Lecture 11: Long-wavelength expansion in the Neel state—Energetic terms

In the last class we derived the low energy effective Hamiltonian for a Mott insulator. This derivation is an example of the kind of analysis that is needed when the single-particle picture breaks down. As we saw, keeping track of many particles is very awkward using our elementary quantum mechanics notation and formalism, and we need a better way of accounting for the presence of many particles and their statistics under exchange—for instance, it would be nice if there was some systematic formalism that would keep track of minus signs in the amplitudes of various processes due to fermion interchange and allow us to complete this kind of calculation without the need to invoke such minus signs by hand at various steps.

In fact, the formalism of "second-quantization" is designed to do precisely this. It provides an indispensible language and calculational tool for many-particle systems, and in the next module, we will provide an introduction to this tool. Before we do that however, we will spend this lecture and the next couple of lectures on developing the low energy description of the antiferromagnetic Mott insulator, starting from the effective Hamiltonian derived in the previous class.

The first step is to start with nearest neighbour Heisenberg antiferromagnet Hamiltonian H^{eff} (and we will drop the superscript from now on, and write it simply as H in what follows below), view it as a classical energy functional for length-S vectors $\vec{S}$, and ask about the *classical* minimum energy configurations of these length-S vectors. This is a useful starting point because the weight of a path in the coherent state path integral for the system has one factor proportional to

$$\exp(-\int_0^\beta d\tau H(S\vec{N}(\vec{r}_i,\tau))) \; . \tag{1}$$

If this were the only factor in the weight, the path integral would be dominated by classical time-independent minimum energy configurations of the length-S vectors $S\vec{N}(\vec{r}_i)$. Each such classical minimum energy configuration is labeled by a unit vector $\hat{n}$, since the vectors $S\vec{N}$ like to all point along or opposite the common axis defined by $\hat{n}$, such that each vector is anti-aligned with all its neighbours on the square lattice. Any such configuration spontaneously breaks the global symmetry of rotations in spin space, and $\hat{n}$ serves

as the order parameter for this symmetry-broken state; this kind of *antiferromagnetic* ordering on bipartite lattices (decomposible into two sublattices such that all neighbours of a site belong to the *other sublattice*) usually goes by the name "Neel order" after the physicist Louis Neel.

Of course, this is *not* the only factor in the weight of a path, and all the quantum mechanics is in the other factor S_B, which we have not analyzed in any detail yet. At a qualitative level, what is the effect of the quantum dynamics encoded by S_B? To get a feel for this, consider two spin-half objects coupled by the Hamiltonian

$$H_{12} \;=\; J_{12}\vec{S}_1 \cdot \vec{S}_2 \tag{2}$$

where the S are now spin-half operators. Viewed as a classical energy functional, this is minimized by any configuration in which the two spins are anti-aligned, say

$$|\uparrow,\downarrow\rangle \tag{3}$$

However, a state like this is *not* an eigenstate of the Hamiltonian, since the action of the "transverse components" of the Hamiltonian, *i.e.*

$$\frac{J_{12}}{2}(S_1^+ S_2^- + S_1^- S_2^+) \tag{4}$$

sends this state to the spin-flipped state

$$|\downarrow,\uparrow\rangle \tag{5}$$

Thus, quantum mechanics leads to *fluctuations* about the classical minimum energy configuration. The real difficulty then is to understand whether such *quantum fluctuations* lead to small corrections to the classical Neel ordered picture of the low energy physics, or whether they can completely overwhelm this tendency for nearest neighbour spins to anti-align and spontaneously break rotational symmetry by forming a Neel ordered state.

To address this question, we begin by parameterizing $\vec{N}(\vec{r}_j,\tau)$ in a way that factors in this classical tendency towards Neel order, but allows for fluctuations:

$$\vec{N}(\vec{r}_j,\tau) \;=\; \eta(\vec{r}_j)\hat{n}(\vec{r}_j,\tau)\sqrt{1-(a^d/S)^2\vec{L}^2} + \frac{a^d}{S}\vec{L}(\vec{r}_j,\tau) \tag{6}$$

where

$$\begin{aligned} \eta(\vec{r}_j) &= +1 \text{ for } \vec{r}_j \in \text{A} - \text{sublattice} \\ &= -1 \text{ for } \vec{r}_j \in \text{B} - \text{sublattice} \end{aligned} \tag{7}$$

on a d-dimensional bipartite hypercubic lattice (which generalizes the two dimensional square lattice example which is of interest to us from the point of view of the cuprate Mott insulator discussed in the previous class).

The idea behind this parameterization is quite simple: $\hat{n}(\vec{r}_j, \tau)$ is the "local" Neel order parameter, *i.e.* the axis along which the system develops alternating spin order on short timescales and short distance scales of order a small number (say five to ten) of lattice spacings a. On the other hand, $\vec{L}(\vec{r}_j, \tau)$ represents fluctuations in space and time, which disrupt this alternating or short-range Neel ordered pattern of spin orientations. Since $\hat{n}$ is a unit vectors at all space-time points, fluctuations about it must be orthogonal to it, *i.e.*

$$\hat{n}(\vec{r}_j, \tau) \cdot \vec{L}(\vec{r}_j, \tau) = 0. \tag{8}$$

Geometrically, one can imagine obtaining $\hat{n}$ by the following coarse-graining procedure. Take a region of size l^d, where l is of order a few lattice spacings a. Define $\hat{n}(\vec{r}, \tau)$ at the center $\vec{r}$ of this region as being the unit vector in the direction of the average of all $\vec{N}(\vec{r}_j, \tau)$ in this region. Now, decompose each $\vec{N}(\vec{r}_j, \tau)$ into a component parallel to and perpendicular to this unit vector $\hat{n}(\vec{r}, \tau)$. This is precisely the decomposition written down in the above formula—the perpendicular component is

$$\frac{a^d}{S} \vec{L}(\vec{r}_j, \tau) \tag{9}$$

and the parallel component is

$$\eta(\vec{r}_j) \hat{n}(\vec{r}_j, \tau) \sqrt{1 - (a^d/S)^2 \vec{L}^2} \tag{10}$$

This "operational definition" makes it clear that the decomposition we have written down involves some coarse-graining or loss of information. We will be a bit sloppy about this aspect of it for now and not worry too much about whether we are defining $\hat{n}$ on the original lattice, or a coarser lattice, or in

the continuum with some restrictions on its momentum content, since this will only *reduce* the upper cut-off in (momentum) Fourier space from the Brillouin zone edge to a smaller value in different ways. The more important thing to keep in mind about this decomposition is that one expects $\hat{n}$ to be a *slowly varying* function of space and imaginary time, and that one expects $\vec{L}$ to be small in magnitude in addition to being slowly varying.

The idea now is to "plug-in" this decomposition for $\vec{N}(\vec{r}_j, \tau)$ into our earlier expression for the partition function:

$$\begin{aligned} Z &= \\ \int \mathcal{D}\vec{N}(\vec{r}_i, \tau) \exp\Big(\int_0^\beta d\tau \Big[\sum_i iS \int_0^1 du\, \vec{N}(\vec{r}_i, u, \tau) \cdot \Big(\frac{\partial \vec{N}(\vec{r}_i, u, \tau)}{\partial u} \times \frac{\partial \vec{N}(\vec{r}_i, u, \tau)}{\partial \tau} \Big) \\ - JS^2 \sum_{\langle i,j \rangle} \vec{N}(\vec{r}_i, \tau) \cdot \vec{N}(\vec{r}_j, \tau) \Big] \Big) \end{aligned}$$

$$\text{with constraint } \vec{N}(\vec{r}_i, 0) = \vec{N}(\vec{r}_i, \beta) \ \ \forall i \tag{11}$$

and expand to leading order in the smallness of $\vec{L}$ and the smallness of space and time derivatives of $\hat{n}$.

We will handle the Berry phase term, and the energetic term separately, since one contains all the information about the quantum dynamics, and the other contains all the information about the classical energetics of the interaction between spins.

In this lecture, let us take the energetic term

$$+JS^2 \int_0^\beta d\tau \sum_{\langle i,j \rangle} \vec{N}(\vec{r}_i, \tau) \cdot \vec{N}(\vec{r}_j, \tau) \tag{12}$$

and express each $\vec{N}$ in terms of $\hat{n}$ and $\vec{L}$, keeping in mind that they are orthogonal to each other at each space-time point. This has three contributions: one obtained by using the $\hat{n}$ piece of each $\vec{N}$, the other obtained by using the $\vec{L}$ part of each $\vec{N}$, and the third being a cross-term in which one uses the $\hat{n}$ part of one $\vec{N}$ and the $\vec{L}$ part of the other $\vec{N}$.

For the first contribution, we expand the square-roots to leading order and note that

$$\eta(\vec{r}_i)\eta(\vec{r}_j) = -1 \tag{13}$$

for all nearest neighbour i and j to obtain

$$\begin{aligned} & +JS^2 \sum_{\langle ij \rangle} \left(\frac{(\hat{n}(\vec{r}_i,\tau) - \hat{n}(\vec{r}_j,\tau))^2}{2} - 1 \right) \left[1 - \frac{1}{2}\frac{a^{2d}}{S^2}\vec{L}^2(\vec{r}_i,\tau) - \frac{1}{2}\frac{a^{2d}}{S^2}\vec{L}^2(\vec{r}_j,\tau) \right] \\ \approx \quad & \frac{JS^2}{2} \sum_{\mu=x,y} \sum_j (\Delta_\mu \hat{n})^2(\vec{r}_j,\tau) + \frac{2dJa^{2d}}{2} \sum_j \vec{L}^2(\vec{r}_j,\tau) \end{aligned}$$

Here and henceforth, we use the convenient notation $(\Delta_\mu \hat{n})^2 = \sum_{\mu,\alpha} (\partial_\mu n_\alpha)^2$.

In the second contribution, one can completely ignore the spatial dependence of $\vec{L}$ to leading order and write it as

$$\frac{Ja^{2d}}{S^2} S^2 \sum_{\langle ij \rangle} \vec{L}(\vec{r}_i,\tau) \cdot \vec{L}(\vec{r}_j,\tau) \quad \approx \quad \frac{2dJa^{2d}}{2} \sum_j \vec{L}^2(\vec{r}_j,\tau) \tag{14}$$

And finally, in the third contribution, we note that the oscillatory spatial dependence due to to one factor of $\eta(\vec{r}_j)$ causes this term to be even higher order in smallness (assuming slowly varying $\hat{n}$ and small $\vec{L}$), and we may therefore ignore it in our leading order expansion.

Collecting terms, and rewriting things in a continuum language by replacing finite differences by derivatives and summations by integrations (with suitable factors of the lattice spacing inserted to make this rewriting possible), we obtain

$$\begin{aligned} & JS^2 \int_0^\beta d\tau \sum_{\langle ij \rangle} \vec{N}(\vec{r}_i,\tau) \cdot \vec{N}(\vec{r}_j,\tau) \;\; = \\ & \qquad \frac{1}{2} \int_\Lambda d^d r \int_0^\beta d\tau [\rho_s (\nabla_\mu \hat{n})^2(\vec{r},\tau) + \frac{1}{\chi_\perp} \vec{L}^2(\vec{r},\tau)] \\ \text{where} \quad \rho_s \quad \approx \quad & JS^2 a^{2-d} \quad \chi_\perp^{-1} \quad \approx \quad 4dJa^d \end{aligned} \tag{15}$$

Here, the repeated index μ is summed over, and the subscript Λ reminds us that although we have used a continuum notation, we are in fact working with variables defined on a coarse-grained lattice with larger "lattice spacing" $l \sim \Lambda^{-1}$, or equivalently, we are working with fields that have a momentum cutoff of Λ which is somewhat smaller than the original linear dimension of the Brillouin zone corersponding to the microscopic lattice.

In the above, ρ_s is clearly the stiffness or rigidity parameter that controls spatial fluctuations about the antiferromagnetic ordering. What is perhaps less obvious from the above is that $\chi_\perp$ controls the macroscopic susceptibility of the system to a uniform external magnetic field applied perpendicular to the direction of the Neel vector. To see that the latter is indeed the case, imagine applying a field $\vec{B}$ perpendicular to the Neel order. From the decomposition of $\vec{N}$ into $\hat{n}$ and $\vec{L}$, it is clear that such a field couples directly to $\vec{L}$ and induces an additional term

$$-\int_0^\beta \int d^d r \vec{B}.\vec{L}(\vec{r},\tau) \tag{16}$$

since the alternating part averages to zero. If the energy functional for $\vec{L}$ is then minimized with respect to $\vec{L}$, the minimum is realized for

$$\vec{L}(\vec{B}) \;=\; \chi_\perp \vec{B} \tag{17}$$

Of course, the numerical prefactors in the expressions for the stiffness and uniform susceptibility should not be taken seriously, since they depend on the coarse-graining procedure employed to write down the continuum action. However, the *qualitative dependence* of ρ_s and $\chi_\perp$ on S, J, a and dimensionality d is expected to be more robust.

Note that although ρ_s penalizes fluctuations of $\hat{n}$ in space, there is no term as yet that controls or penalizes fluctuations of $\hat{n}$ in imaginary time. This is not surprising, since all the quantum dynamics is in the Berry phase term that we have not yet analyzed, and one should expect that such a penalty for rapid temporal fluctuations of $\hat{n}$ will arise naturally from the quantum dynamics, since such rapid fluctuations should cost a large amount of kinetic energy. A second, related point, is the following: $\vec{L}$ has the interpretation of being the total angular momentum density since

$$\begin{aligned} \sum_i S\vec{N}(\vec{r}_i,\tau) \;&=\; a^d \sum_i \vec{L}(\vec{r}_i,\tau) \;\; \text{(oscillatory piece averages out)} \\ &\approx \int d^d r \vec{L}(\vec{r},\tau) \end{aligned} \tag{18}$$

However, the total angular momentum is the generator of rotations of all system vectors, including $\hat{n}$, and is therefore the conjugate variable to $\hat{n}$ in the Hamiltonian dynamics of this problem. However, there is, as yet, no term

that "ties in" $\vec{L}$ to $\hat{n}$ in this manner—again, such a term must arise from the expansion of the Berry phase term whic encodes the quantum dynamics.

[Don't worry if you don't understand what is meant here—it will be clear in the next lecture—this comment is more for people who are familiar with the phase space path integral for a single particle, which has a term in the action that reads

$$i \int dt\ p\ \frac{dx}{dt} \tag{19}$$

and signals the fact that p is the generator of translations in x and is therefore the conjugate variable in the Hamiltonian dynamics of the system]

In the next lecture, we will employ the same procedure to expand the Berry phase term

$$S_B \ = \ \int_0^\beta d\tau \sum_i iS \int_0^1 du\, \vec{N}(\vec{r}_i, u, \tau) \cdot (\frac{\partial \vec{N}(\vec{r}_i, u, \tau)}{\partial u} \times \frac{\partial \vec{N}(\vec{r}_i, u, \tau)}{\partial \tau}) \tag{20}$$

using this decomposition in terms of a slowly varying Neel component $\hat{n}$ and a uniform component $\vec{L}$, and see that such a term tying in $\vec{L}$ to the time derivative of $\hat{n}$ does arise from this expansion, and does control the temporal fluctuations of $\hat{n}$.