Contents

Reading:

1. Chs. 31-33, Ashcroft & Mermin
2. Ch. 4, Kittel
3. For a more detailed discussion of the topic, see e.g. D. Mattis, Theory of Magnetism I & II, Springer 1981

3 Quantum Magnetism

The main purpose of this section is to introduce you to ordered magnetic states in solids and their "spin wave-like" elementary excitations. Magnetism is an enormous field, and reviewing it entirely is beyond the scope of this course.

3.1 Introduction

3.1.1 Atomic magnetic Hamiltonian

The simplest magnetic systems to consider are insulators where electron-electron interactions are weak. If this is the case, the magnetic response of the solid to an applied field is given by the sum of the susceptibilities of the individual atoms. The magnetic susceptibility is defined by the the 2nd derivative of the free energy,[1]

$$\chi = -\frac{\partial^2 F}{\partial H^2}. \tag{1}$$

We would like to know if one can understand (on the basis of an understanding of atomic structure) why some systems (e.g. some

[1] In this section I choose units such that the system volume V = 1.

elements which are insulators) are paramagnetic ($\chi > 0$) and some diamagnetic ($\chi < 0$).

The one-body Hamiltonian for the motion of the electrons in the nuclear Coulomb potential in the presence of the field is

$$H_{atom} = \frac{1}{2m}\sum_i \left(\mathbf{p}_i + \frac{e}{c}\mathbf{A}(\mathbf{r}_i)\right)^2 + \sum_i V(r_i) + g_0\mu_B\mathbf{H}\cdot\mathbf{S}, \quad (2)$$

where $\Sigma_i \mathbf{S}_i$ is the total spin operator, $\mu_B \equiv e/(2mc)$ is the Bohr magneton, $g_0 \simeq 2$ is the gyromagnetic ratio, and $\mathbf{A} = -\frac{1}{2}\mathbf{r}\times\mathbf{H}$ is the vector potential corresponding to the applied uniform field, assumed to point in the $\hat{z}$ direction. Expanding the kinetic energy, H_{atom} may now be expressed in terms of the orbital magnetic moment operator $\mathbf{L} = \Sigma_i \mathbf{r}\times\mathbf{p}_i$ as

$$H_{atom} = \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_i V(r_i) + \delta H, \quad (3)$$

$$\delta H = \mu_B(\mathbf{L} + g_0\mathbf{S})\cdot\mathbf{H} + \frac{e^2}{8mc^2}H^2\sum_i(x_i^2 + y_i^2). \quad (4)$$

Given a set of exact eigenstates for the atomic Hamiltonian in zero field $|n\rangle$ (ignore degeneracies for simplicity), standard perturbation theory in δH gives

$$\delta E_n = \mathbf{H}\cdot\langle n|\vec{\mu}|n\rangle + \sum_{n\neq n'}|\frac{\langle n|\mathbf{H}\cdot\vec{\mu}|n'\rangle|^2}{E_n - E_{n'}} + \frac{e^2}{8mc^2}H^2\langle n|\sum_i(x_i^2 + y_i^2)|n\rangle, (5)$$

where $\vec{\mu} = \mu_B(\mathbf{L} + g_0\mathbf{S})$. It is easy to see that the first term dominates and is of order the cyclotron frequency $\omega_c \equiv eH/(mc)$ unless it vanishes for symmetry reasons. The third term is of order $\omega_c/(e^2/a_0)$ smaller, because typical electron orbits are confined to atomic sizes a_0, and is important in insulators only when the state $|n\rangle$ has zero magnetic moment ($L = S = 0$). Since the coefficient of H^2 is manifestly positive, the susceptibility[2] in the $\mu = 0$ ground state $|0\rangle$ is $\chi = -\partial^2 \delta E_0 / \partial H^2$, which is clearly < 0, i.e. diamagnetic.[3]

In most cases, however, the atomic shells are partially filled, and the ground state is determined by minimizing the atomic energy together with the intraatomic Coulomb interaction needed to break certain degeneracies, i.e. by "Hund's rules".[4] Once this ground state (in particlular the S, L, and J quantum numbers is known, the atomic susceptibility can be calculated. The simplest case is again the $J = 0$ case, where the first term in (5) again vanishes. Then the second term and third term compete, and can result in either a diamagnetic or paramagnetic susceptibility. The 2nd term is called the *Van Vleck* term in the energy. It is paramagnetic since its contribution to χ is

$$\chi_{VV} = -\frac{\partial^2 E_0|_{2nd\ term}}{\partial H^2} = 2\mu_B^2 \sum_n \frac{|\langle 0|\vec{\mu}|n'\rangle|^2}{E_n - E_0} > 0. \tag{6}$$

3.1.2 Curie Law-free spins

In the more usual case of $J \neq 0$, the ground state is $2J + 1$ degenerate, and we have to be more careful about defining the susceptibility. The free energy as $T \to 0$ can no longer be replaced by E_0 as we did above. We have to account for the entropy of the $2J + 1$ degenerate spin states as well. Applying a magnetic field breaks this degeneracy, so we have a small statistical calculation

[2] If the ground state is nondegenerate, we can replace $F = E - TS$ in the definition of the susceptibility by the ground state energy E_0.

[3] This weak diamagnetism in insulators with filled shells is called Larmor diamagnetism

[4] See A&M or any serious quantum mechanics book. I'm not going to lecture on this but ask you about it on homework.

to do. The energies of the "spin" in a field are given by

$$H = -\vec{\mu} \cdot \mathbf{H}, \tag{7}$$

and since $\vec{\mu} = -\gamma \mathbf{J}$ within the subspace of definite J^2,[5] the $2J+1$ degeneracy in H = 0 is completely broken. The free energy is

$$\begin{aligned} F = -T \log Z &= -T \log \sum_{J_z=-J}^{J} e^{\beta\gamma H J_z} \\ &= -T \log \left[\frac{e^{\beta\gamma H(J+1/2)} - e^{-\beta\gamma H(J+1/2)}}{e^{\beta\gamma H/2} - e^{-\beta\gamma H/2}} \right], \end{aligned} \tag{8}$$

so the magnetization of the free spins is

$$M = -\frac{\partial F}{\partial H} = \gamma J B(\beta\gamma J H), \tag{9}$$

where $B(x)$ is the Brillouin function

$$B(x) = \frac{2J+1}{2J} \coth \frac{2J+1}{2J} x - \frac{1}{2J} \coth \frac{1}{2J} x. \tag{10}$$

Note I defined $\gamma = \mu_B g$. We were particularly interested in the $H \to 0$ case, so as to compare with ions with filled shells or $J = 0$; in this case one expands for $T >> \gamma H$, i.e. $\coth x \sim 1/x + x/3 + \ldots$, $B(x) \simeq (J+1)x/(3J)$ to find the susceptibility

$$\boxed{\chi = -\frac{\partial^2 F}{\partial H^2} = \frac{\gamma^2 J(J+1)}{3T}}, \tag{11}$$

i.e. a *Curie law* for the high-temperature susceptibility. A $1/T$ susceptibility at high T is generally taken as evidence for free paramagnetic spins, and the size of the moment given by $\mu^2 = \gamma^2 J(J+1)$.

[5] This is not obvious at first sight. The magnetic moment *operator* $\vec{\mu} \propto \hat{\mathbf{L}} + \mathbf{g_0}\hat{\mathbf{S}}$ is not proportional to the total angular momentum *operator* $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$. However its matrix elements *within the subspace of definite* L, S, J are proportional, due to the Wigner-Eckart theorem. One therefore may assume the proportionality within the subspace, where the proportionality const. $\gamma = g(J, L, S)\mu_B$. g is called Landé g-factor, and is independent of J_z according to Wigner-Eckart. See e.g. Ashcroft & Mermin p. 654.

3.1.3 Magnetic interactions

The most important interactions between magnetic moments in an insulator are electrostatic and inherently quantum mechanical in nature. From a classical perspective, one might expect two such moments to interact via the classical dipolar force, giving a potential energy of configuration of two moments $\vec{\mu}_1$ and $\vec{\mu}_2$ separated by a distance $\mathbf{r}$ of

$$U = \frac{\vec{\mu}_1 \cdot \vec{\mu}_2 - 3(\vec{\mu}_1 \cdot \hat{r})(\vec{\mu}_1 \cdot \hat{r})}{r^3} \tag{12}$$

Putting in typical atomic moments of order $\mu_B = e\hbar/mc$ and distances of order a Bohr radius $r \sim a_0 = \hbar^2/me^2$, we find $U \simeq 10^{-4}eV$, which is very small compared to typical atomic energies of order eV. Quantum-mechanical *exchange* is almost aways a much larger effect, and the dipolar interactions are therefore normally neglected in the discussions of magnetic interactions in solids. Exchange arises because, e.g., two quantum mechanical spins 1/2 (in isolation) can be in either a spin triplet (total $S = 1$), or singlet (total $S = 0$). The spatial part of the two-particle wavefunctions is antisymmetric for triplet and symmetric for singlet, respectively. Since the particles carrying the spins are also charged, there will be a large energetic difference between the two spin configurations due to the different size of the Coulomb matrix elements (the "amount of time the two particles spend close to each other") in the two cases.[6] In terms of hypothetical atomic wave functions $\psi_a(\mathbf{r})$ and $\psi_b(\mathbf{r})$ for the two particles, the singlet and triplet combinations are $\psi_{\substack{0\\1}}(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \pm \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$, so the singlet-triplet

[6]Imagine moving 2 H-atoms together starting from infinite separation. Initially the 3 $S = 1$ states and 1 $S = 0$ states must be degenerate. As the particles start to interact via the Coulomb force at very long range, there will be a splitting between singlet and triplet.

splitting is approximately

$$\begin{aligned} -J \equiv E_0 - E_1 &= \langle 0|H|0\rangle - \langle 1|H|1\rangle \\ &\simeq 2\int d^3r_1 d^3r_2 \psi_a^*(\mathbf{r}_1)\psi_b^*(\mathbf{r}_2) V(\mathbf{r}_1,\mathbf{r}_2)\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1) \end{aligned} \tag{13}$$

where V represents the Coulomb interactions between the particles (and possible other particles in the problem).[7]

We'd now like to write down a simple Hamiltonian for the spins which contains the physics of this exchange splitting. This was done first by Heisenberg (check history here), who suggested

$$H_{2-spin} = J\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 \tag{15}$$

You can easily calculate that the energy of the triplet state in this Hamiltonian is $J/4$, and that of the singlet state $-3J/4$. So the splitting is indeed J. Note that the sign of J in the H_2 case is positive, meaning the $S = 0$ state is favored; the interaction is then said to be antiferromagnetic, meaning it favors antialigning the spins with each other.[8]

The so-called Heitler-London model of exchange just reviewed presented works reasonably well for well-separated molecules, but for N atoms in a real solid, magnetic interactions are much more complicated, and it is in general not sufficient to restrict one's consideration to the 4-state subspace (singlet $\oplus$ 3 components of

[7] For example, in the H molecule,

$$V(\mathbf{r}_1,\mathbf{r}_2) = \frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|} + \frac{e^2}{|\mathbf{R}_1-\mathbf{R}_2|} - \frac{e^2}{|\mathbf{r}_1-\mathbf{R}_1|} - \frac{e^2}{|\mathbf{r}_2-\mathbf{R}_2|}, \tag{14}$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ are the sites of the protons. Note the argument above would suggest naively that the triplet state should be the ground state in the well-separated ion limit, because the Coulomb interaction is minimized in the spatially antisymmetric case. However the true ground state of the H_2 molecule is the Heitler-London *singlet* state $\psi_s(\mathbf{r}_1,\mathbf{r}_2) \simeq \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$

[8] Historically the sign convention for J was the opposite; $J > 0$ was usually taken to be ferromagnetic, i.e. the Hamiltonian was defined with another minus sign. I use the more popular convention $H = J\sum S_i \cdot S_j$. Be careful!

triplet) to calculate the effective exchange. In many cases, particularly when the magnetic ions are reasonably well-separated, it nevertheless turns out to be ok to simply extend the 2-spin form (15) to the entire lattice:

$$H = J \sum_{i\delta} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+\delta} \tag{16}$$

where J is the exchange constant, i runs over sites and δ runs over nearest neighbors.[9] This is the so-called Heisenberg model. The sign of J can be either antiferromagnetic ($J > 0$ in this convention), or ferromagnetic ($J < 0$). This may lead, at sufficiently low temperature, to a quantum ordered state with ferromagnetic or antiferromagnetic-type order. Or, it may not. A good deal depends on the type and dimensionality of the lattice, as we discuss below.

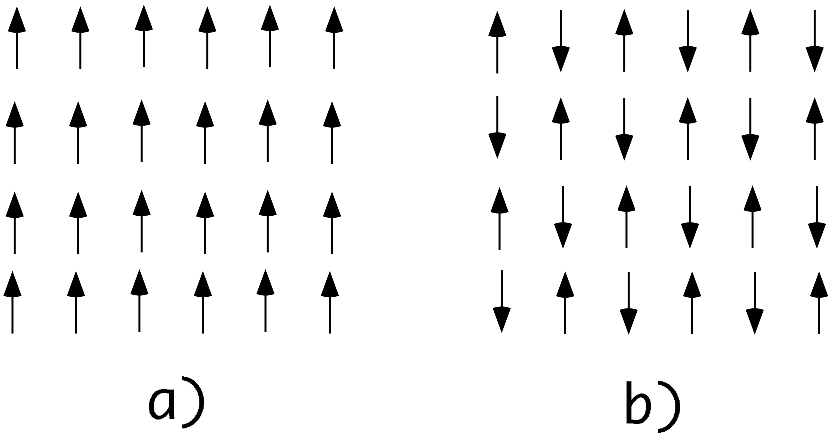


Figure 1: a) Ferromagnetic ordering; b) antiferromagnetic ordering.

Although oversimplified, the Heisenberg model is still very difficult to solve. Fortunately, a good deal has been learned about it, and

[9]One has to be a bit careful about the counting. J is defined conventionally such that there is one term in the Hamiltonian for each *bond* between two sites. Therefore if i runs over all sites, one should have δ only run over, e.g. for the simple cubic lattice, $+\hat{x}$, $+\hat{y}$, and $+\hat{z}$. If it ran over all nearest neighbors, bonds would be double-counted and we would have to multiply by $1/2$.

once one has put in the work it turns out to describe magnetic ordering and magnetic correlations rather well for a wide class of solids, provided one is willing to treat J as a phenomenological parameter to be determined from a fit to experiment.

The simplest thing one can do to take interactions between spins into account is to ask, "What is the average exchange felt by a given spin due to all the other spins?" This is the essence of the *molecular field theory* or *mean field theory* due to Weiss. Split off from the Hamiltonian all terms connecting any spins to a specific spin $\hat{\mathbf{S}}_i$. For the nearest-neighbor exchange model we are considering, this means just look at the nearest neighbors. This part of the Hamiltonian is

$$\delta H_i = \hat{\mathbf{S}}_i \cdot \left[J \sum_\delta \hat{\mathbf{S}}_{i+\delta} \right] - \vec{\mu} \cdot \mathbf{H}, \tag{17}$$

where we have included the coupling of the spin in question to the external field as well. We see that the 1-spin Hamiltonian looks like the Hamiltonian for a spin in an effective field,

$$\delta H_i = -\vec{\mu}_i \cdot \mathbf{H}_{eff}, \tag{18}$$

$$\mathbf{H}_{eff} = \mathbf{H} - \frac{J}{g\mu_B} \sum_\delta \hat{\mathbf{S}}_{i+\delta} \tag{19}$$

Note this "effective field" as currently defined is a complicated *operator*, depending on the neighboring spin operators $\hat{\mathbf{S}}_{i+\delta}$. The mean field theory replaces this quantity with its thermal average

$$\mathbf{H}_{eff} = \mathbf{H} - \frac{J}{g\mu_B} \sum_\delta \langle \hat{\mathbf{S}}_{i+\delta} \rangle \tag{20}$$

$$= \mathbf{H} - \frac{zJ}{(g\mu_B)^2} \mathbf{M}, \tag{21}$$

where the magnetization $M = \langle \mu \rangle$ and z is the number of nearest neighbors again. But since we have an effective one-body Hamiltonian, thermal averages are supposed to be computed just as in the noninteracting system, cf. (9), but in the ensemble with effective magnetic field. Therefore the magnetization is

$$M = \gamma S B(\beta\gamma S H_{eff}) = \gamma S B\left(\beta\gamma S[H - \frac{zJ}{\gamma^2}M]\right). \tag{22}$$

This is now a nonlinear equation for M, which we can solve for any H and T. It should describe a ferromagnet with finite spontaneous ($H \rightarrow 0$) magnetization below a critical temperature T_c. So to search for T_c, set $H = 0$ and expand $B(x)$ for small x:

$$M = -\gamma S B\left(\frac{zJ}{\gamma T_c}M\right) \simeq -\gamma S\,\frac{S+1}{3S}\,\frac{zJ}{\gamma T_c}M \tag{23}$$

$$\Rightarrow\ T_c = \frac{S(S+1)}{3}z(-J), \tag{24}$$

So the critical temperature in this mean field theory (unlike the BCS mean field theory!) is of order the fundamental interaction energy $|J|$. We expect this value to be an upper bound to the true critical temperature, which will be supressed by spin fluctuations about the mean field used in the calculation.

Below T_c, we can calculate how the magnetization varies near the transition by expanding the Brillouin fctn. to one higher power in x. The result is

$$M \sim (T - T_c)^{1/2}. \tag{25}$$

Note this exponent 1/2 is characteristic of the disappearance of the order parameter near the transition of any mean field theory (Landau).

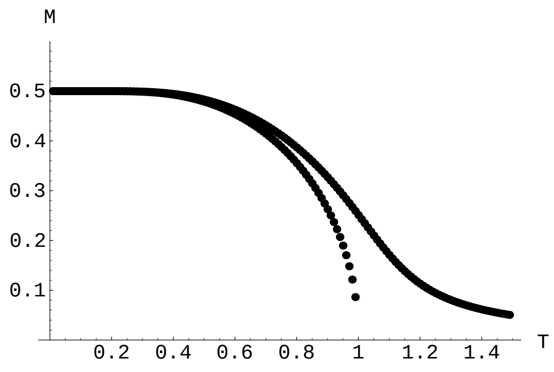


Figure 2: Plot of Mathematica solution to eqn. (22) for M vs. T using $-J = g = 1$; $z = 4; S = 1/2$. T_c=1 for this choice. Upper curve: $H = 0.1$; lower curve: $H = 0$.

3.2 Ising model

The Ising model[10] consists of a set of spins s_i with z-components only localized on lattice sites i interacting via nearest-neighbor exchange $J < 0$:

$$\mathcal{H} = J \sum_{i,j \in n.n.} S_i S_j - 2\mu_B H \sum_i S_i. \tag{26}$$

Note it is an inherently *classical* model, since all spin commutators vanish $[S_i, S_j] = 0$. Its historical importance consisted not so much in its applicability to real ferromagnetic systems as in the role its solutions, particularly the analytical solution of the 2D model published by Onsager in 1944, played in elucidating the nature of phase transitions. Onsager's solution demonstrated clearly that all approximation methods and series expansions heretofore used to attack the ferromagnetic transition failed in the critical regime, and thereby highlighted the fundamental nature of the problem of critical phenomena, not solved (by Wilson and others) until in the early 70's.

[10]The "Ising model" was developed as a model for ferromagnetism by Wilhelm Lenz and his student Ernst Ising in the early '20's.

3.2.1 Phase transition/critical point

We will be interested in properties of the model (26) at low and at high temperatures. What we will find is that there is a temperature T_c below which the system magnetizes spontaneously, just as in mean field theory, but that its true value is in general smaller than that given by mean field theory due to the importance of thermal fluctuations. Qualitatively, the phase diagram looks like this:

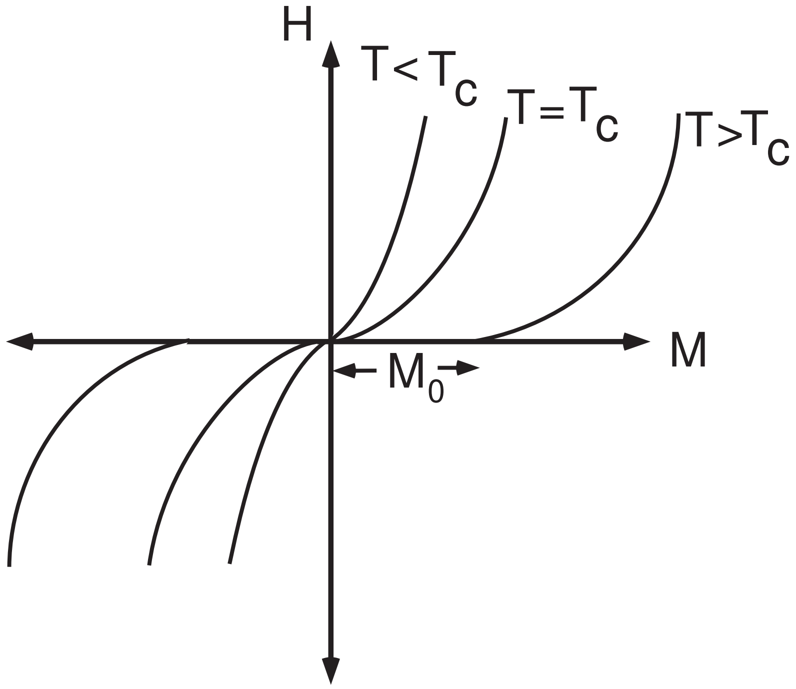


Figure 3: Field-magnetization curves for three cases. M_0 is spontaneous magnetization in ferromagnetic phase $T < T_c$.

Below T_c, the system magnetizes spontaneously even for field $H \to 0$. Instead of investigating the Onsager solution in detail, I will rely on the Monte Carlo simulation of the model developed by Jim Sethna and Bob Silsbee as part of the Solid State Simulation (SSS) project. The idea is as follows. We would like to minimize $F = T \log \mathrm{Tr} \exp -\beta H$ for a given temperature and applied field. Finding the configuration of Ising spins which does so is a complicated task, but we can imagine starting the system at high temperatures, where all configurations are equally likely, and cooling to the desired temperature T.[11] Along the way, we allow the system to relax by "sweeping" through all spins in the finite size lattice, and deciding in the next Monte Carlo "time" step whether the spin will be up or down. Up and down are weighted by a Boltzman probability factor

$$p(S_i = \pm 1/2) = \frac{e^{\pm \mu H_{eff}/T}}{e^{-\mu H_{eff}/T} + e^{\mu H_{eff}/T}}, \tag{27}$$

where H_i^{eff} is the effective field defined in (19). The simulation picks a spin S_i in the next time step randomly, but weighted with these up and down probabilities. A single "sweep" (time step) consists of $L \times L$ such attempts to flip spins, where L is the size of the square sample. Periodic boundary conditions are assumed, and the spin configuration is displayed, with one color for up and one for down.

Here are some remarks on Ising critical phenomena, some of which you can check yourself with the simulation:

- At high temperatures one recovers the expected Curie law $\chi \sim 1/T$
- The susceptibility diverges at a critical temperature below the mean field value.[12] Near, but not too near, the transition χ has the Curie-Weiss form $\chi \sim (T - T_c)^{-1}$.
- With very careful application of the simulation, one should obtain Onsager's result that very near the transition ("critical

[11] This procedure is called simulated annealing

[12] This is given as a homework problem. Note the value of J used in the simulation is 1/4 that defined here, since S's are ± 1.

regime")

$$M \sim (T_c - T)^\beta, \tag{28}$$

with $\beta = 1/8$. The susceptibility actually varies as

$$\chi \sim |T - T_c|^{-\gamma}, \tag{29}$$

with $\gamma = 7/4$. Other physical quantities also diverge near the transition, e.g. the specific heat varies as $|T - T_c|^{-\alpha}$, with $\alpha = 0$ (log divergence).

- There is no real singularity in any physical quantity so long as the system size remains finite.
- The critical exponents α, β, γ... get closer to their mean field values ($\beta = 1/2$, $\alpha =$, $\gamma =$,...) as the number of nearest neighbors in the lattice increases, or if the dimensionality of the system increases.
- The mean square size of themal magnetization fluctuations gets very large close to the transition ("Critical opalescence", so named for the increased scattering of light near the liquid-solid critical point) .
- Magnetization relaxation gets very long near the transition ("Critical slowing down").
- In 1D there is no finite temperature phase transition, although mean field theory predicts one. See next section.

3.2.2 1D solution by transfer matrix

3.2.3 Ferromagnetic domains

3.3 Ferromagnetic magnons

Let's consider the simplest example of an insulating ferromagnet, described by the ferromagnetic Heisenberg Hamiltonian

$$H = J\sum_{i\delta} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+\delta} - 2\mu_B H_0 \sum_i \hat{S}_{iz}, \tag{30}$$

where $J < 0$ is the ferromagnetic exchange constant, i runs over sites and δ runs over nearest neighbors, and H_0 is the magnetic field pointing in the $\hat{z}$ direction. It is clear that the system can minimize its energy by having all the spins $\mathbf{S}$ align along the $\hat{z}$ direction at $T = 0$; i.e. the quantum ground state is identical to the classical ground state. Finding the elementary excitations of the quantum many-body system is not so easy, however, due to the fact that the spin operators do not commute.

3.3.1 Holstein-Primakoff transformation

One can attempt to transform the spin problem to a more standard many-body interacting problem by replacing the spins with boson creation and annihilation operators. This can be done *exactly* by the Holstein-Primakoff transformation[13]

$$\hat{S}_i^+ = \hat{S}_{ix} + i\hat{S}_{iy} = (2S)^{1/2}\left(1 - \frac{a_i^\dagger a_i}{2S}\right)^{1/2} a_i \tag{31}$$

$$\hat{S}_i^- = \hat{S}_{ix} - i\hat{S}_{iy} = (2S)^{1/2} a_i^\dagger \left(1 - \frac{a_i^\dagger a_i}{2S}\right)^{1/2}. \tag{32}$$

[13] T. Holstein and H. Primakoff, Phys. Rev. 58, 1098 (1940).

Verify for yourselves that these definitions $S_i^{\pm}$ give the correct commutation relations $[\hat{S}_x, \hat{S}_y] = i\hat{S}_z$ *if* the bosonic commutation relations $[a, a^\dagger] = 1$ are obeyed on a given lattice site. Note also that the operators which commute with the Hamiltonian are $\hat{S}^2$ and $\hat{S}_z$ as usual, so we can classify all states in terms of their eigenvalues $S(S+1)$ and S_z. To complete the algebra we need a representation for $\hat{S}_z$, which can be obtained by using the identity (on a given site i)

$$\hat{S}_z^2 \;=\; S(S+1) - \frac{1}{2}\left(\hat{S}^+\hat{S}^- + \hat{S}^-\hat{S}^+\right). \tag{33}$$

Using (32) and some tedious applications of the bosonic commutation relations, we find

$$\hat{S}_z = S - a^\dagger a. \tag{34}$$

Now since the system is periodic, we are looking for excitations which can be characterized by a well-defined momentum (crystal momentum) $\mathbf{k}$, so we define the Fourier transformed variables

$$a_i = \frac{1}{N^{1/2}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}_i}\; b_{\mathbf{k}} \quad ; a_i^\dagger = \frac{1}{N^{1/2}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}_i}\; b_{\mathbf{k}}^\dagger, \tag{35}$$

where as usual the F.T. variables also satisfy the bosonic relations $[b_{\mathbf{k}}, b_{\mathbf{k}'}] = \delta_{\mathbf{k}\mathbf{k}'}$, etc. Looking forward a bit, we will see that the operators $b_{\mathbf{k}}^\dagger$ and $b_{\mathbf{k}}$ create and destroy a *magnon* or *spin-wave* excitation of the ferromagnet. These turn out to be excitations where the spins locally deviate only a small amount from their ground state values ($\parallel \hat{z}$) as the "spin wave" passes by. This suggests *a posteriori* that an expansion in the spin *deviations* $a_i^\dagger a_i$ (see (34)) may converge quickly. Holstein and Primakoff therefore suggested expanding the nasty square roots in (32), giving

$$
\begin{aligned}
\hat{S}_i^+ &\simeq (2S)^{1/2}\left[a_i - \left(\frac{a_i^\dagger a_i a_i}{4S}\right) + \ldots\right] \\
&= \left(\frac{2S}{N}\right)^{1/2}\left[\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_i} b_{\mathbf{k}} - \frac{1}{4SN}\sum_{\mathbf{k},\mathbf{k}',\mathbf{k}''} e^{i(\mathbf{k}-\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{R}_i} b_{\mathbf{k}}^\dagger b_{\mathbf{k}'} b_{\mathbf{k}''} + \ldots\right] \quad (36) \\
\hat{S}_i^- &\simeq (2S)^{1/2}\left[a_i^\dagger - \left(\frac{a_i^\dagger a_i^\dagger a_i}{4S}\right) + \ldots\right] \\
&= \left(\frac{2S}{N}\right)^{1/2}\left[\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_i} b_{\mathbf{k}}^\dagger - \frac{1}{4SN}\sum_{\mathbf{k},\mathbf{k}',\mathbf{k}''} e^{i(\mathbf{k}+\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{R}_i} b_{\mathbf{k}}^\dagger b_{\mathbf{k}'}^\dagger b_{\mathbf{k}''} + \ldots\right] \quad (37) \\
\hat{S}_{iz} &= S - a_i^\dagger a_i = S - \frac{1}{N}\sum_{\mathbf{k}\mathbf{k}'} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i} b_{\mathbf{k}}^\dagger b_{\mathbf{k}'}. \quad (38)
\end{aligned}
$$

Note that the expansion is formally an expansion in $1/S$, so we might expect it to converge rapidly in the case of a large-spin system.[14] The last equation is exact, not approximate, and it is useful to note that the total spin of the system along the magnetic field is

$$S_{z,tot} = \sum_i S_z = NS - \frac{1}{N}\sum_{\mathbf{k}} b_{\mathbf{k}}^\dagger b_{\mathbf{k}}, \quad (39)$$

consistent with our picture of magnons tipping the spin away from its $T = 0$ equilibrium direction along the applied field.

3.3.2 Linear spin wave theory

The idea now is to keep initially only the bilinear terms in the magnon operators, leaving us with a soluble Hamiltonian, hoping that we can then treat the 4th-order and higher terms pertur-

[14] For spin-1/2, the case of greatest interest, however, it is far from obvious that this uncontrolled approximation makes any sense, despite the words we have said about spin deviations being small. Why should they be? Yet empirically linear spin wave theory works very well in 3D, and surprisingly well in 2D.

batively.[15] Simply substitute (36)-(38) into (30), and collect the terms first proportional to S^2, S, 1, $1/S$, etc. We find

$$H = \frac{1}{2}JNzS^2 - 2\mu_B H_0 S + H_0^{magnon} + \mathcal{O}(1), \tag{40}$$

where z is the number of nearest neighbors (e.g. 6 for simple cubic lattice), and

$$\begin{aligned} H_0^{magnon} &= \frac{JS}{N}\sum_{i\delta\mathbf{k}\mathbf{k}'}\Big[e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i}e^{i\mathbf{k}'\cdot\delta}b_{\mathbf{k}}b^\dagger_{\mathbf{k}'} + e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i}e^{-i\mathbf{k}'\cdot\delta}b^\dagger_{\mathbf{k}}b_{\mathbf{k}'} \\ &\quad -e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i}b^\dagger_{\mathbf{k}}b_{\mathbf{k}'} - e^{-i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{R}_i+\delta)}b^\dagger_{\mathbf{k}}b_{\mathbf{k}'}\Big] + \frac{2\mu_B H_0}{N}\sum_{i\mathbf{k}\mathbf{k}'}e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i}b^\dagger_{\mathbf{k}}b_{\mathbf{k}'} \\ &= JzS\sum_{\mathbf{k}}\Big[\gamma_{\mathbf{k}}b_{\mathbf{k}}b^\dagger_{\mathbf{k}} + \gamma_{-\mathbf{k}}b^\dagger_{\mathbf{k}}b_{\mathbf{k}} - 2b^\dagger_{\mathbf{k}}b_{\mathbf{k}}\Big] + 2\mu_B H_0\sum_{\mathbf{k}}b^\dagger_{\mathbf{k}}b_{\mathbf{k}} \\ &= \sum_{\mathbf{k}}\left[-2JzS(1-\gamma_{\mathbf{k}}) + 2\mu_B H_0\right]b^\dagger_{\mathbf{k}}b_{\mathbf{k}}, \end{aligned} \tag{41}$$

where

$$\gamma_k = \frac{1}{z}\sum_{\delta}e^{i\mathbf{k}\cdot\delta} \tag{42}$$

is the magnon dispersion function, which in this approximation depends only on the positions of the nearest neighbor spins. Note in the last step of (41), I assumed $\gamma_{\mathbf{k}} = \gamma_{-\mathbf{k}}$, which is true for lattices with inversion symmetry. For example, for the simple cubic lattice in 3D with lattice constant a, $\gamma_{\mathbf{k}} = (\cos k_x a + \cos k_y a + \cos k_z a)/3$, clearly an even function of $\mathbf{k}$. Under these assumptions, the magnon part of the Hamiltonian is remarkably simple, and can be written like a harmonic oscillator or phonon-type Hamiltonian, $H_0^{magnon} = \Sigma_{\mathbf{k}}\, n_{\mathbf{k}}\omega_{\mathbf{k}}$, where $n_{\mathbf{k}} = b^\dagger_{\mathbf{k}}b_{\mathbf{k}}$ is the number of magnons in state $\mathbf{k}$, and

$$\omega_k = -JSz(1-\gamma_{\mathbf{k}}) + 2\mu_B H_0 \tag{43}$$

[15] physically these "nonlinear spin wave" terms represent the interactions of magnons, and resemble closely terms representing interactions of phonons in anharmonic lattice theory

is the magnon dispersion. The most important magnons will be those with momenta close to the center of the Brillouin zone, $k \sim 0$, so we need to examine the small-k dispersion function. For a Bravais lattice, like simple cubic, this expansion gives $1 - \gamma_{\mathbf{k}} \simeq k^2$,[16] i.e. the magnon dispersion vanishes as $k \to 0$. For more complicated lattices, there will be solutions with $\omega_{\mathbf{k}} \to$ const. There is always a "gapless mode" $\omega_{\mathbf{k}} \to 0$ as well, however, since the existence of such a mode is guaranteed by the Goldstone theorem.[17] The figure shows a simple 1D schematic of a spin wave

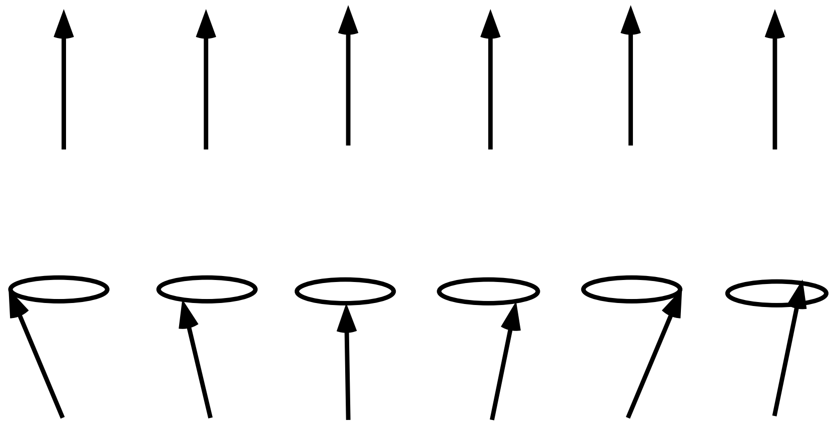


Figure 4: Real space picture of spin deviations in magnon. Top: ordered ground state

with wavelength $\lambda = 2\pi/k$ corresponding to about 10 lattice sites. The picture is supposed to convey the fact that the spin deviations from the ordered state are small, and vary slightly from site to site. Quantum mechanically, the wave function for the spin wave state contains at each site a small amplitude for the spin to be in a state with definite S_x and/or S_y. This can be seen by inverting Eq. (37) to leading order, & noting that the spin wave creation operator $b_{\mathbf{k}}^\dagger$ lowers the spin with $\hat{\mathbf{S}}^- = \hat{S}_x - i\hat{S}_y$ with phase $e^{-i\mathbf{k}\cdot\mathbf{R}_i}$ and amplitude $\sim 1/S$ at each site i.

[16] Check for simple cubic!

[17] For every spontaneously broken continuous symmetry of the Hamiltonian there is a $\omega_{\mathbf{k}\to 0} = 0$ mode.

3.3.3 Dynamical Susceptibility

Experimental interlude

The simple spin wave calculations described above (and below) are uncontrolled for spin-1/2 systems, and it would be nice to know to what extent one can trust them. In recent years, numerical work (exact diagonalization and quantum Monte Carlo) techniques have shown, as noted, that linear spin wave calculations compare suprisingly well with such "exact" results for the Heisenberg model. But we still need to know if there are any physical systems whose effective spin Hamiltonian can really be described by Heisenberg models. In addition, keep in mind that the utility of spin wave theory was recognized long before such numerical calculations were available, mainly through comparison with experiments on simple magnets. The most useful probe of magnetic structure is slow neutron scattering, a technique developed in the 40's by Brockhouse and Schull (Nobel prize 1994). This section is a brief discussion of how one can use neutron scattering techniques to determine the dispersion and lifetimes of magnetic excitations in solids.[18]

Neutrons scatter from solids primarily due to the nuclear strong force, which leads to nonmagnetic neutron-ion scattering and allows structural determinations very similar to x-ray diffraction analysis. In addition, slow neutrons traversing a crystal can emit or absorb phonons, so the inelastic neutron cross-section is also a sensitive measure of the dispersion of the collective modes of the ionic system.[19] There is also a force on the neutron due to

[18]A complete discussion is found in Lovesey, *Theory of Neutron Scattering from Condensed Matter, Oxford 1984, V. 2*

[19]cf. Ashcroft & Mermin ch. 24

the interaction of its small magnetic dipole moment with the spin magnetic moment of the electrons in the solid. There are therefore additional contributions to the peaks in the elastic neutron scattering intensity (at the Bragg angles) corresponding to magnetic scattering if the solid has long-range magnetic order; they can be distinguished from the nonmagnetic scattering because the additional spectral weight is strongly temperature dependent and disappears above the critical temperature, or through application of an external magnetic field. Furthermore, in analogy to the phonon case, inelastic neutron scattering experiments on ferromagnets yield peaks corresponding to processes where the neutron absorbs or emits a spin wave excitation. Thus the dispersion relation for the magnons can be mapped out.[20]

I will not go through the derivation of the inelastic scattering cross section, which can be found in Lovesey's book. It is similar to the elementary derivation given by Ashcroft & Mermin in Appendix N for phonons. The result is

$$\begin{aligned}\left(\frac{d^2\sigma}{d\Omega d\omega}\right)_{inel} &= a_0^2\frac{k'}{k}\left\{\frac{g}{2}F(\mathbf{q})\right\}^2 e^{-2W(\mathbf{q})}\left(1+b(\omega)\right) \\ &\times\frac{-N}{\pi(g\mu_B)^2}\sum_{\alpha\beta}(\delta_{\alpha\beta}-\hat{q}_\alpha\hat{q}_\beta)\ \mathrm{Im}\ \chi_{\alpha\beta}(\mathbf{q},-\omega)\end{aligned} \tag{44}$$

where a_0 is the Bohr radius, $\mathbf{k}$ and $\mathbf{k}'$ are initial and final wave vector, $\mathbf{q}=\mathbf{k}-\mathbf{k}'$, $F(\mathbf{q})$ atomic form factor, $e^{-2W(\mathbf{q})}$ the Debye-Waller factor, and $b(\omega)$ the Bose distribution function, N the number of unit cells, and ω is the energy change of the neuton,

[20] Even in systems without long range magnetic order, neutron experiments provide important information on the correlation length and lifetime of *spin fluctuations*. In strongly correlated systems (e.g. magnets just above their critical temperature, or itinerant magnets close to a magnetic transition as in Sec. xxx), these can be thought of as collective modes with finite lifetime and decay length. This means the correlation function $\langle\hat{\mathbf{S}}_i^\alpha(t)\hat{\mathbf{S}}_j^\alpha\rangle$ is not const. as $t, |\mathbf{R}_i-\mathbf{R}_j|\rightarrow\infty$, but may fall off very slowly, as "power laws" $t^{-\beta}$, $|\mathbf{R}_i-\mathbf{R}_j|^{-\gamma}$.

$k^2/(2m)-k'^2/(2m)$. The physics we are interested in is contained in the imaginary part of the dynamic susceptibility $\chi(\mathbf{q},\omega)$. For $\omega < 0$, this measures the energy loss by neutrons as they slow down while emitting spin waves in the solid; for $\omega > 0$ the neutrons are picking up energy from thermally excited spin waves.

By spin rotational invariance, all components $\chi_{zz} = \chi_{xx} = \chi_{yy}$ are equal. It is then most convenient to calculate—within linear spin wave theory—the transverse spin susceptibility

$$\chi^R_{+-}(\mathbf{R}_i - \mathbf{R}_j, t) \equiv -\mathrm{Tr}\left(\hat{\rho}[\hat{\mathbf{S}}^+_i(t), \hat{\mathbf{S}}^-_j]\right)\theta(t) \tag{45}$$

and then its Fourier transform wrt momentum $\mathbf{q}$ and frequency ω. I won't do this calculation explicitly, but leave it as an exercise. Note it follows exactly the method used to calculate the charge susceptibility on p. 22 of the previous section. You express the S operators in terms of the b_k's, whose time dependence is exactly known since the approximate Hamiltonian is quadratic. At the end, after Fourier transforming, one recovers

$$\chi_{+-}(\mathbf{q},\omega) = (\frac{2S}{N})\left(\frac{1}{\omega + \omega_{\mathbf{q}} + i0^+}\right). \tag{46}$$

Again, as for the Fermi gas, we see that the collective modes of the system are reflected as the poles in the appropriate response function.

The final cross section is now proportional to

$$\mathrm{Im}\ \chi_{+-}(\mathbf{q},-\omega) \sim \delta(\omega + \omega_{\mathbf{q}}), \tag{47}$$

i.e. there is a peak when a magnon is *absorbed* (neutron's energy $k'^2/(2m)$ is larger than initial energy $k^2/(2m) \Rightarrow \omega \equiv k^2/(2m) - k'^2/(2m) < 0.$). There is another similar contribution proportional to $\delta(\omega - \omega_{\mathbf{q}})$ (emission) coming from χ_{-+}. Thus the dispersion $\omega_{\mathbf{q}}$ can be mapped out by careful measurement.

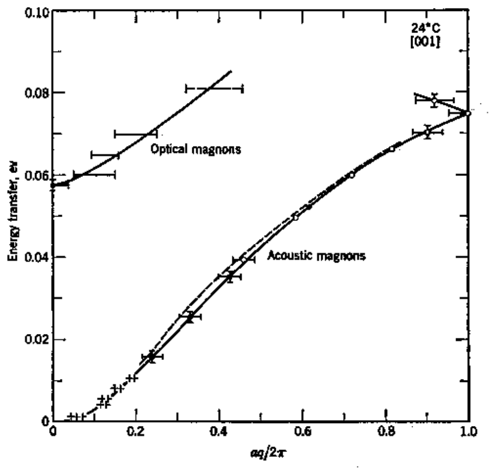


Figure 5: Neutron scattering data on ferromagnet. I searched a bit but couldn't come up with any more modern data than this. This is Fig. 1, Ch. 4 of Kittel, Magnon dispersions in magnetite from inelastic neutron scattering by Brockhouse (Nobel Prize 1994) and Watanabe.

The one-magnon lines are of course broadened by magnon-magnon interactions, and by finite temperatures (Debye-Waller factor). There are also multimagnon absorption and emission processes which contribute in higher order.

3.4 Quantum antiferromagnet

Antiferromagnetic systems are generally approached by analogy with ferromagnetic systems, assuming that the system can be divided up into two or more *sublattices*, i.e. infinite interpenetrating subsets of the lattice whose union is the entire lattice. Classically, it is frequently clear that if we choose all spins on a given judiciously chosen sublattice to be aligned with one another, we must achieve a minimum in the energy. For example, for the classical

AF Heisenberg model $H = J\sum_{i\delta}\mathbf{S}_i \cdot \mathbf{S}_{i+\delta}$ with $J > 0$ on a square lattice, choosing the A-B sublattices in the figure and making all spins up on one and down on another allows each bond to achieve its lowest energy of $-JS^2$. This state, with alternating up and down spins, is referred to as the classical Neél state. Similarly, it may be clear to you that on the triangular lattice the classical lowest energy configuration is achieved when spins are placed at 120° with respect to one another on the sublattices A, B, C. However,

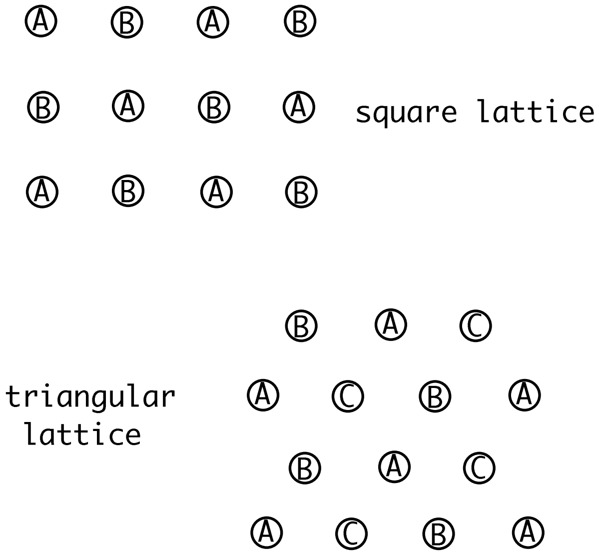


Figure 6: Possible choice of sublattices for antiferromagnet

quantum magnetic systems are not quite so simple. Consider the magnetization M_A on a given sublattice (say the A sites in the figure) of the square lattice; alternatively one can define the *staggered magnetization* as $M_s = \sum_i (-1)^i \langle \hat{\mathbf{S}}_i \rangle$ (Note $(-1)^i$ means $+1$ on the A sites and -1 on the B sites.) Either construct can be used as the order parameter for an antiferromagnet on a bipartite lattice. In the classical Neél state, these is simply $M_A = NS/2$ and $M_s = NS$, respectively, i.e. the sublattice or staggered magnetization are saturated. In the wave function for the ground state of a *quantum* Heisenberg antiferromagnet, however, there is some amplitude for spins to be flipped on a a given sublattice, due to the fact that for a given bond the system can lower its energy by taking advantage of the $\hat{S}_x\hat{S}'_x + \hat{S}_y\hat{S}'_y$ terms. This effect can be seen already by examining the two-spin 1/2 case for the ferromagnet and antiferromagnet. For the ferromagnet, the classical energy is $-|J|S^2 = -|J|/4$, but the quantum energy in the total spin 1 state is also $-|J|/4$. For the antiferromagnet, the classical energy is $-JS^2 = -J/4$, but the energy in the total spin 0 quantum mechanical state is $-3J/4$. So quantum fluctuations–which inevitably depress the magnetization on a given sublattice–lower the energy in the antiferromagnetic case. This can be illustrated in a very simple calculation of magnons in the antiferromagnet following our previous discussion in section 7.2.

3.4.1 Antiferromagnetic magnons

We will follow the same procedure for the ferromagnet on each sublattice A and B, defining

$$\hat{S}_i^{A+} = \hat{S}_{ix}^A + i\hat{S}_{iy}^A = (2S)^{1/2}\left(1-\frac{A_i^\dagger A_i}{2S}\right)^{1/2} A_i \tag{48}$$

$$\hat{S}_i^{A-} = \hat{S}_{ix}^A - i\hat{S}_{iy}^A = (2S)^{1/2}A_i^\dagger\left(1-\frac{A_i^\dagger A_i}{2S}\right)^{1/2} \tag{49}$$

$$\hat{S}_i^{B+} = \hat{S}_{ix}^B + i\hat{S}_{iy}^B = (2S)^{1/2}\left(1-\frac{B_i^\dagger B_i}{2S}\right)^{1/2} B_i \tag{50}$$

$$\hat{S}_i^{B-} = \hat{S}_{ix}^B - i\hat{S}_{iy}^B = (2S)^{1/2}B_i^\dagger\left(1-\frac{B_i^\dagger B_i}{2S}\right)^{1/2} \tag{51}$$

$$
\begin{aligned}
S_{iz}^A &= S - A_i^\dagger A_i && (52)\\
-S_{iz}^B &= S - B_i^\dagger B_i, && (53)
\end{aligned}
$$

i.e. we assume that in the absence of any quantum fluctuations spins on sublattice A are up and those on B are down. Otherwise the formalism on each sublattice is identical to what we did for the ferromagnet. We introduce on each sublattice creation & annihilation operators for spin waves with momentum $\mathbf{k}$:

$$
\begin{aligned}
a_{\mathbf{k}} &= \frac{1}{N^{1/2}} \sum_{i \in A} A_i e^{i\mathbf{k}\cdot\mathbf{R}_i}; \qquad a_{\mathbf{k}}^\dagger = \frac{1}{N^{1/2}} \sum_{i \in A} A_i^\dagger e^{-i\mathbf{k}\cdot\mathbf{R}_i} && (54)\\
b_{\mathbf{k}} &= \frac{1}{N^{1/2}} \sum_{i \in B} B_i e^{i\mathbf{k}\cdot\mathbf{R}_i}; \qquad b_{\mathbf{k}}^\dagger = \frac{1}{N^{1/2}} \sum_{i \in B} B_i^\dagger e^{-i\mathbf{k}\cdot\mathbf{R}_i}. && (55)
\end{aligned}
$$

In principle $\mathbf{k}$ takes values only in the 1st *magnetic Brillouin zone*, or *half-zone*, since the periodicity of the sublattices is twice that of the underlying lattice. The spin operators on a given site are then expanded as

$$
\begin{aligned}
\hat{S}_i^{A+} &\simeq \left(\frac{2S}{N}\right)^{1/2} \left[\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_i} a_{\mathbf{k}} + \ldots\right], && (56)\\
\hat{S}_i^{B+} &\simeq \left(\frac{2S}{N}\right)^{1/2} \left[\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_i} b_{\mathbf{k}} + \ldots\right], && (57)\\
\hat{S}_i^{A-} &\simeq \left(\frac{2S}{N}\right)^{1/2} \left[\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_i} a_{\mathbf{k}}^\dagger + \ldots\right], && (58)\\
\hat{S}_i^{B-} &\simeq \left(\frac{2S}{N}\right)^{1/2} \left[\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_i} b_{\mathbf{k}}^\dagger + \ldots\right], && (59)\\
\hat{S}_{iz}^A &= S - \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i} a_{\mathbf{k}}^\dagger a_{\mathbf{k}'} && (60)\\
\hat{S}_{iz}^B &= -S + \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i} a_{\mathbf{k}}^\dagger a_{\mathbf{k}'}. && (61)
\end{aligned}
$$

The expansion of the Heisenberg Hamiltonian in terms of these variables is now (compare (40))

$$
\begin{aligned}
H &= -NzJS^2 + H_0^{magnon} + \mathcal{O}(1), & (62)\\
H_0^{magnon} &= JzS \sum_{\mathbf{k}} \left[\gamma_{\mathbf{k}}(A_{\mathbf{k}}^\dagger B_{\mathbf{k}}^\dagger + A_{\mathbf{k}} B_{\mathbf{k}}) + (A_{\mathbf{k}}^\dagger A_{\mathbf{k}} + B_{\mathbf{k}}^\dagger B_{\mathbf{k}})\right] & (63)
\end{aligned}
$$

Unlike the ferromagnetic case, merely expressing the Hamiltonian to bilinear order in the magnon variables does not diagonalize it immediately. We can however perform a canonical transformation[21] to bring the Hamiltonian into diagonal form (check!):

$$
\begin{aligned}
\alpha_{\mathbf{k}} &= u_{\mathbf{k}} A_{\mathbf{k}} - v_{\mathbf{k}} B_{\mathbf{k}}^\dagger \ ; \quad \alpha_{\mathbf{k}}^\dagger = u_{\mathbf{k}} A_{\mathbf{k}}^\dagger - v_{\mathbf{k}} B_{\mathbf{k}} & (64)\\
\beta_{\mathbf{k}} &= u_{\mathbf{k}} B_{\mathbf{k}} - v_{\mathbf{k}} A_{\mathbf{k}}^\dagger \ ; \quad \beta_{\mathbf{k}}^\dagger = u_{\mathbf{k}} B_{\mathbf{k}}^\dagger - v_{\mathbf{k}} A_{\mathbf{k}}, & (65)
\end{aligned}
$$

where the coefficients $u_{\mathbf{k}}, v_{\mathbf{k}}$ must be chosen such that $u_{\mathbf{k}}^2 - v_{\mathbf{k}}^2 = 1$. One such choice is $u_{\mathbf{k}} = \cosh\theta_{\mathbf{k}}$ and $v_{\mathbf{k}} = \sinh\theta_{\mathbf{k}}$. For each $\mathbf{k}$, choose the angle $\theta_{\mathbf{k}}$ such that the anomalous terms like $A_{\mathbf{k}}^\dagger B_{\mathbf{k}}^\dagger$ vanish. One then finds the solution

$$\tanh 2\theta_{\mathbf{k}} = -\gamma_{\mathbf{k}}, \tag{66}$$

and

$$H_0^{magnon} = -NE_J + \sum_{\mathbf{k}} \omega_{\mathbf{k}}(\alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + \beta_{\mathbf{k}}^\dagger \beta_{\mathbf{k}} + 1), \tag{67}$$

where

$$\omega_{\mathbf{k}}^2 = E_J^2(1 - \gamma_{\mathbf{k}}^2), \tag{68}$$

and $E_J = JzS$. Whereas in the ferromagnetic case we had $\omega_{\mathbf{k}} \sim (1 - \gamma_{\mathbf{k}}) \sim k^2$, it is noteworthy that in the antiferromagnetic case

[21] The definition of a canonical transformation, I remind you, is one which will give canonical commutation relations for the transformed fields. This is important because it ensures that we can interpret the Hamiltonian represented in terms of the new fields as a (now diagonal) fermion Hamiltonian, read off the energies, etc.

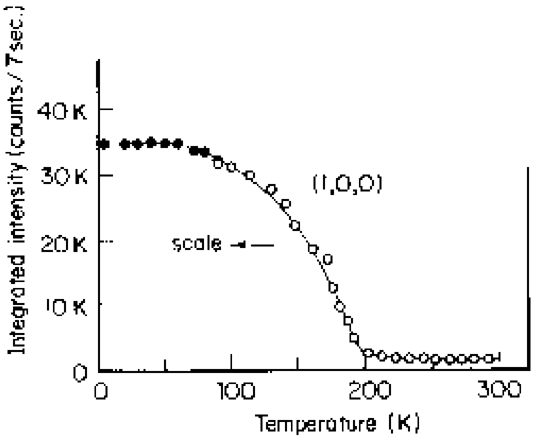


Figure 7: Integrated intensity of (100) Bragg peak vs. temperature for $LaCuO_4$, with $T_N = 195K$. (After Shirane et al. 1987)

the result $\omega_{\mathbf{k}} \sim (1-\gamma_{\mathbf{k}})^{1/2} \sim k$ gives a linear magnon dispersion at long wavelengths (a center of symmetry of the crystal must be assumed to draw this conclusion). Note further that for each $\mathbf{k}$ there are two degenerate modes in H_0^{magnon}.

3.4.2 Quantum fluctuations in the ground state

At $T = 0$, there are no thermally excited spin wave excitations. Nevertheless the spin wave theory yields an decrease of the ground state energy relative to the classical value $-NJzS^2$, but an *increase* over the quantum ferromagnet result of $-N|J|S(S+1)$ due to the zero-point (constant) term in (67).[22] The ground-state energy is

$$E_0 \simeq -NE_J + \sum_{\mathbf{k}} \omega_{\mathbf{k}}. \tag{69}$$

[22]Recall the classical Neél state, which does not contain such fluctuations, is *not* an eigenstate of the quantum Heisenberg Hamiltonian.

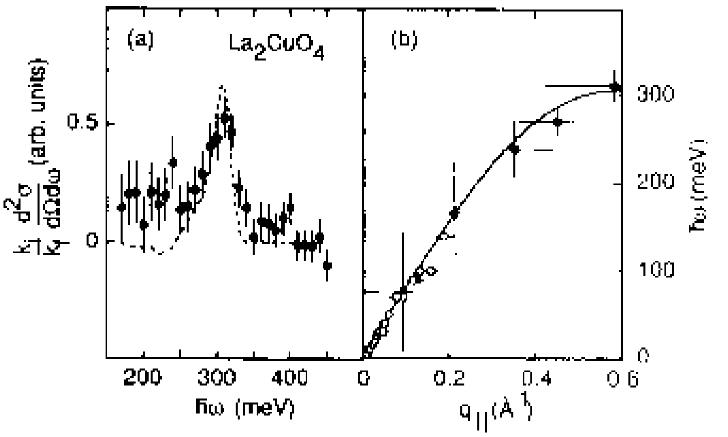


Figure 8: a) Inelastic neutron scattering intensity vs. energy transfer at 296K near zone boundary $Q = (1, k, 0.5)$ for oriented $LaCuO_4$ crystals with Neél temperature 260K; b) spin-wave dispersion $\omega_{\mathbf{q}}$ vs. $q_{\|}$, where $q_{\|}$ is in-plane wavevector. (after Hayden et al. 1991)

The result is usually expressed in terms of a constant β defined by

$$E_0 \equiv -NJzS\left(S + \frac{\beta}{z}\right), \tag{70}$$

and $\beta = 0.58$ for 3D and xxx for 2D.

Quantum fluctuations have the further effect of preventing the staggered magnetization from achieving its full saturated value of S, as in the classical Neél state, as shown first by Anderson.[23] Let us consider the average z-component of spin in equilibrium at temperature T, averaging only over spins on sublattice A of a D-dimensional hypercubic lattice. From (60), we have $\langle \hat{S}_z^A \rangle = S - N^{-1}\sum_{\mathbf{k}} \langle A_{\mathbf{k}}^\dagger A_{\mathbf{k}} \rangle$ within linear spin wave theory. Inverting the transformation (65), we can express the A's in terms of the α's and β's, whose averages we can easily calculate. Note the 0th order in $1/S$ gives the classical result, $\langle \hat{\mathbf{S}}_z^A \rangle$, and the deviation is

[23] P.W. Anderson, Phys. Rev. 86, 694 (1952).

the spin wave reduction of the sublattice moment

$$\begin{aligned}
\frac{\delta M_A}{N} &= \langle \hat{S}_z^A \rangle - S = -\frac{1}{N} \sum_k \langle A_{\mathbf{k}}^\dagger A_{\mathbf{k}} \rangle \\
&= -\frac{1}{N} \sum_{\mathbf{k}} \langle (u_{\mathbf{k}} \alpha_{\mathbf{k}}^\dagger + v_{\mathbf{k}} \beta_{\mathbf{k}})(u_{\mathbf{k}} \alpha_{\mathbf{k}} + v_{\mathbf{k}} \beta_{\mathbf{k}}^\dagger) \rangle \\
&= -\frac{1}{N} \sum_{\mathbf{k}} u_{\mathbf{k}}^2 \langle \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} \rangle + v_{\mathbf{k}}^2 \langle \beta_{\mathbf{k}}^\dagger \beta_{\mathbf{k}} \rangle, \qquad (71)
\end{aligned}$$

where N is the number of spins on sublattice B. We have neglected cross terms like $\langle \alpha_{\mathbf{k}}^\dagger \beta_{\mathbf{k}}^\dagger \rangle$ because the α and β are independent quanta by construction. However the diagonal averages $\langle \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} \rangle$ and $\langle \beta_{\mathbf{k}}^\dagger \beta_{\mathbf{k}} \rangle$ are the expectation values for the number operators of independent bosons with dispersion $\omega_{\mathbf{k}}$ in equiblibrium, and so can be replaced (within linear spin wave theory) by

$$n_{\mathbf{k}} \equiv \langle \beta_{\mathbf{k}}^\dagger \beta_{\mathbf{k}} \rangle = \langle \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} \rangle = b(\omega_{\mathbf{k}}), \qquad (72)$$

where b is the Bose distribution function. The tranformation coefficients $u_{\mathbf{k}} = \cosh \theta_{\mathbf{k}}$ and $v_{\mathbf{k}} = \sinh \theta_{\mathbf{k}}$ are determined by the condition (66) such that

$$\begin{aligned}
u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2 &= \cosh 2\theta_{\mathbf{k}} = \frac{1}{\sqrt{1-\gamma_{\mathbf{k}}^2}} \qquad (73) \\
v_{\mathbf{k}}^2 &= \frac{1}{2} \left(\frac{1}{\sqrt{1-\gamma_{\mathbf{k}}^2}} - 1 \right), \qquad (74)
\end{aligned}$$

such that the sulattice magnetization (71) becomes

$$\boxed{\frac{\delta M_A}{N} = \frac{1}{2} - \frac{1}{N} \sum_{\mathbf{k}} \left(n_{\mathbf{k}} + \frac{1}{2} \right) \frac{1}{\sqrt{1-\gamma_{\mathbf{k}}^2}}} \qquad (75)$$

<u>Remarks:</u>

1. The correction δM_A is independent of S, and negative as it must be (see next point). However relative to the leading classical term S it becomes smaller and smaller as S increases, as expected.

2. The integral in (75) depends on the dimensionality of the system. It has a T-dependent part coming from $n_{\mathbf{k}}$ and a T-independent part coming from the 1/2. At $T = 0$, where there are no spin waves excited thermally, $n_{\mathbf{k}} = 0$, and we find
$$\frac{\delta M_A}{N} \simeq \begin{cases} -0.078 & D = 3 \\ -0.196 & D = 2 \\ \infty & D = 1 \end{cases} \tag{76}$$
The divergence in $D = 1$ indicates the failure of spin-wave theory in one dimension, as discussed further in section xxx.

3. The low temperature behavior of $\delta M(T)$ must be calculated carefully due to the singularities of the bose distribution function when $\omega_k \to 0$. If this divergence is cut off by introducing a scale k_0 near $\mathbf{k} = 0$ and $\mathbf{k} = (\pi/a, \pi/a)$, one finds that δM_A diverges as $1/k_0$ in 1D, and as $\log k_0$ in 2D, whereas it is finite as $\mathbf{k}_0 \to 0$ in 3D. Thus on this basis one does not expect long range spin order at any nonzero temperature in two dimensions (see discussion of Mermin-Wagner theorem below), nor even at $T = 0$ in one dimension.

3.4.3 Nonlinear spin wave theory

3.4.4 Frustrated models

3.5 1D & 2D Heisenberg magnets

3.5.1 Mermin-Wagner theorem

3.5.2 1D: Bethe solution

3.5.3 2D: Brief summary

3.6 Itinerant magnetism

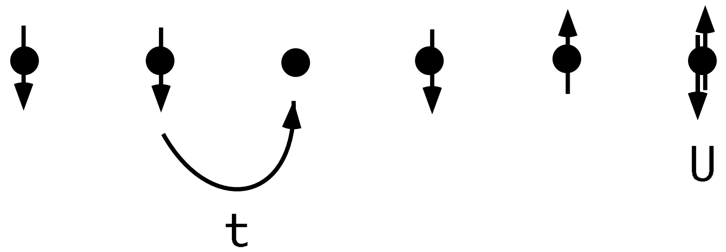
"Itinerant magnetism" is a catch-all phrase which refers to magnetic effects in metallic systems (i.e., with conduction electrons). Most of the above discussion assumes that the spins which interact with each other are *localized*, and there are no mobile electrons. However we may study systems in which the electrons which magnetize, or nearly magnetize are mobile, and situations in which localized impurity spins interact with conduction electron spins in a host metal. The last problem turns out to be a very difficult many-body problem which stimulated Wilson to the develop of renormalization group ideas.

3.7 Stoner model for magnetism in metals

The first question is, can we have ferromagnetism in metallic systems with only one relevant band of electrons. The answer is yes, although the magnetization/electron in the ferromagnetic state is typically reduced drastically with respect to insulating ferromagnets. The simplest model which apparently describes this kind of transition (there is no exact solution in $D > 1$) is the Hubbard model we have already encountered. A great deal of attention has been focussed on the Hubbard model and its variants, particularly because it is the simplest model known to display a metal-insulator (Mott-Hubbard) transition qualitatively similar to what is observed in the high temperature superconductors. To review, the Hubbard model consists of a lattice of sites labelled by i, on which electrons of spin $\uparrow$ or $\downarrow$ may sit. The kinetic energy term in the Hamiltonian allows for electrons to hop between sites with matrix element t, and the potential energy simply requires an energy cost U every time two opposing spins occupy the same site.[24] Longer range interactions are neglected:

$$H = -t \sum_{\substack{\sigma \\ <i,j>}} c_{i\sigma}^{\dagger} c_{j\sigma} + \frac{1}{2} U \sum_{\sigma} n_{i\sigma} n_{i-\sigma}, \tag{77}$$

where $< ij >$ means nearest neighbors only.



In its current form the kinetic energy, when Fourier transformed, corresponds, to a tight binding band in d dimensions of width $4dt$,

$$\epsilon_k = -2t \sum_{\alpha=1}^{d} \cos k_\alpha a, \tag{78}$$

[24]What happened to the long-range part of the Coulomb interaction? Now that we know it is screened, we can hope to describe its effects by including only its matrix elements between electrons in wave functions localized on site i and site j, with $|i - j|$ smaller than a screening length. The largest element is normally the $i = j$ one, so it is frequently retained by itself. Note the Hamiltonian (77) includes only an on-site interaction for opposite spins. Of course like spins are forbidden to occupy the same site anyway by the Pauli principle.

where a is the lattice constant. The physics of the model is as follows. Imagine first that there is one electron per site, i.e. the band is half-filled. If $U = 0$ the system is clearly metallic, but if $U \to \infty$, double occupation of sites will be "frozen out". Since there are no holes, electrons cannot move, so the model must correspond to an insulating state; at some critical U a metal-insulator transition must take place. We are more interested in the case away from half-filling, where the Hubbard model is thought for large U and small doping (deviation of density from 1 particle/site) to have a ferromagnetic ground state.[25] In particular, we would like to investigate the transition from a *paragmagnetic* to a *ferromagnetic* state as T is lowered.

This instability must show up in some quantity we can calculate. In a ferromagnet the susceptibility χ diverges at the transition, i.e. the magnetization produced by the application of an infinitesimal external field is suddenly finite. In many-body language, the static, uniform spin susceptibility is the retarded spin density – spin density correlation function, (for the discussion below I take $g\mu_B = 1$)

$$\chi = \chi(q=0,\omega=0) = \lim_{h_z \to 0} \frac{\langle S_z \rangle}{h_z} = \int d^3r \int_0^\infty dt \langle [S_z(r,t), S_z(0,0)] \rangle, \tag{79}$$

where in terms of electron number operators $n_\sigma = \psi_\sigma^\dagger \psi_\sigma$, the magnetization operators are $S_z = (1/2)[n_\uparrow - n_\downarrow]$, i.e. they just measure the surplus of up over down spins at some point in space.

[25] At 1/2-filling, one electron per site, a great deal is known about the Hubbard model, in particular that the system is metallic for small U (at least for non-nested lattices, otherwise a narrow-gap spin density wave instability is present), but that as we increase U past a critical value $U_c \sim D$ a transition to an antiferromagnetic insulating state occurs (Brinkman-Rice transition). With one single hole present, as $U \to \infty$, the system is however ferromagnetic (Nagaoka state).

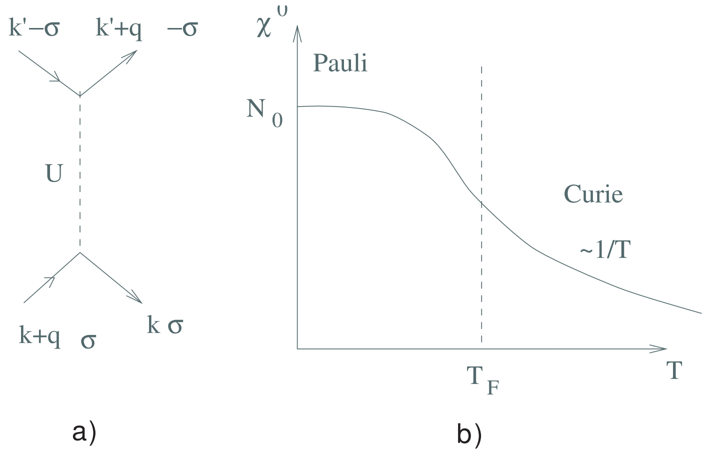


Figure 9: 1a) Hubbard interaction; 1b) Spin susceptibility vs. T for free fermions.

Diagramatically, the Hubbard interaction $H_{int} = U\sum_i n_{i\uparrow}n_{i\downarrow}$ looks like figure 9a); note only electrons of opposite spins interact. The magnetic susceptibility is a correlation function similar to the charge susceptibility we have already calculated. At time t=0, we measure the magnetization S_z of the system, allow the particle and hole thus created to propagate to a later time, scatter in all possible ways, and remeasure S_z. The so-called "Stoner model" of ferromagnetism approximates the perturbation series by the RPA form we have already encountered in our discussion of screening,[26] which gives $\chi = \chi_0/(1-U\chi_0)$. At sufficiently high T (χ_0 varies as $1/T$ in the nondegenerate regime, Fig. 9c)) we will have $U\chi_0(T) < 1$, but as T is lowered, $U\chi_0(T)$ increases. If U is large enough, such that $UN_0 > 1$, there will be a transition at $U\chi_0(T_c) = 1$, where χ diverges. Note for $U\chi_0(T) > 1$ ($T < T_c$), the susceptibility is negative, so the model is unphysical in this region. This problem arises in part because the ferromagnetic state has a spontaneously broken symmetry, $\langle S_z\rangle \neq 0$ even for

[26]In the static, homogeneous case it is equivalent to the self-consistent field (SCF) method of Weiss.

$h_z \to 0$. Nevertheless the approach to the transition from above and the location of the transition appear qualitatively reasonable.

It is also interesting to note that for $U\chi(0) = UN_0 < 1$, there will be no transition, but the magnetic susceptibility will be enhanced at low temperatures. So-called "nearly ferromagnetic" metals like Pd are qualitatively described by this theory. Comparing the RPA form

$$\chi = \frac{\chi_0(T)}{1 - U\chi_0(T)} \tag{80}$$

to the free gas susceptibility in Figure 9b, we see that the system will look a bit like a free gas with enhanced density of states, and reduced degeneracy temperature T_F^*.[27] For Pd, the Fermi temperature calculated just by counting electrons works out to 1200K, but the susceptibility is found experimentally to be $\sim 10\times$ larger than N_0 calculated from band structure, and the susceptibility is already Curie-like around $T \sim T_F^* \simeq$300K.

3.7.1 Moment formation in itinerant systems

We will be interested in asking what happens when we put a localized spin in a metal, but first we should ask how does that local moment form in the first place. If an arbitrary impurity is inserted into a metallic host, it is far from clear that any kind of localized moment will result: a donor electron could take its spin and wander off to join the conduction sea, for example. Fe impurities are known to give rise to a Curie term in the susceptibility when embedded in Cu, for example, but not in Al, suggesting

[27]Compare to the Fermi liquid form

$$\chi = \frac{m^*}{m}\frac{\chi_0}{1 + F_0^a}. \tag{81}$$

that a moment simply does not form in the latter case. Anderson[28] showed under what circumstances an impurity level in an interacting host might give rise to a moment. He considered a model with a band of electrons[29] with energy $\epsilon_{\mathbf{k}}$, with an extra dispersionless impurity level E_0. Suppose there are strong local Coulomb interactions on the impurity site, so that we need to add a Hubbard-type repulsion. And finally suppose the conduction (d) electrons can hop on and off the impurity with some matrix element V. The model then looks like

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + E_0 \sum_{\sigma} n_{0\sigma} + V \sum_{\mathbf{k}\sigma} (c^\dagger_{\mathbf{k}\sigma} c_0 + c^\dagger_0 c_{\mathbf{k}\sigma}) + \frac{1}{2} U \sum_{\sigma} n_{0\sigma} n_{0-\sigma}, \tag{82}$$

where $n_{0\sigma} = c^\dagger_{0\sigma} c_{0\sigma}$ is the number of electrons of spin σ on the impurity site 0. By the Fermi Golden rule the decay rate due to scattering from the impurity of a band state ϵ away from the Fermi level E_F *in the absence of the interaction* U is of order

$$\Delta(\epsilon) = \pi V^2 \sum_k \delta(\epsilon - \epsilon_{\mathbf{k}}) \simeq \pi V^2 N_0 \tag{83}$$

In the "Kondo" case shown in the figure, where E_0 is well below the Fermi level, the scattering processes take place with electrons at the Fermi level $\epsilon = 0$, so the bare width of the impurity state is also $\Delta \simeq \pi V^2 N_0$. So far we still do not have a magnetic moment, since, in the absence of the interaction U, there would be an occupation of 2 antiparallel electrons. If one could effectively prohibit double occupancy, however, i.e. if $U \gg \Delta$, a single spin would remain in the localized with a net moment. Anderson

[28]PW Anderson, Phys. Rev. 124, 41 (1961)

[29]The interesting situation for moment formation is when the bandwidth of the "primary" cond. electron band overlapping the Fermi level is much larger than the bare hybridization width of the impurity state. The two most commonly considered situations are a band of s electrons with d-level impurity (transition metal series) and d-electron band with localized f-level (rare earths/actinides–heavy fermions).

obtained the basic physics (supression of double occupancy) by doing a Hartree-Fock decoupling of the interaction U term. Schrieffer and Wolff in fact showed that in the limit $U \to -\infty$, the virtual charge fluctuations on the impurity site (occasional double occupation) are eliminated, and the only degree of freedom left (In the so-called Kondo regime corresponding to Fig. 10a) is a localized spin interacting with the conduction electrons via an effective Hamiltonian

$$\mathcal{H}_{Kondo} = J\mathbf{S} \cdot \sigma, \tag{84}$$

where J is an antiferromagnetic exchange expressed in terms of the original Anderson model parameters as

$$J = 2\frac{V^2}{E_0}, \tag{85}$$

S is the impurity spin-1/2, and

$$\sigma_i = \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'\alpha\beta} c^{\dagger}_{\mathbf{k}\alpha} \left(\tau_i\right)_{\alpha\beta} c_{\mathbf{k}'\beta}, \tag{86}$$

with τ_i the Pauli matrices, is just the conduction electron spin density at the impurity site.

3.7.2 RKKY Interaction

Kittel p. 360 et seq.

3.7.3 Kondo model

The Hartree-Fock approach to the moment formation problem was able to account for the existence of local moments at defects

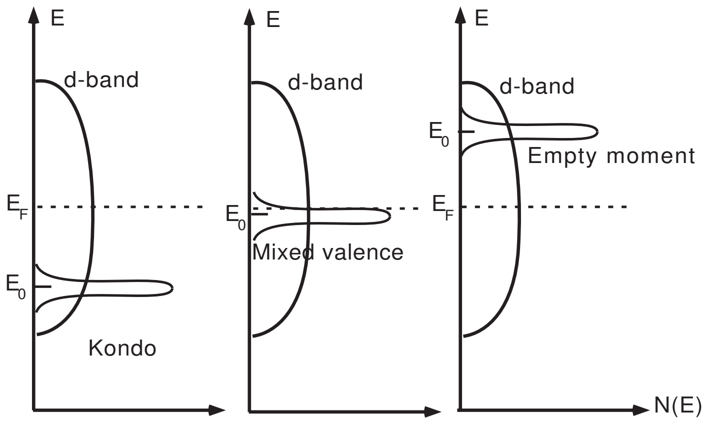


Figure 10: Three different regimes of large U Anderson model depending on position of bare level E_0. In Kondo regime ($E_0 << E_F$), large moments form at high T but are screened at low T. In mixed valent regime, occupancy of impurity level is fractional and moment formation is marginal. For $E_0 > E_F$, level is empty and no moment forms.

in metallic hosts, in particular for large Curie terms in the susceptibility at high temperatures. What it did not predict, however, was that the low-temperature behavior of the model was very strange, and that in fact the moments present at high temperatures *disappear* at low temperatures, i.e. are screened completely by the conduction electrons, one of which effectively binds with the impurity moment to create a local singlet state which acts (at $T = 0$) like a nonmagnetic scatterer. This basic physical picture had been guessed at earlier by J. Kondo,[30] who was trying to explain the existence of a resistance minimum in certain metallic alloys. Normally the resistance of metals is monotonically decreasing as the temperature is lowered, and adding impurities gives rise to a constant offset (Matthiesen's Rule) which does not change the monotonicity. For Fe in Cu, however, the impurity contribu-

[30] J. Kondo, Prog. Theor. Phys. 32, 37 (64).

tion $\delta\rho_{imp}$ *increased* as the temperature was lowered, and eventually saturated. Since Anderson had shown that Fe possessed a moment in this situation, Kondo analyzed the problem perturbatively in terms of a magnetic impurity coupling by exchange to a conduction band, i.e. adopted Eq. (84) as a model. Imagine that the system at $t = -\infty$ consists of a Fermi sea $|0\rangle$ and one additional electron in state $|\mathbf{k}, \sigma\rangle$ at the Fermi level. The impurity spin has initial spin projection M, so we write the initial state as $|i\rangle = c^\dagger_{\mathbf{k}\sigma}|0; M\rangle$ Now turn on an interaction $\mathcal{H}_1$ adiabatically and look at the scattering amplitude[31]

between $|i\rangle$ and $|f\rangle = c^\dagger_{\mathbf{k}'\sigma'}|0; M'\rangle$

$$\langle f|T|i\rangle \;=\; -2\pi i\langle f|\mathcal{H}_1 + \mathcal{H}_1\frac{1}{\epsilon_{\mathbf{k}} - \mathcal{H}_0}\mathcal{H}_1 + \ldots|i\rangle \qquad (90)$$

If $\mathcal{H}_1$ were just ordinary impurity (potential) scattering, we would have $\mathcal{H}_1 = \Sigma_{\mathbf{k}\mathbf{k}'\sigma\sigma'}\, c^\dagger_{\mathbf{k}\sigma} V_{\mathbf{k}\mathbf{k}'} c_{\mathbf{k}'\sigma'}$, and there would be two distinct second-order processes $\mathbf{k} \to \mathbf{k}'$ contributing to Eq. (90), as shown schematically at the top of Figure 11, of type a),

$$\langle 0|c_{\mathbf{k}'}c^\dagger_{\mathbf{k}'}V_{\mathbf{k}'\mathbf{p}}c_{\mathbf{p}}\frac{1}{\epsilon_{\mathbf{k}} - \mathcal{H}_0}c^\dagger_{\mathbf{p}}V_{\mathbf{p}\mathbf{k}}c_{\mathbf{k}}c^\dagger_{\mathbf{k}}|0\rangle = V_{\mathbf{k}'\mathbf{p}}\frac{1 - f(\epsilon_{\mathbf{p}})}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{p}}}V_{\mathbf{p}\mathbf{k}} \qquad (91)$$

---

[31] Reminder: when using Fermi's Golden Rule (see e.g., Messiah, Quantum Mechanics p.807):

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar v}|T|^2\rho(E) \qquad (87)$$

we frequently are able to get away with replacing the full T-matrix by its perhaps more familiar 1st-order expansion

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar v}|\mathcal{H}_1|^2\rho(E) \qquad (88)$$

(Recall the T matrix is defined by $\langle\phi_f|T|\phi_i\rangle = \langle\phi_f|T|\psi_i^+\rangle$, where the ϕ's are plane waves and ψ^+ is the scattering state with outgoing boundary condition.) In this case, however, we will not find the interesting $\log T$ divergence until we go to 2nd order! So we take matrix elements of

$$T = \mathcal{H}_1 + \mathcal{H}_1\frac{1}{\epsilon_{\mathbf{k}} - \mathcal{H}_0}\mathcal{H}_1 + \ldots \qquad (89)$$

This is equivalent and, I hope, clearer than the transition amplitude I calculated in class.

and type b),

$$
\begin{aligned}
\langle 0|c_{\mathbf{k}'}c^\dagger_{\mathbf{p}}V_{\mathbf{pk}}c_{\mathbf{k}}\ \frac{1}{\epsilon_{\mathbf{k}}-\mathcal{H}_0}\ c^\dagger_{\mathbf{k}'}V_{\mathbf{k}'\mathbf{p}}c_{\mathbf{p}}c^\dagger_{\mathbf{k}}|0\rangle & \\
&= -V_{\mathbf{pk}}\frac{f(\epsilon_{\mathbf{p}})}{\epsilon_{\mathbf{k}}-(\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{p}}+\epsilon_{\mathbf{k}'})}V_{\mathbf{k}'\mathbf{p}} \\
&= V_{\mathbf{pk}}\frac{f(\epsilon_{\mathbf{p}})}{\epsilon_{\mathbf{k}'}-\epsilon_{\mathbf{p}}}V_{\mathbf{k}'\mathbf{p}} \qquad (92)
\end{aligned}
$$

where I have assumed $\mathbf{k}$ is initially occupied, and $\mathbf{k}'$ empty, with $\epsilon_{\mathbf{k}}=\epsilon_{\mathbf{k}'}$, whereas $\mathbf{p}$ can be either; the equalities then follow from the usual application of $c^\dagger c$ and $cc^\dagger$ to $|0\rangle$. Careful checking of the order of the c's in the two matrix elements will show you that the first process only takes place if the intermediate state $\mathbf{p}$ is unoccupied, and the second only if it is unoccupied.

Nnow when one sums the two processes, the Fermi function cancels. This means there is no significant T dependence to this order (ρ_{imp} due to impurities is a constant at low T), and thus the exclusion principle does not play an important role in ordinary potential scattering.

Now consider the analagous processes for spin-scattering. The perturbing Hamiltonian is Eq. 84. Let's examine the amplitude for *spin-flip* transitions caused by $\mathcal{H}_1$, first of type a),

$$
\begin{aligned}
&\langle 0M_{s'}|\mathcal{H}_1|0M_s\rangle \qquad (93)\\
&= \frac{J^2}{4}S_\mu S_\nu\langle 0M_{s'}|c_{\mathbf{k}'\sigma'}c^\dagger_{\mathbf{k}'\sigma'}\tau^\nu_{\sigma'\sigma''}c_{\mathbf{p}\sigma''}\frac{1}{\epsilon_{\mathbf{k}}-\mathcal{H}_0}c^\dagger_{\mathbf{p}\sigma''}\tau^\mu_{\sigma''\sigma}c_{\mathbf{k}\sigma}c^\dagger_{\mathbf{k}\sigma}|0M_s\rangle\\
&= \frac{J^2}{4}\frac{1-f(\epsilon_{\mathbf{p}})}{\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{p}}}\langle M'_S|S_\nu S_\mu\tau^\nu_{\sigma'\sigma''}\tau^\mu_{\sigma''\sigma}|M_S\rangle
\end{aligned}
$$

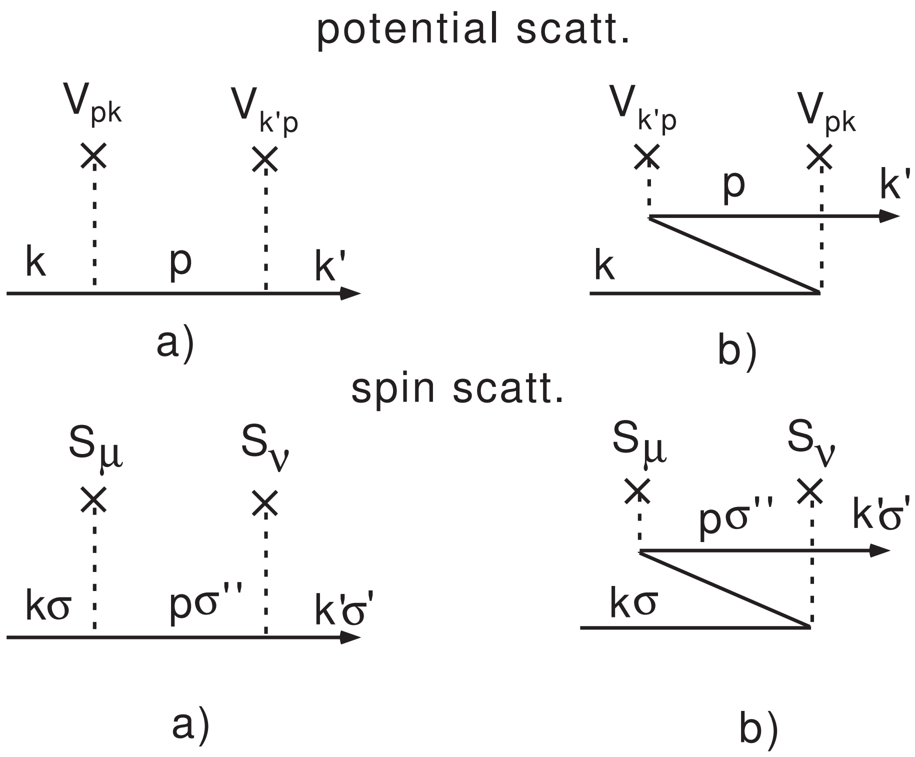


Figure 11: 2nd-order scattering processes: a) direct and b) exchange scattering. Top: potential scattering; bottom: spin scattering. Two terms corresponding to Eq. 90

and then of type b),

$$\frac{J^2}{4}\frac{f(\epsilon_{\mathbf{p}})}{\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{p}}}\langle M_S'|S_\nu S_\mu \tau^\nu_{\sigma''\sigma}\tau^\mu_{\sigma'\sigma''}|M_S\rangle. \tag{94}$$

Now note that $\tau^\mu_{\sigma'\sigma''}\tau^\nu_{\sigma''\sigma} = (\tau^\mu\tau^\nu)_{\sigma'\sigma}$, and use the identity $\tau_\mu\tau_\nu = \delta_{\nu\mu} + i\tau_\alpha\epsilon_{\alpha\mu\nu}$. The $\delta_{\mu\nu}$ pieces clearly will only give contributions proportional to $\mathbf{S}^2$, so they aren't the important ones which will distinguish between a) and b) processes compared to the potential scattering case. The $\epsilon_{\alpha\nu\mu}$ terms give results of differing sign, since a) gives $\epsilon_{\alpha\nu\mu}$ and b) gives $\epsilon_{\alpha\mu\nu}$. Note the basic difference between the 2nd-order potential scattering and spin scattering is that the matrix elements in the spin scattering case, i.e. the S_μ, didn't commute! When we add a) and b) again the result is

$$\frac{J^2}{\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{p}}}\left\{\left[\frac{1}{4}S(S+1) - \langle M_S'\sigma'|\mathbf{S}\cdot\sigma|M_S,\sigma\rangle\right]\right.$$
$$\left. + f(\epsilon_{\mathbf{p}})\langle M_S'\sigma'|\mathbf{S}\cdot\ \sigma|M_S,\sigma\rangle\right\} \tag{95}$$

so the spin-scattering amplitude does depend on T through $f(\epsilon_{\mathbf{p}})$. Summing over the intermediate states $\mathbf{p}\sigma''$ gives a factor

$$\sum_{\mathbf{p}} \frac{f(\epsilon_{\mathbf{p}})}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{p}}} \simeq N_0 \int d\xi_{\mathbf{p}} \frac{f(\xi_{\mathbf{p}})}{\xi_{\mathbf{k}} - \xi_{\mathbf{p}}}$$
$$= N_0 \int d\xi_{\mathbf{p}} \left(-\frac{\partial f}{\partial \xi_{\mathbf{p}}} \right) \log |\xi_{\mathbf{k}} - \xi_{\mathbf{p}}|, \qquad (96)$$

which is of order $\log T$ for states $\xi_{\mathbf{k}}$ at the Fermi surface! Thus the spin part of the interaction, to a first approximation, makes a contribution to the resistance of order $J^3 \log T$ (ρ involves the square of the scattering amplitude, and the cross terms between the 1st and 2nd-order terms in perturbation theory give this result). Kondo pointed to this result and said, "Aha!", here is a contribution which gets big as T gets small. However this can't be the final answer. The divergence signals that perturbation theory is breaking down, so this is one of those very singular problems where we have to find a way to sum all the processes. We have discovered that the Kondo problem, despite the fact that only a single impurity is involved, is a complicated many-body problem. Why? Because the spin induces correlations between the electrons in the Fermi sea. Example: suppose two electrons, both with spin up, try to spin flip scatter from a spin-down impurity. The first electron can exchange its spin with the impurity and leave it spin up. The second electron therefore cannot spin-flip scatter by spin conservation. Thus the electrons of the conduction band can't be treated as independent objects.

Summing an infinite subset of the processes depicted in Figure 11 or a variety of other techniques give a picture of a singlet bound state where the impurity spin binds one electron from the conduction electron sea with binding energy

$$T_K = De^{-1/(JN_0)}, \tag{97}$$

where D is the conduction electron bandwidth. The renormalization group picture developed by Wilson in the 70s and the exact Bethe ansatz solution of Wiegman/Tsvelick and Andrei/Lowenstein in 198? give a picture of a free spin at high temperatures, which hybridizes more and more strongly with the conduction electron sea as the temperature is lowered. Below the singlet formation temperature of T_K, the moment is screened and the impurity acts like a potential scatterer with large phase shift, which approaches $\pi/2$ at $T = 0$.