbf{k}} A_{\mathbf{k}}^\dagger A_{\mathbf{k}} \tag{2.27}$$

where in the first expression we replaced the sum over nearest neighbour sites j by the number of such neighbours, z, and the second expression replaced the sum over sites i by a delta function in momentum space.

Similarly:

$$\begin{aligned}\sum_{\langle ij\rangle} a_i^\dagger b_j^\dagger &= \sum_i \frac{2}{N}\sum_{\mathbf{k},\mathbf{k}'} A_{\mathbf{k}}^\dagger B_{\mathbf{k}'}^\dagger e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}_i} \sum_{j\in\mathrm{NN}[i]} e^{i\mathbf{k}'\cdot(\mathbf{r}_j-\mathbf{r}_i)} \\ &= \sum_{\mathbf{k}} \gamma_{\mathbf{k}} A_{\mathbf{k}}^\dagger B_{-\mathbf{k}}^\dagger, \qquad \gamma_{\mathbf{k}} = \sum_{j\in\mathrm{NN}[i]} e^{i\mathbf{k}'\cdot(\mathbf{r}_j-\mathbf{r}_i)} \end{aligned} \tag{2.28}$$

The function $\gamma_{\mathbf{k}}$ is a tight-binding dispersion relation, which for small $\mathbf{k}$ can be seen to have the form $\gamma_{\mathbf{k}} = z - \mathcal{O}(k^2)$.

Putting these parts together we now have Eq. (2.25) becomes:

$$H = -\frac{JS^2Nz}{2} + JS\sum_{\mathbf{k}}\left[z\left(A_{\mathbf{k}}^\dagger A_{\mathbf{k}} + B_{\mathbf{k}}^\dagger B_{\mathbf{k}}\right) + \gamma_{\mathbf{k}}\left(A_{\mathbf{k}}^\dagger B_{-\mathbf{k}}^\dagger + A_{\mathbf{k}} B_{-\mathbf{k}}\right)\right] \tag{2.29}$$

Bogoliubov transform Writing $H = E_0 + H_{fluct}$, where $E_0 = -JS^2Nz/2$ is the Néel energy, we may relabel $\mathbf{k} \to -\mathbf{k}$ for $B_{\mathbf{k}}^\dagger, B_{\mathbf{k}}$, and use the bosonic commutator $B_{\mathbf{k}}^\dagger B_{\mathbf{k}} = B_{\mathbf{k}} B_{\mathbf{k}}^\dagger - 1$ to write:

$$H_{fluct} = JS\sum_{\mathbf{k}}\left[\begin{pmatrix} A_{\mathbf{k}}^\dagger & B_{-\mathbf{k}}\end{pmatrix}\begin{pmatrix} z & \gamma_{\mathbf{k}} \\ \gamma_{\mathbf{k}} & z\end{pmatrix}\begin{pmatrix} A_{\mathbf{k}} \\ B_{-\mathbf{k}}^\dagger\end{pmatrix} - z\right] \tag{2.30}$$

and then make a Bogoliubov transform:

$$\begin{pmatrix} A_{\mathbf{k}} \\ B_{-\mathbf{k}}^\dagger\end{pmatrix} = \begin{pmatrix} \cosh(\theta_{\mathbf{k}}) & \sinh(\theta_{\mathbf{k}}) \\ \sinh(\theta_{\mathbf{k}}) & \cosh(\theta_{\mathbf{k}})\end{pmatrix}\begin{pmatrix} \alpha_{\mathbf{k}} \\ \beta_{-\mathbf{k}}^\dagger\end{pmatrix}. \tag{2.31}$$

It can be shown that if α, β obey bosonic commutation relations then this reproduces the desired commutation relations of A, B. After transforming the matrix in Eq. (2.30), the diagonal and off-diagonal elements are:

$$\begin{aligned}\text{Diagonal:}\quad & z(\cosh^2(\theta_{\mathbf{k}}) + \sinh^2(\theta_{\mathbf{k}})) + \gamma_{\mathbf{k}} 2\cosh(\theta_{\mathbf{k}})\sinh(\theta_{\mathbf{k}}) \\ \text{Off-diagonal:}\quad & z2\cosh(\theta_{\mathbf{k}})\sinh(\theta_{\mathbf{k}}) + \gamma_{\mathbf{k}}(\cosh^2(\theta_{\mathbf{k}}) + \sinh^2(\theta_{\mathbf{k}}))\end{aligned}$$

thus we choose:

$$\sinh(2\theta_{\mathbf{k}}) = \frac{-\gamma_{\mathbf{k}}}{\sqrt{z^2 - \gamma_{\mathbf{k}}^2}}, \qquad \cosh(2\theta_{\mathbf{k}}) = \frac{z}{\sqrt{z^2 - \gamma_{\mathbf{k}}^2}} \tag{2.32}$$

so:

$$\begin{aligned} H_{fluct} &= JS\sum_{\mathbf{k}}\left[\begin{pmatrix} \alpha_{\mathbf{k}}^\dagger & \beta_{-\mathbf{k}}\end{pmatrix}\begin{pmatrix} \sqrt{z^2-\gamma_{\mathbf{k}}^2} & 0 \\ 0 & \sqrt{z^2-\gamma_{\mathbf{k}}^2}\end{pmatrix}\begin{pmatrix} \alpha_{\mathbf{k}} \\ \beta_{-\mathbf{k}}^\dagger\end{pmatrix} - z\right] \\ &= JS\sum_{\mathbf{k}}\left[\sqrt{z^2-\gamma_{\mathbf{k}}^2}\left(\alpha_{\mathbf{k}}^\dagger\alpha_{\mathbf{k}} + \beta_{\mathbf{k}}^\dagger\beta_{\mathbf{k}}\right) + \sqrt{z^2-\gamma_{\mathbf{k}}^2} - z\right] \end{aligned} \tag{2.33}$$

The ground state of H_{fluct} corresponds to the state with no excitations as measured by $\alpha_{\mathbf{k}}^\dagger\alpha_{\mathbf{k}} + \beta_{\mathbf{k}}^\dagger\beta_{\mathbf{k}} = 0$. Since these operators are, according to

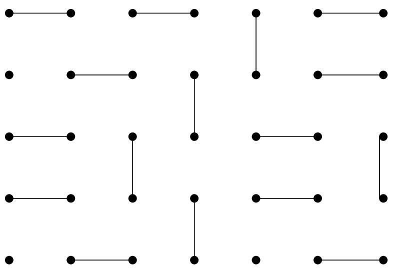


Figure 2.2: A possible configuration of singlet bonds in 2D

Eq. (2.31), superpositions of the original excitation creation and annihilation operators, the vacuum of excitations is not the vacuum of the original operators — i.e. the Néel state is not the ground state, but rather one should construct the state $|\Psi_0\rangle$ by demanding $\beta_{\mathbf{k}}|\Psi_0\rangle = 0$.

Having done this, the ground state energy as found by this approach is:

$$E_{gs} < -\frac{JS^2Nz}{2} - JSz\sum_{\mathbf{k}}\left[1-\sqrt{1-\frac{\gamma_{\mathbf{k}}^2}{z^2}}\right] \tag{2.34}$$

Comparing this to the forms in Eq. (2.23), we may write:

$$E_{gs} < -\frac{JS(S+\delta)Nz}{2}, \quad \delta = \frac{2}{N}\sum_{\mathbf{k}}\left[1-\sqrt{1-\frac{\gamma_{\mathbf{k}}^2}{z^2}}\right] \tag{2.35}$$

where it is clear that the value of this number, δ depends on the coordination number z, and so for certain values, this and Eq. (2.23) provide rather tight bounds on the ground state energy.[1].

2.3 Valence bond states

We have seen that the Néel state is not the ground state, and that singlets of pairs can in small systems provide a lower energy. We discuss here a competing set of ground states to the Néel state, the valence bond states.

Consider a state in which all atoms are in a singlet state with another atom, which we represent as a set of singlet bonds between sites (Fig. 2.2). When the Heisenberg Hamiltonian acts on a pair of spins that are joined by a singlet, one finds $H\psi = -JS(S+1)\psi$.

Let us for the moment restrict to $S = 1/2$, and consider what happens when H acts on a pair not joined by a singlet bond. Using Eq. (1.22), we have:

$$H_{ij} = -J\mathcal{P}_{\text{Singlet}} + \text{const} \tag{2.36}$$

---

[1] In one dimension, one should be aware that the approximation of neglecting interactions between modes becomes invalid, because the population of low energy modes in 1D is not small; this is a signature of the general statement that no continuous symmetry can be spontaneously broke in one dimension, even at zero temperature

and if this acts between spins ij where singlet bonds Ii and jJ exist, we have:

$$\begin{aligned}H_{ij}\left|(iI)(jJ)\right\rangle &= -J\left|\frac{1}{\sqrt{2}}(\uparrow_i\downarrow_j - \downarrow_i\uparrow_j)\right\rangle\left\langle\frac{1}{\sqrt{2}}(\uparrow_i\downarrow_j - \downarrow_i\uparrow_j)\right| \\ &\quad\times\left|\frac{1}{\sqrt{2}}(\uparrow_I\downarrow_i - \downarrow_I\uparrow_i)\cdot\frac{1}{\sqrt{2}}(\uparrow_j\downarrow_J - \downarrow_j\uparrow_J)\right\rangle \\ &= -\frac{J}{2}\left|\frac{1}{\sqrt{2}}(\uparrow_i\downarrow_j - \downarrow_i\uparrow_j)\cdot\frac{1}{\sqrt{2}}(\uparrow_I\downarrow_J - \downarrow_I\uparrow_J)\right\rangle = -\frac{J}{2}\left|(ij)(IJ)\right\rangle\end{aligned} \tag{2.37}$$

hence, H acting on a state such as that in Fig. 2.2 causes bonds to resonate, and swap between atoms.

The general valence bond state is thus a liquid, containing a superposition of all such states. It turns out that such configurations are not generally the ground state of antiferromagnetic materials, and that the Néel state is instead a better description.

In addition to valence bond liquids, one can also consider valence bond solids, for which there is a fixed configuration of bonds. Examples of these require higher spins (or alternatively strange forms of lattices such that a collection of lattice sites collectively behave as a single larger spin); this requirement is so that a site can be involved in multiple singlet bonds at once.

Rather than constructing the state, we will first construct a variant on the Heisenberg Hamiltonian whose ground state is clearly a valence bond solid — the one dimensional problem:

$$H_{AKLT} = \sum_i \mathbf{S}_i\cdot\mathbf{S}_{i+1} + \frac{1}{3}(\mathbf{S}_i\cdot\mathbf{S}_{i+1})^2 \tag{2.38}$$

with $|S| = 1$. One can rewrite $\mathbf{S}_i\cdot\mathbf{S}_j = (S_{tot}(S_{tot}+1)-4)/2$, and then note that:

$$\begin{array}{rrr} S_{tot} = 0 & \mathbf{S}_i\cdot\mathbf{S}_j = -2 & H_{ij} = -2/3 \\ 1 & -1 & -2/3 \\ 2 & 1 & 4/3 \end{array}$$

hence the ground state requires the total spin projection is not 2, but does not care whether it is 0 or 1. Question 2.2 demonstrates how to construct such a state using Schwinger boson parameterisation, and thus to show it is a valence bond solid.

Questions

* Question 2.1: Schwinger bosons

The Schwinger boson parameterisation of a spin involves writing:

$$S^z = \frac{1}{2}\left(b^\dagger b - a^\dagger a\right), \quad S^+ = b^\dagger a, \quad S^- = a^\dagger b. \tag{2.39}$$

along with the requirement $b^\dagger b + a^\dagger a = 2S$

2.1.(a) Find how the operators $S^{x,y,z}$ transform under the transformations:

$$\begin{pmatrix} b^\dagger \\ a^\dagger \end{pmatrix} \to \begin{pmatrix} \cos(\chi) & -\sin(\chi) \\ \sin(\chi) & \cos(\chi) \end{pmatrix} \begin{pmatrix} b^\dagger \\ a^\dagger \end{pmatrix}$$
$$\begin{pmatrix} b^\dagger \\ a^\dagger \end{pmatrix} \to \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix} \begin{pmatrix} b^\dagger \\ a^\dagger \end{pmatrix}$$

and hence show that these transformations allow one to produce all possible rotations in space.

2.1.(b) Consider the state of two spins $(a_i^\dagger b_j^\dagger - a_j^\dagger b_i^\dagger)^{2S}|0\rangle$. Show that this state is rotationally invariant, and that it satisfies the constraint $b^\dagger b + a^\dagger a = 2S$ for both spins. Find the modulus of total spin.

Question 2.2: Schwinger bosons and valence bond solid state

Consider the state on the infinite 1D lattice:

$$|\Psi\rangle = \prod_i \left(a_i^\dagger b_{i+1}^\dagger - a_{i+1}^\dagger b_i^\dagger\right)^M |0\rangle \tag{2.40}$$

2.2.(a) Show that one requires $M = S$ in order to satisfy the constraint.

2.2.(b) Find the maximum possible value of $S^z_{i,i+1} = S^z_i + S^z_{i+1}$. Show by analogy to the previous question that this state is rotationally invariant, and hence that this state is a ground state of the Hamiltonian in Eq. (2.38).

2.2.(c) Show that the state in Eq. (2.40) corresponds to making singlet bonds between each adjacent spin — note that shared singlets are possible for $|S| = 1$ in 1D.

Question 2.3: Four-site Heisenberg model

Consider the Heisenberg Hamiltonian on a square, with each site connected to its two nearest neighbours. By considering resonating valence bond states or otherwise, find the ground state of this model.

Lecture 3

Real insulators

The previous lecture discussed in some detail the Hubbard model, and its special case at half filling, the antiferromagnetic Heisenberg model, motivated by the aim of describing magnetic insulators. This lecture puts back in some of the details present in real insulators, such as: interactions beyond nearest neighbours; the chemistry of insulators — typically containing magnetic cations with neighbouring ligand orbitals, hence kinetic exchange due to overlap between magnetic ion orbitals is small but exchange via the oxygen can be large; the effects of multiple orbitals per atom, and their splitting in the crystal field; and anisotropy of interactions due to anisotropy in the crystal structure.

We will see that for most of these effects, the Hubbard model still captures well the essential physics, but identifying the parameters in it becomes more complicated. Anisotropy however can lead to extra behaviour, including ferrimagnetism — canted antiferromagnetism leading to a parasitic ferromagnetic order parameter. The aim of this lecture is therefore not to calculate the real parameters, but rather to understand what features of a magnetic insulator may effect them, and where possible to understand the general trends.

3.1 Extended Hubbard model: non-local interaction effects

We start by continuing to consider a regular lattice of magnetic ions, and ask how the Hubbard model relates to a microscopic picture of electrons in localised Wannier orbitals.

If we consider a single band, written in the Wannier basis, then we may write the interaction term as:

$$H = \sum_{ijkl} a^\dagger_{i\sigma} a^\dagger_{j\sigma'} V_{ijkl} a_{k\sigma'} a_{l\sigma} \tag{3.1}$$

$$V_{ijkl} = \iint d\mathbf{r}_1 d\mathbf{r}_2 \phi^*(\mathbf{r}_1 - \mathbf{R}_i)\phi^*(\mathbf{r}_2 - \mathbf{R}_j) V(\mathbf{r}_1 - \mathbf{r}_2)\phi(\mathbf{r}_2 - \mathbf{R}_k)\phi(\mathbf{r}_1 - \mathbf{R}_l) \tag{3.2}$$

If the wavefunctions are reasonably well localised, then we may restrict the sum to on-site and nearest-neighbour terms; i.e. for i, j, k, l we take at most two different sites. We will first write down the various possible forms these terms may take, and then discuss their characteristic size — as each term depends different on the overlap matrix elements between nearest neighbour sites — and then discuss how these terms affect the magnetic state of the system.

Choice of site	Contribution	Description
$i = j = k = l$	$U n_{i\uparrow} n_{j\downarrow}$	Basic Hubbard term.
$i = j = k, l$ $i = j = l, k$ $\dots$	$A(a^\dagger_{i\sigma} a_{j\sigma} + \text{H.c.})(n_{i,-\sigma} + n_{j,-\sigma})$	Assisted hopping
$i = l, k = j$	$V n_{i\sigma} n_{j\sigma'}$	NN interaction
$i = k, j = l$	$F a^\dagger_{i\sigma} a^\dagger_{j\sigma'} a_{i\sigma'} a_{j\sigma}$	NN exchange
$i = j, k = l$	$P a^\dagger_{i\uparrow} a^\dagger_{i\downarrow} a_{j\downarrow} a_{j\uparrow}$	Pair hopping

Relative size of coefficients

The size of the coefficients U, A, V, F, P in the above expressions can be determined by calculating the overlap matrix elements. Compared to the Hubbard term, the other expressions are reduced either because they involved interactions between different positions, or because they involve overlap of wavefunctions centred on different sites. Generally the latter effect causes a faster decrease — i.e. the decay of wavefunction is faster than the decay of Coulomb interaction — and so those terms that involve more off-site overlaps are smaller. Using this basis, one can list the terms in decreasing order of size as:

$$\begin{aligned}
U &\sim \iint d\mathbf{r}_1 d\mathbf{r}_2 V(\mathbf{r}_1 - \mathbf{r}_2) \left|\phi(\mathbf{r}_1)\right|^2 \left|\phi(\mathbf{r}_2)\right|^2 \\
V &\sim \iint d\mathbf{r}_1 d\mathbf{r}_2 V(\mathbf{r}_1 - \mathbf{r}_2) \left|\phi\left[\mathbf{r}_1 - \frac{\mathbf{R}}{2}\right]\right|^2 \left|\phi\left[\mathbf{r}_2 + \frac{\mathbf{R}}{2}\right]\right|^2 \\
A &\sim \iint d\mathbf{r}_1 d\mathbf{r}_2 V(\mathbf{r}_1 - \mathbf{r}_2) \phi^*\left[\mathbf{r}_1 - \frac{\mathbf{R}}{2}\right] \phi\left[\mathbf{r}_1 + \frac{\mathbf{R}}{2}\right] \left|\phi\left[\mathbf{r}_2 + \frac{\mathbf{R}}{2}\right]\right|^2 \\
F, P &\sim \iint d\mathbf{r}_1 d\mathbf{r}_2 V(\mathbf{r}_1 - \mathbf{r}_2) \phi^*\left[\mathbf{r}_1 - \frac{\mathbf{R}}{2}\right] \phi\left[\mathbf{r}_1 + \frac{\mathbf{R}}{2}\right] \phi^*\left[\mathbf{r}_2 - \frac{\mathbf{R}}{2}\right] \phi\left[\mathbf{r}_2 + \frac{\mathbf{R}}{2}\right]
\end{aligned}$$

(NB; the last two terms, F, P are both of the same order of magnitude, but are not the same; the expression above is correct for P; the term for F should differ by taking complex conjugate of the wavefunctions involving $\mathbf{r}_2$).

An order of magnitude estimate based on reduction in Coulomb interaction or on overlap implies:

$$\frac{V}{U} \sim \frac{a_{\text{Bohr}}}{R}, \qquad \frac{A}{V} \sim \frac{F}{A} \sim \frac{P}{A} \sim \exp\left(-\frac{R}{a_{\text{Bohr}}}\right).$$

Calculations by Hubbard [8] for 3d electrons, and accounting for screening in the Coulomb interaction yield numbers:

$$U \sim 10\text{eV} > V \sim 2\text{eV} > A \sim 1/2\text{eV} > F, P \sim 1/40\text{eV}$$

Effect of nearest neighbour terms on magnetic phases

Considering the above terms in decreasing order of magnitude, we may now discuss what effect they have on magnetic phases.

V The term V does not directly affect magnetism, but it does allow for the possibility of other, possibly competing, charge ordering. In particular, near filling 1/4, it favours a strong charge density wave, which may coexist with antiferromagnetism, of the form:

$$\underline{\uparrow}\ \underline{\ \ }\ \underline{\downarrow}\ \underline{\ \ }\ \underline{\uparrow}\ \underline{\ \ }\ \underline{\downarrow}\ \underline{\ \ }\ \underline{\uparrow}\ \underline{\ \ }\ \underline{\downarrow}\ \underline{\ \ }$$

A This term also has no particularly strong affect on magnetic ordering, but one should note that this term is the only term in the extended Hubbard model that does not respect electron-hole symmetry, and leads to larger effective hopping when hole doped. For a mean electron density ρ, the hopping element can be schematically written as $-t \rightarrow -t + A(\rho - 1/2)$.

P This also has no significant effect on magnetic phases, but may have an effect on superconductivity.

F This term, although weak, has a significant effect, in that it favour ferromagnetism. What is particularly important is that in the limit of strong interactions, i.e. if $U, V, A, P, F \gg t$, one has a kinetic exchange mediated antiferromagnetic interaction t^2/U, which will become smaller than F; hence the previous results on antiferromagnetism in the infinite U limit of the Hubbard model may not necessarily describe real materials in the strong interaction regime.

3.2 Anions and superexchange

The above descriptions assumed direct overlap and direct hopping between magnetic ions. In most real insulators of interest, the crystal structure has oxygen anions (ligands) between the magnetic ions, as in the octahedral perovskite structure shown in Fig. 3.1(a).

The direct overlap between magnetic ion orbitals is therefore small, and the relevant question involves hopping via the oxygen orbitals, as illustrated in Fig. 3.1(b). Superexchange is the name given to this exchange via intermediate ions — in some books, superexchange is used more generally to describe what in these notes we refer to as kinetic exchange. The usage is normally clear from context; the reason for this confusion is that in almost all relevant cases, kinetic exchange directly between magnetic ions does not exist, due to lack of overlap, and the real process is the superexchange discussed here.

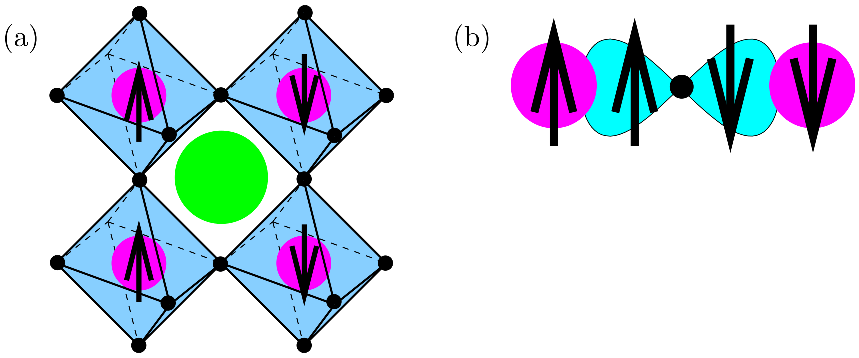


Figure 3.1: (a)Perovskite structure, indicating magnetic cations (marked as having spins) surrounded by octahedral configuration of oxygen anions (ligands). (b) Schematic overlap of oxygen p-orbitals and magnetic cation orbitals [See also Fig. 3.2 for more precise picture of orbitals]

Possible superexchange elements

Let us study how the energy of the magnetic ion configurations ↑ . . . ↓ and ↑ . . . ↑ differ, by considering virtual hopping processes via the oxygen ion that may lower the energy of the antiferromagnetic combination.

The oxygen orbitals are typically lower in energy than the magnetic ion orbitals — oxygen being more electronegative, and so there is a significant single particle energy cost Δ to transferring an electron from the oxygen orbitals to the magnetic ion. One should thus consider higher order perturbative processes such as:

$$\underline{\uparrow}\ \ \underline{\uparrow\downarrow}\ \ \underline{\downarrow} \overset{t}{\Longleftrightarrow} \overbrace{\underline{\uparrow}\ \ \underline{\downarrow}\ \ \underline{\uparrow\downarrow}}^{1/\Delta} \overset{t}{\Longleftrightarrow} \overbrace{\underline{}\ \ \underline{\uparrow\downarrow}\ \ \underline{\uparrow\downarrow}}^{1/U}$$

This is the lowest order process that allows exchange of the magnetic ion spins, and so it leads to an antiferromagnetic coupling that is of the order:

$$J \approx t\frac{1}{\Delta}t\frac{1}{U}t\frac{1}{\Delta}t = \left(\frac{t^2}{\Delta}\right)^2 \frac{1}{U} \tag{3.3}$$

The second rewriting of this expression highlights another way of viewing this term; t/Δ describes the degree of admixture (in first order perturbation theory) between the ligand and magnetic ion orbitals:

$$|\psi'_M\rangle = |\psi_M\rangle + \left(\frac{t}{\Delta}\right)|\psi_{\text{ligand}}\rangle.$$

One can then consider $t' = t^2/\Delta$ as a characteristic size of the indirect hopping between the magnetic ions, and hence $J = t'^2/U$ is the expected superexchange contribution.

At the same order in perturbation theory, one may note that there is an overlap of the electronic wavefunctions of the cations because they include admixture on the ligand. This leads to a repulsive interaction between electronic states, corresponding to the standard ferromagnetic direct exchange. This can be written as:

$$U_{\text{eff}} \approx \left(\frac{t}{\Delta}\right)^4 U_a \tag{3.4}$$

where U_a is the Hubbard element calculated on the anion. This expression is purely schematic as an accurate calculation of this term should take into account also the excited (anti-bonding) p orbital states of the oxygen ion There are also several more terms one might consider writing down, and question 3.1 discusses another way to see that Eq. (3.3) is the leading order kinetic exchange. The two terms we have considered — superexchange when a full hopping pathway between ions exist, and direct exchange whenever electron orbitals overlap on the ligand — are however sufficient to describe in basic terms how alignment of orbitals can change magnetic properties.

Effects of geometry and Goodenough-Kanamori rules

The basic idea of the effect of geometry is illustrated in Fig. 3.2. Because the antiferromagnetic superexchange process relied on a full hopping pathway from cation to ligand to cation, it can only exist between cations which overlap with the *same p* orbital of the ligand. On the other hand, the direct exchange relies on spatial overlap, not hopping, and so still exists even if the cations overlap with different ligand orbitals. Since the superexchange is typically larger (as $\Delta^2/UU_a > 1$), if it exists, it generally wins, and cations which share a common ligand overlap couple antiferromagnetically, while those which do not share a common ligand orbital couple ferromagnetically. This statement should of course be read with the caveat expressed above that a real calculation ought to include effects of higher excited orbitals, which might change the relative strengths of these terms.

3.3 Crystal field and anisotropy

The preceeding discussion made use of the existence of multiple orbitals, and their energy splitting, however we have not yet discussed the fact that magnetic ions — which are generally either transition metals with partially filled d orbitals, or rare earth ions with partially filled f orbitals — have multiple orbital states per ion, and so one must understand why considering only the spin $\mathbf{S}$ rather than angular momentum $\mathbf{J}$ of electrons can be appropriate.

For an isolated partially filled shell atom, there already exists a splitting due to electrostatic interaction between the electrons within an atom leading to Hund's rules, which require maximal $|\mathbf{S}|$ and $|\mathbf{L}|$ for each atom. This however leaves a remaining degeneracy of which direction the spin and angular momentum point in, giving $(2|\mathbf{S}|+1)(2|\mathbf{L}|+1)$ states.

For an atom in isolation, this remaining degeneracy is then partially reduced by spin-orbit coupling $H_{SO} = \lambda \mathbf{L} \cdot \mathbf{S}$ leaving a degeneracy $(2|\mathbf{J}|+1)$,

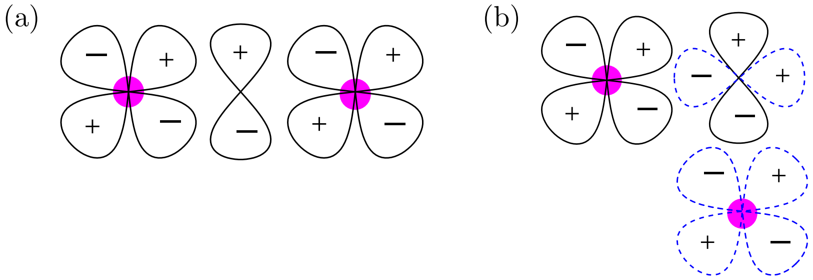


Figure 3.2: Possible relative configurations of d orbitals of magnetic cations and oxygen p-orbitals. In configuration (a) hopping between the magnetic ions via superexchange is possible, and so this leads to a dominant antiferromagnetic coupling. In configuration (b) hopping is not possible because of the failure of overlap between the bottom cation and the oxygen; spatial overlap of electron orbitals does however still exist, and so an direct interaction exists, favouring ferromagnetism.

which would lead to each atom having a spin larger than 1/2 in general. However, we are not interested in ions in isolation, but rather in a solid, and so we have in addition a crystal field — i.e. an anisotropic electrostatic field due to the neighbouring ions, and so $\delta H = H_{\text{crystal}} + H_{SO}$.

The primary distinction between magnetic properties of transition metal compounds and those of rare earth compounds is that in transition metals, spin orbit effects are weaker, so one first considers crystal fields, and then treats spin orbit as a perturbation. For rare earths, the electrons are more localised so crystal field effects are weaker, and spin orbit is stronger, and so spin orbit splitting dominates. For rare earths, it is thus often appropriate to consider writing Heisenberg models but with $\mathbf{S} \to \mathbf{J}$, as we discussed in the previous lecture when considering Heisenberg models with higher spins. What happens in transition metals is the subject of the next section.

Quenching of angular momentum

The crystal field Hamiltonian is a real potential, favouring particular orientation of wavefunctions. Because it is a real, time symmetric perturbation, one *may* choose eigenstates which are also real, while the angular momentum is a purely imaginary operator, which would imply that for such an eigenstate:

$$\langle \psi | \mathbf{L} | \psi \rangle = 0,$$

i.e. angular momentum is quenched, and only the spin degree of freedom remains. There is however a caveat — if the crystal field does not entirely remove degeneracy, then from a pair of degenerate real wavefunctions, one may construct combinations for which the above equality does not hold. It is however usually the case that such degeneracies do not survive, since it becomes energetically favourable to pay the (quadratic) elastic cost of deforming a symmetric lattice to gain a linear increase in electronic energy

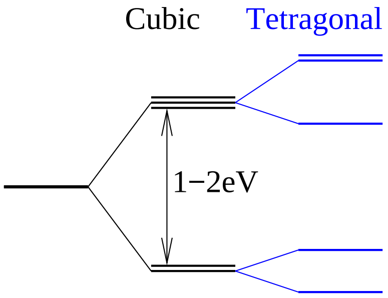


Figure 3.3: e-t splitting in an octahedral crystal field

from splitting a degenerate state. Thus, while the following will principally discuss the dominant splitting due to octahedrally symmetric terms, one should bear in mind that an additional reduction to tetragonal, or lower, symmetry will also occur.

The next three sections, Sec. 3.3– 3.3 discuss three further consequences of the crystal field splitting:

- Splitting of energy levels, and what symmetry remains.
- Anisotropy of single spin terms.
- Anisotropy of exchange interactions.

Structure of typical (octahedrally split) d-orbitals

Given d orbitals, which have representative functions like $x^2, y^2, z^2, xy, xz, yz$, and the representation table for cubic or octahedral symmetry, one finds that (for both cases), the splitting due to a field with such symmetry is between terms like:

$$\begin{cases} e_g : & x^2-y^2, x^2-z^2 \\ t_{2g} : & xy, xz, yz \end{cases}$$

thus, the ground state would have either 2-fold or 3-fold symmetry due to orbital symmetry, as illustrated in Fig. 3.3 . The typical scale of this splitting is $1-2$eV. As discussed above, further symmetry breaking by tetragonal distortion is likely, but will be on a smaller energy scale.

Anisotropy of on-site interaction

Assuming complete quenching of angular momentum, i.e. $\langle\psi|\mathbf{L}|\psi\rangle = 0$, it is clear that the spin-orbit perturbation $H_{SO} = \lambda\mathbf{L}\cdot\mathbf{S}$ has no effect at first order in perturbation theory, as there is no remaining degeneracy to break. However, this term may still have an effect in second order perturbation theory. Considering an unknown spin state, but integrating out orbital excitations perturbatively, one may write:

$$\delta H = |\lambda|^2 \sum_n \frac{\langle 0|L_\mu|n\rangle\langle n|L_\nu|0\rangle}{E_0 - E_n} S_\mu S_\nu = -S_\mu \Lambda_{\mu\nu} S_\nu \tag{3.5}$$

where generally $\Lambda_{\mu\nu}$ is not proportional to the identity matrix, but depends on the ordering of the excited orbitals. For single electron ions, where $|\mathbf{S}| = 1/2$, this term can have no effect, since the above expression has time reversal symmetry and so the two configurations must have identical energies; this means that for $|\mathbf{S}| = 1/2$ the matrix must be proportional to the identity. This anisotropy is discussed further in question 3.2.

Anisotropic exchange: Dzyaloshinskii-Moriya term

For $|\mathbf{S}| = 1/2$, since the above anisotropy plays no role, a weaker anisotropy of exchange interactions may be relevant — such a term always exists, but is not as large as the on-site anisotropy, hence only becomes relevant when the on-site anisotropy vanishes because of symmetry of time reversed states. This term is possible because it involves two adjacent spins, allowing time reversal to be preserved between their two pairs of two eigenstates, but no longer requiring complete equality of all eigenvalues.

We thus consider:

$$H = \lambda \mathbf{L}_1 \cdot \mathbf{S}_1 + \lambda \mathbf{L}_2 \cdot \mathbf{S}_2 + \hat{J} \mathbf{S}_1 \cdot \mathbf{S}_2$$

where $\hat{J}$ is an operator that depends on what orbital state the electron is in; normally this fact is irrelevant because the orbital states are split, and so one may normally consider $\hat{J} \to J = \langle 0|\hat{J}|0\rangle$ but combined with the spin-orbit coupling we may write:

$$\delta H = -\lambda \sum_{n_1} \left[\frac{\langle 0_1 0_2|\mathbf{L}_1 \cdot \mathbf{S}_1|n_1 0_2\rangle\langle n_1 0_2|\hat{J}\mathbf{S}_1 \cdot \mathbf{S}_1|0_1 0_2\rangle}{E_{n_1 0_2} - E_{0_1 0_2}} \right. \\ \left. + \frac{\langle 0_1 0_2|\hat{J}\mathbf{S}_1 \cdot \mathbf{S}_1|n_1 0_2\rangle\langle n_1 0_2|\mathbf{L}_1 \cdot \mathbf{S}_1|0_1 0_2\rangle}{E_{n_1 0_2} - E_{0_1 0_2}} \right] + (1 \leftrightarrow 2) \tag{3.6}$$

This may be simplified because, for real wavefunctions one has $\langle\psi|\mathbf{L}|\phi\rangle = -\langle\phi|\mathbf{L}|\psi\rangle$, and $\langle\psi|\hat{J}|\phi\rangle = \langle\phi|\hat{J}|\psi\rangle$, hence:

$$\delta H = -\lambda \sum_{n_1} \left[\frac{\langle 0_1|\mathbf{L}_{1\mu}|n_1\rangle\langle n_1 0_2|\hat{J}|0_1 0_2\rangle}{E_{n_1 0_2} - E_{0_1 0_2}} \right] [S_{1\mu}, \mathbf{S}_1 \cdot \mathbf{S}_2] + (1 \leftrightarrow 2) \tag{3.7}$$

Writing the sum over n_1 as η_1, and using the commutation relations to write, e.g.:

$$[S_{1x}, S_{1x}S_{2x} + S_{1y}S_{2y} + S_{1z}S_{2z}] = i\left(0 + S_{1z}S_{2y} - S_{1y}S_{2z}\right) = -i\left(\mathbf{S}_1 \times \mathbf{S}_2\right)_x \tag{3.8}$$

one can write:

$$\delta H = (\eta_1 - \eta_2) \cdot (\mathbf{S}_1 \times \mathbf{S}_2), \quad i\eta_1 = -\lambda \sum_{n_1} \left[\frac{\langle 0_1|\mathbf{L}_1|n_1\rangle\langle n_1 0_2|\hat{J}|0_1 0_2\rangle}{E_{n_1 0_2} - E_{0_1 0_2}} \right] \tag{3.9}$$

It is thus clear that such a term exists only between two *different* atoms. However, "*different*" need not mean different species of ion, but can just

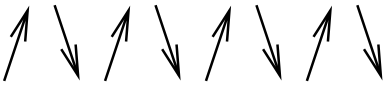


Figure 3.4: Canted antiferromagnet, producing ferrimagnetism

mean different chemical environments, such as in a bipartite lattice, where the symmetry of orbitals may be different between the two types of lattice site.

This interaction is known as the Dzyaloshinskii-Moriya interaction, and it favours an alignment of adjacent atoms that is neither parallel nor antiparallel. When combined with an antiferromagnetic interaction, it will cause the spins to cant away from their antiferromagnetic configuration, as illustrated in Fig. 3.4, so as to give a non-vanishing expectation of the net magnetisation. This ferromagnetism parasitic upon antiferromagnetism is a form of ferrimagnetism.

3.4 Orbital degeneracy and competing symmetries

To complete this exploration of other considerations, let us mention finally another way in which orbital degeneracy can affect magnetic configuration. If one has degenerate orbitals, then as well as considering virtual hopping affecting the energy of two spin states in one orbital, it can also allow hopping between different orbitals, i.e.

$$\begin{array}{cc} \underline{} & \underline{\uparrow} \\ \underline{\uparrow} & \underline{} \end{array} \Longleftrightarrow \begin{array}{cc} \underline{\uparrow} & \underline{} \\ \underline{\uparrow} & \underline{} \end{array}$$

We may then consider a number of different "antiferromagnetic" configurations:

$$|A\rangle = \frac{1}{\sqrt{2}}\left(\begin{array}{cc} \underline{} & \underline{} \\ \underline{\uparrow} & \underline{\downarrow} \end{array} - \begin{array}{cc} \underline{} & \underline{} \\ \underline{\downarrow} & \underline{\uparrow} \end{array} \right)$$

$$|B\rangle = \frac{1}{\sqrt{2}}\left(\begin{array}{cc} \underline{} & \underline{\downarrow} \\ \underline{\uparrow} & \underline{} \end{array} - \begin{array}{cc} \underline{\downarrow} & \underline{} \\ \underline{} & \underline{\uparrow} \end{array} \right)$$

$$|C\rangle = \frac{1}{\sqrt{2}}\left(\begin{array}{cc} \underline{} & \underline{\uparrow} \\ \underline{\uparrow} & \underline{} \end{array} - \begin{array}{cc} \underline{\uparrow} & \underline{} \\ \underline{} & \underline{\uparrow} \end{array} \right)$$

The last of these is antiferromagnetic in orbital ordering, but ferromagnetic in spin ordering.

We now consider the energy of these states, assuming that all hoppings are the same. For the first two, the energy is exactly equal (assuming the same on-site repulsion, independent of orbital wavefunction), i.e. $E = -2t^2/U$. However the last has a lower energy, since the intermediate state, having two aligned spins, in different orbitals, will, according to

Hund's rules, be the ground state configuration of two spins on one atom, maximising total spin. If we call this Hund's rule splitting J, then this term has energy $E = -2t^2/(U - J)$, and is thus the ground state.

Hence, orbital degeneracy allows the antiferromagnetic gain of hopping to be transfered to orbital order, while also gaining the ferromagnetic interaction energy as described by on-site Hund's rule terms, hence showing that ferromagnetism is favoured by degeneracy of orbitals. This idea will appear in a rather different guise toward the end of the next lecture too.

Questions

* **Question 3.1: The importance of interactions in superexchange**

The aim of this question is to study the difference between the strength of superexchange with and without on-site interactions. We consider in both cases a model magnetic ion – ligand – magnetic ion, i.e.:

3.1.(a) Non-interacting case: Consider the free-fermion Hamiltonian for these three states:

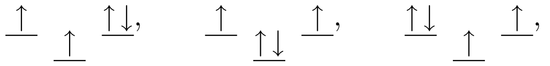
$$H = \begin{pmatrix} 0 & t & 0 \\ t & -\Delta & t \\ 0 & t & 0 \end{pmatrix} \tag{3.10}$$

Find the energies of the ferromagnetic and antiferromagnetic configurations of this model — i.e. either populate these eigenstates with spins $\uparrow\uparrow\uparrow\downarrow$ or $\uparrow\uparrow\downarrow\downarrow$, and hence produce the non-interacting exchange splitting to leading non-zero order in t.

3.1.(b) Interacting: Ferromagnetism.

In the interacting case, we add an on-site energy U if two electrons occupy the magnetic ion orbitals, but neglect the weaker on-site interaction in the ligand orbital.

By working in the basis of states:



find the ground state energy, and show that to leading order in the limit of large $U \gg t, \Delta$.

$$E_{\mathrm{FM}} \simeq -2\Delta + V - \frac{2t^2}{\Delta + U} \tag{3.11}$$

3.1.(c) Interacting: Antiferromagnetism

In this case, there are nine basis states, however by noticing symmetries it is possible to greatly simplify this problem. Thus or otherwise, show that to leading order in $U \gg t, \Delta$, the energy is the same as Eq. (3.11), and hence that the non-interacting estimate is wrong for the interacting case.

3.1.(d) Interacting: Actual energy difference.

By considering the next order term, proportional to t^4/U^3, show that the ferromagnetic configuration is in fact higher energy. The fact that this difference exists only at order t^4 justifies the cavalier approach used in Sec. 3.2.

Question 3.2: Anisotropic spins in a magnetic field

Consider combining the anisotropy due to matrix elements of L described in Eq. (3.5) with a magnetic field:

$$\delta H = \lambda \mathbf{L} \cdot \mathbf{S} + \mu_B \mathbf{H} \cdot (2\mathbf{S} + \mathbf{L}) \tag{3.12}$$

Following the same logic as before, show that the Zeeman term becomes anisotropic.

Lecture 4

Metals, itinerant electrons, and the possibility of Ferromagnetism

In the previous chapters, we studied magnetisation by starting from localised moments — i.e. electrons were locked on a given site by their strong repulsive interactions preventing hopping to already occupied sites, and magnetism arose by ordering of these localised spins. The aim of this lecture is instead to consider magnetisation when electrons are not localised, and so a given site need not have a well defined spin moment. For simplicity, we will still however consider the limit of good screening, in which interactions between electrons can be approximated by an on-site interaction, thus our starting Hamiltonian will be:

$$H \sum_{k\sigma} \epsilon_k a^\dagger_{k\sigma} a_{k\sigma} + \frac{U}{2} \int d^3 r a^\dagger_\sigma(r) a^\dagger_{\sigma'}(r) a_{\sigma'}(r) a_\sigma(r) \tag{4.1}$$

Despite the warnings in lecture 1 about the dangers of mean-field theory for ferromagnetism, we will begin by discussing the results of a mean-field (Stoner) theory for itinerant electrons, and then spend the rest of the lecture discussing the obstacles to realisation of this mean-field theory, thereby discovering the conditions under which the Stoner theory is of relevance to real systems.

4.1 Mean-field theory: Stoner instability

Before writing the mean-field theory description of the transition, it is convenient to re-arrange the interaction part of Eq. (4.1), making use of the fact that on-site interactions only exist between opposite spins, so:

$$H_{\text{int}} = U \int d^3 r n_\uparrow(r) n_\downarrow(r) = \frac{U}{4} \int d^3 r \left[(n_\uparrow(r) + n_\downarrow(r))^2 - (n_\uparrow(r) - n_\downarrow(r))^2 \right] \tag{4.2}$$

We may thus consider the second term to be the energy gain associated with ferromagnetism.

Using a mean-field decoupling — i.e. writing

$$\begin{aligned}
x &= x - \langle x\rangle + \langle x\rangle \\
x^2 &\simeq \langle x\rangle^2 + 2\langle x\rangle(x - \langle x\rangle) + \mathcal{O}[(x - \langle x\rangle)^2] = 2x\langle x\rangle - \langle x\rangle^2,
\end{aligned}$$

where we neglected the small fluctuation term, we may decompose our Hamiltonian by assuming an order parameter:

$$m = \frac{1}{2}\langle n_\uparrow - n_\downarrow\rangle$$

and so:

$$H = \sum_{k\sigma}(\epsilon_k - \sigma m U)a^\dagger_{k\sigma}a_{k\sigma} + VUm^2, \tag{4.3}$$

where V is the system volume. We then study the conditions under which a self-consistent solution of Eq. (4.2) and Eq. (4.3) exist (other than the always present solution $m = 0$).

To have $m \neq 0$, we need the chemical potentials of $n_\uparrow, n_\downarrow$ to be different, so we consider a solution in which the chemical potentials for the two spins differ by δ, then:

$$n_{\uparrow\downarrow} = \int_0^{\epsilon_F \pm \delta} d\epsilon\rho(\epsilon) \simeq n_0 \pm \rho(\epsilon_F)\delta, \qquad m = \rho(\epsilon_F)\delta. \tag{4.4}$$

Then, considering the total energy of this state as:

$$\begin{aligned}
\frac{E}{V} &= \sum_\sigma \int_0^{\epsilon_F + \sigma\delta} d\epsilon(\epsilon - \sigma m U)\rho(\epsilon) + Um^2 \\
&\simeq \int_0^{\epsilon_F} d\epsilon\epsilon\rho(\epsilon) + 2\int_0^\delta dx(x - Um)\rho(\epsilon_F + x)dx + Um^2 \\
&\simeq \text{const} + \left(\delta^2 - 2Um\delta\right)\rho(\epsilon_F) + Um^2 \qquad (4.5) \\
&= \text{const} + \frac{m^2}{\rho(\epsilon_F)} - Um^2. \qquad (4.6)
\end{aligned}$$

In the last line, we used the self consistency condition relating m and δ from Eq. (4.4). It is thus clear that the condition for the $m = 0$ state to become unstable, and thus replaced by the $m \neq 0$ state is that $U\rho(\epsilon_F) > 1$.

Interpreting instability via susceptibility

The above working may appear unnecessarily complicated — one could assume from the beginning that the shift of chemical potential is seen from the Hamiltonian to be given by: $\delta = mU$, and then note that the condition for a non-trivial solution is that $m = \delta/U = \rho(\epsilon_F)\delta$. Such an approach does not however make clear which side of this equality the $m = 0$ state becomes unstable.

Another rewriting of the same instability, which is useful for connecting to the subsequent discussions of other instabilities, is to note that, at a second order transition to a magnetised state, the susceptibility to a magnetic field will diverge. The Hamiltonian with a magnetic field is:

$$H = \sum_{k\sigma} [\epsilon_k - \sigma(mU + \frac{g\mu_B}{2} H)] a^\dagger_{k\sigma} a_{k\sigma} + VUm^2. \tag{4.7}$$

Taking the straightforward "self-consistency" route advocated a moment before, we find:

$$\begin{aligned} \delta = mU + \frac{g\mu_B}{2} H &= \frac{m}{\rho(\epsilon_F)}, \\ m\,[1 - U\rho(\epsilon_F)] &= \frac{g\mu_B}{2} H \rho(\epsilon_F) \end{aligned} \tag{4.8}$$

hence, the paramagnetic susceptibility has become:

$$\chi = (g\mu_B) \frac{g\mu_B}{2} \rho(\epsilon_F) \frac{1}{[1 - U\rho(\epsilon_F)]} = \frac{\chi_{\text{Pauli}}}{1 - U\rho(\epsilon_F)}. \tag{4.9}$$

where $\chi_{\text{Pauli}} = (g\mu_B)^2 \rho(\epsilon_F)/2$ is the Pauli susceptibility. Considered at $T \neq 0$, the Stoner instability is associated with exactly the kind of "molecular field" discussed in the Curie-Weiss approach, as applied to the Pauli paramagnetism of a Fermi gas.

Susceptibility via free energy

For completeness, note that one can also generalise the approach in Sec. 4.1 to find the susceptibility. The change to the Hamiltonian is:

$$\langle \delta H \rangle = -\frac{g\mu_B}{2} V (n_\uparrow - n_\downarrow)$$

and so:

$$\frac{E}{V} = \text{const} + \frac{m^2}{\rho(\epsilon_F)} - Um^2 - g\mu_B H m \tag{4.10}$$

which is minimised by the solution as in Eq. (4.8).

Outline of complications to ferromagnetism

The above concludes the mean-field theory; there are however at least three serious concerns with the above approach:

1. We only studied the expansion for small magnetisations; this is valid assuming there is indeed a second order transition, however without investigating the free energy more fully, one may find that rather than an instability associated with $\delta \simeq 0$ becoming unstable, one instead has an instability when $E(\delta) < E(0)$ at a finite δ, i.e. a first order transition.

2. We only considered instability to a ferromagnetic state, and so, were a different magnetic order to also be possible, such an ordering might intervene before the ferromagnetic state arises. This idea is what we address in section 4.2.

3. As stated in section 1.3, if ferromagnetism can exist, this means interactions are so strong as to compete with kinetic energy costs, and so one must allow for the possibility of wavefunction renormalisation — i.e. one cannot use the bare interaction, but should account for the reduced interaction energy when one allows for the tendency of electrons to avoid each other in a strongly interacting system. This will be addressed in Sec. 4.3.

4.2 Competing instabilities

As yet we have only considered the possibility of ferromagnetism, and so would miss any other instability that might develop. This section will therefore consider instability to a competing finite momentum magnetic ordering. Following the logic of section 4.1, we will proceed by considering a perturbation:

$$\delta H = -g\mu_B \int d^3r \mathbf{S}(r) \cdot \mathbf{H} e^{i\mathbf{Q}\cdot\mathbf{r}} \tag{4.11}$$

where the spin density operator $\mathbf{S}(r) = a^\dagger_\sigma(r)\mathbf{S}_{\sigma\sigma'}a_{\sigma'}(r)$, and then study the susceptibility to magnetisation with the same wavevector; a finite $\mathbf{Q}$ instability is signalled by a diverging susceptibility.

Finite momentum susceptibility

Since the Hubbard interaction of the Hamiltonian in Eq. (4.1) is isotropic — i.e. invariant under rotations of spin [see Eq. (1.15)] — it makes no difference which direction of $\mathbf{H}$ we choose. We therefore choose $\mathbf{H} = H\hat{\mathbf{x}}$, and so using:

$$S_x(\mathbf{Q}) = \int d^3r S_x e^{i\mathbf{Q}\cdot\mathbf{r}} = \sum_p \frac{1}{2}\left(a^\dagger_{\mathbf{p}+\mathbf{Q},\uparrow}a_{\mathbf{p},\downarrow} + a^\dagger_{\mathbf{p}+\mathbf{Q},\downarrow}a_{\mathbf{p},\uparrow}\right) \tag{4.12}$$

we will write $\delta H = -g\mu_B S_x(\mathbf{Q})$.

Perturbation theory

Our aim will be to decouple the interactions in terms of an order parameter $\tilde{m}_\mathbf{Q}$, and then ensure the self-consistency condition by finding the expectation of $\tilde{m}_\mathbf{Q}$ using the first order perturbation result for the wavefunction:

$$\langle \tilde{m}_\mathbf{Q} \rangle = 2\Re\left[\sum_j \frac{\langle 0|\tilde{m}_\mathbf{Q}|j\rangle\langle j|\delta H_{\text{tot}}|0\rangle}{E_0 - E_j}\right]. \tag{4.13}$$

Here, the fact this is twice the real part comes from perturbing either the left or the right wavefunction to give a first order contribution. The perturbation δH_{tot} includes both δH due to the applied field as well as the decoupled interaction term.

Decoupling interaction to find total perturbation

We now proceed to decoupling the interaction; it is easiest to work from Eq. (4.2) in momentum space:

$$\begin{aligned} H_{int} &= U \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} a^\dagger_{\mathbf{k}+\mathbf{q}\uparrow} a^\dagger_{\mathbf{k}'-\mathbf{q}\downarrow} a_{\mathbf{k}'\downarrow} a_{\mathbf{k}\uparrow} \\ &\simeq -U \sum_{\mathbf{k}\mathbf{k}'} \left[\left(a^\dagger_{\mathbf{k}'+\mathbf{Q}\uparrow} a_{\mathbf{k}'\downarrow} \right) \left(a^\dagger_{\mathbf{k}-\mathbf{Q}\downarrow} a_{\mathbf{k}\uparrow} \right) + \left(a^\dagger_{\mathbf{k}'+\mathbf{Q}\downarrow} a_{\mathbf{k}'\uparrow} \right) \left(a^\dagger_{\mathbf{k}-\mathbf{Q}\uparrow} a_{\mathbf{k}\downarrow} \right) \right]. \end{aligned} \tag{4.14}$$

In the second line we took $\mathbf{q} = \pm\mathbf{Q} + \mathbf{k}' - \mathbf{k}$ and so restricted the sum to terms having the same form as the perturbation in Eq. (4.12) — this is analogous to considering uniform order parameters in the ferromagnetic case. We may similarly also assume that

$$\left\langle \sum_k a^\dagger_{\mathbf{k}-\mathbf{Q}\downarrow} a_{\mathbf{k}\uparrow} \right\rangle = \left\langle \sum_k a^\dagger_{\mathbf{k}-\mathbf{Q}\downarrow} a_{\mathbf{k}\uparrow} \right\rangle = \langle S_x(-\mathbf{Q}) \rangle \tag{4.15}$$

which is equivalent to the statement $\langle S_y(-\mathbf{Q})\rangle = 0$. In terms of this parameterisation, the decoupled interaction becomes:

$$H_{\text{int}} \to -2U \langle S_x(-\mathbf{Q})\rangle S_x(\mathbf{Q}). \tag{4.16}$$

This suggests — as can be confirmed by looking at Eq. (4.13) — that a perturbation like $S_x(\mathbf{Q})$ causes an order parameter like $S_x(-\mathbf{Q})$; the sum in Eq. (4.13) requires that the same momentum is created and destroyed, otherwise overlap with the ground state will vanish.

Combining to find the susceptibility.

Putting the above ingredients together:

$$\delta H_{\text{tot}} = -\left[g\mu_B H + 2U\langle S_x(-\mathbf{Q})\rangle \right] S_x(\mathbf{Q})$$

and so the general consistency condition in Eq. (4.13) can be written explicity here as:

$$\langle S_x(-\mathbf{Q})\rangle = 2\left[g\mu_B H + 2U\langle S_x(-\mathbf{Q})\rangle \right] \sum_j \frac{\langle 0|S_x(-\mathbf{Q})|j\rangle\langle j|S_x(\mathbf{Q})|0\rangle}{E_j - E_0}. \tag{4.17}$$

The sum over excited states has two possible routes; one either flips a spin down, then the same spin up, or *vice-versa*; in either case, the spin that is flipped transfers from momentum $\mathbf{k} \to \mathbf{k} + \mathbf{Q}$, and for such a process to be possible it is necessary that $\mathbf{k}$ is within the Fermi surface while $\mathbf{k} + \mathbf{Q}$ is not. Hence:

$$\sum_j \frac{\langle 0|S_x(-\mathbf{Q})|j\rangle\langle j|S_x(\mathbf{Q})|0\rangle}{E_j - E_0} = \left(\frac{1}{2}\right)^2 2 \sum_{\substack{\mathbf{k}\in F.S. \\ \mathbf{k}+\mathbf{Q}\notin F.S.}} \frac{1}{\epsilon_{\mathbf{k}+\mathbf{Q}} - \epsilon_{\mathbf{k}}}$$

Let us define:

$$\chi_0(\mathbf{Q}) = 2 \sum_{\substack{\mathbf{k}\in F.S.\\ \mathbf{k}+\mathbf{Q}\notin F.S.}} \frac{1}{\epsilon_{\mathbf{k}+\mathbf{Q}} - \epsilon_{\mathbf{k}}} = 2\sum_{\mathbf{k}} \frac{f_{\mathbf{k}}(1-f_{\mathbf{k}+\mathbf{Q}})}{\epsilon_{\mathbf{k}+\mathbf{Q}} - \epsilon_{\mathbf{k}}} \tag{4.18}$$

where $f_{\mathbf{k}}$ is the occupation of that momentum state. We may then write:

$$\langle S_x(-\mathbf{Q})\rangle = \left[\frac{g\mu_B H}{2} + U\langle S_x(-\mathbf{Q})\rangle\right]\chi_0(\mathbf{Q}) \tag{4.19}$$

or equivalently:

$$\chi(\mathbf{Q})H = (g\mu_B)\langle S_x(-\mathbf{Q})\rangle = \chi_{\text{Pauli}}\frac{\chi_0(\mathbf{Q})/\rho(\epsilon_F)}{1-U\chi_0(\mathbf{Q})}. \tag{4.20}$$

This shows exactly the same kind of enhancement as the ferromagnetic susceptibility, showing that interactions are not preferential about "which" magnetic ordering — to have any ordering requires a local moment to exist, which requires having only a single electron per site, hence reducing the interaction energy. The differences of susceptibility depend only on $\chi_0(\mathbf{Q})$, i.e. on the kinetic energy cost associated with such an ordering. An instability will develop at $\mathbf{Q}$ when $U\chi_0(\mathbf{Q}) = 1$, so we next turn to understand this momentum dependent χ_0 better.

Momentum dependence susceptibility

By a series of relabellings, $\mathbf{k} \to -\mathbf{k}$ and noting that $f_{\mathbf{k}} = f(\epsilon_{\mathbf{k}})$ along with $\epsilon_{-\mathbf{k}} = \epsilon_{\mathbf{k}}$ we may write Eq. (4.18) as:

$$\chi_0(\mathbf{Q}) = \sum_{\mathbf{k}}\left[\frac{f_{\mathbf{k}}(1-f_{\mathbf{k}+\mathbf{Q}})}{\epsilon_{\mathbf{k}+\mathbf{Q}} - \epsilon_{\mathbf{k}}} + \frac{f_{\mathbf{k}}(1-f_{\mathbf{k}-\mathbf{Q}})}{\epsilon_{\mathbf{k}-\mathbf{Q}} - \epsilon_{\mathbf{k}}}\right]. \tag{4.21}$$

This then allows yet another relabelling $\mathbf{k} \to \mathbf{k}+\mathbf{Q}$ to give[1]:

$$\chi_0(\mathbf{Q}) = \sum_{\mathbf{k}}\left[\frac{f_{\mathbf{k}}(1-f_{\mathbf{k}+\mathbf{Q}})}{\epsilon_{\mathbf{k}+\mathbf{Q}} - \epsilon_{\mathbf{k}}} + \frac{f_{\mathbf{k}+\mathbf{Q}}(1-f_{\mathbf{k}})}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{Q}}}\right] = \sum_{\mathbf{k}}\frac{f_{\mathbf{k}} - f_{\mathbf{k}+\mathbf{Q}}}{\epsilon_{\mathbf{k}+\mathbf{Q}} - \epsilon_{\mathbf{k}}}. \tag{4.22}$$

The advantage of this final form is that one can connect this result to our previous result for ferromagnetism. If $\mathbf{Q}$ is small, then $\epsilon_{\mathbf{k}+\mathbf{Q}} = \epsilon_{\mathbf{k}} + \mathbf{Q}\cdot\mathbf{v}_{\mathbf{k}}$ and so:

$$\chi_0(\mathbf{Q}) = \sum_{\mathbf{k}}\frac{-(\partial f/\partial\epsilon)\mathbf{Q}\cdot\mathbf{v}_{\mathbf{k}}}{\mathbf{Q}\cdot\mathbf{v}_{\mathbf{k}}} = \sum_{k}\delta(\epsilon-\epsilon_F) = \rho(\epsilon_F). \tag{4.23}$$

where the final steps used the fact that the occupation decreases from 1 to 0 across the Fermi surface.

For $\mathbf{Q} \simeq 0$, it is clear that the denominator will grow, while the numerator changes little (this is particularly clear at zero temperature), thus $\mathbf{Q} = 0$ is more unstable than other vectors near 0, however it remains possible for a general band structure — i.e. general form $\epsilon_{\mathbf{k}}$, that a finite $\mathbf{Q}$ is even more unstable. We will discuss below some conditions under which this certainly does occur, but first we discuss what a finite $\mathbf{Q}$ instability would mean

---

[1] The complete relabelling was $\mathbf{k} \to -\mathbf{k} - \mathbf{Q}$ so that $\mathbf{k}+\mathbf{Q} \to -\mathbf{k}$.

Spin density wave vs anti-ferromagnetism

A finite $\mathbf{Q}$ instability means that there is a spatially varying magnetisation, and since the most unstable $\mathbf{Q}$ is controlled by the dispersion, this vector is likely to be commensurate with the lattice. The antiferromagnetic behaviour of a half filled insulator can also be seen as a spatially varying magnetisation, however there is a notable difference — that antiferromagnetism was strong, in the sense that each site had its complete magnetic moment pointing in alternating directions. The "spin-density wave" that the instability in this section describes grows continuously from zero, and so generally represents a small, partial magnetisation of each site.

It is however worth noticing that, as seen below, half filling is a state particularly susceptible to spin density weave ordering, and with increasing interaction strength, there is a crossover from the spin-density wave ordering of a weakly interacting half-filled itinerant electron system to the antiferromagnetism of a half filled Mott insulator.

Nesting and spin density waves

An instability is likely to develop if there exists a large region where $\epsilon_{\mathbf{k}} = \epsilon_{\mathbf{k}+\mathbf{Q}}$ for a constant vector $\mathbf{Q}$. In one dimension, the problem is particularly simple; if we consider a tight binding band structure, $\epsilon_k = -t\cos(ka)$, then as shown in Fig. 4.1, there is an instability at $Q = 2k_F$.

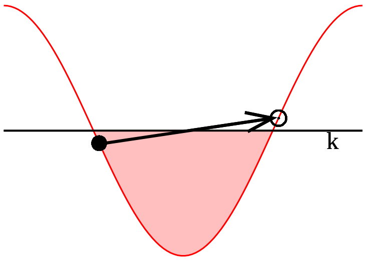


Figure 4.1: Nesting in one dimension, showing wavevector connecting equal energy states

It is unsurprising that in 1D ferromagnetism is overcome by spin density wave ordering, since we already know that one dimensional systems should not show ferromagnetic ordering. In higher dimensions, such an instability relies on there being a part of the Fermi surface which is parallel to another part, so that a constant wavevector can connect particles and holes across the Fermi sea. Examples of how this may occur are shown in Fig. 4.2; note that at half filling (at least for tight binding band structures) one will have nesting across the entire Fermi surface.

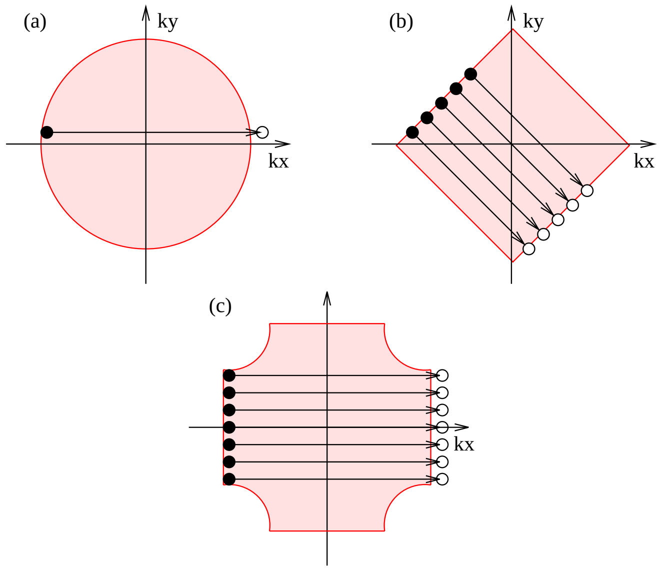


Figure 4.2: Possible scenarios for nesting in 2D. (a) Typical, circular Fermi surface, for which a given wavevector only couples isolated points on the Fermi surface. (b) Half filling, so $(\pm\pi, \pm\pi)/a_{\text{lattice}}$ wavevector generically leads to nesting. (c) Fermi surface in which a region is nested, so that the same spin-density-wave wavevector couples many points on the Fermi surface

4.3 Reduction of interaction for strong correlations

The other main objection to mean field theory ferromagnetism is that, as discussed in lecture 1, interactions that are strong enough to favour ferromagnetism should also be strong enough to distort the wavefunction. What we have done in our mean field theory is like first order perturbation theory, we must therefore now consider the next order effect, which means mixing of wavefunctions.

We will proceed in two steps, first considering the changes of wavefunction for two electrons, and then using the method found in this simple problem to discuss the many electron case.

Two-particle scattering matrix

We consider the same Hamiltonian as always, but choose yet another way to rewrite the interaction term for convenience:

$$H_{int} = \frac{U}{V} \sum_{\mathbf{q}} n_{\mathbf{q}\uparrow} n_{-\mathbf{q}\downarrow}, \qquad n_{\mathbf{q}\sigma} = \sum_{\mathbf{k}} a^\dagger_{\mathbf{k}+\mathbf{q}\sigma} a_{\mathbf{k}\sigma}.$$

We are interested in the possible energetic gain of a ferromagnetic state, and although for two particles the ground state is a singlet, to understand ferromagnetic type wavefunctions we should instead consider triplet wavefunctions. Let us thus define the triplet operator:

$$T^\dagger_{\mathbf{k},\mathbf{k}'} = \frac{1}{\sqrt{2}}\left(a^\dagger_{\mathbf{k}\uparrow}a^\dagger_{\mathbf{k}'\downarrow} + a^\dagger_{\mathbf{k}\downarrow}a^\dagger_{\mathbf{k}'\uparrow}\right). \tag{4.24}$$

This does of course still create an overall antisymmetric wavefunction, but one which has $\langle S^2\rangle = 2$.

The triplet that this operator creates has a bare energy $\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'}$; we want to find the energy including interactions. It is however clear that the interaction term, by changing particle momenta, can scatter this triplet to another triplet state.

We therefore make the (unnormalised) ansatz:

$$|\Psi(\mathbf{k},\mathbf{k}')\rangle = \left(T^\dagger_{\mathbf{k},\mathbf{k}'} + {\sum_{\mathbf{q}}}' \psi_{\mathbf{q}} T^\dagger_{\mathbf{k}+\mathbf{q},\mathbf{k}'-\mathbf{q}}\right)|0\rangle. \tag{4.25}$$

The choice of this scattering part is given by the momentum dependence in the interaction, which clearly conserves centre of mass momentum.

It turns out that this ansatz is in fact an eigenstate, so we can write down the secular equation, and attempt to solve it for the eigenvalue.

$$\begin{aligned}
E(\mathbf{k},\mathbf{k}')|\Psi(\mathbf{k},\mathbf{k}')\rangle &= H|\Psi(\mathbf{k},\mathbf{k}')\rangle \\
&= \left[(\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'}) T^\dagger_{\mathbf{k},\mathbf{k}'} + {\sum_{\mathbf{q}}}' \psi_{\mathbf{q}} \left(\epsilon_{\mathbf{k}+\mathbf{q}} + \epsilon_{\mathbf{k}'-\mathbf{q}}\right) T^\dagger_{\mathbf{k}+\mathbf{q},\mathbf{k}'-\mathbf{q}}\right]|0\rangle \\
&+ \frac{U}{V}\sum_{\mathbf{q}}\left[T^\dagger_{\mathbf{k}+\mathbf{q},\mathbf{k}'-\mathbf{q}} + {\sum_{\mathbf{q}'}}' \psi_{\mathbf{q}'} T^\dagger_{\mathbf{k}+\mathbf{q},\mathbf{k}'-\mathbf{q}}\right]|0\rangle.
\end{aligned} \tag{4.26}$$

The final term takes a particularly simple form since, as the scattering is independent of $\mathbf{q}$, the effect of scattering is to make the triplet momentum independent of the wavefunction parameterisation ψ_q. Identifying terms with the same triplet momentum one finds:

$$(E - \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}) = \frac{U}{V}\left[1 + \sum_{\mathbf{q}} \psi_{\mathbf{q}}\right] \tag{4.27}$$

$$\left(E - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}\right)\psi_{\mathbf{q}} = \frac{U}{V}\left[1 + \sum_{\mathbf{q}'} \psi_{\mathbf{q}'}\right] \tag{4.28}$$

If we define $\bar{\psi} = \sum_{\mathbf{q}} \psi_{\mathbf{q}}$ we may rearrange the second sum in Eq. (4.28) to write:

$$\bar{\psi} = \sum_q \frac{1}{E - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}}\frac{U}{V}(1 + \bar{\psi}) \tag{4.29}$$

Then, defining:

$$\mathcal{G}(\mathbf{k},\mathbf{k}',E) = \frac{1}{V}\sum_q \frac{1}{\epsilon_{\mathbf{k}+\mathbf{q}} + \epsilon_{\mathbf{k}'-\mathbf{q}} - E} \tag{4.30}$$

(note the choice of sign) we have the result:

$$\bar{\psi}\left[1 + U\mathcal{G}(\mathbf{k}, \mathbf{k}', E)\right] = -U\mathcal{G}(\mathbf{k}, \mathbf{k}', E). \tag{4.31}$$

Substituting this into Eq. (4.27) we finally find:

$$E = \epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} + \frac{U/V}{1 + U\mathcal{G}(\mathbf{k}, \mathbf{k}', E)} \tag{4.32}$$

Interpretation

Eq. (4.32) is of a very similar form to the naive energy we might have written using the Hartree-Fock approximation, but it has replaced:

$$U \to \frac{U}{1 + U\mathcal{G}(\mathbf{k}, \mathbf{k}', E)} \tag{4.33}$$

This expression can be identified as the T-matrix, describing repeated scattering, which is entirely equivalent to allowing for wavefunction renormalisation — the T-matrix accounts for particles reduced scattering amplitude when they are able to distort their paths to avoid one another.

Many particle scattering matrix

We may now use this renormalised interaction as a first attempt at understanding how changing wavefunctions affects magnetic ordering. As long as E is not far outside the range of possible energies of two electrons (which will be true when considering the ground state energy of a ferromagnetic configuration), we may neglect the energy dependence of $\mathcal{G}$, by taking $E = \epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'}$. There is one important difference for considering triplet states in a many particle system — we consider triplet formation on top of a filled Fermi sea, and so have:

$$\mathcal{G} = \sum_{\substack{\mathbf{q} \in BZ \\ |\mathbf{k}+\mathbf{q}|, |\mathbf{k}'-\mathbf{q}| \notin FS}} \frac{1}{\epsilon_{\mathbf{k}+\mathbf{q}} + \epsilon_{\mathbf{k}'-\mathbf{q}} - \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}} \tag{4.34}$$

We combine the renormalisation in Eq. (4.33) with the susceptibility enhancement in Eq. (4.9) to get:

$$\chi = \frac{\chi_{\text{Pauli}}}{1 - \frac{U\rho(\epsilon_F)}{1+U\mathcal{G}}} = \frac{\chi_{\text{Pauli}}(1 + U\mathcal{G})}{1 + U\mathcal{G} - U\rho(\epsilon_F)} \tag{4.35}$$

and so the condition for instability is that:

$$U\rho(\epsilon_F) \geq 1 + U\mathcal{G}. \tag{4.36}$$

Conditions on the band structure for Ferromagnetism

As illustrated in Fig. 4.3, for an otherwise unremarkable band structure, this condition is unlikely, even if strongly interacting, since $\mathcal{G} \sim 1/W$, using the bandwidth as a characteristic scale of the denominator, and $\rho(\epsilon_F) \sim 1/W$

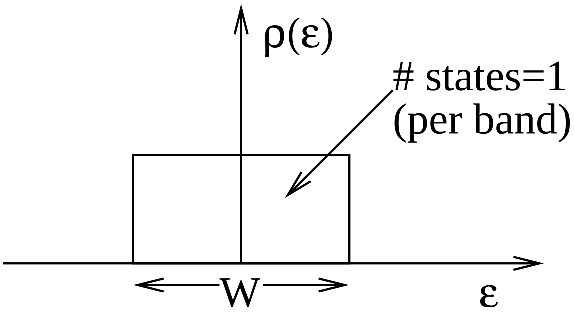


Figure 4.3: Typical density of states for a single band, which will be normalised so that the total number of electrons is one per atom

is also likely to hold as shown in Fig. 4.3, making ferromagnetism hard to achieve.

To overcome this problem requires conditions where the density of states is especially higher, higher than $1/W$. In a 3D free-electron band, as discussed in question 4.1, this condition is achieved if sufficiently close to a filled band, since the density of states increases with filling. Similarly, in many band structures, there can be an enhancement of density of states at the band maximum — e.g. van-Hove singularities, which can favour ferromagnetism in nearly filled bands. In addition, hybridisation between wide and narrow bands, as illustrated in Fig. 4.4 allows one to have simultaneously a large bandwidth and high density of states over an extended range of fillings.

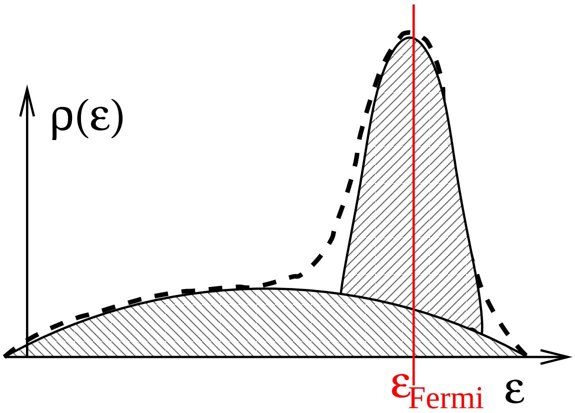


Figure 4.4: Density of states resulting from hybridisation of s- and d-orbitals; this provides simultaneously a high density of states at the Fermi surface along with a large bandwidth, and hence strongly favours Ferromagnetism. Such a density of states is seen, for example, in Nickel.

It is worth noting that the condition here, that a high density of states favours ferromagnetism, is closely connected to the statement at the end of the last lecture that orbital degeneracy in localised electron systems favours ferromagnetism — a high density of states means having many states of similar (almost degenerate) energies, and this in turn allows one to construct ferromagnetic configurations without paying significant kinetic energy costs.

Questions

Question 4.1: Kanamori condition for quasi-free electrons
The aim of this question is to evaluate the quantity $\mathcal{G}$ for an explicit band-structure and hence the conditions required for ferromagnetism. The structure is not particularly realistic, but is instructive as to what conditions favour ferromagnetism.

Consider free electrons, i.e. $\epsilon_k = \hbar^2 k^2/2m$, but with a maximum value $\epsilon_0 = \hbar^2 k_0^2/2m$, such that

$$V \int_0^{k_0} \rho(k) dk = 1$$

and define the bandwidth $W = \hbar^2 k_0^2/2m$, and the partial filling by $x = (k_F/k_0)^3$,

4.1.(a) Show that with these definitions one may write:

$$\rho(\epsilon_F) = \frac{mk_F}{2\pi^2\hbar^2} = \frac{3}{2}\frac{1}{WV}x^{1/3} \tag{4.37}$$

4.1.(b) Show that, if $|\mathbf{k}|, |\mathbf{k}'| \simeq k_f \ll k_0$, i.e. if the system is dilute, then the two-particle Green's function is given by:

$$\mathcal{G} = \sum_{\mathbf{q}} \frac{1}{\epsilon_{\mathbf{k}+\mathbf{q}} + \epsilon_{\mathbf{k}'-\mathbf{q}} - \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}} \simeq \frac{1}{2\pi^2}\frac{m}{\hbar^2}(k_0 - k_F) \tag{4.38}$$

4.1.(c) Hence show that ferromagnetism requires:

$$x > \left(\frac{1+\alpha}{2\alpha}\right)^3, \qquad \alpha = \frac{3}{2}\frac{U}{WV}. \tag{4.39}$$

What is the smallest fractional band filling at which ferromagnetism could arise?

Lecture 5

Magnetic alloys: mixing localised and conducting electrons

This lecture will mainly discuss the interaction between a single localised spin and a continuum of delocalised electrons. There are two motivations for this problem; one concerns isolated impurity magnetic atoms in an otherwise non-magnetic metal, the other concerns compounds in which a lattice of localised d or f orbitals interact with delocalised s electron bands.

The former problem, that of magnetic impurity atoms, is the historic origin of the problem we primarily study, the Kondo problem. This problem is interesting in showing a large separation of energy scales between the scales in the microscopic Hamiltonian and the effective binding energy of the ground state, hence producing a wide range of temperatures in which response properties are dominated by excited, rather than ground state configurations.

The latter part of the lecture will study another related problem; that of the interaction between pairs of localised magnetic impurities, leading to an effective spin-spin coupling that can be either ferromagnetic or anti-ferromagnetic dependent on separation between the spins. By combining interactions between many localised spins, and the competition of direct interaction, and "Kondo screening", there is a rich variety of possible behaviour, describing Kondo liquids and heavy fermions. Such a discussion is however beyond the scope of these lectures.

5.1 A localised electron state coupled to delocalised electrons: the Anderson Model

We begin from a model of localised states, written $d^\dagger, d$ coupled to delocalised conduction electrons. The localised states, due to their spatial confinement, will have a strong Hubbard U; we neglect in comparison the weaker electrostatic interaction between conduction electrons. Hence we may write:

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}\sigma} a_{\mathbf{k}\sigma} + \epsilon_d d^\dagger_\sigma d_\sigma + U d^\dagger_\uparrow d_\uparrow d^\dagger_\downarrow d_\downarrow + \frac{\eta}{\sqrt{V}} \left(d^\dagger_\sigma a_{\mathbf{k}\sigma} + a^\dagger_{\mathbf{k}\sigma} d_\sigma \right) \tag{5.1}$$

Our aim with this, the Anderson model, is first to understand when a localised spin exists — i.e. under what conditions the occupation of the localised level is one electron, rather than zero or two. Our second concern, in Sec. 5.2 is then to see when this localised magnetic moment remains free, rather than being screened by spin fluctuations of the surrounding conductor. It is important to note the energy scales of these two different processes; the interesting physics is that two very different energy scales arise from these two problems.

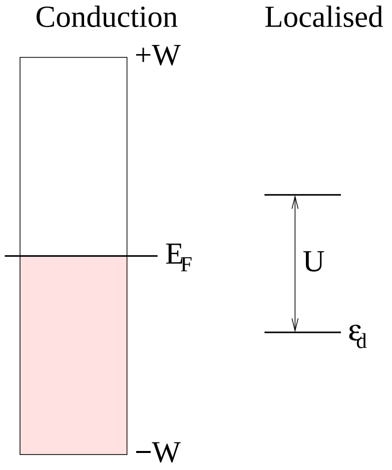


Figure 5.1: Schematic of energy levels of singly occupied and doubly occupied localised levels compared to the band of conduction states

The case in which a magnetic moment can arise, as illustrated in Fig. 5.1 is when the localised level has energy less than the Fermi level, so will be occupied by at least one electron, but the interaction strength is large enough that the energy to have two electrons is greater than the Fermi level, i.e. $\epsilon_d < E_F < \epsilon_d + U$. One can more formally study the conditions under which there is an isolated moment, by considering a mean-field theory decoupling with $n_{d\uparrow} - n_{d\downarrow}$ as the order parameter; see question 5.1.

Relating the Anderson and Kondo models

Let us suppose that the energies are as described in the previous section, and then consider an effective Hamiltonian restricted to the subspace in which the localised level is singly occupied. We may apply the Schrieffer-Wolff transformation discussed earlier (even though only two degenerate states exist). We thus consider the Hamiltonian in Eq. (5.1) as $H = H_0 + \Delta H$, in which we take $\Delta H = (\eta/\sqrt{V})(d^\dagger_\sigma a_{\mathbf{k}\sigma} + \text{H.c.})$. Following the general scheme for the Schrieffer-Wolff transformation given in Sec. 2.1, we begin by expanding ΔH into parts that change between different Hilbert subspaces.

$$\Delta H = \frac{\eta}{\sqrt{V}} \sum_{\mathbf{k}} \left[H_{21}(\mathbf{k}) + H_{12}(\mathbf{k}) + H_{01}(\mathbf{k}) + H_{10}(\mathbf{k}) \right] \tag{5.2}$$

$$H_{21}(\mathbf{k}) = \sum_{\sigma} n_{-\sigma} d^{\dagger}_{\sigma} a_{\mathbf{k}\sigma}, \qquad H_{10}(\mathbf{k}) = \sum_{\sigma} (1 - n_{-\sigma}) d^{\dagger}_{\sigma} a_{\mathbf{k}\sigma}, \tag{5.3}$$

where n is the population on the d level. H_{12}, H_{10} are the appropriate Hermitian conjugates. These terms are such that; H_{21} takes the localised state from singly occupied to doubly occupied, $n_d = 1 \to 2$, and H_{10} takes it from empty to singly occupied.

As a reminder; expanding the in transformation S to second order:

$$\tilde{H} = e^{iS} H e^{-iS} = H_0 + i[S, H_0] + \Delta H + i[S, \Delta H] - \frac{1}{2} \left[S, [S, H_0] \right],$$

and demanding $[S, H_0] = i \Delta H$ one thus has

$$\tilde{H} = H_0 + \frac{i}{2}[S, \Delta H]. \tag{5.4}$$

To find the appropriate transform, we note that:

$$[H_{21}(\mathbf{k}), H_0] = (\epsilon_{\mathbf{k}} - \epsilon_0 - U) \sum_{\sigma} n_{-\sigma} d^{\dagger}_{\sigma} a_{\mathbf{k}\sigma}, \tag{5.5}$$

$$[H_{10}(\mathbf{k}), H_0] = (\epsilon_{\mathbf{k}} - \epsilon_0) \sum_{\sigma} (1 - n_{-\sigma}) d^{\dagger}_{\sigma} a_{\mathbf{k}\sigma}. \tag{5.6}$$

In both cases the prefactor represents the difference in energy between the state before the part of ΔH acts and afterward. The Hermitian conjugate terms will obviously have opposite signs, and so we may write:

$$S = i \frac{\eta}{\sqrt{V}} \sum_{\mathbf{k}} \left[\frac{H_{21}(\mathbf{k}) - H_{12}(\mathbf{k})}{\epsilon_{\mathbf{k}} - \epsilon_d - U} + \frac{H_{10}(\mathbf{k}) - H_{01}(\mathbf{k})}{\epsilon_{\mathbf{k}} - \epsilon_d} \right]. \tag{5.7}$$

When we then evaluate $[S, \Delta H]$, the only terms of interest in the Hilbert subspace of occupation 1 are those pairs like $H_{12}(\mathbf{k}')H_{21}(\mathbf{k})$ and $H_{10}(\mathbf{k}')H_{01}(\mathbf{k})$ — i.e. virtual double or single occupancy terms. One thus gets:

$$\begin{aligned} \frac{i}{2}[S, \Delta H] = -\frac{\eta^2}{2V} \sum_{\mathbf{k},\mathbf{k}'} \Bigg[&- \frac{H_{12}(\mathbf{k}')H_{21}(\mathbf{k}) + H_{12}(\mathbf{k})H_{21}(\mathbf{k}')}{\epsilon_{\mathbf{k}} - \epsilon_d - U} \\ &+ \frac{H_{10}(\mathbf{k}')H_{01}(\mathbf{k}) + H_{10}(\mathbf{k})H_{01}(\mathbf{k}')}{\epsilon_{\mathbf{k}} - \epsilon_d} \Bigg]. \end{aligned} \tag{5.8}$$

Note the opposite signs coming from the opposite signs of H_{21} and H_{01} in Eq. (5.7).

To complete the analysis, we then need to consider the products of operators and how they act on the state of the spin and of the conduction electrons. We will make the simplifying assumption that the dependence of the denominator on $\mathbf{k}$ is slow — i.e. that ϵ_d is far below the Fermi surface, and $\epsilon_d + U$ far above. In this case we can approximately write:

$$\frac{i}{2}[S, \Delta H] = \frac{1}{2V} \sum_{\mathbf{k},\mathbf{k}'} \left[J_{12} H_{12}(\mathbf{k}') H_{21}(\mathbf{k}) - J_{10} H_{10}(\mathbf{k}') H_{01}(\mathbf{k}) \right], \tag{5.9}$$

$$J_{12} = 2\eta^2 \frac{1}{V} \sum_{\mathbf{k}} \frac{1}{\epsilon_{\mathbf{k}} - \epsilon_d - U}, \qquad J_{10} = 2\eta^2 \frac{1}{V} \sum_{\mathbf{k}} \frac{1}{\epsilon_{\mathbf{k}} - \epsilon_d} \tag{5.10}$$

Hence the elements we must evaluate are (using Eq. (5.3):

$$\begin{aligned} H_{12}(\mathbf{k}') H_{21}(\mathbf{k}) &= \sum_\sigma n_{-\sigma} a^\dagger_{\mathbf{k}'\sigma} d_\sigma n_{-\sigma} d^\dagger_\sigma a_{\mathbf{k}\sigma} + n_\sigma a^\dagger_{\mathbf{k}'-\sigma} d_{-\sigma} n_{-\sigma} d^\dagger_\sigma a_{\mathbf{k}\sigma} \\ &= \sum_\sigma n_{-\sigma}(1 - n_\sigma) a^\dagger_{\mathbf{k}'\sigma} a_{\mathbf{k}\sigma} - d^\dagger_\sigma d_{-\sigma} a^\dagger_{\mathbf{k}'-\sigma} a_{\mathbf{k}\sigma}, \qquad (5.11) \\ H_{10}(\mathbf{k}') H_{01}(\mathbf{k}) &= \sum_\sigma (1 - n_{-\sigma}) d^\dagger_\sigma a_{\mathbf{k}'\sigma} (1 - n_{-\sigma}) a^\dagger_{\mathbf{k}\sigma} d_\sigma \\ &\quad + (1 - n_\sigma) d^\dagger_{-\sigma} a_{\mathbf{k}'-\sigma} (1 - n_{-\sigma}) a^\dagger_{\mathbf{k}\sigma} d_\sigma \\ &= \sum_\sigma -n_\sigma (1 - n_{-\sigma}) a^\dagger_{\mathbf{k}\sigma} a_{\mathbf{k}'\sigma} - d^\dagger_{-\sigma} d_\sigma a^\dagger_{\mathbf{k}\sigma} a_{\mathbf{k}'-\sigma}. \qquad (5.12) \end{aligned}$$

Finally, by considering $n_\uparrow(1 - n_\downarrow) = 1/2 + S^z$ and $n_\downarrow(1 - n_\uparrow) = 1/2 - S^z$ one can then write Eq. (5.11) as:

$$\begin{aligned} (12) &= n_\downarrow(1 - n_\uparrow) a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{k}\uparrow} - d^\dagger_\uparrow d_\downarrow a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}\uparrow} + n_\uparrow(1 - n_\downarrow) a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}\downarrow} - d^\dagger_\downarrow d_\uparrow a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{k}\downarrow} \\ &= \left(\frac{1}{2} - S^z \right) a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{k}\uparrow} - S^+ a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}\uparrow} + \left(\frac{1}{2} + S^z \right) a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}\downarrow} - S^- a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{k}\downarrow} \\ &= \frac{1}{2} \left(a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{k}\uparrow} + a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}\downarrow} \right) - 2\mathbf{S} \cdot a^\dagger_{\mathbf{k}'\sigma} \mathbf{s}^{\sigma\sigma'} a_{\mathbf{k}\sigma'} \qquad (5.13) \end{aligned}$$

where we have used the same vector of fermion spin operators before Eq. (1.14). A similar calculation for Eq. (5.12) gives:

$$(10) = -\frac{1}{2} \left(a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{k}\uparrow} + a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}\downarrow} \right) - 2\mathbf{S} \cdot a^\dagger_{\mathbf{k}'\sigma} \mathbf{s}^{\sigma\sigma'} a_{\mathbf{k}\sigma'}. \tag{5.14}$$

Hence, the scattering divides into two parts; a spin independent scattering term – regular disorder scattering (with a prefactor $J_{10} + J_{12}$), and a part that depends on the state of the spin. Ignoring the background part that is independent of the magnetisation, one finally reaches:

$$H_{Kondo} = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}\sigma} a_{\mathbf{k}\sigma} + \frac{J}{V} \mathbf{S} \cdot \sum_{\mathbf{k},\mathbf{k}'} a^\dagger_{\mathbf{k}'\sigma} \mathbf{s}^{\sigma\sigma'} a_{\mathbf{k}\sigma'}. \tag{5.15}$$

Here $J = J_{10} - J_{12}$, which can clearly be seen to be positive. In the strong interaction limit, the typical scale is $J \simeq \eta^2/(E_F - \epsilon_d)$. Thus, the localised spin couples to the spins of the electrons that overlap with that spin, and can cause spin flips, which may enhance scattering. Since the coupling is antiferromagnetic, it favours the formation of a singlet, but this competes with the kinetic energy cost of single occupation of conduction electron levels.

5.2 A localised spin coupled to delocalised electrons: the Kondo model

We have reduced the model of an isolated impurity electron to a problem of a spin coupled to the conduction electron spins. Because coupling is at a point, only s-wave scattering takes place, and thus the problem is actually one dimensional, hence we immediately know the ground state should be a singlet. At low enough temperatures, the effect of the localised spin is screened by the formation of a singlet with the conduction electrons. We next consider the binding energy of this singlet, the Kondo temperature $k_B T_K$, and see it is remarkably small, so that for a wide range of temperature, an unscreened spin can be seen, but at low enough temperatures, it disappears as Kondo screening takes place.

Ground state and Kondo energy scale

If the delocalised electron system has N electrons before addition of the localised state, and the localised state contributes both the level and an extra electron, then the simplest ansatz for spin singlet formation would be to consider a filled Fermi sea of $N-1$ electrons, and then create a singlet from the last conduction electron and the localised state. For the purpose of understanding the binding energy of the singlet, we must also compare this to the energy of the triplet, so our singlet/triplet ansatz may be written as:

$$\begin{aligned}|\psi\rangle &= \sum_{\mathbf{k}} \psi_{\mathbf{k}} \left(d^\dagger_\uparrow a^\dagger_{\mathbf{k}\downarrow} \mp d^\dagger_\downarrow a^\dagger_{\mathbf{k}\uparrow} \right) |FS(N-1)\rangle \\ &= \sum_{\mathbf{k}} \psi_{\mathbf{k}} \left(d^\dagger_\uparrow a_{\mathbf{k}\uparrow} \pm d^\dagger_\downarrow a_{\mathbf{k}\downarrow} \right) |FS(N+1)\rangle \end{aligned} \tag{5.16}$$

where $|FS(N)\rangle$ describes the filled Fermi sea with N electrons in it. The relative signs in the first term can be determined by determining the total spin projection for the singlet/triplet. In the second line, the change of signs can be understood via:

$$\begin{aligned} a^\dagger_{\mathbf{k}\downarrow}|FS(N-1)\rangle &= a_{\mathbf{k}\uparrow} a^\dagger_{\mathbf{k}\uparrow} a^\dagger_{\mathbf{k}\downarrow}|FS(N-1)\rangle \\ a^\dagger_{\mathbf{k}\uparrow}|FS(N-1)\rangle &= -a_{\mathbf{k}\downarrow} a^\dagger_{\mathbf{k}\uparrow} a^\dagger_{\mathbf{k}\downarrow}|FS(N-1)\rangle, \end{aligned}$$

i.e., the relative signs originate from a consistent choice of ordering for the filled Fermi sea.

We should next use ψ_k as variational parameters to find estimates of the ground state energy of the singlet and triplet configurations. To do so, it is most straightforward to directly determine the action of Hamiltonian Eq. (5.15) on this state, and thus the expectation of the Hamiltonian. Let us introduce the states:

$$|X_{\mathbf{k}\uparrow}\rangle = d^\dagger_\uparrow a_{\mathbf{k}\uparrow}|FS(N+1)\rangle, \qquad |\psi\rangle = \sum_{\mathbf{k}} \psi_{\mathbf{k}} (|X_{\mathbf{k}\uparrow}\rangle \pm |X_{\mathbf{k}\downarrow}\rangle). \tag{5.17}$$

Compared to the energy of the filled Fermi sea with $N+1$ particles E_0, the energy due to the non spin-flip part of the Kondo Hamiltonian can be written:

$$H_0|X_{\mathbf{k}\sigma}\rangle = E_0 + \epsilon_d - \epsilon_{\mathbf{k}}|X_{\mathbf{k}\sigma}\rangle. \tag{5.18}$$

Note that for consistency with the notation in Eq. (5.16)– (5.17), we have included also the additional energy from occupying the localised level, so that the energy gain on transferring an electron from the Fermi sea to the filled state is $\epsilon_d - \epsilon_{\mathbf{k}}$. Since the energy scale that matters, the binding energy of the singlet, is unaffected by such a shift (both the singlet and the triplet states have the localised level occupied), this is a minor point.

The important point is the effect of the spin-flip operator:

$$\begin{aligned} H_{\text{flip}} &= \frac{J}{V}\mathbf{S}\cdot\sum_{\mathbf{k},\mathbf{k}'} a^\dagger_{\mathbf{k}'\sigma}\mathbf{s}^{\sigma\sigma'}a_{\mathbf{k}\sigma'}. \\ &= \frac{J}{2V}\sum_{\mathbf{k},\mathbf{k}'}\left[S^+a^\dagger_{\mathbf{k}\downarrow}a_{\mathbf{k}'\uparrow} + S^-a^\dagger_{\mathbf{k}\uparrow}a_{\mathbf{k}'\downarrow} + S^z\left(a^\dagger_{\mathbf{k}\uparrow}a_{\mathbf{k}'\uparrow} - a^\dagger_{\mathbf{k}\downarrow}a_{\mathbf{k}'\downarrow}\right)\right]. \end{aligned} \tag{5.19}$$

One may write:

$$\begin{aligned} H_{\text{flip}}|X_{\mathbf{k}\uparrow}\rangle &= \sum_{k'}\frac{J}{2V}\left[d^\dagger_\downarrow a^\dagger_{\mathbf{k}\uparrow}a_{\mathbf{k}'\downarrow}a_{\mathbf{k}\uparrow} + \frac{1}{2}d^\dagger_\uparrow\left(a^\dagger_{\mathbf{k}\uparrow}a_{\mathbf{k}'\uparrow}a_{\mathbf{k}\uparrow} - \ldots\right)\right]|FS(N+1)\rangle \\ &= \sum_{k'}\frac{J}{2V}\left[-|X_{\mathbf{k}'\downarrow}\rangle - \frac{1}{2}|X_{\mathbf{k}'\uparrow}\rangle + \ldots\right]. \end{aligned} \tag{5.20}$$

The ellipsis here represents terms where extra electron hole pairs are created; the terms we keep are those in which there is a single hole in the $N+1$ particle Fermi sea, and hence the only terms which can have an overlap with $|\psi\rangle$. We may use this to write:

$$H_{\text{flip}}|\psi\rangle = \frac{1}{2}\left(-\frac{1}{2}\mp 1\right)\frac{J}{V}\sum_{\mathbf{k}\mathbf{k}'}\psi_k\left[|X_{\mathbf{k}'\uparrow}\rangle \pm |X_{\mathbf{k}'\downarrow}\rangle\right]. \tag{5.21}$$

The result is similar for singlet and triplet states, but for the singlet the prefactor is $-3/4$ while for the triplet it is $+1/4$. We can write this is as $Z = \mathbf{S}_d\cdot\mathbf{S}_a = (1/4) - \mathcal{P}_{\text{Singlet}}$ and so:

$$\langle\psi|H|\psi\rangle = \sum_{\mathbf{k}}(E_0 + \epsilon_d - \epsilon_{\mathbf{k}})|\psi_{\mathbf{k}}|^2 + \frac{ZJ}{V}\sum_{\mathbf{k}\mathbf{k}'}\psi^*_{\mathbf{k}}\psi_{\mathbf{k}'}. \tag{5.22}$$

Then, minimising with respect to $\psi_{\mathbf{k}}$ at fixed normalisation yields the equation:

$$(E_0 + \epsilon_d - \epsilon_{\mathbf{k}} - E)\psi_{\mathbf{k}} = -\frac{ZJ}{V}\sum_{\mathbf{k}'}\psi_{\mathbf{k}'}. \tag{5.23}$$

This equation can be solved for the energy by writing:

$$\Sigma = \sum_{\mathbf{k}}\psi_{\mathbf{k}}, \quad \Sigma = -\frac{ZJ}{V}\sum_{\mathbf{k}}\frac{1}{E_0 + \epsilon_d - \epsilon_{\mathbf{k}} - E}\Sigma, \tag{5.24}$$

and so rewriting sums as integrals over the density of states:

$$\frac{1}{ZJ} = \int_{-W}^{\epsilon_F} d\epsilon \frac{\rho(\epsilon)}{E - E_0 - \epsilon_d + \epsilon}. \tag{5.25}$$

It is convenient at this point to introduce ΔE via:

$$E = E_0 + \epsilon_d - \epsilon_F - \Delta E, \tag{5.26}$$

and $\epsilon' = \epsilon - \epsilon_F$, hence:

$$\frac{1}{ZJ} = \int_{-\epsilon_F - W}^{0} d\epsilon' \frac{\rho(\epsilon_F + \epsilon')}{\epsilon' - \Delta E}. \tag{5.27}$$

This integral has the form of the branch-cut limit of a sum of poles, i.e. poles exist whenever $\Delta E = \epsilon'$. Since $\epsilon' < 0$, and we are interested in ground states, i.e. maximising ΔE, then Fig. 5.2 shows that only for negative ZJ does an isolated bound state exist.

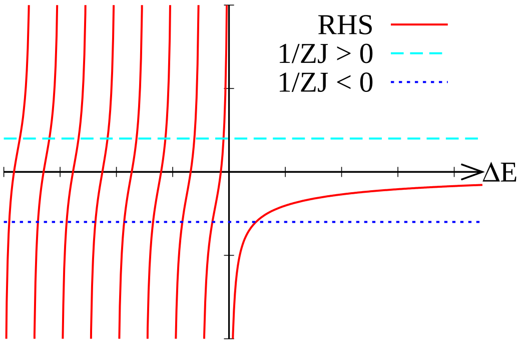


Figure 5.2: Schematic sum of poles in integral for singlet/triplet energy, showing how the difference of sign of Z controls what value the integral may take; only for negative Z does an isolated pole at positive ΔE exist.

Since $J > 0$ for the Kondo model as derived from the Anderson model, we thus have that, for the singlet state there is a binding energy that may be determined from the above integral if we assume $\Delta E > 0$, and hence find:

$$-\frac{4}{3J} \simeq \rho(\epsilon_F) \ln \left[\frac{\Delta E}{W + \epsilon_F + \Delta E} \right],$$

$$\Delta E = W \exp \left[\frac{-4}{3\rho(\epsilon_F) J} \right]. \tag{5.28}$$

This ΔE gives the binding energy, since the triplet state does not have a corresponding bound state outside the continuum.

The above is an upper bound on the ground state triplet and singlet energies, because the variational wavefunction was not an eigenstate — we had to throw away terms with extra electron hole pairs. A fuller calculation including all electron hole pairs does not significantly change the answer, but gives:

$$\Delta E = W \exp\left[\frac{-1}{\rho(\epsilon_F)J}\right], \tag{5.29}$$

There are thus two very different energy scales associated with the original Anderson problem ($\sim \epsilon_d$) and with the Kondo problem $k_B T_k \sim W \exp[-1/\rho(\epsilon_F)J]$. The Kondo temperature can become small, particular when J is large, and a typical scale is around 10K. Since this is relatively low, it becomes relevant to ask what occurs at temperatures where a localised moment exists, but the Kondo singlet is thermally dissociated, which we will address next.

Finite temperature spin susceptibility

We consider how the spin flip term in the Kondo model modifies the spin susceptibility, if we start from an initial state in which the localised spin exists, and is not screened out by a singlet.

We are thus interested in the problem:

$$H = H_0 + H_{\text{flip}} + H_{\text{field}} \tag{5.30}$$

and want to see how, at leading order H_{flip} modifies the linear response to an applied B field. The susceptibility can be found from the expectation of S^z, given by:

$$\langle S^z \rangle = \frac{1}{\mathcal{Z}}\left(\langle \uparrow' | S^z | \uparrow' \rangle e^{-\beta E'_\uparrow} + \langle \downarrow' | S^z | \downarrow' \rangle e^{-\beta E'_\downarrow}\right), \tag{5.31}$$

where the primes indicate changes to the states and energies of the states due to the Kondo spin flip term. Without this change, the energies would be $E_\sigma = E_0 + \sigma g \mu_B H_{\text{ext}}$, i.e. the only energy difference is due to the external field. Two contributions from H_{flip} will arise; a second order change of energy, and a first and second order change of the wavefunction. To see the characteristic behaviour, let us examine the first order change of the wavefunction; because this misses other equivalent sized terms, considering this term alone will have incorrect numerical prefactors, but as an illustration of what H_{flip} does, this will not matter.

At first order in H_{flip} we will get $|\uparrow'\rangle = |\uparrow\rangle + J \ldots |\downarrow\rangle$ and a similar change for $|\downarrow'\rangle$, i.e. the eigenstates will become "less polarised", hence reducing the effective magnetic response, and the susceptibility. This reduction is given by:

$$\begin{aligned} |\uparrow'\rangle = {} & |\uparrow\rangle |FS(N)\rangle + \\ & \frac{J}{2V} \sum_{\substack{|\mathbf{k}|>k_F \\ |\mathbf{k}'|<k_F}} \left[|\downarrow\rangle \frac{a^\dagger_{\mathbf{k}\uparrow} a_{\mathbf{k}'\downarrow}}{\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}}} + \frac{1}{2} |\uparrow\rangle \frac{a^\dagger_{\mathbf{k}\uparrow} a_{\mathbf{k}'\uparrow} - a^\dagger_{\mathbf{k}\downarrow} a_{\mathbf{k}'\downarrow}}{\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}}} \right] |FS(N)\rangle. \end{aligned} \tag{5.32}$$

Considering then the change to the expectation of S^z, since overlap will exist only between identical states of the conduction electrons, we get:

$$\begin{aligned}\langle\uparrow' |S^z| \uparrow'\rangle &\simeq \frac{1}{2}\left[1+\left(\frac{J}{2V}\right)^2\left(-1+\frac{1}{4}+\frac{1}{4}\right)\sum_{\mathbf{k}\mathbf{k}'}\frac{f_{\mathbf{k}'}(1-f_{\mathbf{k}})}{(\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}'})^2}\right]\\ &\simeq \frac{1}{2}\left[1-\frac{1}{2}\left(\frac{J\rho(\epsilon_F)}{2}\right)^2\iint d\epsilon d\epsilon' \frac{f(\epsilon')(1-f(\epsilon))}{(\epsilon-\epsilon')^2}\right]\\ &\simeq \frac{1}{2}\left[1-\left(\frac{J\rho(\epsilon_F)}{2}\right)^2\ln\left(\frac{W}{k_BT}\right)\right]. \end{aligned} \tag{5.33}$$

Including similar corrections at second order in J in the wavefunction, as well as corrections in the energy, one finds:

$$\chi(J)=\chi(J=0)\left[1-2\left(\frac{J\rho(\epsilon_F)}{2}\right)^2\ln\left(\frac{W}{k_BT}\right)\right]. \tag{5.34}$$

Note that in the above calculation we used the finite temperature distribution function, and that as $T \to 0$, the integral diverges logarithmically. For finite frequency susceptibility, the frequency can play a similar role as a cutoff on the susceptibility, but the logarithmic divergence as $T \to 0, \omega \to 0$ remains. Question 5.1 discusses how to evaluate such an integral to see the logarithmic divergence.

This divergence can be connected to our previous calculation of the binding energy, $k_BT_K = W\exp[-1/J\rho(\epsilon_F)]$, by writing:

$$\begin{aligned}\ln\left(\frac{T}{T_K}\right) &= \ln\left(\frac{k_BT}{W}\right)+\frac{1}{J\rho(\epsilon_F)}\\ J\rho(\epsilon_F)\ln\left(\frac{k_BT}{W}\right) &= J\rho(\epsilon_F)\left[\ln\left(\frac{T}{T_K}\right)-1\right].\end{aligned} \tag{5.35}$$

This identification, along with higher order corrections in J, can be shown to give the susceptibility in the form:

$$\chi(J)=\chi(J=0)\left[1-\frac{J\rho(\epsilon_F)}{1+J\rho(\epsilon_F)\ln\left(\frac{k_BT}{W}\right)}\right]=\chi(J=0)\left[1-\frac{1}{\ln(T/T_K)}\right]. \tag{5.36}$$

The significance of this form is that what we had considered as a perturbative expansion in J is really a perturbative expansion in high T, i.e. in the requirement $\ln(T/T_K) \gg 1$ — since for $J \to 0^+$, $T_K \to 0$, this is unsurprising. It is however important to note that since this perturbative expansion in high T diverges at low T, it indicates a transition to new behaviour — this is the formation of singlets.

At low temperatures, the susceptibility $\chi \sim (g\mu_B)^2/k_BT$ is eventually replaced by $\chi \sim (g\mu_B)^2/k_BT_K$, since the excitations required to break the singlet are gapped, while the $1/T$ was a feature of a gapless susceptibility. Many other physical properties share a similar logarithmic divergence at low temperatures. Historically, conductivity was the first such observed property, and the Kondo problem was first encountered as a resistance minimum on decreasing temperature. Resistance generally decreases in a metal, but at low temperatures, screening clouds develop around magnetic impurities leading to a logarithmically increasing resistance.

5.3 Interaction between two localised spins via conduction electrons: RKKY Interaction

We now consider the same model as in Sec. 5.2, generalised to multiple impurity atoms, but approach it in a quite different way; we will see another kind of behaviour which may compete with the singlet formation described above. We thus start from:

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}\sigma} a_{\mathbf{k}\sigma} + \frac{J}{V} \sum_j \mathbf{S}_j \cdot \sum_{\mathbf{k},\mathbf{k}'} e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}_j} a^\dagger_{\mathbf{k}'\sigma} \mathbf{s}^{\sigma\sigma'} a_{\mathbf{k}\sigma'}, \tag{5.37}$$

where we have added phase factors associated with the location of each different impurity. Assuming $T > T_K$, we ignore singlet formation, and so start from a “ground state” of a filled Fermi sea and localised spin impurities. We may then use the Schrieffer-Wolff approach to eliminate excitations of the filled Fermi sea, and hence find an effective interaction between the localised spins due to creation and annihilation of electron-hole spin flip pairs.

As always, we aim to find S such that $[S, H_0] = i\Delta H$ and then want $(i/2)[S, \Delta H]$. In this case, ΔH is the spin flip term and so:

$$S = \frac{iJ}{V} \sum_j \mathbf{S}_j \cdot \sum_{\mathbf{k},\mathbf{k}'} \frac{e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}_j}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}} a^\dagger_{\mathbf{k}'\sigma} \mathbf{s}^{\sigma\sigma'} a_{\mathbf{k}\sigma'}, \tag{5.38}$$

can be seen to have the desired property. It is clear both this and ΔH are isotropic, and so we may appeal to this isotropy to reduce the number of terms we must find — we will just find those terms like $S^+_j S^-_{j'}$, and noting that we only keep terms that return the Fermi sea to its original state, this term will be indicative of the full spin-spin interaction. Thus we want:

$$\begin{aligned} &\left(\frac{i}{2}\right)\left(\frac{J}{V}\right)\left(\frac{iJ}{V}\right) \sum_{jj'} \sum_{\mathbf{k}\mathbf{k}'} \frac{e^{i(\mathbf{k}'-\mathbf{k})\cdot(\mathbf{R}_j - \mathbf{R}_{j'})}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}} \left[S^+_j a^\dagger_{\mathbf{k}'\downarrow} a_{\mathbf{k}\uparrow}, S^-_{j'} a^\dagger_{\mathbf{k}\uparrow} a_{\mathbf{k}'\downarrow}\right] \\ &= \frac{J^2}{2V^2} \sum_{jj'} \sum_{\mathbf{k}\mathbf{k}'} e^{i(\mathbf{k}'-\mathbf{k})\cdot(\mathbf{R}_j - \mathbf{R}_{j'})} S^+_j S^-_{j'} \frac{f_{\mathbf{k}'}(1 - f_{\mathbf{k}}) - (1 - f_{\mathbf{k}'}) f_{\mathbf{k}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}}. \end{aligned} \tag{5.39}$$

Hence, putting this and equivalent terms together yields:

$$H = H_0 + \sum_{jj'} \mathbf{S}_j \cdot \mathbf{S}_{j'} I_{\text{RKKY}} \left(\mathbf{R}_j - \mathbf{R}_{j'}\right), \tag{5.40}$$

$$I_{\text{RKKY}}(\mathbf{R}) = -J^2 \frac{1}{V^2} \sum_{\mathbf{k}\mathbf{k}'} e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}} \frac{f_{\mathbf{k}'} - f_{\mathbf{k}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}}. \tag{5.41}$$

It can be shown (Question 5.3) that this sum over states takes the form:

$$I_{\text{RKKY}}(\mathbf{R}) = -J^2 \frac{[\sin(2k_F R) - 2k_F R\cos(2k_F R)]}{(k_F R)^4}. \tag{5.42}$$

For small $k_F R$ this has the form $I \sim -J^2/3k_F R$ and so is ferromagnetic, but at larger distances its sign alternates. This means that for randomly located impurities, there is an interaction term $H = -J_{jj'}\mathbf{S}_j \cdot \mathbf{S}_{j'}$ with random sign and magnitude; this disordered spin-spin coupling can lead to glassy behaviour, with very many nearly degenerate ground states.

5.4 Interplay between Kondo screening and RKKY

In compounds where a regular lattice of localised spins is coupled to a conducting band of electrons, the two effects discussed so far interact. As a characteristic estimate of the relative sizes of terms favouring singlet formation vs RKKY interactions, one can plot the dependence of these terms in $J\rho(\epsilon_F)$, as done in Fig. 5.3. The RKKY term is of order J^2/W, since the denominator in Eq. (5.41) clearly involves the characteristic energy scale of the conduction band. By writing $W\rho(\epsilon_F) \sim 1$, both effects depend only on the combination $J\rho(\epsilon_F)$.

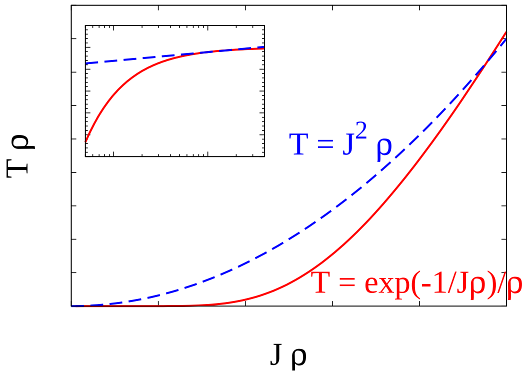


Figure 5.3: Doniach phase diagram, based on characteristic temperatures associated with inter-impurity energy scale and singlet formation energy scale. The inset shows the same data on a logarithmic scale.

Naively, between T_K and T_{RKKY}, one might expect a magnetically ordered phase with local moments, however real behaviour is more complicated. One important source of the complication is that “competition” of the two effects is moderated by the fact that Kondo singlets are very large (since their binding energies are small), and so many impurity ions may be within the screening cloud of each ion, hence a picture of separated screened ions breaks down.

Questions

* Question 5.1: Mean-Field Theory of Anderson model

This question uses the Anderson model as an example of mean-field theory. Consider the order parameter $S = d^\dagger_\uparrow d_\uparrow - d^\dagger_\downarrow d_\downarrow$ for the model in Eq. (5.1).

5.1.(a) By using the fermionic commutation relations of $d, d^\dagger$, show one may write the model as:

$$H = \sum_\sigma \left[\sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}\sigma} a_{\mathbf{k}\sigma} + \tilde{\epsilon}_d d^\dagger_\sigma d_\sigma + \frac{\eta}{\sqrt{V}} \left(d^\dagger_\sigma a_{\mathbf{k}\sigma} + a^\dagger_{\mathbf{k}\sigma} d_\sigma \right) \right] - \frac{US^2}{2} \tag{5.43}$$

$$\tilde{\epsilon}_d = \left(\epsilon_d + \frac{U}{2} - U\sigma S \right)$$

where

We attempt to find the conditions under which a self-consistent solution $S \neq 0$ exists. We thus need to find:

$$n_d = \sum_{\epsilon_n < \epsilon_F} |\langle \psi_n | d \rangle|^2 \tag{5.44}$$

for each spin, where (ϵ_n, ψ_N) are the eigenvalue, eigenstate pair of the single particle problem in Eq. (5.43).

5.1.(b) Show that this may be written as:

$$n_{d\sigma} = \int^{\epsilon_F} dE \frac{1}{\pi} \Im \left[\langle d | \mathcal{G}(E - i0) | d \rangle \right], \qquad (E - H)\mathcal{G}(E) = \mathbb{1} \tag{5.45}$$

5.1.(c) By considering the structure of the single particle matrix H, and making reasonable approximations, show:

$$n_d \int^{\epsilon_F} dE \frac{\eta^2 \rho(E)}{(E - \tilde{\epsilon}_d)^2 + V^2 \rho(E) \pi}. \tag{5.46}$$

5.1.(d) By evaluating this integral, find the the self consistency condition on S.

Question 5.2: Finite temperature evaluation of Eq. (5.33)

By considering:

$$\iint d\epsilon d\epsilon' \frac{d^2}{d\epsilon d\epsilon'} \ln(\epsilon - \epsilon') f(\epsilon)[1 - f(\epsilon')] \tag{5.47}$$

or otherwise, show that the temperature dependence of the integral in Eq. (5.33) is like $2\ln(W/k_B T)$ where W is characteristic of the bandwidth.

Question 5.3: Explicit form of RKKY interaction

Show that, for a free electron dispersion, Eq. (5.41) leads to a spatial dependence as in Eq. (5.42).

Lecture 6

Frustrated spin systems and magnetic domains

The contents of this lecture cover briefly a variety of important topics in the theory of magnetism, but which have little direct connection to the quantum nature of magnetism. Spin glasses arise from random locations of magnetic impurities, leading to random signs and strengths of RKKY interactions; this produces a combination of disorder and frustration that lead to complicated equilibrium states, as well as to very slow relaxation and response to a wide range of frequencies.

Frustration alone, without disorder, can be studied in regularly frustrated systems, such as non-bipartite antiferromagnetic systems. We discuss here instead cases where it is possible for frustration to arise from ferromagnetic interactions, when combined with strong anisotropy: $Ho_2Ti_2O_7$ and $Dy_2Ti_2O_7$, both known as spin ice. Strong anisotropy arises from large on-site magnetic moments, which also increase the relative effectiveness of direct magnetostatic interactions, and the above compounds do also have important effects of magnetic fields.

Another case in which it is important that magnetic degrees of freedom are not just another internal symmetry, but couple to long-range magnetostatic fields, is domain formation. How ferromagnetic domains are arranged to minimise this energy will be the final subject of these lectures.

6.1 Spin glasses

This brief discussion of spin glasses is intended merely to mention some of the physical concepts involved, rather than to discuss the techniques needed for their modelling. Rather than deal with spatially disordered spins and their RKKY interactions, a simpler model can be written as:

$$H = \sum_{ij} J_{ij} s_i s_j, \tag{6.1}$$

where the quantum mechanical spins have been replaced by Ising spins, so $s_i = \pm 1$, and J_{ij} is assumed to be a random variable drawn from a known distribution. We will consider a Gaussian distribution with $\overline{J_{ij}} = 0$, and $\overline{J_{ij}J_{kl}} = \delta_{ik}\delta_{jl}J^2/N$, where N is the number of sites. This model, the Sherrington-Kirkpatrick model, is thus effectively infinite dimensional, since all notion of proximity has been removed.

Spin glasses combine both frustration and disorder; frustration is the inability to simultaneously satisfy all bonds, i.e. to make $J_{ij}s_is_j < 0$ for all i, j. Disorder without frustration is possible, consider for example the model:

$$J_{ij} = -J\xi_i\xi_j \quad \xi_i = \pm 1 \tag{6.2}$$

In this case, the signs of the bonds are random, but there are correlations around any closed plaquette, since one necessarily has that:

$$J_{12}J_{23}J_{34}J_{41} = J^4(\xi_1)^2(\xi_2)^2(\xi_3)^2(\xi_4)^2 = J^4 > 0. \tag{6.3}$$

This model is therefore a ferromagnet in disguise; by choosing $s_i = \xi_i\tilde{s}_i$, the model is transformed to a ferromagnetic problem[1].

Frustration leads to a large degeneracy of the ground state, since there are multiple configurations that manage to satisfy the maximum number of bonds. Disorder means that different parts of this degenerate subspace become disconnected by large barriers.

Susceptibility and spin glass order parameter

When considering averages over disorder realisations, an order parameter such as $\langle s_i\rangle$ no longer works. In the following, it will be useful to distinguish averaging over the thermal ensemble, denoted $\langle\ldots\rangle$ and averaging over realisations of disorder, denoted by a bar. While $\overline{\langle s_i\rangle}$ vanishes, one can consider

$$q = \overline{\langle s_i\rangle^2} \tag{6.4}$$

as an order parameter for the spin glass phase, since this expression will survive averaging over the different signed ground states.

The above order parameter also connects directly to the experimentally measurable susceptibility. Putting aside for the moment the problems of applying mean-field theory (discussed in Sec. 6.1), we suppose we may write:

$$\langle s_i\rangle = \tanh(\beta[h + \tilde{h}_i]), \tag{6.5}$$

$$\chi_i = \frac{d\langle s_i\rangle}{dh} = \beta\text{sech}^2(\beta[h + \tilde{h}_i]) = \beta(1 - \langle s_i\rangle^2). \tag{6.6}$$

Hence, one sees that $\chi = \overline{\chi_i} = \beta(1 - q)$. Thus, as illustrated in Fig. 6.1, the susceptibility reduces below the transition temperature, as local moments freeze into their optimal configurations.

Mean field theory of spin glasses

For a ferromagnet, we could construct a mean-field theory by writing:

$$m_i = \langle s_i\rangle, \qquad m_i = \tanh(\beta\tilde{h}_i), \qquad \tilde{h}_i = \sum_k J_{ik}m_k. \tag{6.7}$$

---

[1] Because these are Ising, not quantum, spins the ferromagnetic and unfrustrated antiferromagnetic models are in fact the same.

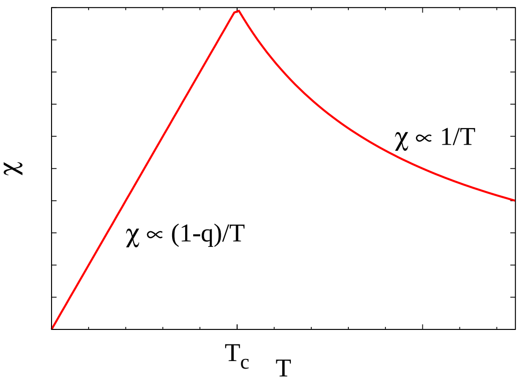


Figure 6.1: Schematic picture of susceptibility for paramagnetic high temperature phase and spin glass low temperature phase.

For the spin glass, this approach has its first problem because the effective local field is a sum of terms with different signs, and so fluctuations are much more important. Fluctuations alone are not the end of mean field theory, as long as one has a controlled probability distribution for $\tilde{h}_i$, one could then average $\tanh(\beta\tilde{h}_i)$ over this distribution. The problem comes in producing this probability distribution. If $J_{ik}m_k$ were independent random variables with bounded variance, then the central limit theorem would dictate that in the large N limit, the probability distribution of their sum becomes Gaussian. However, they are not independent — the probability distribution for values of the spin on a site is correlated with the choice of J_{ij}, because the disorder is quenched — one first chooses J_{ij}, and then finds the probability distribution in this background.

There does however exist a way out of this problem. If one divides the spins into spin 0, and spins $1\dots N$, and then first solves the problem for the latter N spins, the solution for $N+1$ spins can be constructed iteratively, as follows.

Given the state of spins $1\dots N$, the variables $J_{0k}s_k$ are independent, as long as s_k is drawn from the thermal ensemble of spins $1\dots N$, neglecting the effect of coupling to spin 0. Thus, we have the on-site field,

$$h_0^{(N)} = \sum_{k=1}^{N} J_{0k}s_k \tag{6.8}$$

with mean $\langle h_0^{(N)}\rangle$ and variance σ_N^2. We may then write the joint probability distribution for the field on site 0 and the state of site 0, as:

$$P(h_0, s_0) \propto \exp\left[-\frac{\left(h_0 - \langle h_0^{(N)}\rangle\right)^2}{2\sigma_N^2} + \beta h_0 s_0\right]. \tag{6.9}$$

While the prior distribution for h_0 is controlled by the other spins and is Gaussian, the full distribution for h_0 is no longer Gaussian — large values of $|h_0|$ are more common than would be expected, since these large values allow particularly low energies on site 0; this effect is represented by the extra term in the exponent, $\beta h_0 s_0$, enhancing the probability of states with a large on site energy.

We may now use this probability to determine the actual expectation of s_0, h_0, including this correlation between local field and state. We thus find:

$$\begin{aligned}
\langle s_0^{(N+1)} \rangle &= \sum_{s_0} \int dh_0 P(h_0, s_0) s_0 = \frac{\sum_{s_0} \exp\left[\frac{\left(\beta s_0 \sigma_N^2 + \langle h_0^{(N)} \rangle\right)^2}{2\sigma_N^2} \right] s_0}{\sum_{s_0} \exp\left[\frac{\left(\beta s_0 \sigma_N^2 + \langle h_0^{(N)} \rangle\right)^2}{2\sigma_N^2} \right]} \\
&= \tanh\left(\beta \langle h_0^{(N)} \rangle\right) \qquad (6.10) \\
\langle h_0^{(N+1)} \rangle &= \sum_{s_0} \int dh_0 P(h_0, s_0) h_0 \qquad (6.11) \\
&= \frac{\sum_{s_0} \exp\left[\frac{\left(\beta s_0 \sigma_N^2 + \langle h_0^{(N)} \rangle\right)^2}{2\sigma_N^2} \right] \left(\beta s_0 \sigma_N^2 + \langle h_0^{(N)} \rangle\right)}{\sum_{s_0} \exp\left[\frac{\left(\beta s_0 \sigma_N^2 + \langle h_0^{(N)} \rangle\right)^2}{2\sigma_N^2} \right]} \\
&= \beta \langle s_0^{(N+1)} \rangle \sigma_N^2 + \langle h_0^{(N)} \rangle. \qquad (6.12)
\end{aligned}$$

Identifying $\langle s_0^{(N+1)} \rangle = m_0$, and $\langle h_0^{(N+1)} \rangle = \sum_{0k} J_{0k} m_k$, these equations then mean we may write the mean-field equation as:

$$m_i = \tanh(\beta \tilde{h}_i), \qquad \tilde{h}_i = \sum_{ik} J_{ik} m_k - \beta m_i \sigma_N^2. \qquad (6.13)$$

Physically, this may be interpreted as $\tilde{h}_i$ being the average field that would have been acting on site i if site i were not there.

To complete this problem, we must find an expression for

$$\sigma_N^2 = \left\langle \left(h_0^{(N)} - \langle h_0^{(N)} \rangle \right)^2 \right\rangle. \qquad (6.14)$$

It can be shown that, due to the fluctuation dissipation theorem, this will be related to χ_i, the response of a site to the change of field on that site, and using Eq. (6.6), we may write:

$$\sigma_N^2 = J^2(1-q). \qquad (6.15)$$

The final form of the mean-field equations (Thouless–Anderson–Palmer equations) are:

$$m_i = \tanh\left(\beta \left[\sum_k J_{ik} m_k - \beta J^2 (1-q) m_i \right]\right), \qquad q = \overline{m_i^2}. \qquad (6.16)$$

The critical temperature can now be straightforwardly identified; if we consider $m_i = m j_i$ where $\mathbf{j}$ is the eigenvector of J_{ik} with maximum eigenvalue, λ_J, the critical equation (at $T = T_c$, $q = 0$) becomes:

$$m = \beta_c \left[\lambda_J m - \beta_c J^2 m \right]. \tag{6.17}$$

For a Gaussian random matrix we are considering, it can be shown that $\lambda_J = 2J$, and so this becomes:

$$1 = \beta_c J (2 - \beta_c J) \tag{6.18}$$

with a critical temperature $\beta_c J = 1$, rather than $\beta_c J = 1/2$ which a naive mean-field theory would have predicted.

This discussion is merely an outline of one method of approaching spin glasses, a fuller discussion can be found for example in Mezard et al. [11].

6.2 Spin Ice

A cleaner example of frustration can be seen in periodically frustrated systems. While the most obvious realisation of such a system is an antiferromagnet on a non-bipartite lattice, such as a triangular lattice, this section discusses a less obvious case, that of spin ice (so called because of the relation between its statistical mechanics and that of water ice).

Frustrated ferromagnetism

We start with a Hamiltonian:

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + H_{\text{anisotropy}} \tag{6.19}$$

for spins on a Pyrochlore lattice — i.e. a lattice of corner sharing tetrahedra, as shown in Fig. 6.2. The basic interaction is ferromagnetic, but because of a strong anisotropy term, the directions of $\mathbf{S}_i$ are limited in such a way that this becomes a frustrated Ising model.

The crystal field anisotropy is such that the easy axis of the spins is along the direction into or out of the tetrahedron. We therefore write $\mathbf{S}_i = s_i |S| \hat{\mathbf{z}}_i$, where $\hat{\mathbf{z}}_i$ is the direction set by anisotropy, and $s_i = \pm 1$. Choosing directions into the tetrahedron, the four directions for the sites labelled in Fig. 6.2 are:

$$\begin{array}{rl} A: & (\ \ 1,\ \ 1,\ \ 1)/\sqrt{3} \\ B: & (-1,-1,\ \ 1)/\sqrt{3} \\ C: & (-1,\ \ 1,-1)/\sqrt{3} \\ D: & (\ \ 1,-1,-1)/\sqrt{3} \end{array} \tag{6.20}$$

Hence, one has that:

$$\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B = \hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_C = \hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_D = \hat{\mathbf{z}}_B \cdot \hat{\mathbf{z}}_C = \hat{\mathbf{z}}_B \cdot \hat{\mathbf{z}}_D = \hat{\mathbf{z}}_C \cdot \hat{\mathbf{z}}_D = -\frac{1}{3}, \tag{6.21}$$

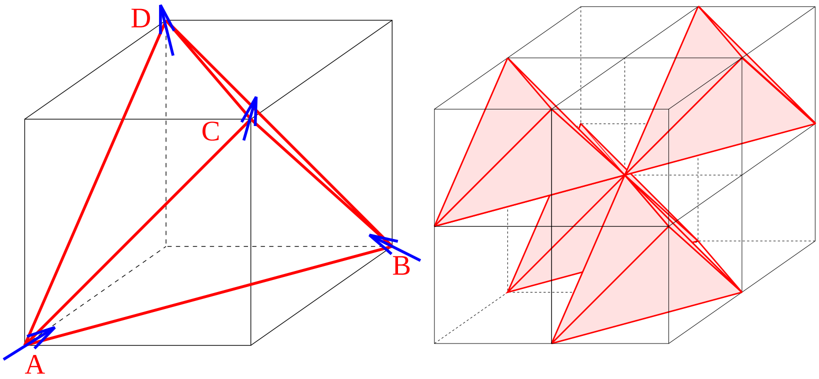


Figure 6.2: Left: A single tetrahedron, labelling the locations of the four interacting spins within that tetrahedron, and the cubic coordinate system referred to in the text. Right: The Pyrochlore lattice, formed by a network of corner sharing tetrahedra.

i.e. the problem becomes an effectively antiferromagnetic model, and can be rewritten as:

$$H = \frac{J|S|^2}{6}\left[(s_A + s_B + s_C + s_D)^2 - 4\right] \tag{6.22}$$

within a given tetrahedron. Hence the interactions inside each tetrahedron are minimised by choosing two s_i to be positive and two to be negative.

This rule is repeated in each tetrahedron, and since only corners are shared, there is a massive degeneracy in the number of ways to satisfy these rules. [The rules are analogous to the relative separations of protons from oxygen nuclei in water ice, with protons on the vertices of the tetrahedra and oxygen on the diamond lattice of centres of tetrahedra.] The ground state entropy of this configuration was estimated by Pauling to be $k_B(1/2)\ln(3/2)$ per spin.

Magnetostatic interaction

The above configuration, with strong anisotropy, is seen in both $\mathrm{Ho_2Ti_2O_7}$ and $\mathrm{Dy_2Ti_2O_7}$, but because the anisotropy is associated with large on-site moments, e.g. $10\mu_B$ for the Ho^{3+} ions, there is also a substantial magnetic dipole interaction — magnetostatic interactions scale like μ^2, and combined with the reduced exchange interactions for the localised moments of rare earth ions, magnetostatic and exchange terms can become comparable in some rare earth systems such as these.

Thus, one should consider in addition:

$$\begin{aligned} H^{ij}_{\text{magnetostatic}} &= (g\mu_B)^2 \sum_{ij} \frac{\mathbf{S}_i \cdot \mathbf{S}_j - 3\,(\mathbf{S}_i \cdot \hat{\mathbf{r}}_{ij})\,(\mathbf{S}_j \cdot \hat{\mathbf{r}}_{ij})}{|\mathbf{r}_{ij}|^3} \\ &= (g\mu_B|S|)^2 \sum_{ij} s_i s_j \left(\frac{\hat{\mathbf{z}}_i \cdot \hat{\mathbf{z}}_j - 3\,(\hat{\mathbf{z}}_i \cdot \hat{\mathbf{r}}_{ij})\,(\hat{\mathbf{z}}_j \cdot \hat{\mathbf{r}}_{ij})}{|\mathbf{r}_{ij}|^3}\right) \end{aligned} \tag{6.23}$$

Considering for example the interaction between sites A and B, the second term can be written using $\hat{\mathbf{r}}_{AB} = (1,1,0)/\sqrt{2}$ giving $\hat{\mathbf{z}}_{A,B} \cdot \hat{\mathbf{r}}_{AB} = \pm\sqrt{2/3}$

$$H^{AB}_{\text{magnetostatic}} = (g\mu_B|S|)^2 \frac{s_A s_B}{|\mathbf{r}_{AB}|^3}\left[-\frac{1}{3} - 3\left(-\frac{2}{3}\right)\right] = (g\mu_B|S|)^2 \frac{5}{3}\frac{s_A s_B}{|\mathbf{r}_{AB}|^3}, \tag{6.24}$$

hence this interaction has the same effect as the ferromagnetic exchange, and reinforces the nearest neighbour frustration. Taking into account longer range interactions from this dipole moment can complicate this picture, however the spin ice state, with residual entropy at low temperatures can be seen experimentally from measurements of e.g. heat capacity. For further discussion of these materials, see the review by Bramwell and Gringras [12].

6.3 Magnetic domains

Another situation in which magnetostatics is important is in the domain structure of ferromagnets. This large scale structure can be described without a detailed picture of the ferromagnetic exchange underlying the magnetic ordering, but it is nevertheless the result of the strength of that exchange, and also of the magnetic anisotropy that arose from combining crystal field splitting and spin-orbit coupling. For further discussion, see the review by Kittel [13].

The basic driving force behind the formation of domains is reduction of the magnetic field energy. For a completely closed domain structure, the magnetic field should in fact vanish. Magnetic fields will however exist if the magnetisation ends on any surface normal to the magnetisation, i.e. from:

$$\mathbf{B} = \mu_0 (\mathbf{M} + \mathbf{H}), \quad \nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M} \simeq -\hat{\mathbf{S}} \cdot \mathbf{M} \tag{6.25}$$

where $\hat{\mathbf{S}}$ is a surface normal. However, before one can address the size and arrangement of domains, it is first necessary to consider the energy of a domain wall, which will reduce the number of domains.

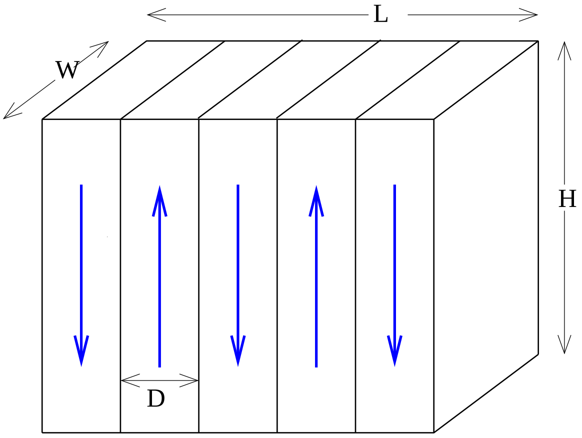


Figure 6.3: Domain structure with emergence of magnetic field (appropriate for large magnetic anisotropy).

Domain wall energy

The energy of a domain wall comes from a combination of exchange and anisotropy energy. We will for the following discussion assume that magnetic anisotropy leads to an easy axis; the case of easy plane or cubic anisotropy will be discussed briefly later. Anisotropy is necessary since otherwise, exchange energy favours alignment of spins, and hence favours as large a domain wall as possible to reduce the degree of misalignment.

For an easy axis in the z direction, we can write down a gradient expansion theory for the magnetic energy as:

$$E = \int d^3 r U(|\mathbf{M}|) + \frac{\alpha}{2} \sum_{ik} \frac{dM_i}{dr_k} \frac{dM_i}{dr_k} + \frac{\beta}{2}(M_x^2 + M_y^2). \tag{6.26}$$

The first term is assumed particularly strong, thus fixing the modulus of the local magnetisation; the second term favours alignment, while the final term is the anisotropy. The alignment term could in principle be anisotropic, however this would not affect the following discussion.

Let us consider a wall between a region where $\mathbf{M} = M(0, 0, -1)$ for $x < 0$ and $\mathbf{M} = M(0, 0, 1)$ for $x > 0$. It might at first appear that there is a general parameterisation $\mathbf{M} = M(\sin(\chi)\sin(\theta), \cos(\chi)\sin(\theta), \cos(\theta))$, however the choice of χ is restricted by the requirement that $\nabla \cdot \mathbf{M}$ should be as small as possible to minimise magnetic energy. This means $\chi = 0$ is required, so there is no component of $\mathbf{M}$ in the direction the magnetisation is changing.

Hence, we may write $\mathbf{M} = M(0, \sin(\theta), \cos(\theta))$ with $\theta(x)$, and substitute into Eq. (6.26) to give:

$$E[\theta(x)] = E_0 + M^2 \iint dz dy \int dx \left[\frac{\alpha}{2} (\partial_x \theta)^2 + \frac{\beta}{2} \sin(\theta)^2 \right] \tag{6.27}$$

This is easily solved via the first integral of the Euler-Lagrange equations:

$$\alpha \frac{d^2\theta}{dx^2} = \beta \sin(\theta)\cos(\theta) \quad \Rightarrow \quad \alpha(\partial_x \theta)^2 = \beta \sin(\theta)^2 + C, \tag{6.28}$$

and by the boundary conditions $\theta \to \pi, 0$, one has $C = 0$, and the remaining equation can be integrated to give the solution:

$$\cos(\theta) = \tanh\left(x\sqrt{\beta}\alpha \right) \tag{6.29}$$

and thus the width of the domain wall is $\sqrt{\alpha/\beta}$, and the energy given by:

$$\sigma = \frac{E_{\text{wall}}}{\text{Area}} = M^2 \beta \int dx \text{sech}^2 \left(x\sqrt{\beta}\alpha \right) = 2M^2 \sqrt{\alpha\beta}. \tag{6.30}$$

We may now employ this expression to work out the optimum number of domains.

Magnetic energy and domains of closure

Considering the arrangement in Fig. 6.3, there are two sources of energy that compete: magnetic energy is associated with the region in space above and below the crystal where magnetic fields emerge, and wall energy with the number of walls.

From the above expression for the energy density of a wall, we have that the energy of the walls is given by:

$$E_{\text{wall}} = (\sigma W H) N_{\text{walls}} = (2M^2\sqrt{\alpha\beta} W H)\left(\frac{L}{D}\right). \tag{6.31}$$

The energy of the magnetic field is given by the energy density $\mu_0 \mathbf{H}^2/2$ multiplied by the volume of space in which it exists. Formally, one should use $\mathbf{H} = -\nabla\phi_M$ and solve:

$$\nabla^2\phi_M = \nabla\cdot\mathbf{M} \simeq \delta(z) M \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{\sin[(2n+1)\pi x/D)]}{2n+1} - \delta(z-L)\ldots \tag{6.32}$$

i.e., taking a square wave source. This can be straightforwardly solved in the Fourier domain. However, it is immediately clear that the only lengthscale in the problem is D, and hence D controls both the periodicity of the solution, and its extent in the z direction, above the crystal. Thus, up to a numerical factor, we have that:

$$E_{\text{magnetic}} = \#(WLD)(\mu_0 M^2). \tag{6.33}$$

Minimising the total energy with respect to domain size then yields:

$$\frac{d}{dD}\left(\frac{E}{WL}\right) = 0 = -2M^2\sqrt{\alpha\beta}\frac{H}{D^2} + \#\mu_0 M^2$$

giving:

$$D \propto \sqrt{H\sqrt{\alpha\beta}/\mu_0}. \tag{6.34}$$

An alternative possibility, if anisotropy energy is smaller, is that the emergence of the magnetic field can be entirely avoided by creating *domains of closure*, as illustrated in Fig. 6.4. In this case, there is no magnetostatic energy cost, as $\nabla\cdot\mathbf{M}$ is everywhere zero (the perpendicular magnetisation is constant at the domain wall). However, there is energy due to anisotropy, dependent on the volume of the closure domains, hence

$$E_{\text{closure}} = \frac{\beta}{2}M^2 \times \left(\frac{D^2}{4}W\right)\frac{L}{D} \propto \beta M^2 DWL. \tag{6.35}$$

This equation is very similar to Eq. (6.33) with $\#\mu_0 \to \beta$, and so the effect of such a term on domain size is identical. It is also clear that closure domains will arise if $\beta > \#\mu_0$, and that the numerical factor depends on the geometry of the problem.

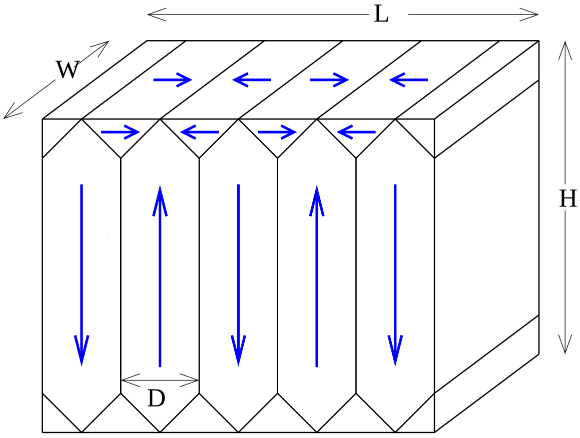


Figure 6.4: Domain structure with closure domains (appropriate for small magnetic anisotropy).

Cubic symmetry and effects of magnetoelastic energy

One question that may arise from the above is whether, since the previous arguments relied on magnetic anisotropy, in a cubic crystal should there always be equal sized closure domains, as illustrated in Fig. 6.5(a). In a cubic structure, the anisotropy energy can have no quadratic term, and the lowest order term consistent with the symmetry would be:

$$E_{\text{anisotropy}} = \beta' \left(M_x^2 M_y^2 + M_y^2 M_z^2 + M_x^2 M_z^2 \right) \tag{6.36}$$

which favours alignment along one of the three cubic axes. However, in such symmetric cases, there can be magnetoelastic effects that become relevant, i.e. a coupling between elastic deformation of the lattice and the direction of magnetisation, so that the domain structure would lead to deformation as in Fig. 6.5(b). To resolve such tension in the crystal, a domain structure with long domains in the extended direction becomes favourable, restoring Fig. 6.5(c).

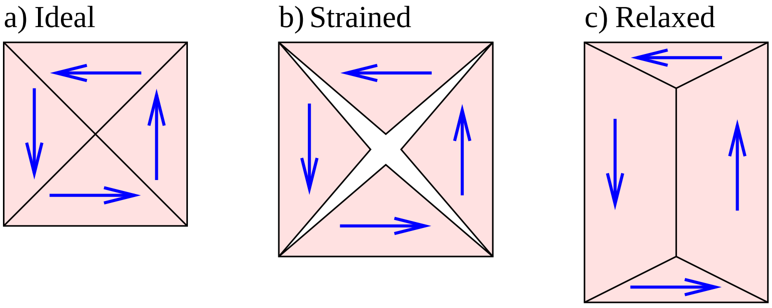


Figure 6.5: Deformation of domains due to magnetoelastic stress, leading to strain energy cost existing when domains are made to keep in contact. (a) Shows the ideal case with no stress; (b) the natural deformation of that structure; (c) a structure where domains join up, but elastic energy favours non-equal domain sizes.

Branching domains

Another concern with the above argument is that if $D \propto \sqrt{H}$, it seems domain size would grow without limit. While it is true that equilibrium domain size does grow with crystals, for large enough domains, another effect changes the size and shape of the domains.

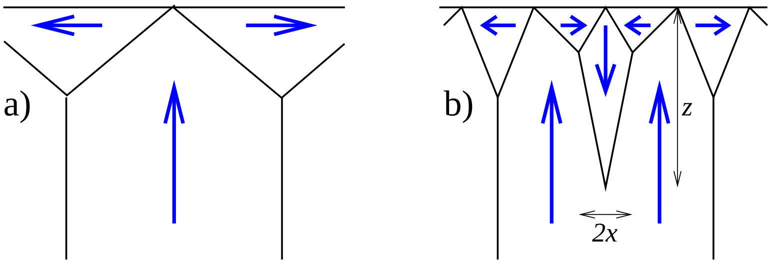


Figure 6.6: For large enough crystals, the width of domains is such that it becomes energetically favourable to form branching domains (right) at the surface. These domains have boundaries non parallel to the magnetic field, and hence a cost from magnetic field energy, but can reduce the energy of emergence or the closure domain energy

Figure 6.6(b) shows schematically the idea of a branching domain, which either acts to reduce anisotropy energy (as shown here) or magnetostatic energy above the surface. This branching domain differs from the previous examples because its boundaries are not parallel to the field, hence:

- It has a height z not controlled by the crystal geometry, but by the angle $\sim x/z$.
- It has an associated magnetostatic energy inside the crystal.

We proceed in two steps; first finding the optimal z for given x, and then the optimal x. The optimal z comes from competition of the additional wall energy and magnetostatic energy. For this one branching domain the wall cost is approximately:

$$E_{\text{wall}} = 2M^2\sqrt{\alpha\beta}(2Wz). \tag{6.37}$$

The magnetostatic energy arises due to a horizontal H field driven by poles on the surface of the branch domain, $\hat{\mathbf{S}} \cdot \mathbf{M} = -\hat{\mathbf{S}} \cdot \mathbf{H}$, giving $Mx/z \simeq H_x$, and so:

$$E_{\text{magnetostatic}} \simeq \frac{\mu_0}{2}\left(\frac{Mx}{z}\right)^2 (DWz) = \frac{\mu_0}{2}M^2\frac{x^2DW}{z} \tag{6.38}$$

This expression assumes the branching domain is repeated in other domains, giving cancelling poles on the next surface. We will however write down and compare the energies per domain. Minimising the above combination yields:

$$z = \frac{x}{2}\sqrt{\frac{D\mu_0}{2\sqrt{\alpha\beta}}}, \qquad E_{\text{branch}} = 2\sqrt{2}WM^2x\sqrt{D\mu_0\sqrt{\alpha\beta}}. \tag{6.39}$$

This energy cost for increasing x is balanced against the reduction in closure domain volume; the closure domains in Fig. 6.6(b) occupy a volume:

$$V = W\left[2x^2 + \left(\frac{D}{2} - 2x\right)^2\right] \tag{6.40}$$

(taking all angles to by $\pi/4$ as is optimal for continuity of fields, and assuming $x < D/4$). The reduction in energy for non-zero x is thus:

$$\delta E_{\text{closure}} = \frac{\beta M^2}{2} W(-2xD + 6x^2). \tag{6.41}$$

Combining and differentiating with respect to x gives:

$$\frac{dE_{\text{total}}}{dx} = WM^2\left[\frac{\beta}{2}(-2D + 12x) + 2\sqrt{2D\mu_0\sqrt{\alpha\beta}}\right], \tag{6.42}$$

hence:

$$x = \frac{1}{12}\left[2D - \frac{4}{\beta}\sqrt{2D\mu_0\sqrt{\alpha\beta}}\right] \tag{6.43}$$

meaning that such a domain (non-zero x) is favourable once:

$$D > D_0 \propto \frac{\mu_0\sqrt{\alpha\beta}}{\beta^2}. \tag{6.44}$$

Hints for the questions

Lecture 1: Quantum magnetism and exchange

Question 1.1: Lieb-Mattis theorem: Absence of ferromagnetism in one dimension

a The Hamiltonian is invariant under rotations, and so one has $[H, S^x] = [H, S^y] = 0 = [H, S^{\pm}]$, thus raising or lowering operators cannot affect energy.

b This is very similar to the two-electron problem discussed in lectures.

c Consider writing the spin state involving the larger $|S|$ in terms of lowering operators acting on the maximum state, i.e. $\propto \langle S_2, M|(S^-)^{S_1-M}|S_1, S_1\rangle$, and then consider how S_- acts on the opposite state. Why does this argument not apply if $S_1 = S_2$?

d To have an antisymmetric spatial wavefunction, the spin wavefunction of such coordinates must be symmetric; hence show that the spin raising operator anihilates this state.

e Consider the steps in the above argument that differentiated ground and excited states.

Question 1.2: Three electron atom and Hund's rules

The matrix is written out in a basis:

$$\begin{gathered}
\psi_1(x)\psi_0(y)\psi_{-1}(z)\\
\psi_1(y)\psi_0(x)\psi_{-1}(z)\\
\psi_1(z)\psi_0(y)\psi_{-1}(x)\\
\psi_1(x)\psi_0(z)\psi_{-1}(y)\\
\psi_1(y)\psi_0(z)\psi_{-1}(x)\\
\psi_1(z)\psi_0(x)\psi_{-1}(y)
\end{gathered}$$

where x, y, z are the coordinates of the three particles. Consider an arbitrary matrix element of the interaction, and note each such matrix element can be split into matrix elements of terms such as $1/r_{12}$ alone. For $1/r_{12}$, it is clear that the element will vanish unless the third particle is in the same initial/final state, hence causing many cancellations. Particles $1, 2$ must be in states $\pm 1, \mp 1$ or $\pm 1, 0$ — note that reversing the sign of momentum should not affect the matrix element.

Consider the x integral for A and B; the parts in the sum over modes n of the interaction due corresponding to the first coordinate, i.e. $A = |\sum_n a_n|^2, B = |\sum_n b_n|^2$, are given by

$$a_n = C_n \int d\theta d\phi [Y_1^{-1}]^2 Y_n^2$$

$$b_n = C_n \int d\theta d\phi Y_1^{-1} Y_1^0 Y_n^1$$

where C_n is a common constant to both terms. The ϕ integral has been anticipated, and so only superscripts m that do not cancel have been kept, C_n is common to both integrals. By writing the θ dependence of the spherical harmonics in terms of legendre polynomials, and using the orthogonality of legendre polynomials, one may show (by partial integration) that a_n, b_n vanish unless $n = 2$, and that $a_2 = \sqrt{2} b_2$, hence $A = 2B$.

Lecture 2: Hubbard and Heisenberg models

Question 2.1: Schwinger bosons

a Under the first transformation, one gets:

$$S^x \to \sin(2\chi)S^z + \cos(2\chi)S^x, \quad S^y \to S^y, \quad S^z \to \cos(2\chi)S^z - \sin(2\chi)S^z. \tag{6.45}$$

and under the second:

$$S^x \to \cos(2\phi)S^x - \sin(2\phi)S^y, \quad S^y \to \cos(2\phi)S^y + \sin(2\phi)S^x, \quad S^z \to S^z. \tag{6.46}$$

Hence together these span the possible rotations.

b To show this state is rotationally invariant, one need only show that under the two global rotations in the previous part the state is not changed. This is trivial for the second rotation. For the first:

$$\begin{aligned}(\cos(\chi)a_i^\dagger &+ \sin(\chi)b_i^\dagger)(\cos(\chi)b_j^\dagger - \sin(\chi)a_j^\dagger) \\ &- (\cos(\chi)b_i^\dagger - \sin(\chi)a_i^\dagger)(\cos(\chi)a_j^\dagger + \sin(\chi)b_j^\dagger) \\ &= \cos^2(\chi)(a_i^\dagger b_j^\dagger - b_i^\dagger a_j^\dagger) - \sin^2(\chi)(b_i^\dagger a_j^\dagger - a_i^\dagger b_j^\dagger) \\ &\quad + \sin(\chi)\cos(\chi)(b_i^\dagger b_j^\dagger - a_i^\dagger a_j^\dagger - b_i^\dagger b_j^\dagger + a_i^\dagger a_j^\dagger).\end{aligned} \tag{6.47}$$

For the constraint, one need only count the number of creation operators of each type in any given term

Question 2.2: Schwinger bosons and valence bond solid state

a As in the previous part, simple counting suffices, noting that each site will appear twice in the product.

b The maximum value of S^z will come from a term like:

$$\dots (a_{i-1}^\dagger b_i^\dagger)^M \left(a_i^\dagger b_{i+1}^\dagger - a_{i+1}^\dagger b_i^\dagger\right)^M (b_{i+1}^\dagger a_{i+2}^\dagger)^M \dots |0\rangle.$$

Since the terms on the bond $(i, i+1)$ always have equivalent numbers of a and b operators, it does not matter which term we consider, on the relevant pair of sites it will always have $3M$ operators $b^\dagger$, and M operators $a^\dagger$, giving spin $2M/2 = M$. Because rotationally invariant, $|S_{i,i+1}| \leq M$.

c If $S = M = 1$, it is clear that this state, having no pairs of sites with total spin projection greater than 1 eliminates the non-ground state parts, and hence is in the ground state manifold.

Question 2.3: Four-site Heisenberg model

Defining states $|a\rangle = (12)(34), |b\rangle = (14)(23)$ where (ij) indicates a singlet bond between i and j, one has:

$$\begin{aligned}H|a\rangle &= -J|a\rangle + J|b\rangle, \\ H|b\rangle &= -J|b\rangle + J|a\rangle.\end{aligned}$$

and (if normalised), then:

$$\langle a|b\rangle = 1/2 = \langle b|a\rangle. \tag{6.48}$$

Hence, one must solve:

$$\mathrm{Det}\left[E\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} - J\begin{pmatrix} -1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}\right] = 0 \tag{6.49}$$

having solutions $E = 0, -2J$, corresponding to $(12)(34) \pm (14)(23)$.

Lecture 3: Real insulators

Question 3.1: The importance of interactions in superexchange

Question 3.2: Anisotropic spins in a magnetic field

Using the same method as in Sec. 3.3 one may eliminate the spin-orbit term to leading order in $1/(E_n - E_0)$ and so get:

$$\delta H = 2\mu_B \mathbf{H} \cdot \mathbf{S} - (S_\mu + \mu_B H_\mu)\Lambda_{\mu\nu}(S_\nu + \mu_B H_\nu). \tag{6.50}$$

The Zeeman term is identified by the linear coupling between **H** and **S**.

Lecture 4: Itinerant electrons

Question 4.1: Kanamori condition for quasi-free electrons

Bibliography

[1] P. Fazekas, *Lecture notes on electron correlation and magnetism* (World Scientific, Singapore, 1999).

[2] K. Yosida, *Theory of Magnetism* (Springer, Heidleberg, 1998).

[3] D. C. Mattis, *The theory of magnetism* (Springer-Verlag, Heidleberg, 1981).

[4] D. C. Mattis, *The Theory of Magnetism Made Simple* (World Scientific, Singapore, 2004).

[5] P. W. Anderson, *Theory of Magnetic Exchange Interactions: Exchange in Insulators and semiconductors* (Academic Press, New York, 1963), vol. 14 of *Solid State Physics*, pp. 99–214.

[6] A. Auerbach, *Interacting electrons and quantum magnetism* (Springer, Heidleberg, 1998).

[7] N. Nagaosa, *Quantum Field Theory in Strongly Correlated Electronic Systems* (Springer, Heidleberg, 1999).

[8] J. Hubbard, Proc. Roy. Soc.: A **276**, 238 (1963).

[9] E. Lieb and D. Mattis, Phys. Rev. **125**, 164 (1962).

[10] Y. Nagaoka, Phys. Rev. **147**, 392 (1966).

[11] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond* (World Scientific, Singapore, 1987).

[12] S. T. Bramwell and M. J. P. Gringras, Science **294**, 1495 (2001).

[13] C. Kittel, Rev. Mod. Phys. **21**, 541 (1949).