Chapter 8: Models of magnetic materials: symmetry in quantum mechanics

I will provide in class further information clarifying which parts of these notes we will apply to a computing project. The written exercises and problems from these notes are for background and need not be turned in.

Interactions involving localized magnetic moments appear to play an important role in a number of new and still ill-understood materials, notably high-temperature superconductors, heavy-Fermion metals, and colossal-magneto-resistive perovskites. While one usually wishes to learn the effect of magnetism on itinerant (conduction) electrons, much can often be gained by looking at the fixed magnetic spins, whereas including the conduction electrons generally increases the complexity enormously. We shall examine some of the simplest spin problems and briefly discuss how to extend them into current research.

As a spin chain models a highly anisotropic (directional) material with a line of discretely-spaced magnetic moments, each interacting with its immediate neighbors. The magnetic moment at each site (which for convenience we’ll call spin) comes from a combination of electron spin and electron orbital angular momentum; nuclear moments play no important role. When the spin is very large, for instance 15/2 in the rare-earth metals dysprosium and erbium, classical models are appropriate.¹ For small spins, such as ½ or 1, however, quantum effects dominate.

We begin with a review of spin. Because spins can take on only particular values, their quantum description is actually simpler than that of moving particles and particularly amenable to numerical treatment. The student needs, however, to absorb some amount of formalism. We then examine how spatial and spin symmetries make the system computationally tractable and finally plan our computer assault.

1. Quantum-mechanical background

Some of you may not yet have taken a full-blown quantum-mechanics course. This section aims to give you the equipment you need to understand the problem of a magnetic spin chain; for derivations and a detailed understanding of the physics, see any elementary quantum-mechanics textbook, e.g., Liboff §11.6 or Feynman vol. 3. Those with weaker backgrounds especially should try the exercises, solutions to which will follow.

First consider just two interacting spins of any sort (classical or quantum). The very simplest energy we can imagine is just their dot product times some constant. Calling the spins $S_1$ and $S_2$,

$$H = J S_1 \cdot S_2 = J (S_{1x} S_{2x} + S_{1y} S_{2y} + S_{1z} S_{2z}).$$

(1)

This is the Heisenberg Hamiltonian; it takes its microscopic derivation from processes such as indirect exchange, superexchange, and double exchange, depending on the system, but we shall not concern ourselves with those processes here.² Certainly, (1) seems intuitively plausible if whatever interaction exists between spins can be Taylor expanded. If there is to be no externally-imposed directional dependence, (1) is the first term in this expansion, aside from a constant. (Indeed, for the case of quantum spin ½, it is the only term.)

At zero temperature, a system contains as little energy as quantum mechanics allows; all the rest has flowed to a reservoir. Thus we try to minimize $H$ in (1). A negative $J$ therefore means

¹ That large angular momenta approach classical behavior is most evident in the coherent-state representation.
that the spins wish to point in the same direction and so models ferromagnetism; a positive $J$
models antiferromagnetism, in which the spins prefer to point oppositely.

**Exercise.** The energy function (1) is solved trivially for classical continuous spins at zero
temperature. Convince yourself that for ferromagnetic (negative) $J$, the lowest-energy state has
all the spins parallel, while for antiferromagnetic (positive) $J$, spins alternate. The classical problem
starts to become interesting when we both turn on temperature and also add preferred spin
directions. The quantum problem is interesting all by itself.

There are two things you need to know about quantum mechanics:

1. When we measure a spin, it is either $\uparrow$ or $\downarrow$ (Stern-Gerlach).
2. Before we measure a spin, it has some probability amplitude to be $\uparrow$ and some probability
   amplitude to be $\downarrow$. The probability amplitude is a complex number, the modulus squared
   of which gives the probability.

We represent quantum spins $\frac{1}{2}$ as vectors (“spinors”) in the two-component space of $\hat{z}$
projections:

$$\psi = \begin{pmatrix} a \\ b \end{pmatrix} = a\psi_\uparrow + b\psi_\downarrow.$$  \hspace{1cm} (2)

An electron in state $\psi$ is in a superposition of states “spin up” and “spin down.” If we were
to measure the $\hat{z}$ component of spin, we would find it spin up ($+1/2$) with probability $|a|^2$ and
spin down ($-1/2$) with probability $|b|^2$. The expected spin is $+(1/2)|a|^2 - (1/2)|b|^2$. Evidently,
normalization requires $|a|^2 + |b|^2 = 1$, so the expected spin is $|a|^2 - 1/2$. After the measurement,
the spin would be in one of the two $S^z$ basis states, either $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Recall Dirac’s formalism, in which a measurement corresponding to an operator is effected by
taking the bracket of the complex conjugate transpose of the state vector, the operator, and then
the state vector:

$$\langle S^z \rangle = (a^* b^*) S^z \begin{pmatrix} a \\ b \end{pmatrix}.$$  \hspace{1cm} (3)

(The asterisk (*) represents complex conjugation, the action of reversing the imaginary part’s sign.)

**Exercise.** comparing (3) with the expected probabilities of a spin being up or down, show
that $S^z$ for a single spin has the matrix representation

$$S^z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$  \hspace{1cm} (4)

(Noted that we’re ignoring $\hbar$; really $S^z$ is $\hbar$ times what we just said, because electrons have angular
momentum $\pm \frac{1}{2}\hbar$.)

**Exercise.** Convince yourself that an $S^z$ basis state, either up, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, or down, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, is an
eigenvector (proper or “own” state) of the matrix $S^z$, while general linear combinations of these
basis states are not. An eigenvector $\mathbf{v}$ of a matrix $M$ satisfies $M\mathbf{v} = \lambda\mathbf{v}$, where $\lambda$ is just a number,
not a matrix, and is called the eigenvalue. The eigenvalues of the matrix $S^z$ are $\pm \frac{1}{2}$, the two
possible values for the $m$ quantum number.

Because any measurement leaves a system in an eigenstate of the corresponding operator,
these eigenstates take on a central significance in quantum mechanics. We have already seen that
the two possible eigenstates of $S^z$ form a basis for the vector space of possible configurations of
a single spin $\frac{1}{2}$ (see equation (2)). As another example, the eigenstates $E$ of an energy operator
(Hamiltonian) $H$ of a system, such as (1), are the only allowed levels of a system: the Schrödinger
equation

$$H\psi = E\psi$$  \hspace{1cm} (5)

is an eigenvector equation. Because spin systems have only finitely many possible states, $H$ is
easily represented as a matrix. In a “modern physics” course, you may have seen the Schrödinger
equation for a free particle. There, \( H = \frac{\hbar^2}{2m} \frac{d^2}{dx^2} \) plays the role of a matrix of infinite dimension, so what we are doing here really is enormously simpler.

(We review some of the correspondences between continuum wave mechanics and the simpler spin case. For a particle, we call \( \psi(x) \) the probability amplitude for finding the particle at position \( x \), because, as Feynman put it, “we don’t know what it means.” For each \( x \), \( \psi(x) \) is a complex number; since there are uncountably infinitely many positions \( x \), \( \psi \) corresponds (poetically) to an infinite-dimensional column vector. The probability of finding the particle in some interval \( [x_1,x_2] \) is the integral from \( x_1 \) to \( x_2 \) of the absolute value squared of \( \psi \): \( P(x_1,x_2) = \int_{x_1}^{x_2} |\psi(x)|^2 \, dx \). Since the probability of finding the particle somewhere is unity, \( \int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = 1 \). Roughly speaking, this adds the probability of finding the particle at position \( 0 \) plus that at position \( 0+\epsilon \) plus that at position \( 0+2\epsilon \) and so forth, including negative positions, too. In the case of spin-\( \frac{1}{2} \), there are only two possibilities, up and down, instead of infinitely many \( x \)’s. Therefore, we add just the probabilities that the spin is up and that it is down, arriving at the normalization condition \( \psi^* \psi + \psi\psi^* = \psi^\dagger \psi = 1 \), where \( \psi^\dagger \) means the row-vector version of \( \psi \), complex conjugated.)

While successive consecutive measurements of any one component of spin (\( e.g., S^z \)) will always yield the same result, consecutive measurements of perpendicular spin components are uncorrelated: if I measure \( S^z \) as spin up, a subsequent measurement of \( S^x \) will have a fifty-fifty chance of being “spin right” or being “spin left.” This means that if \( |z\rangle \) is any eigenvector of \( S^z \), either \( e^{i\theta} |\uparrow\rangle \) or \( e^{-i\theta} |\downarrow\rangle \) (but not a combination of them), where \( \theta \) is an arbitrary real number, the expectation value \( \langle z|S^z|z\rangle \) is zero. One can show that this implies that the expectation of \( S^z S^z + S^z S^x \) is zero for any state:

\[
\langle S^z S^z + S^z S^x \rangle = (a^* \ b^*) (S^z S^z + S^z S^x) \begin{pmatrix} a \\ b \end{pmatrix} = 0
\]

for any choice of \( a \) and \( b \).

Extending this argument\(^4\) lets us write, for spins \( \frac{1}{2} \),

\[
S^x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \quad S^y = \frac{i}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} .
\]

**Exercise.** Verify that the choice of spin matrices in equations (7) and (4) satisfies (6). It is likewise true that the expectation of \( S^z S^y + S^y S^z \) is zero, but the expectation of the square of any spin component is always +\( \frac{1}{4} \).

The following combinations of spin matrices are often very convenient:

\[
S^+ = S^x + iS^y ; \quad S^- = S^x - iS^y
\]

For spins-\( \frac{1}{2} \),

\[
S^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} , \quad S^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} .
\]

**Exercise.** To see why these are called raising and lowering operators, apply \( S^+ \) and \( S^- \) to the basis states \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \), representing an up spin, and \( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \), representing a down spin. By this, I mean to look at the vectors \( S^+ \psi \) and \( S^- \psi \) for each of the two basis states \( \psi \).

\(^3\) I have used Dirac’s “bracket” notation, in which \( |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \psi_\uparrow \) and \( |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \psi_\downarrow \).

\(^4\) A careful development of this subject would include a discussion of spin rotations and the group \( SU(2) \). For now, note that the “commutator” \( S^z S^x - S^x S^z = iS^y \) is non-zero (and its expectation value is imaginary, not real). If we could measure \( S^z \) and \( S^x \) simultaneously, the result would correspond to the operator \( \hat{S} \hat{S} \), which would be Hermitian and equal to \( S^z S^z \). It isn’t, they don’t, and we can’t.
Exercise. Use the matrix forms (4) and (7) and the definition (8) to rewrite the Heisenberg energy function (Hamiltonian) (1)

\[ H = J(S^z_1 S^z_2 + \frac{1}{2}(S^+_1 S^-_2 + S^-_1 S^+_2)) \]  

(10)

Let’s examine what each term in (10) does to a combination of spin states, \( \psi = \psi_1 \psi_2 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \). Note that the subscript-1 matrices act only on the subscript-1 vector, and likewise with the subscript-2 matrices. Matrices acting on the two spaces are not to be multiplied together. If you prefer, you can think of a four dimensional spin space\(^5\) spanned by basis vectors meaning “up\(_1\) up\(_2\),” “up\(_1\) down\(_2\),” “down\(_1\) up\(_2\),” and “down\(_1\) down\(_2\).” In fact, we’ll have to move to this way of thinking soon enough, so you’ll be ahead of the game.

The first term in (10) is called the Ising term. It says there’s an energy cost \( J/4 \) whenever the two spins point the same way, either both up or both down. When they point oppositely, the cost is \(-J/4\). If we were to stop here, we would say (for positive \( J \)) that the ground states \( \uparrow_1 \downarrow_2 \) and \( \downarrow_1 \uparrow_2 \) were degenerate with energy \(-J/4\), the excited states \( \uparrow_1 \uparrow_2 \) and \( \downarrow_1 \downarrow_2 \) likewise degenerate with energy \(+J/4\). This is essentially classical behavior. Note that \( S^z \) basis states (up or down) are eigenstates of the first term: it does not change either spin.

The next two terms allow for the spins to flip. If both spins point the same way, these two terms kill the vector and so fail to contribute (e.g., \( S^+_1 S^-_2 \uparrow_1 \downarrow_2 = 0 \) because \( S^+_1 \uparrow_1 = 0 \)). If one is up and the other is down, however, the action of these two terms will yield a state in which the one that was up is down and the one that was down is up. The spins have exchanged sign. This spoils the old conservation law of just the Ising piece: it is no longer true that the energy function (Hamiltonian) leaves the \( S^z \) components of the two spins invariant. In consequence, eigenstates of \( S^z_1 \) and \( S^z_2 \) will no longer be eigenvectors. However, a more general conservation law still applies: the total \( z \) component of spin, \( S^z_{\text{total}} = S^z_1 + S^z_2 \), is conserved, because \( H \) leaves alone the total number of up and down spins\(^6\).

We’ve been speaking rather abstractly; it’s better to try everything out explicitly. I mentioned above that we could think of the states of the two spins as vectors with four components, the first representing the probability amplitude for both spins being up, the second for \( \uparrow_1 \downarrow_2 \), etc. Expressing the four-dimensional basis in terms of the two two-dimensional ones,

\(^5\) This is the tensor product of the two original spin spaces.

\(^6\) There’s an additional, more subtle, conservation: the Hamiltonian does not change the total spin represented by the operator \( S \cdot S \). We’ll mention this again but not make use of it in this class.
\[ |1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 = |\uparrow\uparrow\rangle \]

\[ |2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 = |\uparrow\downarrow\rangle \]  \[(11)\]

\[ |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 = |\downarrow\uparrow\rangle \]

\[ |4\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}_2 = |\downarrow\downarrow\rangle . \]

We’re going to need four-dimensional matrices, too. By the rules of multiplying a row vector times a matrix times a column vector, the element of the matrix in the \(i\)th row and \(j\)th column is the result of multiplying the \(i\)th basis row vector times the operator times the \(j\)th basis column vector:

\[ S^z_1 S^z_2 = \begin{pmatrix} \langle 1 | S^z_1 S^z_2 | 1 \rangle & \langle 1 | S^z_1 S^z_2 | 2 \rangle & \langle 1 | S^z_1 S^z_2 | 3 \rangle & \langle 1 | S^z_1 S^z_2 | 4 \rangle \\ \langle 2 | S^z_1 S^z_2 | 1 \rangle & \langle 2 | S^z_1 S^z_2 | 2 \rangle & \langle 2 | S^z_1 S^z_2 | 3 \rangle & \langle 2 | S^z_1 S^z_2 | 4 \rangle \\ \langle 3 | S^z_1 S^z_2 | 1 \rangle & \langle 3 | S^z_1 S^z_2 | 2 \rangle & \langle 3 | S^z_1 S^z_2 | 3 \rangle & \langle 3 | S^z_1 S^z_2 | 4 \rangle \\ \langle 4 | S^z_1 S^z_2 | 1 \rangle & \langle 4 | S^z_1 S^z_2 | 2 \rangle & \langle 4 | S^z_1 S^z_2 | 3 \rangle & \langle 4 | S^z_1 S^z_2 | 4 \rangle \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \]  \[(12)\]

The off-diagonal elements of \( S^z_1 S^z_2 \) all vanish because the \( S^z \) operators flip no spins.\(^7\) The matrix (12) is the outer product of the matrices \( S^z_1 \) and \( S^z_2 \) given by (4). You may wish to look at our outer-product functions for Mathematica, Maple, or Matlab, available in the directory \home/5156/assignments/spins.

The remaining matrices are very easy to calculate. \( S^+_1 S^-_2 \) gives zero unless acting to the right on \( |3\rangle = |\downarrow\uparrow\rangle \) and to the left on \( \langle 2 | = \langle\uparrow\downarrow | \), so

\[ S^+_1 S^-_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ; \]  \[(13)\]

\( S^-_1 S^+_2 \) is similarly easily calculated (it’s the Hermitian conjugate, \((S^+_1 S^-_2)^\dagger \)), and we combine the matrices to get (from (10))

\[ H = \frac{J}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \]  \[(14)\]

\(^7\) This is because we are representing spins in the \( S^z \) basis. If we were representing spins in the basis consisting of “spin right” and “spin left,” this operator would flip them.
Problem 1. By hand or using a package such as Maple, find the eigenvectors and eigenvalues of this matrix. Compare with those for the Ising problem. If you choose to diagonalize the matrix by hand, note that it is block diagonal.

Notice the symmetry in every solution: under interchange of the spins, the wavefunction is either invariant or else changes sign. We could have seen this from the original Hamiltonian, (10); it is a consequence of Bloch’s or Floquet’s theorem or more deeply of Nöther’s. We call the symmetric solution momentum 0, the antisymmetric momentum $\pi$. The Hamiltonian never mixes the two momentum sectors, so we can solve them separately. The momentum-0 sector is spanned by the three vectors

$$
\begin{align*}
|s_1\rangle &= |\uparrow_1\uparrow_2\rangle = |1\rangle \\
|s_2\rangle &= \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle + |\downarrow_1\uparrow_2\rangle) = \frac{1}{\sqrt{2}}(|2\rangle + |3\rangle) \\
|s_3\rangle &= |\downarrow_1\downarrow_2\rangle = |4\rangle,
\end{align*}
$$

(15)

while the antisymmetric space is “spanned” by the single vector

$$
|a_1\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle) = \frac{1}{\sqrt{2}}(|2\rangle - |3\rangle).
$$

(16)

The Hamiltonian in the basis of the symmetric space is

$$
H_s = \begin{pmatrix}
\langle s_1|H|s_1\rangle & \langle s_1|H|s_2\rangle & \langle s_1|H|s_3\rangle \\
\langle s_2|H|s_1\rangle & \langle s_2|H|s_2\rangle & \langle s_2|H|s_3\rangle \\
\langle s_3|H|s_1\rangle & \langle s_3|H|s_2\rangle & \langle s_3|H|s_3\rangle
\end{pmatrix} = \frac{J}{4} \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix},
$$

(17)

while in the antisymmetric space

$$
H_a = \begin{pmatrix}
-\frac{3J}{4}
\end{pmatrix}.
$$

(18)

These two matrices are extremely easy to diagonalize: verify that the results agree with what you found in problem 1. Note furthermore the spin values of the three degenerate symmetric eigenstates, which are ground states for ferromagnetic ($J < 0$) interactions. The $H$-eigenstates $|s_1\rangle$, $|s_2\rangle$, and $|s_3\rangle$ are also eigenstates of $S^z_{\text{total}}$ (verify this!) with eigenvalues $+1$, $0$, and $-1$.\(^8\)

The fact that $H_s$ is already diagonal tips us off to another symmetry in the problem: as noted before, $H$ never alters $S^z_{\text{total}}$. Therefore we can split the symmetric space further into $S^z_{\text{total}} = +1$, $S^z_{\text{total}} = 0$, and $S^z_{\text{total}} = -1$ subspaces, each one-dimensional and therefore completely trivial. Note the overall effect: we have reduced the problem from that of diagonalizing one $4 \times 4$ matrix to that of four independent $1 \times 1$ “matrices.” Since matrix diagonalization is in general an $O(N^3)$ process, the former problem is $4^3/4 = 16$ times harder numerically. As we go to larger chains, the difference becomes enormous: without symmetries we simply could not calculate beyond a few sites.

2. More sites

The degenerate ground state of the two-spin antiferromagnetic ($J > 0$) Ising (classical) system included $|2\rangle = |\uparrow_1\downarrow_2\rangle$ and $|3\rangle = |\downarrow_1\uparrow_2\rangle$. These are not eigenstates of the quantum Hamiltonian,\(^8\) It takes not much more work to confirm that they are also all eigenstates of the operator $S^z_{\text{total}} = (S_1 + S_2) \cdot (S_1 + S_2)$, all with eigenvalue 2. You may recall that the combination of two spins-$\frac{1}{2}$ can have either spin $s = 0$ or $s = 1$, and that the eigenvalue of $S^2_{\text{total}} = s(s + 1)$. Our three states, therefore, lie all in the $s_{\text{total}} = 1$ sector and represent the three possible $m$ quantum numbers for a spin 1. We’ll not pursue this much further in this course, but it’s nice to see it working.
(10), and their energy expectation value \( \langle 2|H|2 \rangle = \langle 3|H|3 \rangle = -J/4 \) is larger than the ground-state energy, \(-3J/4\). Each state, \( |2 \rangle \) or \( |3 \rangle \), appears antiferromagnetic, but the peculiar nature of quantum mechanics permits a lower energy when the two combine antisymmetrically. Interestingly, the symmetric “antiferromagnetic” configuration \( (1/\sqrt{2})(|2 \rangle + |3 \rangle) \) has a high energy.

It comes as no surprise, therefore, that our first naive guess of the ground state of a large or an infinite chain of antiferromagnetically coupled spins, the Néel state \( |\uparrow \downarrow | \ldots \rangle \), is wrong. Of course, we see straight away that the \( S^+S^- \) terms in (10) mean the Néel configuration is not an eigenstate.\(^9\)

The surprise is the magnitude of the problem. Consider first a simulation of \( N = 36 \) classical vector spins.\(^{10}\) We would need to store two angles, real numbers, for each spin. The interaction energy between neighboring spins \( i \) and \( i+1 \), \( S_i \cdot S_{i+1} \), is the cosine of their relative angle. In some detached sense, the classical problem is larger than the quantum, for the classical spins can assume any of infinitely many configurations, the quantum only finitely many. Actually solving the classical case (which we can do by hand, but let’s imagine it numerically) is, however, a straightforward and fast exercise in minimization. In one approach, we start with a random configuration of the spins, then pick a spin (at random). Its neighbors put it in a potential, which we quickly minimize by changing that spin’s two angles. We repeat the exercise until the spins stop changing. If we double the number of sites, there are twice as many angles to twiddle, and the problem takes twice as long.\(^{11}\)

Contrast this to the spin-\(1/2\) quantum case, where at first we ignore symmetry. Each spin can point either up or down, two choices. With \( N \) sites, we have \( 2^N \) possible configurations. To diagonalize the resulting matrix takes of order \( (2^N)^3 \) operations; for \( N = 36 \), this is \( 2^{108} \sim 10^{32} \), which on the world’s fastest computer would take longer than the age of the universe.\(^{12}\)

Yet we can learn a great deal about the 36-site system. Our only hope is to reduce the number of configurations. That is where symmetry comes in. Let’s reverse the order of symmetries from our two-site example and try \( S_z \) total first. Our \( N \)-site chain (36 for concreteness) can have a total spin equal to any integer from \(-N/2 = -18\) to \(+N/2 = +18\), 37 sectors. The \( \pm 18 \) cases are pretty easy but not of interest. Indeed, for an even number of sites, we need not consider any but the \( S_z^{\text{total}} = 0 \) sector. Just as we saw in the two-site case, a state with some other \( S_z^{\text{total}} \) quantum number will be a member of a multiplet of states, all with the same \( S_z^{\text{total}} = s(s+1) \), energy, and other interesting properties but with \( S_z^{\text{total}} \) ranging from \(-s\) to \(+s\).\(^{13}\)

\(^9\) Bethe (1931) calculated the true ground state of an infinite spin-\(1/2\) antiferromagnetic chain: his elegant and very useful technique goes well beyond the level of the present course and is restricted to one-dimensional problems, while the numerical approach we follow here can be extended to two and three dimensions.

\(^{10}\) These spins have a fixed length but can point anywhere on the surface of a sphere. Do not confuse them with classical Ising spins, which, like quantum spins \(1/2\), take only two values.

\(^{11}\) Of course, we wouldn’t really solve this problem this way. However, we might imagine a more complicated classical problem in two or three dimensions where spins have bonds to close neighbors, not just nearest, some of which are ferromagnetic, others antiferromagnetic. Then such an approach becomes plausible. Add temperature and anisotropy, and it becomes necessary.

\(^{12}\) We’ve overstated the case by assuming, impossibly, the use of dense matrices. Even sparse-matrix techniques do not suffice. That’s just 36 spins. Yet nature manages \(10^{23}\) without difficulty. Thoughts such as these lead naturally to contemplation of a quantum computer, one whose spin elements would be applied to problems equally difficult for classical logic.

\(^{13}\) In equations (15) and (17), we had a triplet with \( s = 1 \), \( S_z^{\text{total}} = s(s+1) = 2 \), and energy \( E = J/4 \), while \( S_z^{\text{total}} \) took values \( \pm 1 \) and 0. If we need to know the degeneracy of a given state \( |x \rangle \) of energy \( E \), we need merely calculate its total spin using \( S_z^{\text{total}}|x \rangle = s(s+1)|x \rangle \). The degeneracy is the
Let’s measure the size of the reduced Hilbert space. Instead of $2^N$ configurations, we now count only those chains that have total spin projection $S_{\text{total}}^z = 0$. We have $S_{\text{total}}^z = 0$ if and only if the total number of up spins equals the total number of down spins. Think of the $N$ sites as $N$ boxes. We can fill $N/2$ of them with up spins; the other $N/2$ will (by default) end up with down spins. The order in which we fill half the boxes with up spins doesn’t matter, so the appropriate combination is

$$\binom{N}{N/2} = \frac{N!}{(N/2)!^2};$$

(19)

for $N = 36$, this is about $10^{10}$, more than seven times smaller than what we had before but still too big.\(^{14}\)

Next, we consider translational symmetry. Strictly speaking, a chain of 36 spins doesn’t have any, but we’re really interested in macroscopic, “infinite” chains. Such a chain differs from the one at hand most notably in lacking endpoints. What happens in a finite chain in the vicinity of the dangling ends does not usually reflect the physics we wish to understand. Therefore, we seek a way to get rid of the endpoints. The most typical technique links them up so that the chain becomes a circle. No point is any closer to the “end” than any other. It isn’t a perfect simulation of an infinite chain, but it’s often better than what we had.

It also has a translational symmetry we can use to reduce further the size of the Hilbert space. The Hamiltonian

$$H = J \sum_{i=0}^{N-1} (\mathbf{S}_i \cdot \mathbf{S}_{i+1})$$

(20)

(with the understanding that site $N$ is the same as site 0) is invariant under translation by one unit to the right. This implies a translational symmetry of the eigenvectors, too. Take for example a six-site chain. The state

$$| \uparrow_1 \uparrow_2 \downarrow_3 \uparrow_4 \downarrow_5 \downarrow_6 \rangle$$

(21)

should not appear in our basis, since it lacks the full translational symmetry of the Hamiltonian. However, the combination of states

$$\frac{1}{\sqrt{6}} (| \uparrow_1 \downarrow_2 \uparrow_3 \downarrow_4 \uparrow_5 \downarrow_6 \rangle + | \downarrow_1 \uparrow_2 \downarrow_3 \downarrow_4 \uparrow_5 \downarrow_6 \rangle + | \downarrow_1 \downarrow_2 \uparrow_3 \downarrow_4 \downarrow_5 \uparrow_6 \rangle + | \uparrow_1 \downarrow_2 \downarrow_3 \downarrow_4 \uparrow_5 \downarrow_6 \rangle + | \uparrow_1 \downarrow_2 \downarrow_3 \uparrow_4 \downarrow_5 \uparrow_6 \rangle + | \uparrow_1 \downarrow_2 \downarrow_3 \downarrow_4 \downarrow_5 \uparrow_6 \rangle),$$

(22)

should appear. The first term in (22) is (21). The second term is the result of starting with (21) and translating the whole chain one step to the right; similarly, the third term is the result of letting the translation operator act on the second. Eventually, we come back where we started.

\(^{14}\) We’ll mention briefly that we could look at the different $S_\text{total}^2$ sectors one at a time. The largest of these, $s = 0$, has a Hilbert space (for $N = 36$) another factor of 19 smaller than the $S_\text{total}^z = 0$ sector. As a practical matter, a single-$s$ subspace is awkward to deal with. The technique outlined by Ramasesha and Soos (reference supplied on request) applies well to one-dimensional chains when all one wants is the ground-state energy but loses its advantages in higher-dimensional problems, problems with long-range interactions, and when one needs eigenvectors, not just energies.
Unlike (21), (22) is invariant under translation. The numerical factor in front is just normalization, so that the dot product of the transposed version of (22) with itself is 1.

Now the symmetry of a wavefunction is a little more complicated than that of a Hamiltonian: because a measurable quantity is given by the bracket
\[
\langle v | \text{operator} | v \rangle = (\text{row-vector}\cdot v^*^T)(\text{matrix-operator})(\text{column-vector} \cdot v) ,
\]
the wavefunction \( v \) is determined only up to a complex factor of unit length, most generally
\[
e^{i\theta}
\]
for any real number \( \theta \). Thus, while (22) is invariant under a translation of one unit to the right, the action of translation by one unit to the right on a general basis vector may multiply the vector by a number of the form (24). With only six sites, \( \theta \) is restricted to \( \frac{2\pi n}{6} \) for integers \( n \). Because it derives from a translational symmetry, \( \theta \) is called the momentum.\(^\text{15}\) From the original state (21), therefore, we construct, in addition to (22) \((n = 0)\), basis states for three additional momenta,
\[
|\uparrow\uparrow\downarrow\uparrow\downarrow\uparrow, n\rangle = \frac{1}{\sqrt{6}} \left( |\uparrow\uparrow\downarrow\downarrow\uparrow\uparrow, n\rangle + e^{i\pi/3} |\downarrow\downarrow\uparrow\uparrow\downarrow\downarrow, n\rangle + e^{i2\pi/3} |\downarrow\downarrow\uparrow\downarrow\uparrow\uparrow, n\rangle + e^{i3\pi/3} |\downarrow\downarrow\uparrow\uparrow\downarrow\downarrow, n\rangle \right) ,
\]
for \( n = 1, 2, 3 \). Momenta \( 4\pi/3 \) and \( 5\pi/3 \) are the complex conjugates of \( 2\pi/3 \) and \( \pi/3 \). Since the spin chain has a mirror symmetry about the origin, the physics of negative momenta looks just like the physics at positive momenta.\(^\text{16}\)

The zero-momentum sector contains the largest number of basis vectors, the others fewer, since not every starting spin configuration, \( \text{e.g., } |\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\rangle \), admits every momentum. Let’s see how far translational symmetry has reduced the size of our Hilbert space. The configuration (21) we used as our example has a period equal to the length of the chain (six): it doesn’t coincide with itself until translated by six units. If every configuration had period equal to the chain length, \( N \), the reduction in Hilbert-space size would be a factor of \( N \). Since some configurations have periods less than \( N \), the reduction isn’t quite so large. However, it is still very significant. Remember that matrix diagonalization is an order \( N^3 \) process, so any way we can manage to reduce the Hilbert space is welcome.

We will use one more symmetry, one that works only because we are restricting ourselves to \( S^z_{\text{total}} = 0 \). Consider the parity operator, \( P \), which flips all the spins in a configuration. Parity is very much like a translation by half the length of the chain: since \( P^2 = 1 \), its only eigenvalues are \( e^{0i\pi} \) and \( e^{1i\pi} \), or \( \pm 1 \). The ket (25) is not an eigenvector of \( P \). To construct such an eigenvector, we act with \( P \) on (25):
\[
|\uparrow\downarrow\downarrow\uparrow\downarrow\downarrow, n, p\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\uparrow\downarrow\downarrow\downarrow, n\rangle \pm P|\uparrow\downarrow\downarrow\uparrow\downarrow\downarrow, n\rangle) = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\downarrow\uparrow\downarrow\downarrow, n\rangle \pm |\downarrow\downarrow\uparrow\uparrow\uparrow\uparrow, n\rangle) ,
\]

\(^{15}\) Let \( X \) be the translation operator. Then if \( X|v\rangle = e^{i\theta}|v\rangle \), we say that \( |v\rangle \) has momentum \( \theta \). The relationship between \( \theta \) and momentum as ordinarily construed is the subject of solid-state physics.

\(^{16}\) In a problem with an external magnetic field (which breaks time-reversal symmetry), we have to consider these conjugate momentum sectors, too.
where the eigenvalue \( p \) of \( P \) is \( \pm 1 \). Note (from equation (25)) that this expression has twelve terms.

Since the energy is the same for a flipped as for an unflipped configuration, we can find simultaneous eigenvectors of \( H \) and of \( P \), and \( H \) will not mix sectors of different parity. In fact, we can find simultaneous eigenvectors of translation (momentum), energy, \( S_{\text{total}}^z \), and parity. The use of parity symmetry results in a further modest reduction in the size of the Hilbert space.

Now that we have constructed a symmetry-reduced Hilbert space (for \( S^z = 0 \), a particular momentum, and parity plus or minus one), we need to build a Hamiltonian matrix. Let us label the various kets in the reduced space, of which (26) serves as an example, \( |0\rangle \), \( |1\rangle \), \( |2\rangle \), etc., and the various spin configurations of which they are composed \( |0\rangle \), \( |1\rangle \), \( |2\rangle \), etc. There are substantially more unprimed kets than primed. The Hamiltonian in the primed basis has an element in row \( i \), column \( j \) given by \( H'_{ij} = ' (i | H | j)' \). Most entries will be zero, so it’s not efficient to consider every pair of \( i \) and \( j \). Rather, we consider each ket \( j \) in turn, then act with the Hamiltonian (20) on each of the spin configurations in the state \( |j\rangle \). The \( S_k^z S_{k+1}^z \) term in the dot-product expansion (remember that \( k \) is a site index) will either add or subtract \( J/4 \) times a coefficient from the \( j\)-\( j\)-th diagonal term of the matrix \( H' \) for each unprimed ket in the expansion of \( |j\rangle \), depending on whether the unprimed ket has spins up-up or down-down, or up-down or down-up, on the sites \( k \) and \( k + 1 \). The coefficient of \( \pm J/4 \) is the absolute value squared of the coefficient with which the unprimed ket appears in the expansion of the primed ket.

Evaluating contributions from the off-diagonal terms \( S_k^z S_{k+1}^z \) and \( S_k^z S_{k+1}^- \) goes faster with a trick. Because we deal with only one set of symmetry numbers (momentum and parity, while spin-\( z \) always is zero) at a time, an unprimed ket that appears in the expansion of one primed ket will never appear in any other (in this momentum-parity sector). In shorthand, let’s call the expansion of a primed ket in terms of unprimed kets the orbit of the first unprimed ket in the expansion. (It resembles an orbit because under the action of translation and parity it comes back to the same place.) Then orbits do not intersect. Now when we consider the action of the off-diagonal elements on each of the unprimed kets in the \( j \)-th orbit, we check to see if the resulting unprimed ket is the first element of some new orbit (which may be the same orbit). If it is, we know by symmetry that the off-diagonal terms acting on our current orbit will generate the whole new orbit. If the state is not the first element of some orbit, we ignore it; it’s in some orbit certainly, but the off-diagonal operators acting on some other ket in the current orbit will give the first element of that orbit, so we’ll get it eventually. Naturally, we need to keep track of coefficients throughout.

### 3. Correlation functions

We have already seen that the ferromagnetic problem (\( J < 0 \)) has the classically expected ground state, \( |\uparrow\uparrow\ldots\uparrow\rangle \). How are we going to find this in the \( S^z = 0 \) sector, however? No state in this sector, after all, has any net magnetization: \( \langle \text{state} | S_{\text{total}}^z | \text{state} \rangle = 0 \). In the two-site case, there were three ferromagnetic ground states, (15), of which \( |s_2\rangle \) fell in the \( S^z = 0 \) sector. While it had no net magnetization in the \( \hat{z} \) direction, it did have magnetization in the \( \hat{x} \) direction. For a chain with many sites, the ferromagnetic state would seem rather more difficult to identify. The problem, however, is easily overcome. Instead of looking for net magnetization in any direction, we examine the correlation between two distant spins at sites \( i \) and \( j \). Since the chain is translationally invariant, this must be a function only of the difference between the two sites:

\[
C(i, j) = C(|i - j|) = \langle \text{ground state} | S_i \cdot S_j | \text{ground state} \rangle .
\]

---

\(^{17}\) Such operators are said to commute: expressed as matrices, \( X \) representing translation, \( H P X S_{\text{total}}^z = P S_{\text{total}}^z H X \) etc. This is not true of matrices or quantum operators in general, as we have already seen with spin components \( S^z \) and \( S^y \).

\(^{18}\) In a more sophisticated treatment, we would implement \( H' \) as a sparse matrix, storing only the non-zero values and diagonalizing with the Lanczos technique.
Such a function is called a site-site (or spin-spin) correlation. If $C$ is zero, it means the two sites are completely independent. So long as the sites $i$ and $j$ are distinct, the correlation (27) will be some number between $-3/4$ and $+1/4$.

As an example, consider the two-site ferromagnetic ground state $|s_2\rangle$ from equation (15). We calculate

$$C(1, 2) = \langle s_2 | S_1^+ S_2^- + \frac{1}{2} [S_1^+ S_2^- + S_1^- S_2^+] | s_2 \rangle$$

$$= \frac{1}{2} (\langle \uparrow \downarrow | + \langle \downarrow \uparrow |) [S_1^+ S_2^- (\uparrow \uparrow | + \downarrow \downarrow |) + S_1^+ S_2^- (0 + \downarrow \downarrow |) + \frac{1}{2} S_1^+ S_2^+ (| \uparrow \downarrow \rangle + 0)]$$

$$= \frac{1}{2} \left( \frac{1}{2} \langle \uparrow \downarrow | \uparrow \downarrow \rangle + \frac{1}{2} \langle \downarrow \uparrow | \downarrow \uparrow \rangle \right)$$

$$= \frac{1}{4}$$

The positive sign tells us, as we expected, that the two spins point in the same direction; that the magnitude is as large as possible tells us that they always point in the same direction in the ground state.

**Exercise.** Following the calculation (28), show that the spin-spin correlation for the state $|a_1\rangle$ in equation (16) is $-3/4$, indicating a perfect antiferromagnetic correlation.

The correlation is a function of distance. It seems reasonable to say that if $\lim_{x \to \infty} C(x) \neq 0$, the chain has long-range order, since no matter how far away the two spins, what happens on one site tells us what’s happening on the other site. If, on the other hand, $C(x)$ is non-zero only for small $x$ but goes to zero in the limit of a large distance, we would say that the chain has only short-range order. When looking for antiferromagnetic correlations, we need to account for the fact that the sign of $C$ will change depending on whether the distance between two sites is odd or even; the total number of sites in the chain should additionally be even for an antiferromagnetic chain unless we’re interested in studying effects of frustration.

We can never actually take the limit $x \to \infty$, but for any finite chain of $N$ sites, we can graph $C$ for $x = 1, 2, \ldots, N/2$. (Why can we go only so far as $N/2$?) We can then fit the resulting curve to a form

$$C_N(x) = C_N(\infty) + a_N e^{-x/r_0}$$

where $C_N(\infty)$ is the long-range correlation and $r_0$ the characteristic decay length. We graph our best fits for $C_N(\infty)$ against $N$ and fit that graph to a similar form to estimate $C_{\infty}(\infty)$. With only a few points, we shall need not only the best estimate of $C_{\infty}(\infty)$ but also a statistical estimate of its range. If $C_{\infty}(\infty)$ is further away from zero than our uncertainty in its measurement, the chain would appear to have long-range order.

4. Conclusions

Once we have diagonalized a spin chain, we have an enormous number of things to calculate, beginning with the energy density of states (a smoothed version of a scatter plot with one data point at the energy of each eigenvalue) and the dispersion relation giving energy as a function of momentum. With enough spins, we must abandon dense matrices and go to Lanczos techniques. We could also go to a higher dimension than one and consider further-than-nearest-neighbor bonds or random bonds. These possibilities all bring up the issue of frustration, in which the classical

---

19 Sometimes one makes a finer distinction based on the speed with which the correlation decays: $\lim_{x \to \infty} C(x)/e^{-x/r_0} = \text{constant}$ is short-range ordered (exponentially decaying), but $\lim_{x \to \infty} C(x)/x^{-\alpha} = \text{constant}$ is critical (algebraically decaying).

20 A technical word, not a description of the student’s efforts.
problem’s bonds cannot all be satisfied simultaneously. Alternatively, we could consider spins greater than one-half and Haldane’s results on spin order in such chains.\textsuperscript{21}

We should not leave you with the idea of exact diagonalization as the only approach to collections of spins. We have already mentioned the Bethe-Ansatz ground state as an exact solution. Sometimes one can construct a Hamiltonian for which an exact solution is known, then expand in some parameter to learn about a wider class of problems. Spin-wave theory, mean-field theory, the density-matrix renormalization group, Monte Carlo, and variational approaches, among others, have all been applied to spin chains.