1 The Heisenberg model

1.1 Definition of the model

The model we will focus on is called the Heisenberg model. It has the following Hamiltonian:

$$\mathcal{H} = \frac{1}{2} \sum_{\substack{i,j \\ i \neq j}} J_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j. \tag{1}$$

Here i and j refer to sites on a lattice. The model can be defined on any lattice, but for concreteness (and to keep things simple) we will here limit ourselves to a hypercubic lattice in d dimensions ($d = 1, 2, 3$).[1,2] The $\boldsymbol{S}_i$ are spin operators which live on the lattice sites. Spin components on the same lattice site obey the standard angular momentum commutation relations

$$[S_j^\alpha, S_j^\beta] = i \sum_\gamma \epsilon_{\alpha\beta\gamma} S_j^\gamma \quad (\alpha, \beta, \gamma = x, y, z) \tag{2}$$

and spins on different sites commute with each other. The spin operators all have spin S, i.e. the operators $\boldsymbol{S}_i^2$ have eigenvalue $S(S+1)$ where S is an integer or half-integer. The spin interaction in (1), which is of the form $\boldsymbol{S}_i \cdot \boldsymbol{S}_j$, is called an **exchange interaction**, and the coefficients J_{ij} are called **exchange constants**. We will make the simplifying assumption (which is often realistic) that the spin interactions are negligible between spins that are not nearest-neighbors, i.e. J_{ij} is nonzero only if i and j are nearest-neighbor lattice sites, in which case we further assume $J_{ij} = J$, where J is a constant. There are then two different cases to consider, $J < 0$ and $J > 0$. For $J < 0$ the interaction energy of two spins favors them to be parallel; this is the ferromagnetic case. For $J > 0$ antiparallel orientation is instead favored; this is the antiferromagnetic case.

1.2 The $S = 1/2$ Heisenberg antiferromagnet as an effective low-energy description of the half-filled Hubbard model for $U \gg t$

It turns out that the magnetic properties of many insulating crystals can be quite well described by Heisenberg-type models of interacting spins. Let us consider an example based on assuming that the electrons can be described in terms of the so-called Hubbard model, with Hamiltonian

$$H = -t \sum_{\langle i,j \rangle, \sigma} (c_{i\sigma}^\dagger c_{j\sigma} + \text{h.c.}) + U \sum_i n_{i\uparrow} n_{i\downarrow}. \tag{3}$$

This is probably the most important (and famous) lattice model of interacting electrons. $c_{i\sigma}^\dagger$ creates an electron on site i with spin σ, and $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$ counts the number of electrons with spin σ on site i. In this model there is therefore one electronic orbital per site. The first term in (3) is the kinetic energy describing electrons hopping between nearest-neighbor sites i and j, and the second term is the interaction energy describing the energy cost $U > 0$

[1] In 2 dimensions a hypercubic lattice is a square lattice and in 3 dimensions it is a cubic lattice.

[2] Periodic boundary conditions will be used in all spatial directions, so that the Hamiltonian is fully translationally invariant.

associated with having two electrons on the same site (these electrons must have opposite spin, as having two electrons on the same site with the same spin would violate the Pauli principle). Note that the interaction energy between electrons which are not on the same site is completely neglected in this model.[3] The Hubbard model is the simplest model describing the fundamental competition between the kinetic energy and the interaction energy of electrons on a lattice. Despite much research, there is still a great deal of controversy about many of its properties in two and three dimensions (the model has been solved exactly in one dimension, but even in that case the solution is extremely complicated).

1.2.1 Physical picture

Now consider a system with N lattice sites, and assume that there are also N electrons in the system, so the average number of electrons per site is 1. This is called the **half-filled** case because maximally the system could contain 2 electrons per site (one for each spin projection $\uparrow$, $\downarrow$) and thus a total electron number of $2N$. Assume further that $U \gg t$. This suggests that to find the low-energy states we should first minimize the interaction energy and then treat the kinetic energy as a perturbation. The interaction energy is minimized by putting exactly one electron on each site; then no site is doubly occupied so the total interaction energy is zero. Furthermore, whether the electron on a given site has $\uparrow$ or $\downarrow$-spin is clearly unimportant; thus any spin distribution gives the same interaction energy, leading to a large (2^N-fold) degeneracy. Moving an electron to a different site containing an electron with opposite spin creates a doubly occupied site which is penalized by a large energy cost U. Thus as long as we're only interested in understanding the physics of the system for energies (and/or temperatures) much less than U, we can neglect configurations with double occupancies. Thus the interaction energy completely determines the charge distribution of the electrons while putting no constraints on their spin distribution. However, if we now consider the kinetic energy term ($\propto t$) as a perturbation, it is clear that (see Fig. 1(a)) if neighboring electrons have opposite spins, an electron can hop (virtually, in the sense of 2nd order perturbation theory) to a neighboring site and back; this virtual delocalization reduces the kinetic energy.[4] In contrast (see Fig. 1(b)), if neighboring spins are parallel such hopping is forbidden, as the intermediate state with two electrons on the same site with the same spin would violate the Pauli principle. Therefore in this situation an effective interaction is generated which favours neighbouring electrons to have opposite spin, i.e. antiparallel orientation. It can be shown (see next subsection) that the resulting effective model, valid at temperatures and energies $\ll U$, is the Heisenberg *antiferromagnet* for the $S = 1/2$ electron spins, with $J = 4t^2/U$ (this expression for J can be understood from 2nd order perturbation theory: there is one factor of $-t$ for each of the two hops (first to the neighboring site, then back) and a factor U coming from the energy denominator due to the larger energy of the intermediate state). Therefore, the antiferromagnetic exhange interaction J comes about due to an interplay between electron hopping, the electron-electron interactions, and the

[3] One can use the general formalism for two-particle operators in second quantization (discussed in the notes on second quantization) to express the electron-electron (Coulomb) interaction in second quantization using the Wannier basis. For details, see the discussion in Altland & Simons, Sec. 2.2., pp. 59-61. The "on-site" term in the resulting expression is the interaction term included in the Hubbard model.

[4] That the kinetic energy is reduced by this process can be understood from the formula for the energy correction in 2nd order perturbation theory, which is always negative for the ground state.

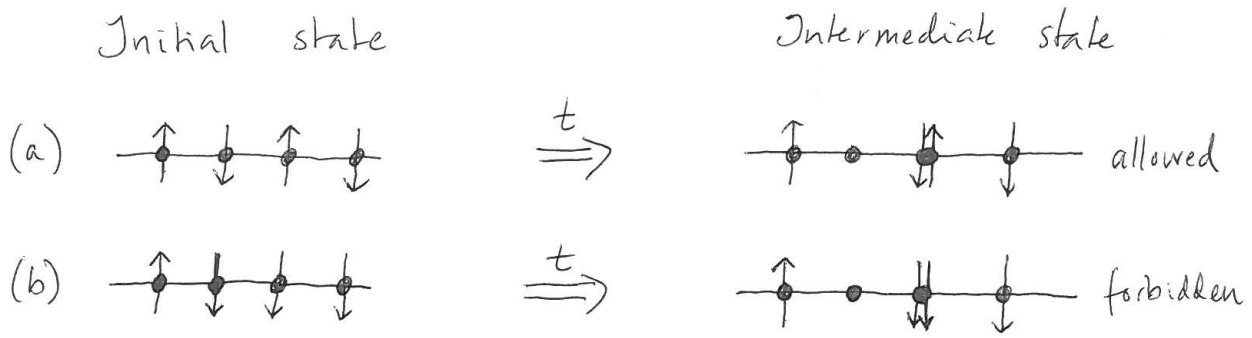


Figure 1: Illustration of why an effective antiferromagnetic interaction is generated in the half-filled Hubbard model with $U \gg t$. In both (a) and (b) the transition shown is between an initial state with an electron on each site and an intermediate state where the electron initially on site 2 has hopped to site 3. See text for further explanation.

Pauli principle; the effect of the antiferromagnetic exchange is to reduce the kinetic energy of the electrons.

1.2.2 Mathematical derivation

In the following we will show how this result can be derived mathematically.[5] We start out by considering a general model with Hamiltonian H whose eigenstates can be classified according to whether they lie in the low-energy subspace (LES) that we want to find an effective Hamiltonian for, or in the complementary high-energy subspace (HES). For the half-filled Hubbard model with $U \gg t$, the LES would consist of the states with no doubly occupied sites and the HES would consist of the states with at least one doubly occupied site (i.e. having energies of order U). Let P and Q be projection operators which project onto LES and HES respectively. Like any projection operators they satisfy $P^2 = P$ and $Q^2 = Q$. And since LES and HES are orthogonal subspaces (i.e. a state in LES has no component in HES and vice versa) we also have $QP = PQ = 0$. The time-independent Schrodinger equation for the system is $H\Psi = E\Psi$, where E is an energy eigenvalue and Ψ is an eigenstate. Since $P + Q = I$ we can write this as

$$H(P+Q)\Psi = E(P+Q)\Psi. \tag{4}$$

If we act with Q from the left on this equation, use $QP = 0$ and $Q^2 = Q$, and rearrange, we get

$$(QHQ - E)(Q\Psi) = -QH(P\Psi) \quad \Rightarrow Q\Psi = -(QHQ - E)^{-1}QH(P\Psi). \tag{5}$$

This expression relates $Q\Psi$ in HES to $P\Psi$ in LES. Since we want to find an effective Hamiltonian for LES we want to eliminate $Q\Psi$ from the description. Hence we insert this expression for $Q\Psi$ back into (4), which then becomes

$$[H - H(QHQ - E)^{-1}QH]P\Psi = E(P+Q)\Psi. \tag{6}$$

[5] This derivation follows Nagaosa's book "Quantum field theory in strongly correlated electronic systems", Sec. 3.1.

Next we act with P from the left to project this onto LES. Using $PQ = 0$ on the rhs and $P^2 = P$ on the lhs we get

$$[PHP - PHQ(QHQ - E)^{-1}QHP](P\Psi) = E(P\Psi). \tag{7}$$

(Note that in arriving at this result we have also used $PH(QHQ - E)^{-1}Q = PHQ(QHQ - E)^{-1}Q$, i.e. we inserted a Q to the right of PH because the operator to its right already projects onto the HES subspace so inserting another Q won't change this.)

So far everything has been exact and valid for a general Hamiltonian H. We now specialize to the half-filled Hubbard model. We can write its Hamiltonian as $H = H_K + H_U$, where the kinetic energy is

$$H_K = -\sum_{i,j}\sum_{\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} \tag{8}$$

and the interaction energy is

$$H_U = U\sum_{i} n_{i\uparrow} n_{i\downarrow}. \tag{9}$$

(Note that the kinetic energy operator H_K we consider here is slightly more general than the one in (3) as the hopping in H_K is not necessarily restricted to be between nearest-neighbor sites.) Furthermore, the projection operator P for the LES can be written $P = \prod_i (1 - n_{i\uparrow} n_{i\downarrow})$ and $Q = I - P$. We now need to find more explicit expressions for the operators PHP, PHQ, $(QHQ - E)^{-1}$ and QHP appearing in (7).

- $\underline{PHP.}$ $PHP = PH_KP + PH_UP$. $PH_KP = 0$. Why? Because first the rightmost P projects onto LES. Then H_K creates a state with one doubly occupied site (or rather a linear combination of such states) which therefore lies in HES. Consequently, when the leftmost P acts this state is killed. Next, $PH_UP = 0$ is zero as well, because $H_UP = 0$ since a state in LES has zero interaction energy. Hence $PHP = 0$.

- $\underline{PHQ.}$ $PHQ = PH_KQ + PH_UQ$. Here, $PH_UQ = 0$ because the state created by H_UQ lies in HES, and therefore is killed by P. Hence $PHQ = PH_KQ$.

- $\underline{QHP.}$ $QHP = QH_KP + QH_UP$. Here, $QH_UP = 0$ because $H_UP = 0$ (see above). Hence $QHP = QH_KP$.

- $\underline{(QHQ - E)^{-1}.}$ Using the results so far, the operator acting on $P\Psi$ on the lhs in (7) can be written $-PH_KQ(QHQ - E)^{-1}QH_KP$. As $H_K \propto t_{ij}$ (the hopping matrix elements) this operator is already of order t_{ij}^2. As we want to treat the effects of H_K on the states in LES to lowest order in t_{ij} (since we are considering the limit $t \ll U$), we may therefore replace $(QHQ - E)^{-1}$ by its value to zeroth order in t_{ij}. Thus in $QHQ = QH_KQ + QH_UQ$, we may set $t_{ij} = 0$ in H_K giving $QH_KQ \to 0$. Furthermore, QH_UQ can be replaced by U. This is because before this operator acts, the operator QH_KP to its right has created a state in HES with only one doubly occupied site, which thus has energy U. Finally, in $(QHQ - E)^{-1}$, the eigenvalue E in LES can also be replaced by its value to zeroth order in t_{ij}, which is 0. Hence $(QHQ - E)^{-1}$ can be replaced by $1/U$.

Using these results, the operator acting on $P\Psi$ on the lhs in (7) can to lowest order in t_{ij} be replaced by $-PH_KQ(1/U)QH_KP$. This can be simplified further: As the state created by H_KP is entirely in HES, the Q's to its left are unnecessary and can be removed. Hence the effective Hamiltonian for the states in LES is

$$H_{\text{eff}} = -P\frac{H_K^2}{U}P. \tag{10}$$

Inserting the expression (8) for H_K we find

$$PH_K^2P = \sum_{i,j}\sum_{i',j'}\sum_{\sigma,\sigma'} t_{ij}t_{i'j'}\ Pc_{i\sigma}^\dagger c_{j\sigma}c_{i'\sigma'}^\dagger c_{j'\sigma'}P. \tag{11}$$

We now analyze this in more detail. First the rightmost P projects onto LES. Subsequently, the composite operator $c_{i\sigma}^\dagger c_{j\sigma}c_{i'\sigma'}^\dagger c_{j'\sigma'}$ acts. The resulting state must be in LES, otherwise it will be killed by the leftmost P. Hence the composite operator cannot create doubly occupied sites, so the sites j, j' where electrons are annihilated must match the sites i, i' where electrons are created. Having $i = j$ and $i' = j'$ will give zero result since $t_{ii} = 0$ by definition of the electron hopping. Hence a nonzero result is only obtained if $i = j'$ and $i' = j$. Hence

$$\begin{aligned} PH_K^2P &= \sum_{ij}\sum_{\sigma\sigma'} |t_{ij}|^2 Pc_{i\sigma}^\dagger c_{j\sigma}c_{j\sigma'}^\dagger c_{i\sigma'}P \\ &= \sum_{ij}\sum_{\sigma\sigma'} |t_{ij}|^2 Pc_{i\sigma}^\dagger c_{i\sigma'}c_{j\sigma}c_{j\sigma'}^\dagger P, \end{aligned} \tag{12}$$

where to arrive at the last expression we anticommuted twice (and used $t_{ii} = 0$). Since $c_{j\sigma}c_{j\sigma'}^\dagger$ doesn't take us out of LES, we can insert $P = P^2$ to its left without any effect, giving

$$PH_K^2P = \sum_{ij}\sum_{\sigma\sigma'} |t_{ij}|^2 Pc_{i\sigma}^\dagger c_{i\sigma'}P \cdot Pc_{j\sigma}c_{j\sigma'}^\dagger P. \tag{13}$$

At this point we use the identities

$$c_{i\sigma}^\dagger c_{i\sigma'} = \frac{1}{2}\delta_{\sigma\sigma'}(n_{i\uparrow} + n_{i\downarrow}) + \boldsymbol{S}_i \cdot \boldsymbol{\sigma}_{\sigma'\sigma}, \tag{14}$$

$$c_{i\sigma}c_{i\sigma'}^\dagger = \delta_{\sigma\sigma'}\left(1 - \frac{n_{i\uparrow} + n_{i\downarrow}}{2}\right) - \boldsymbol{S}_i \cdot \boldsymbol{\sigma}_{\sigma\sigma'}, \tag{15}$$

where $\boldsymbol{S}_i$ is the spin operator at site i, given by

$$\boldsymbol{S}_i = \frac{1}{2}\sum_{\sigma,\sigma'} c_{i\sigma}^\dagger \boldsymbol{\sigma}_{\sigma,\sigma'} c_{i\sigma'}, \tag{16}$$

and $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ is the vector of Pauli matrices given by

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{17}$$

where the first row/column corresponds to the index $\sigma = \uparrow = +1/2$ and the second row/column corresponds to $\sigma = \downarrow = -1/2$. When the number of electrons per site is exactly 1, as it is in LES, the operators defined in (16) satisfy all properties of spin operators for spin $S = 1/2$. Inserting the identities (14)-(15) into (13) and using $\mathrm{Tr}(\sigma^a \sigma^b) = 2\delta_{ab}$ where $a, b = x, y, z$, it can be shown (you will be asked to do this in a tutorial) that the effective Hamiltonian for LES can be written

$$H_{\text{eff}} = \sum_{i,j} J_{ij}(\boldsymbol{S}_i \cdot \boldsymbol{S}_j - 1/4), \tag{18}$$

where

$$J_{ij} = \frac{2|t_{ij}|^2}{U} > 0. \tag{19}$$

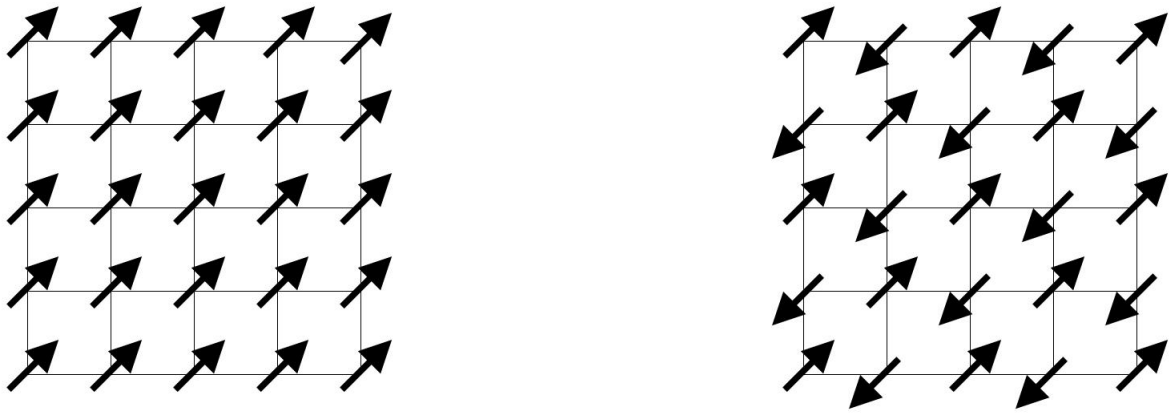
The second term $-(1/4)\sum_{i,j} J_{ij}$ in (18) is just a constant and can be neglected. Hence the effective Hamiltonian for LES is nothing but the *antiferromagnetic* Heisenberg model for spin-1/2.[6]

We will not here go into how other examples of Heisenberg models (e.g. with spin $S > 1/2$ on each site and/or with ferromagnetic interactions) can arise as low-energy descriptions of different systems. Thus in the remainder of these notes we will simply consider the Heisenberg model as an interesting effective model of interacting spins and explore its properties (such as the nature of its ground state and excitations) without dwelling more on the origin of the model itself.

1.3 Ferro- and antiferromagnetic order

Let us discuss the behavior of the spins, as a function of temperature, in materials described by the Heisenberg model. At high temperatures there are strong thermal fluctuations so that the spins are disordered, meaning that the expectation value of each spin vanishes: $\langle \hat{\boldsymbol{S}}_i \rangle = 0$. (Here the brackets represent both a thermal and quantum-mechanical expectation value.) However, below some critical temperature T_c it may be that the spins order magnetically, meaning that the spins on average point in some definite direction in spin space, $\langle \hat{\boldsymbol{S}}_i \rangle \neq 0$. Whether or not such magnetic order occurs depends on the dimensionality and type of lattice, and the range of the interactions (we will limit ourselves to hypercubic lattices and nearest-neighbor interactions in our explicit investigation of the Heisenberg model). If magnetic order occurs with $T_c > 0$, then, as the temperature is lowered from T_c down to zero, $\langle \boldsymbol{S}_i \rangle$ will increase and reach some maximum value at zero temperature. The critical temperature T_c is called the Curie temperature (T_C) in ferromagnets and the Neel temperature (T_N) in antiferromagnets. The spin ordering pattern in two magnetically ordered phases are illustrated for a square lattice in Fig. 2: In the ferromagnetic case ($J < 0$) all spins point in

[6] Electrons have both charge and spin. As the charge degrees of freedom of the electrons are "frozen" in this limit of large U/t (as reflected in the fact that the low-energy effective model (18) only involves the spin degrees of freedom), the system is an insulator. Note, however, that since the system is half-filled, band theory would predict metallic, not insulating behaviour. While "conventional" insulators, also called **band insulators**, occur as a consequence of having only completely filled/empty bands and are predicted by band theory, we here have an example of something different: a **Mott insulator**. A Mott insulator is a system that band theory would predict to be metallic, but which nevertheless is an insulator, due to electron interaction effects.



,

Figure 2: Average spin directions in phases with (a) ferromagnetic and (b) antiferromagnetic order on the square lattice. (The ordering direction is arbitrary.)

the same direction (Fig. 2(a)) while in the antiferromagnetic case ($J > 0$) neighboring spins point in opposite directions (Fig. 2(b)).

2 The Holstein-Primakoff representation

Note that the commutator of two spin operators is itself an operator (see (2)), rather than just a complex number (c-number). This makes it much more complicated[7] to work directly with spin operators than with canonical bosonic (fermionic) creation and annihilation operators whose commutators (anticommutators) are just c-numbers. It would therefore be advantageous if one could represent the spin operators in terms of such canonical bosonic or fermionic operators and work with these instead. Fortunately quite a few such representations are known. In these lectures we will make use of the so-called **Holstein-Primakoff (HP) representation** which expresses the spin operators on a site j in terms of canonical boson creation and annihilation operators $a_j^\dagger$ and a_j as follows:

$$S_j^+ = \sqrt{2S - \hat{n}_j}\, a_j, \tag{20}$$

$$S_j^- = a_j^\dagger \sqrt{2S - \hat{n}_j}, \tag{21}$$

$$S_j^z = S - \hat{n}_j. \tag{22}$$

Here we have introduced the raising and lowering operators $S_j^\pm = S_j^x \pm i S_j^y$. The operator $\hat{n}_j \equiv a_j^\dagger a_j$ is the number operator for site j, i.e. it counts the number of bosons on this site. As the allowed eigenvalues of S_j^z are $-S, -S+1, \ldots, S$ you can see from the equation for S_j^z that this boson number must satisfy the constraint

$$\langle \hat{n}_j \rangle \leq 2S. \tag{23}$$

[7] At least as far as analytical (as opposed to numerical) approaches are concerned.

Also note that the expressions for S_j^+ and S_j^- are Hermitian conjugates of each other, as they should be. It can be shown that the spin commutation relations (2) and the relation $\boldsymbol{S}^2 = S(S+1)$ follow from the HP representation and the bosonic commutation relations $[a_j, a_j^\dagger] = 1$ etc. (You will be asked to verify this in a tutorial). As will hopefully become clear in the next couple of sections, the HP representation is very useful for studying magnetically ordered states and their excitations. The reason for this is related to the fact that the vacuum of the a_j bosons (i.e. the state with no such bosons) is the state corresponding to the maximum eigenvalue S of S_j^z.

3 Spin-wave theory of ferromagnets

We will now begin our study of the Heisenberg model (1) using spin-wave theory. We will first consider the ferromagnetic case, with $J < 0$ (which can be written $J = -|J|$). It will be convenient to rewrite the Hamiltonian in terms of the raising and lowering operators $\hat{S}_j^\pm = \hat{S}_j^x \pm i\hat{S}_j^y$. Also since we're only considering nearest neighbor interactions we can write $j = i + \boldsymbol{\delta}$ where $\boldsymbol{\delta}$ is a vector connecting nearest-neighbor sites. Actually to avoid counting each interaction twice we will take δ to run over only *half* the nearest neighbor vectors. We will consider hypercubic lattices only. Thus in 1D, we take $\boldsymbol{\delta} = +\hat{x}$, in 2D $\boldsymbol{\delta} = +\hat{x}, +\hat{y}$ and in 3D $\boldsymbol{\delta} = +\hat{x}, +\hat{y}, +\hat{z}$. This gives

$$\mathcal{H} = J\sum_{i,\boldsymbol{\delta}} \left[\frac{1}{2}(S_i^+ S_{i+\boldsymbol{\delta}}^- + S_i^- S_{i+\boldsymbol{\delta}}^+) + S_i^z S_{i+\boldsymbol{\delta}}^z\right]. \tag{24}$$

The state with all spins pointing along z, i.e. $S_j^z = S$ is a ground state of the Heisenberg ferromagnet. This should be intuitively quite reasonable, and it can be proven easily. First let us show that it is an eigenstate. When we apply (24) to this state, the part involving the ladder operators give exactly zero because the operators S_j^+ kill the ground state since it already has maximum S_j^z which therefore cannot be increased further. Acting with the $S_i^z S_j^z$ gives back an energy $-|J|S^2Nz/2$ times the same state, where z is the number of nearest neighbors. This shows that this state is an eigenstate. To show that it is also the ground state, we note that the minimal energy possible is given by

$$E_0 = -|J|\sum_{i,\boldsymbol{\delta}} \max\langle \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+\boldsymbol{\delta}}\rangle. \tag{25}$$

It can be shown[8] that $\max\langle \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+\boldsymbol{\delta}}\rangle = S^2$ which gives $E_0 = -|J|S^2Nz/2$, the same as above. Hence the state in question is indeed a ground state.

[8] *The material in this footnote was not discussed in the lectures.* Let O be a Hermitian operator and let $\{|\Phi_n\rangle\}$ be its complete set of orthonormal eigenstates, with eigenvalues O_n. We will consider the expectation value of O in an arbitrary state $|\Psi\rangle$. We can expand $|\Psi\rangle$ in the set $\{|\Phi_n\rangle\}$: $|\Psi\rangle = \sum_n c_n|\Phi_n\rangle$. Normalization of $|\Psi\rangle$ gives that $\sum_n |c_n|^2 = 1$. The expectation value of O can then be written

$$\langle O\rangle \equiv \langle\Psi|O|\Psi\rangle = \sum_{m,n} c_m^* c_n \underbrace{\langle\Phi_m|O|\Phi_n\rangle}_{O_n\delta_{mn}} = \sum_n |c_n|^2 O_n.$$

This shows that the largest value of $\langle O\rangle$ is given by the largest eigenvalue of O, obtained by taking $|\Psi\rangle$ to be the eigenstate of O with the largest eigenvalue.

Now we consider spins on two different sites $i \neq j$. Since $(\boldsymbol{S}_i + \boldsymbol{S}_j)^2 = \boldsymbol{S}_i^2 + \boldsymbol{S}_j^2 + 2\boldsymbol{S}_i \cdot \boldsymbol{S}_j$, we can write

$$\boldsymbol{S}_i \cdot \boldsymbol{S}_j = \frac{1}{2}\Big[(\boldsymbol{S}_i + \boldsymbol{S}_j)^2 - \boldsymbol{S}_i^2 - \boldsymbol{S}_j^2\Big].$$

The largest expectation value of $\boldsymbol{S}_i \cdot \boldsymbol{S}_j$ is, from (8), given by the largest eigenvalue of the operator on the rhs of this equation. The three terms on the rhs commute with each other, so we can find simultaneous eigenstates for them. The eigenvalue of $\boldsymbol{S}_i^2$ and $\boldsymbol{S}_j^2$ is $S(S+1)$. The eigenvalues of $(\boldsymbol{S}_i + \boldsymbol{S}_j)^2$ are $S_{\rm tot}(S_{\rm tot}+1)$

We will now use the HP representation to study the Heisenberg ferromagnet, especially its excitations. As noted earlier we will consider a hypercubic lattice in d spatial dimensions, i.e. a standard lattice in one dimension, a square lattice in 2 dimensions, and a cubic lattice in 3 dimensions. The theory that will be developed is known as **spin-wave theory**.

A natural guess for the low-energy excitations would be that they just correspond to small collective oscillations of the spins around the ordering direction (which we choose to be the z direction). Thus these oscillations, which are called **spin waves**, make $\langle S_j^z \rangle$ less than the maximum value S. In terms of the HP representation this means that the boson number $\langle \hat{n}_j \rangle$ is nonzero, see (22). If this boson number $\langle \hat{n}_j \rangle$ is much smaller than S the reduction in S_j^z is small (i.e. very weak oscillations) and one might expect that an expansion in a small parameter proportional to $\langle \hat{n}_j \rangle / S$ would make sense. One might expect this to work better the larger S is, since one might guess that increasing S would make this parameter smaller (we'll verify this explicitly later). On the other hand, if it should turn out that $\langle \hat{n}_j \rangle / S$ is not small (this is something we'll have to check at the end of our calculation), then our basic assumption, that spin-waves are just weak oscillations around an ordered state, is wrong or at least questionable, and we may have to conclude that the system is not magnetically ordered after all. For example, this will be the conclusion if $\langle \hat{n}_j \rangle / S$ turns out to be divergent, which we'll see some examples of later.

The above indicates that spin-wave theory is essentially a $1/S$ **expansion**. It is **semi-classical** in nature, which follows since the limit $S \to \infty$ corresponds to classical spins, which can be seen e.g. from the fact that the eigenvalues of $\hat{\boldsymbol{S}}_i^2$ are $S(S+1) = S^2(1+1/S)$. If the spins were just classical vectors of length S, the square of their length should be just S^2. Instead we see that there is a correction factor $(1+1/S)$ due to the quantum nature of the spins. As the correction factor goes to 1 in the limit $S \to \infty$, this limit corresponds to classical spins.

Let us now discuss the spin-wave theory for the Heisenberg ferromagnet in detail. As noted earlier, in both the quantum and classical ground state all the spins point along the same direction, which we will take to be the z direction. We then rewrite the spin operators in the Heisenberg Hamiltonian in terms of boson operators using the HP representation. We write $\sqrt{2S - \hat{n}_j} = \sqrt{2S}\sqrt{1 - \hat{n}_j/(2S)}$ and expand the last square root here in a series in the operator $\hat{n}_j/(2S)$. This gives

$$\mathcal{H} = -|J|NS^2 z/2 - |J|S \sum_{i,\boldsymbol{\delta}} [a_{i+\boldsymbol{\delta}}^\dagger a_i + a_i^\dagger a_{i+\boldsymbol{\delta}} - a_i^\dagger a_i - a_{i+\boldsymbol{\delta}}^\dagger a_{i+\boldsymbol{\delta}}] + O(S^0). \tag{26}$$

Here N is the total number of lattice sites and z is the number of nearest neighbors and given by $z = 2d$ for a hypercubic lattice in d spatial dimensions. (Note that after doing the summations the two last terms inside the square brackets are in fact identical). We have only included terms of $O(S^2)$ and $O(S)$ in (26). To this order, the HP representation reduces to[9] $S_j^+ \approx \sqrt{2S} a_j$, $S_j^- \approx \sqrt{2S} a_j^\dagger$ (and $S_j^z = S - a_j^\dagger a_j$). The term of $O(S)$ is quadratic in the boson operators and can be straightforwardly diagonalized. As it is quadratic it is equivalent to noninteracting bosons. Terms in the Hamiltonian which are higher order in the $1/S$

[9]Note that when the HP expressions for $S_j^\pm$ are expanded in a series in $\hat{n}_j/(2S)$, and this series is truncated, the result is just approximate expressions for $S_j^\pm$. In particular they do not exactly obey the spin commutation relations.

expansion (not shown) contain four or more boson operators and thus represent interactions between bosons. However, these are suppressed at least by a factor $1/S$ compared to the $O(S)$ noninteracting term and one can thus hope that their effects are small (at least at large S and when the boson number is small) so that they can either be neglected to a first approximation or be treated as weak perturbations on the noninteracting theory.

Let us next diagonalize the quadratic, $O(S)$ term. In this ferromagnetic case, this can be accomplished simply by introducing Fourier-transformed boson operators as follows:

$$a_{\boldsymbol{k}} = \frac{1}{\sqrt{N}} \sum_i e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i} a_i. \tag{27}$$

This is just a variable transformation, so there are as many operators $a_{\boldsymbol{k}}$ as there are operators a_i. The inverse transformation is

$$a_i = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} a_{\boldsymbol{k}}. \tag{28}$$

Periodic boundary conditions imply that e.g. $a_i = a_{i+N_x\hat{x}}$ where N_x is the number of sites in the x direction. This is satisfied if $e^{ik_xN_x} = 1$, i.e. k_x takes the form $k_x = 2\pi n_x/N_x$ where n_x is an integer. It is customary to choose n_x to take the N_x successive values $-N_x/2, -N_x/2+1, \ldots, N_x/2-1$ (here we have assumed for simplicity that N_x is even so that $N_x/2$ is in fact an integer). Then k_x takes values in the interval $[-\pi, \pi\rangle$. Doing the same for all directions, the resulting values of $\boldsymbol{k}$ lie within what is called **the first Brillouin zone**. An important aspect of the transformation (27) is that it is canonical, i.e. it preserves the commutation relations in the sense that the operators $a_{\boldsymbol{k}}$ obey the same kind of commutation relations as the original boson operators: $[a_{\boldsymbol{k}}, a^\dagger_{\boldsymbol{k}'}] = \delta_{\boldsymbol{k},\boldsymbol{k}'}$ etc. Inserting (28) in (26) and using that $\sum_i e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_i} = N\delta_{\boldsymbol{k},\boldsymbol{k}'}$ one gets

$$\mathcal{H} = E_0 + \sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}} a^\dagger_{\boldsymbol{k}} a_{\boldsymbol{k}}, \tag{29}$$

where

$$E_0 = -|J|NS^2z/2 \tag{30}$$

and

$$\omega_{\boldsymbol{k}} = 2|J|S \sum_{\boldsymbol{\delta}} (1 - \cos \boldsymbol{k}\cdot\boldsymbol{\delta}) \equiv S|J|z(1-\gamma_{\boldsymbol{k}}), \tag{31}$$

where we have defined

$$\gamma_{\boldsymbol{k}} = \frac{2}{z} \sum_{\boldsymbol{\delta}} \cos \boldsymbol{k}\cdot\boldsymbol{\delta}. \tag{32}$$

Eq. (29) describes a Hamiltonian which is just a bunch of independent harmonic oscillators, each labeled by a wavevector $\boldsymbol{k}$. The quanta of the harmonic oscillators are called **magnons**; they are the quantized spin wave excitations (just like phonons are the quantized lattice vibrations in a crystal) with energy $\omega_{\boldsymbol{k}}$. In the limit $\boldsymbol{k} \to 0$, we have

$$\omega_{\boldsymbol{k}} \approx |J|S|\boldsymbol{k}|^2. \tag{33}$$

As a magnon with wavevector $\boldsymbol{k}$ costs an energy $\omega_{\boldsymbol{k}} > 0$,[10] the ground state has no magnons, i.e. $\langle \hat{n}_{\boldsymbol{k}} \rangle = 0$ at zero temperature (here $\hat{n}_k = a_k^\dagger a_k$). The ground state energy is therefore simply E_0 which is the interaction energy of all spins pointing in the same direction with maximal projection S along the z axis. As the temperature is increased, magnons will be thermally excited. Since they are just noninteracting bosons (when $O(S^0)$ terms and higher are neglected in the Hamiltonian, as done so far), the mean number of magnons with momentum $\boldsymbol{k}$ is given by the Bose-Einstein distribution function,

$$\langle \hat{n}_{\boldsymbol{k}} \rangle = \frac{1}{e^{\beta \omega_{\boldsymbol{k}}} - 1}. \tag{34}$$

The magnetization $\boldsymbol{M} \equiv (1/N) \sum_i \langle \boldsymbol{S}_i \rangle$ is a natural measure of the strength of the putative magnetic order in the system (cf. the discussion in Sec. 1.3). If $M \equiv |\boldsymbol{M}|$ is positive (zero) we say that the system is (is not) ferromagnetically ordered. The larger M is, the stronger is the ferromagnetic order. We say that the magnetization is an **order parameter** for the ferromagnetic phase. By definition, an order parameter for a given type of order is a quantity that is nonzero in the phase(s) where that order is present and is zero in other phases. With the ordering direction being the z direction, we get

$$M = \frac{1}{N} \sum_i \langle S_i^z \rangle = S - \frac{1}{N} \sum_i \langle \hat{n}_i \rangle = S - \frac{1}{N} \sum_{\boldsymbol{k}} \langle \hat{n}_{\boldsymbol{k}} \rangle \equiv S - \Delta M. \tag{35}$$

We'd like to look at how ΔM depends on temperature for low temperatures. We first introduce an artificial wavevector cutoff k_0 which is the smallest wavevector in the $\boldsymbol{k}$ sum; the real system is described by the limit $k_0 \to 0$. We also introduce another wavevector $\bar{k} > k_0$ which is chosen such that $\omega_{\bar{\boldsymbol{k}}} \ll k_B T \ll |J| S$; this means in particular that for $|k| < \bar{k}$ the quadratic form (33) is valid. This gives

$$\Delta M = \frac{1}{N} \left(\sum_{k_0 < |\boldsymbol{k}| < \bar{k}} \frac{1}{e^{|J| S k^2 / k_B T} - 1} + \sum_{|\boldsymbol{k}| > \bar{k}} \frac{1}{e^{\omega_{\boldsymbol{k}} / k_B T} - 1} \right). \tag{36}$$

The second term is independent of k_0 and finite. For reasons that soon will become clear, we will therefore neglect it and focus on the first term. Converting the sum to an integral and expanding the exponential (using $|J| S^2 k^2 \ll k_B T$) we get

$$\Delta M \propto \int_{k_0}^{\bar{k}} dk \; k^{d-1} \frac{k_B T}{|J| S k^2} \propto \frac{k_B T}{|J| S} \cdot \begin{cases} 1/k_0 + \ldots, & d = 1 \\ -\log k_0 + \ldots, & d = 2 \end{cases} \tag{37}$$

Therefore we see that, at nonzero temperatures in one and two dimensions, ΔM *diverges* as the cutoff k_0 is sent to zero. Therefore our initial assumption that this correction is small, is found to be wrong for these cases (note that the quantity $\Delta M/(2S)$ is the expectation value of the average over all sites of our original expansion parameter $\hat{n}_j/(2S)$, which we assumed to be small when we expanded the square roots in the HP expression). Thus we conclude

[10] Here (and also in the antiferromagnetic case to be considered later) we gloss over a minor subtlety associated with the $\boldsymbol{k} = 0$ wavevector which comes with an energy $\omega_{\boldsymbol{k}} = 0$.

that $M = 0$ (i.e. there is no ferromagnetic order) at finite (i.e. nonzero) temperatures for the Heisenberg model in one and two dimensions.

For the case of three dimensions, it can be shown[11] that $\Delta M \propto T^{3/2}$ as $T \to 0$. Thus spin-wave theory predicts that ferromagnetic order is stable (i.e. $M > 0$) at sufficiently low temperatures in three dimensions.

4 Spin-wave theory of antiferromagnets

We next turn to the antiferromagnetic case ($J > 0$ in (1)). As the ground state for classical spins has the spins on neighbouring sites pointing in opposite directions, one might naively guess that the ground state in the quantum case is analogous, thus having maximal and opposite spin projections $\pm S$ on neighboring sites. Such a state can be written

$$\prod_{j\in A} |S\rangle_j \prod_{l\in B} |-S\rangle_l \equiv |N\rangle. \tag{38}$$

Here A and B denote the two sublattices such that the spins on A (B) sites have spin projection S ($-S$), i.e. the states $|\pm S\rangle$ are eigenstates of S_z for the given lattice site with eigenvalue $\pm S$. It is however easy to see that $|N\rangle$ can not be the ground state, and is in fact not even an eigenstate, for the Heisenberg model for finite values of S: Acting with H in (1) on this state, the quantum fluctuation terms involving the spin raising and lowering operators change the state so that $H|N\rangle$ is not proportional to $|N\rangle$. Note that this did not happen in the ferromagnetic case because then the S^+ operators always killed the ferromagnetic ground state, leaving only the contribution from the S^zS^z part of the Hamiltonian. Consequently, quantum fluctuations play a much more important role in the antiferromagnetic case, as they change the ground state (and its energy) away from the "classical" result.

Although the ground state is not given by $|N\rangle$, it may still be that the ground state has antiferromagnetic order, i.e. that the spins on sublattice A point predominantly in one

[11] *The material in this footnote was not covered in the lectures.* In this case one can find the leading temperature dependence at low temperatures by writing the Bose-Einstein function in terms of a geometric sum as follows ($x \equiv |J|Sk^2/k_BT$):

$$\frac{1}{e^x - 1} = e^{-x}\frac{1}{1-e^{-x}} = e^{-x}\sum_{n=0}^{\infty}(e^{-x})^n = \sum_{n=1}^{\infty} e^{-nx},$$

and integrating over wavevectors up to infinity (i.e. $k_0 \to 0$ and $\bar{k} \to \infty$). This gives

$$\Delta M \approx \frac{1}{(2\pi)^3}\cdot 2\cdot 2\pi \int_0^{\infty} dk\; k^2 \sum_{n=1}^{\infty} e^{-n|J|Sk^2/k_BT}.$$

Changing integration variable to $u = n|J|Sk^2/(k_BT)$ the integral to solve is $\int du\, \sqrt{u}e^{-u} = (1/2)\sqrt{\pi}\mathrm{erf}(\sqrt{u}) - e^{-u}\sqrt{u}$ where erf is the so-called error function which satisfies $\mathrm{erf}(0) = 0$ and $\mathrm{erf}(\infty) = 1$. This gives finally

$$\Delta M \approx \frac{1}{8}\left(\frac{k_BT}{\pi|J|S}\right)^{3/2}\zeta(3/2),$$

where $\zeta(s) = \sum_n n^{-s}$ is the Riemann zeta function.

direction (taken to be the z direction here) and the spins on sublattice B point predominantly in the opposite direction. (If so the state $|N\rangle$ captures the structure of the true ground state at least in a qualitative sense.) To investigate this possibility we again develop a spin-wave theory based on expanding the square roots in the HP expansion. On the A sublattice where the spin projection in (38) is $+S$ we use the standard expressions:

$$S^+_{Aj} = \sqrt{2S - a^\dagger_j a_j}\; a_j, \tag{39}$$

$$S^-_{Aj} = a^\dagger_j \sqrt{2S - a^\dagger_j a_j}, \tag{40}$$

$$S^z_{Aj} = S - a^\dagger_j a_j. \tag{41}$$

However, on the B sublattice where the spin projection in (38) is $-S$ we must modify the HP expressions accordingly to reflect this:

$$S^+_{Bl} = b^\dagger_l \sqrt{2S - b^\dagger_l b_l}, \tag{42}$$

$$S^-_{Bl} = \sqrt{2S - b^\dagger_l b_l}\; b_l, \tag{43}$$

$$S^z_{Bl} = -S + b^\dagger_l b_l. \tag{44}$$

Compared to the expressions on the A sublattice, these modified expressions correspond to the changes $S^+ \leftrightarrow S^-$, $S^z \to -S^z$, which preserve the commutation relations, which shows that the HP expressions for sublattice B are indeed correct. Note that different boson operators a_j and b_l have been introduced for sublattices A and B respectively. The indices j and l run over the sites in A and B respectively. Inserting the HP expressions in the Hamiltonian, expanding the square roots and keeping terms to order S in the Hamiltonian, we get

$$\begin{aligned}\mathcal{H} &= J\sum_{j\in A}\sum_{\boldsymbol{\delta}}\left[\frac{2S}{2}(a_j b_{j+\boldsymbol{\delta}} + \text{h.c.}) + S(a^\dagger_j a_j + b^\dagger_{j+\boldsymbol{\delta}} b_{j+\boldsymbol{\delta}}) - S^2\right] \\ &+ J\sum_{l\in B}\sum_{\boldsymbol{\delta}}\left[\frac{2S}{2}(b_l a_{l+\boldsymbol{\delta}} + \text{h.c.}) + S(b^\dagger_l b_l + a^\dagger_{l+\boldsymbol{\delta}} a_{l+\boldsymbol{\delta}}) - S^2\right].\end{aligned} \tag{45}$$

Next we introduce Fourier-transformed operators

$$a_{\boldsymbol{k}} = \frac{1}{\sqrt{N_A}}\sum_{j\in A} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_j} a_j, \tag{46}$$

$$b_{\boldsymbol{k}} = \frac{1}{\sqrt{N_B}}\sum_{l\in B} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_l} b_l. \tag{47}$$

where $N_A = N_B = N/2$ is the number of lattice sites in each sublattice. The inverse transformation is

$$a_j = \frac{1}{\sqrt{N_A}}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j} a_{\boldsymbol{k}}, \tag{48}$$

$$b_l = \frac{1}{\sqrt{N_B}}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_l} b_{\boldsymbol{k}}. \tag{49}$$

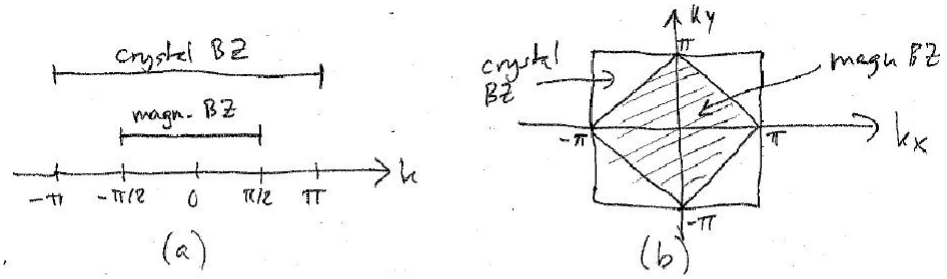


Figure 3: Crystal Brillouin zone and magnetic Brillouin zone in one dimension (a) and two dimensions (b).

The commutation relations are standard bosonic, i.e. the only nonzero commutators are $[a_{\boldsymbol{k}}, a^\dagger_{\boldsymbol{k}'}] = [b_{\boldsymbol{k}}, b^\dagger_{\boldsymbol{k}'}] = \delta_{\boldsymbol{k},\boldsymbol{k}'}$. We choose the $\boldsymbol{k}$-vectors to lie in the Brillouin zone associated with each sublattice. (Since the two sublattices are identical, their Brillouin zone is also identical). This Brillouin zone is called the magnetic Brillouin zone to distinguish it from the Brillouin zone associated with the full lattice which is called the crystal Brillouin zone. In one dimension periodic boundary conditions gives $a_{j+N_A\cdot 2\hat{x}} = a_j$ (the factor of 2 comes from the fact that the spacing between neigboring sites in the sublattice is 2, not 1) which gives $k_x = 2\pi n_x/(2N_A)$. Choosing the N_A values of n_x which lie closest to 0 (i.e. $n_x = -N_A/2, \ldots, N_A/2-1$) then gives $k_x \in [\pi/2, \pi/2\rangle$. Thus the length of the magnetic Brillouin zone is half the length of the crystal Brillouin zone $[\pi, \pi\rangle$. In two dimensions the two sublattices are oriented at a 45 degree angle with respect to the full lattice and have a lattice spacing which is $\sqrt{2}$ larger. It follows from this that the magnetic Brillouin zone is a square oriented at a 45 degree angle with respect to the crystal Brillouin zone and has half its area. These facts are illustrated in Fig. 3. In the following you should keep in mind that when we are discussing antiferromagnets, it is implicitly understood that all $\boldsymbol{k}$-sums are over the magnetic Brillouin zone.

Inserting (48)-(49) in the Hamiltonian and using $\sum_j e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_j} = N_A\delta_{\boldsymbol{k},\boldsymbol{k}'}$ etc., and also renaming the summation variable $\boldsymbol{k}$ as $-\boldsymbol{k}$ where needed, we get

$$\mathcal{H} = -NJS^2z/2 + JSz\sum_{\boldsymbol{k}}[\gamma_{\boldsymbol{k}}(a_{\boldsymbol{k}}b_{-\boldsymbol{k}} + a^\dagger_{\boldsymbol{k}}b^\dagger_{-\boldsymbol{k}}) + a^\dagger_{\boldsymbol{k}}a_{\boldsymbol{k}} + b^\dagger_{\boldsymbol{k}}b_{\boldsymbol{k}}], \tag{50}$$

where $\gamma_{\boldsymbol{k}}$ was defined earlier in (32). While the last two terms in the $O(S)$ part are in diagonal form, the first two terms are not; thus in contrast to the ferromagnetic case, Fourier transformation alone does not diagonalize the Hamiltonian in the antiferromagnetic case. To bring the Hamiltonian into diagonal form, we will perform another canonical transformation, known as a Bogoliubov transformation:

$$\alpha_{\boldsymbol{k}} = u_{\boldsymbol{k}}a_{\boldsymbol{k}} - v_{\boldsymbol{k}}b^\dagger_{-\boldsymbol{k}}, \tag{51}$$

$$\beta_{\boldsymbol{k}} = u_{\boldsymbol{k}}b_{\boldsymbol{k}} - v_{\boldsymbol{k}}a^\dagger_{-\boldsymbol{k}}. \tag{52}$$

Here $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ are two real functions of $\boldsymbol{k}$ which are to be determined. From the requirement that $[\alpha_{\boldsymbol{k}}, \alpha^\dagger_{\boldsymbol{k}'}] = [\beta_{\boldsymbol{k}}, \beta^\dagger_{\boldsymbol{k}'}] = \delta_{\boldsymbol{k},\boldsymbol{k}'}$ we get the condition

$$u^2_{\boldsymbol{k}} - v^2_{\boldsymbol{k}} = 1. \tag{53}$$

From the requirement $[\alpha_{\boldsymbol{k}}, \beta_{\boldsymbol{k}'}] = 0$ we get another condition, $u_{\boldsymbol{k}} v_{-\boldsymbol{k}} = u_{-\boldsymbol{k}} v_{\boldsymbol{k}}$. This is satisfied if $u_{-\boldsymbol{k}} = u_{\boldsymbol{k}}$ and $v_{-\boldsymbol{k}} = v_{\boldsymbol{k}}$, which will be assumed to hold in the following. The inverse transformation is given by

$$a_{\boldsymbol{k}} = u_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} + v_{\boldsymbol{k}} \beta^\dagger_{-\boldsymbol{k}}, \tag{54}$$

$$b_{\boldsymbol{k}} = u_{\boldsymbol{k}} \beta_{\boldsymbol{k}} + v_{\boldsymbol{k}} \alpha^\dagger_{-\boldsymbol{k}}. \tag{55}$$

By expressing the Hamiltonian in terms of the α and β bosons it can be rewritten as

$$\begin{aligned} \mathcal{H} &= -NJS^2 z/2 + JSz \sum_{\boldsymbol{k}} \{(2\gamma_{\boldsymbol{k}} u_{\boldsymbol{k}} v_{\boldsymbol{k}} + u^2_{\boldsymbol{k}} + v^2_{\boldsymbol{k}})(\alpha^\dagger_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} + \beta^\dagger_{\boldsymbol{k}} \beta_{\boldsymbol{k}}) + 2(\gamma_{\boldsymbol{k}} u_{\boldsymbol{k}} v_{\boldsymbol{k}} + v^2_{\boldsymbol{k}}) \\ &+ [\gamma_{\boldsymbol{k}} (u^2_{\boldsymbol{k}} + v^2_{\boldsymbol{k}}) + 2u_{\boldsymbol{k}} v_{\boldsymbol{k}}](\alpha_{\boldsymbol{k}} \beta_{-\boldsymbol{k}} + \alpha^\dagger_{\boldsymbol{k}} \beta^\dagger_{-\boldsymbol{k}})\}. \end{aligned} \tag{56}$$

Now we will choose $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ such that the term which is not on diagonal form, i.e. the term proportional to $\alpha_{\boldsymbol{k}} \beta_{-\boldsymbol{k}} + \alpha^\dagger_{\boldsymbol{k}} \beta^\dagger_{-\boldsymbol{k}}$, vanishes. Thus we require that

$$\gamma_{\boldsymbol{k}} (u^2_{\boldsymbol{k}} + v^2_{\boldsymbol{k}}) + 2u_{\boldsymbol{k}} v_{\boldsymbol{k}} = 0. \tag{57}$$

Furthermore we note that the condition (53) is automatically satisfied if we set

$$u_{\boldsymbol{k}} = \cosh \theta_{\boldsymbol{k}}, \quad v_{\boldsymbol{k}} = \sinh \theta_{\boldsymbol{k}}. \tag{58}$$

Inserting this into (57) gives[12] an equation which determines $\theta_{\boldsymbol{k}}$:

$$\tanh 2\theta_{\boldsymbol{k}} = -\gamma_{\boldsymbol{k}}. \tag{59}$$

Since $\gamma_{-\boldsymbol{k}} = \gamma_{\boldsymbol{k}}$ it follows that $\theta_{-\boldsymbol{k}} = \theta_{\boldsymbol{k}}$ which is consistent with our earlier assumption that $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ were even functions of $\boldsymbol{k}$. Using (58) and (59), after some manipulations we get

$$\begin{aligned} \mathcal{H} &= -NJS^2 z/2 - NJSz/2 + \sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}} (\alpha^\dagger_{\boldsymbol{k}} a_{\boldsymbol{k}} + \beta^\dagger_{\boldsymbol{k}} \beta_{\boldsymbol{k}} + 1) \\ &= E_0 + \sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}} (\alpha^\dagger_{\boldsymbol{k}} a_{\boldsymbol{k}} + \beta^\dagger_{\boldsymbol{k}} \beta_{\boldsymbol{k}}), \end{aligned} \tag{60}$$

where we have defined

$$\omega_{\boldsymbol{k}} = JSz\sqrt{1 - \gamma^2_{\boldsymbol{k}}} \tag{61}$$

and

$$E_0 = -NJS^2 z/2 - NJSz/2 + \sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}}. \tag{62}$$

Eq. (60) expresses the Hamiltonian in its final, diagonal form. The operators $\alpha^\dagger_{\boldsymbol{k}}$ and $\beta^\dagger_{\boldsymbol{k}}$ create magnon excitations with wavevector $\boldsymbol{k}$ and energy $\omega_{\boldsymbol{k}}$. These magnons are bosons, and are noninteracting in the approximation used here (i.e. when only including terms of $O(S^2)$ and $O(S)$ in the Hamiltonian). Note that in this antiferromagnetic case, for each $\boldsymbol{k}$ there are two types of magnons (α and β) which are degenerate in energy. On the other hand, the $\boldsymbol{k}$ sum goes over the magnetic Brillouin zone which only has $N/2$ $\boldsymbol{k}$ vectors, so

[12] We use that $\cosh^2 x + \sinh^2 x = \cosh 2x$ and $2 \cosh x \sinh x = \sinh 2x$.

the total number of magnon modes is $2 \cdot N/2 = N$, the same as for the ferromagnetic case. Note that as $\boldsymbol{k} \to 0$, $\omega_{\boldsymbol{k}} \to 0$ as in the ferromagnetic case, but unlike the ferromagnetic case, for which a quadratic dispersion $\omega_{\boldsymbol{k}} \propto |\boldsymbol{k}|^2$ was found in this limit, in the antiferromagnetic case we have instead a linear dispersion,

$$\omega_{\boldsymbol{k}} \propto |\boldsymbol{k}| \text{ as } \boldsymbol{k} \to 0. \tag{63}$$

In the ground state of $\mathcal{H}$ (call it $|G\rangle$) there are neither $\alpha_{\boldsymbol{k}}$ nor $\beta_{\boldsymbol{k}}$ magnons as these cost an energy $\omega_{\boldsymbol{k}} > 0$. Thus $|G\rangle$ can be defined by the relations

$$\alpha_{\boldsymbol{k}}|G\rangle = 0, \quad \beta_{\boldsymbol{k}}|G\rangle = 0, \text{ for all } \boldsymbol{k}. \tag{64}$$

This gives $\mathcal{H}|G\rangle = E_0|G\rangle$, i.e. the ground state energy is E_0, given in Eq. (62). The first term $-NJS^2z/2$, i.e. the term $\propto S^2$, is just the ground state energy E_{class} of a classical nearest-neighbor antiferromagnet of spins with length S. The other terms are $\propto S$ and represent quantum corrections to the classical ground state energy. Note that this quantum correction ΔE is *negative*:

$$\Delta E = E_0 - E_{\text{class}} = \sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}} - NJSz/2 = JSz \sum_{\boldsymbol{k}} \underbrace{[\sqrt{1-\gamma_{\boldsymbol{k}}^2} - 1]}_{<0}. \tag{65}$$

Thus quantum fluctuations lower the energy of the system. (Note that ΔE was 0 in the ferromagnetic case, i.e. there were no quantum fluctuations in the ferromagnetic ground state.)

Let us next investigate the amount of magnetic order in the system. Thus we need to identify an order parameter for antiferromagnetic order. Note that the magnetization $\boldsymbol{M} = (1/N)\sum_i \langle \boldsymbol{S}_i \rangle$ can not be used since it is zero in the presence of antiferromagnetic order, because the two sublattices give equal-magnitude but opposite-sign contributions to $\boldsymbol{M}$. Instead the natural order parameter is the so-called *sublattice magnetization*, defined by averaging $\langle \boldsymbol{S}_i \rangle$ only over the sites of one of the two sublattices. Without loss of generality, let's pick sublattice A, where the putative ordering is in the z direction. The magnitude of the sublattice magnetization is thus

$$M_A = \frac{1}{N_A} \sum_{j \in A} \langle S_j^z \rangle = S - \frac{1}{N_A} \sum_{j \in A} \langle a_j^\dagger a_j \rangle = S - \frac{1}{N_A} \sum_{\boldsymbol{k}} \langle a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}} \rangle. \tag{66}$$

Writing $M_A = S - \Delta M_A$, the correction ΔM_A to the classical result S is therefore given by

$$\Delta M_A = \frac{1}{N_A} \sum_{\boldsymbol{k}} \langle a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}} \rangle. \tag{67}$$

We need to rewrite this further in terms of the α and β type magnon operators, as it is in terms of them that the Hamiltonian takes its simple harmonic oscillator form. This gives

$$\Delta M_A = \frac{2}{N} \sum_{\boldsymbol{k}} [u_{\boldsymbol{k}}^2 \langle \alpha_{\boldsymbol{k}}^\dagger \alpha_{\boldsymbol{k}} \rangle + v_{\boldsymbol{k}}^2 \langle \beta_{-\boldsymbol{k}} \beta_{-\boldsymbol{k}}^\dagger \rangle + u_{\boldsymbol{k}} v_{\boldsymbol{k}} \langle \alpha_{\boldsymbol{k}} \beta_{-\boldsymbol{k}} + \text{h.c.} \rangle]. \tag{68}$$

Now we can calculate the expectation values. They are both thermal and quantum, i.e. $\langle O\rangle \equiv Z^{-1}\sum_m \langle m|O|m\rangle e^{-\beta E_m}$ where $Z = \sum_m e^{-\beta E_m}$. Here the sum $\sum_m$ is over the eigenstates $\{|m\rangle\}$ of $\mathcal{H}$ with energy eigenvalues E_m and $\beta = 1/(k_B T)$ where T is the temperature and k_B is Boltzmann's constant. First consider

$$\langle \alpha_{\boldsymbol{k}}\beta_{-\boldsymbol{k}}\rangle = Z^{-1}\sum_m \langle m|\alpha_{\boldsymbol{k}}\beta_{-\boldsymbol{k}}|m\rangle e^{-\beta E_m} \tag{69}$$

and its complex conjugate. The summand involves $\langle m|\alpha_{\boldsymbol{k}}\beta_{-\boldsymbol{k}}|m\rangle$, which is the overlap of the states $\beta_{-\boldsymbol{k}}|m\rangle$ and $\alpha^\dagger_{\boldsymbol{k}}|m\rangle$. Both states are eigenstates of $\mathcal{H}$, but as they clearly do not have identical sets of occupation numbers of the various α and β bosons, their overlap is zero. Hence $\langle \alpha_{\boldsymbol{k}}\beta_{-\boldsymbol{k}}\rangle$ and its c.c. are zero. Thus

$$\begin{aligned}
\Delta M_A &= \frac{2}{N}\sum_{\boldsymbol{k}}[u^2_{\boldsymbol{k}}\langle \alpha^\dagger_{\boldsymbol{k}}\alpha_{\boldsymbol{k}}\rangle + v^2_{\boldsymbol{k}}\langle \beta^\dagger_{\boldsymbol{k}}\beta_{\boldsymbol{k}}\rangle + v^2_{\boldsymbol{k}}] &(70)\\
&= \frac{2}{N}\sum_{\boldsymbol{k}}[n_{\boldsymbol{k}}\cosh 2\theta_{\boldsymbol{k}} + \frac{1}{2}(\cosh 2\theta_{\boldsymbol{k}} - 1)]\\
&= -\frac{1}{2} + \frac{2}{N}\sum_{\boldsymbol{k}}\left(n_{\boldsymbol{k}} + \frac{1}{2}\right)\frac{1}{\sqrt{1-\gamma^2_{\boldsymbol{k}}}}. &(71)
\end{aligned}$$

Here we used that $\langle \alpha^\dagger_{\boldsymbol{k}}\alpha_{\boldsymbol{k}}\rangle = \langle \beta^\dagger_{\boldsymbol{k}}\beta_{\boldsymbol{k}}\rangle = 1/(e^{\beta\omega_{\boldsymbol{k}}} - 1) \equiv n_{\boldsymbol{k}}$ and $\cosh 2\theta_{\boldsymbol{k}} = 1/\sqrt{1-\gamma^2_{\boldsymbol{k}}}$ (the last result follows from $\cosh^2 x = 1/(1-\tanh^2 x)$).

We will not analyze Eq. (71) in its full glory, but briefly consider a few important points. This expression has a temperature-dependent part coming from n_k (note that $n_k = 0$ at $T = 0$) and a temperature-independent part coming from the two terms containing the factor $1/2$. Let us first consider the case of zero temperature. In one dimension the $\boldsymbol{k}$-sum becomes (note $\gamma_k = \cos k$ in one dimension)

$$\frac{1}{N}\sum_k \frac{1}{\sqrt{1-\gamma^2_k}} \propto \lim_{k_0\to 0}\int_{k_0}^{\pi/2}\frac{dk}{\sin k}. \tag{72}$$

The most important contribution to this integral comes from the small-k region where we can approximate $\sin k \approx k$. Thus the leading term becomes $\int_{k_0} dk/k = -\log k_0 \to \infty$ as $k_0 \to 0$. Thus ΔM_A diverges even at zero temperature. We must therefore conclude that for this one-dimensional case our assumption that the system was magnetically ordered is invalid and so is our truncated spin-wave expansion. Note that this conclusion holds for any S. In two dimensions at zero temperature ΔM_A is still nonzero so quantum fluctuations do reduce the magnetization, but the correction turns out to be small enough, $\Delta M \approx 0.2$, so that even for the lowest spin, $S = 1/2$, spin-wave theory indicates that the system is ordered at zero temperature for a square lattice. This is also in agreement with other methods. In three dimensions ΔM is even smaller so the order is more robust then.

Next we consider finite nonzero temperatures. We just summarize the results that are obtained by analyzing (71) (which, we stress, is valid for a hypercubic lattice). In one dimension there is of course no antiferromagnetic order since none existed even at zero temperature. In two dimensions it turns out that the order does not survive at finite temperatures, so the spin-wave approach again is invalid. In three dimensions the system is ordered at sufficiently low temperatures.

5 Summary of spin-wave results

In the table below we summarize the main conclusions we have obtained from applying spin-wave theory to the Heisenberg ferromagnet and antiferromagnet on a hypercubic lattice in d spatial dimensions ($d = 1, 2, 3$). "Ordered" means spin-wave theory predicts magnetic ordering of the ferromagnetic/antiferromagnetic type, "disordered" means spin-wave theory predicts the absence of such order.

	Ferromagnet	Antiferromagnet
$d = 1$, $T = 0$	Ordered	Disordered
$d = 1$, $T > 0$	Disordered	Disordered
$d = 2$, $T = 0$	Ordered	Ordered
$d = 2$, $T > 0$	Disordered	Disordered
$d = 3$, $T = 0$	Ordered	Ordered
$d = 3$, $T > 0$	Ordered (at low T)	Ordered (at low T)

6 Broken symmetry and Goldstone modes

6.1 Broken symmetry

By definition, a symmetry transformation of a Hamiltonian is a transformation that leaves the Hamiltonian invariant (for more details, see the lecture notes named "Transformations and symmetries in quantum mechanics"). As an example, the Heisenberg Hamiltonian (1) considered in these notes is invariant under *global spin rotations*. Therefore we say that the Heisenberg Hamiltonian has a global spin rotation symmetry. A global spin rotation $(\boldsymbol{w}, \phi)$ is a transformation in which all the spin operators $\boldsymbol{S}_i$ are rotated by the angle ϕ around the axis $\boldsymbol{w}$ (more precisely, $\boldsymbol{w}$ is a unit vector that points in the direction of the rotation axis). The word 'global' refers to the fact that all spin operators are rotated in the same way. An intuitive argument for the invariance of the Heisenberg Hamiltonian under global spin rotations can be made by noting that if the spins had been just classical vectors, rotating them all in the same way preserves the angles between them and thus it also preserves the scalar products $\boldsymbol{S}_i \cdot \boldsymbol{S}_j$ in the Heisenberg Hamiltonian. We can show the invariance rigorously by making use of some results from the theory of transformations and symmetries in quantum mechanics (again, see the lecture notes with the same name). Since the rotation axis can be arbitrary, H is invariant provided that it commutes with the generator $\boldsymbol{S}_{\rm tot} \cdot \boldsymbol{w}$ for global spin rotations around any axis $\boldsymbol{w}$, which it will do if it commutes with the generators $S^x_{\rm tot}$, $S^y_{\rm tot}$, $S^z_{\rm tot}$ for global spin rotations around the x, y, and z axes. Here $\boldsymbol{S}_{\rm tot} = \sum_{i=1}^{N} \boldsymbol{S}_i$ is the total spin operator for the system. Let us consider the proof involving $S^z_{\rm tot}$; the proofs involving $S^x_{\rm tot}$ and $S^y_{\rm tot}$ can be done in exactly the same way. So we wish to show that $[H, S^z_{\rm tot}] = 0$. Consider any two spins 1 and 2 that interact with each other via a term $J_{12}\, \boldsymbol{S}_1 \cdot \boldsymbol{S}_2$ in the Heisenberg Hamiltonian. Their contribution to $[H, S^z_{\rm tot}]$ is proportional to

$$\begin{aligned}
[\boldsymbol{S}_1 \cdot \boldsymbol{S}_2, S^z_{\rm tot}] &= [S^x_1 S^x_2 + S^y_1 S^y_2 + S^z_1 S^z_2, S^z_1 + S^z_2] \\
&= [S^x_1, S^z_1] S^x_2 + [S^x_2, S^z_2] S^x_1 + [S^y_1, S^z_1] S^y_2 + [S^y_2, S^z_2] S^y_1 \\
&= -i S^y_1 S^x_2 - i S^y_2 S^x_1 + i S^x_1 S^y_2 + i S^x_2 S^y_1 = 0,
\end{aligned} \tag{73}$$

so the result $[H, S^z_{\text{tot}}] = 0$ follows immediately.

When the ground state of the Heisenberg Hamiltonian is magnetically ordered (either ferromagnetically for $J < 0$ or antiferromagnetically for $J > 0$), so that $\langle \boldsymbol{S}_i \rangle \neq 0$, this ground state is **not** invariant under a global spin rotation. For example, if we have a ferromagnetically ordered ground state with all the spins pointing in the z direction, rotating this state by an angle ϕ around a general axis $\boldsymbol{w}$ leads to a *different* ferromagnetically ordered ground state with all spins pointing in a different direction. As a concrete example, if we rotate all spins by 90 degrees around the negative x direction, after the rotation all the spins would point in the y direction. Clearly the nonzero order parameter $\boldsymbol{M}$ is also not invariant; it undergoes exactly the same rotation as the spins. The same conclusion also holds for the antiferromagnetic ground state and its order parameter.

To describe this situation we say that the ground state "spontaneously" **breaks** the global spin rotation symmetry of the Hamiltonian. More generally, spontaneous symmetry breaking refers to the situation when the ground state of the Hamiltonian is less symmetric than the Hamiltonian, i.e. the ground state is not invariant under all transformations that leave the Hamiltonian invariant.

For simplicity we focused on the ground state here. The notion of spontaneous symmetry breaking can however be generalized to systems at finite temperature: the equilibrium state of the Heisenberg model is said to exhibit spontaneous symmetry breaking if it has magnetic order, i.e. if its order parameter is nonzero. More generally, most known ordered phases of matter can be described in terms of broken symmetries.[13]

The fact that the symmetry can be spontaneously broken is actually a subtle issue in itself, and here we will only present a brief (and rather handwaving) discussion to hopefully shed some light on this. At first sight it may seem strange that symmetry breaking can happen at all. Any global spin rotation of a spin configuration will lead to a different, rotated configuration with exactly the same energy. Thus, as no directions are energetically more preferred than others, a statistical average over all directions should give $\langle \boldsymbol{S}_i \rangle = 0$ and thus no broken symmetry. The shortcoming of this argument is that it is only based on considering the energy of different states and does not take into account the extent to which the system can appreciably change its state during the relevant time scale for experimental measurements. Broken symmetry occurs when the system effectively gets trapped in a certain subset of states corresponding to a particular nonzero value of the order parameter. In this situation, to change the value of its order parameter would require essentially all spins to behave "in unison", which for a macroscopically large number of spins below the critical temperature becomes prohibitively unlikely, at least on the time scales of the measurement. Broken symmetry is thus related to the breakdown of the ergodic hypothesis (so-called ergodicity breaking).[14]

[13] Much research, both theoretical and experimental, is currently being done looking for novel phases of matter with more subtle types of order that cannot be described in terms of broken symmetries.

[14] For a more detailed discussion of these issues, see e.g. the discussion in Ch. 2 in N. Goldenfeld, "Lectures on phase transitions and the renormalization group", Westview Press, 1992.

6.2 Goldstone modes

Symmetries of a Hamiltonian can be classified as *discrete* or *continuous*. A symmetry is discrete if the associated symmetry transformations form a discrete set. In contrast, a symmetry is continuous if the associated symmetry transformations form a continuous set. The global spin-flip symmetry of an Ising model is an example of a discrete symmetry. The global spin-rotation symmetry of the Heisenberg model, on the other hand, is a continuous symmetry, since the rotations form a continuous set.

We found that both in the ferromagnetic and antiferromagnetic Heisenberg model, the magnon excitations have an energy $\omega_{\boldsymbol{k}}$ that goes to zero as $\boldsymbol{k} \to 0$. This implies that there is no *gap* in the excitation spectrum, i.e. there is no minimum nonzero energy cost to creating an excitation. We say that the magnon excitations are *gapless*. When the ground state of the Heisenberg model is magnetically ordered and thus breaks the continuous symmetry of the Heisenberg model, these gapless excitations are a consequence of the so-called Goldstone theorem,[15] and are therefore also sometimes referred to as Goldstone modes or Goldstone bosons. The Goldstone theorem implies that when the ground state breaks a *continuous* symmetry of the Hamiltonian, gapless bosonic excitations will exist in the energy spectrum.

[15] More precisely, a generalized version of this theorem due to Nielsen and Chadha.