

Notes for “Geometry and topology in many-particle systems’ Physics 250, University of California, Berkeley

(Dated: March 20, 2009)

I. TOPOLOGICAL PHASES I: THOULESS PHASES ARISING FROM BERRY PHASES, CONTINUED

We will give a very quick introduction to the band structure invariants that allowed generalization of the previous discussion of topological insulators to three dimensions. However, most of our discussion of the three-dimensional topological insulator will be in terms of emergent properties that are difficult to perceive directly from the bulk band structure invariant.

A. 3D band structure invariants and topological insulators

We start by asking to what extent the two-dimensional integer quantum Hall effect can be generalized to three dimensions. A generalization of the previous homotopy argument (from Avron, Seiler, and Simon, 1983) can be used to show that there are three Chern numbers per band in three dimensions, associated with the xy , yz , and xz planes of the Brillouin zone. A more physical way to view this is that a three-dimensional integer quantum Hall system consists of a single Chern number and a reciprocal lattice vector that describes the “stacking” of integer quantum Hall layers. The edge of this three-dimensional IQHE is quite interesting: it can form a two-dimensional chiral metal, as the chiral modes from each IQHE combine and point in the same direction.

Consider the Brillouin zone of a three-dimensional time-reversal-invariant material. Our approach will be to build on our understanding of the two-dimensional case: concentrating on a single band pair, there is a \mathbb{Z}_2 topological invariant defined in the two-dimensional problem with time-reversal invariance. Taking the Brillouin zone to be a torus, there are two inequivalent xy planes that are distinguished from others by the way time-reversal acts: the $k_z = 0$ and $k_z = \pm\pi/a$ planes are taken to themselves by time-reversal (note that $\pm\pi/a$ are equivalent because of the periodic boundary conditions). These special planes are essentially copies of the two-dimensional problem, and we can label them by \mathbb{Z}_2 invariants $z_0 = \pm 1$, $z_{\pm 1} = \pm 1$, where $+1$ denotes “even Chern parity” or ordinary 2D insulator and -1 denotes “odd Chern parity” or topological 2D insulator. Other xy planes are not constrained by time-reversal and hence do not have to have a \mathbb{Z}_2 invariant.

The most interesting 3D topological insulator phase (the “strong topological insulator”) results when the z_0 and $z_{\pm 1}$ planes are in different 2D classes. This can occur if, moving in the z direction between these two planes, one has a series of 2D problems that interpolate between ordinary and topological insulators by breaking time-reversal. We will concentrate on this type of 3D topological insulator here. Another way to make a 3D topological insulator is to stack 2D topological insulators, but considering the edge of such a system shows that it will not be very stable: since two “odd” edges combine to make an “even” edge, which is unstable in the presence of T -invariant backscattering, we call such a stacked system a “weak topological insulator”.

Above we found two xy planes with two-dimensional \mathbb{Z}_2 invariants. By the same logic, we could identify four other such invariants $x_0, x_{\pm 1}, y_0, y_{\pm 1}$. However, not all six of these invariants are independent: some geometry (exercise) shows that there are two relations, reducing the number of independent invariants to four:

$$x_0 x_{\pm 1} = y_0 y_{\pm 1} = z_0 z_{\pm 1}. \tag{1}$$

(Sketch of geometry: to establish the first equality above, consider evaluating the Fu-Kane 2D formula on the four EBZs described by the four invariants $x_0, x_{\pm 1}, y_0, y_{\pm 1}$. These define a torus, on whose interior the Chern two-form F is well-defined. Arranging the four invariants so that all have the same orientation, the A terms drop out, and the F integral vanishes as the torus can be shrunk to a loop. In other words, for some gauge choice the difference $x_0 - x_{\pm 1}$ is equal to $y_0 - y_{\pm 1}$.) We can take these four invariants in three dimensions as $(x_0, y_0, z_0, x_0 x_{\pm 1})$, where the first three describe layered “weak” topological insulators, and the last describes the Alternately, the “axion electrodynamics” field theory in the next subsection can be viewed as suggesting that there should be only one genuinely three-dimensional \mathbb{Z}_2 invariant.

For example, the strong topological insulator cannot be realized in any model with S_z conservation, while, as explained earlier, a useful example of the 2D topological insulator (a.k.a. “quantum spin Hall effect”) can be obtained from combining IQHE phases of up and down electrons. The impossibility of making an STI with S_z conservation follows from noting that all planes normal to z have the same Chern number, as Chern number is a topological

invariant whether or not the plane is preserved by time-reversal. In particular, the $k_z = 0$ and $k_z = \pm\pi/a$ phases have the same Chern number for up electrons, say, which means that these two planes are either both 2D ordinary or 2D topological insulators.

While the above argument is rigorous, it doesn't give much insight into what sort of gapless surface states we should expect at the surface of a strong topological insulator. The answer can be obtained by other means (some properties can be found via the field-theory approach given in the next section): the spin-resolved surface Fermi surface encloses an odd number of Dirac points. In the simplest case of a single Dirac point, believed to be realized in Bi_2Se_3 , the surface state can be pictured as "one-quarter of graphene." Graphene, a single layer of carbon atoms that form a honeycomb lattice, has two Dirac points and two spin states at each k ; spin-orbit coupling is quite weak since carbon is a relatively light element. The surface state of a three-dimensional topological insulator can have a single Dirac point and a single spin state at each k . As in the edge of the 2D topological insulator, time-reversal invariance implies that the spin state at k must be the T conjugate of the spin state at $-k$.

B. Axion electrodynamics, second Chern number, and magnetoelectric polarizability

The three-dimensional topological insulator turns out to be connected to a basic electromagnetic property of solids. We know that in an insulating solid, Maxwell's equations can be modified because the dielectric constant ϵ and magnetic permeability μ need not take their vacuum values. Another effect is that solids can generate the electromagnetic term

$$\Delta\mathcal{L}_{EM} = \frac{\theta e^2}{2\pi\hbar} \mathbf{E} \cdot \mathbf{B} = \frac{\theta e^2}{16\pi\hbar} \epsilon^{\alpha\beta\gamma\delta} F_{\alpha\beta} F_{\gamma\delta}. \quad (2)$$

This term describes a magnetoelectric polarizability: an applied electrical field generates a magnetic dipole, and vice versa. An essential feature of the above "axion electrodynamics" theory (cf. Wilczek PRL 1987) is that, when the axion field $\theta(t)$ is constant, it plays no role in electrodynamics; this follows because θ couples to a total derivative, $\epsilon^{\alpha\beta\gamma\delta} F_{\alpha\beta} F_{\gamma\delta} = 2\epsilon^{\alpha\beta\gamma\delta} \partial_\alpha (A_\beta F_{\gamma\delta})$ (here we used that F is closed, i.e., $dF = 0$), and so does not modify the equations of motion. However, the presence of the axion field can have profound consequences at surfaces and interfaces, where gradients in $\theta(x)$ appear.

A bit of work shows that, at a surface where θ changes, there is a surface quantum Hall layer of magnitude

$$\sigma_{xy} = \frac{e^2(\Delta\theta)}{2\pi\hbar}. \quad (3)$$

(This can be obtained by moving the derivative from one of the A fields to act on θ , leading to a Chern-Simons term for the EM field at the surface. The connection between Chern-Simons terms and the quantum Hall effect will be a major subject of the last part of this course.) The magnetoelectric polarizability described above can be obtained from these layers: for example, an applied electric field generates circulating surface currents, which in turn generate a magnetic dipole moment. In a sense, σ_{xy} is what accumulates at surfaces because of the magnetoelectric polarizability, in the same way as charge is what accumulates at surfaces because of ordinary polarization.

We are jumping ahead a bit in writing θ as an angle: we will see that, like polarization, θ is only well defined as a bulk property modulo 2π (for an alternate picture on why θ is periodic, see Wilczek, 1987). The integer multiple of 2π is only specified once we specify a particular way to make the boundary. How does this connect to the 3D topological insulator? At first glance, $\theta = 0$ in any time-reversal-invariant system, since $\theta \rightarrow -\theta$ under time-reversal. However, since θ is periodic, $\theta = \pi$ also works, as $-\theta$ and θ are equivalent because of the periodicity, and is inequivalent to $\theta = 0$.

Here we will not give a microscopic derivation of how θ can be obtained, for a band structure of noninteracting electrons, as an integral of the Chern-Simons form:

$$\theta = \frac{1}{2\pi} \int_{\text{BZ}} d^3k \epsilon_{ijk} \text{Tr}[\mathcal{A}_i \partial_j \mathcal{A}_k - i \frac{2}{3} \mathcal{A}_i \mathcal{A}_j \mathcal{A}_k], \quad (4)$$

which can be done by imitating our previous derivation of the polarization formula; for details see either Qi, Hughes, Zhang (2008) or Essin, Moore, Vanderbilt (2008). Instead we will focus on understanding the physical and mathematical meaning of the Chern-Simons form that constitutes the integrand, chiefly by discussing analogies with our previous treatment of polarization in one dimension and the IQHE in two dimensions. These analogies are summarized in Table I.

Throughout this section,

$$\mathcal{F}_{ij} = \partial_i \mathcal{A}_j - \partial_j \mathcal{A}_i - i[\mathcal{A}_i, \mathcal{A}_j] \quad (5)$$

	Polarization	Magnetoelectric polarizability
d_{\min}	1	3
Observable	$\mathbf{P} = \partial\langle H \rangle / \partial E$	$M_{ij} = \partial\langle H \rangle / \partial E_i \partial B_j$ $= \delta_{ij} \theta e^2 / (2\pi h)$
Quantum	$\Delta\mathbf{P} = e\mathbf{R} / \Omega$	$\Delta M = e^2 / h$
Surface	$q = (\mathbf{P}_1 - \mathbf{P}_2) \cdot \hat{\mathbf{n}}$	$\sigma_{xy} = (M_1 - M_2)$
EM coupling	$\mathbf{P} \cdot \mathbf{E}$	$M\mathbf{E} \cdot \mathbf{B}$
CS form	\mathcal{A}_i	$\epsilon_{ijk}(\mathcal{A}_i \mathcal{F}_{jk} + i\mathcal{A}_i \mathcal{A}_j \mathcal{A}_k / 3)$
Chern form	$\epsilon_{ij} \partial_i \mathcal{A}_j$	$\epsilon_{ijkl} \mathcal{F}_{ij} \mathcal{F}_{kl}$

TABLE I Comparison of Berry-phase theories of polarization and magnetoelectric polarizability.

is the (generally non-Abelian) Berry curvature tensor ($\mathcal{A}_\lambda = i\langle u | \partial_\lambda | u \rangle$), and the trace and commutator refer to band indices. We will understand the Chern-Simons form $K = \text{Tr}[\mathcal{A}_i \partial_j \mathcal{A}_k - i\frac{2}{3} \mathcal{A}_i \mathcal{A}_j \mathcal{A}_k]$ above starting from the second Chern form $\text{Tr}[\mathcal{F} \wedge \mathcal{F}]$; the relationship between the two is that

$$dK = \text{Tr}[\mathcal{F} \wedge \mathcal{F}], \quad (6)$$

just as \mathcal{A} is related to the first Chern form: $d(\text{Tr}\mathcal{A}) = \text{Tr}\mathcal{F}$. These relationships hold locally (this is known as Poincaré’s lemma, that given a closed form, it is *locally* an exact form) but not globally, unless the first or second Chern form generates the trivial cohomology class. For example, we saw that the existence of a nonzero first Chern number on the sphere prevented us from finding globally defined wavefunctions that would give an \mathcal{A} with $d\mathcal{A} = \mathcal{F}$. We are assuming in even writing the Chern-Simons formula for θ that the ordinary Chern numbers are zero, so that an \mathcal{A} can be defined in the 3D Brillouin zone. We would run into trouble if we assumed that an \mathcal{A} could be defined in the 4D Brillouin zone if the *first or second* Chern number were nonzero. Note that the electromagnetic action above is just the second Chern form of the (Abelian) electromagnetic field.

The second Chern form is closed and hence generates an element of the de Rham cohomology we studied earlier. There are higher Chern forms as well: the key is that symmetric polynomials can be used to construct closed forms, by the antisymmetry properties of the exterior derivative. In physics, we typically keep the manifold fixed (in our Brillouin zone examples, it is usually a torus T^n), and are interested in classifying different fiber bundles on the manifold. In mathematical language, we want to use a properly normalized cohomology form to compute a homotopy invariant (i.e., with respect to changing the connection, not the manifold). This is exactly what we did with the Chern number in the IQHE, which was argued to compute certain integer-valued homotopy π_2 invariants of nondegenerate Hermitian matrices.

More precisely, we saw that the $U(1)$ gauge-dependence of polarization was connected to the homotopy group $\pi_1(U(1)) = \mathbb{Z}$, but that this is connected also to the existence of integer-valued Chern numbers, which we explained in terms of π_2 . (These statements are not as inconsistent as they might seem, because our calculation of π_2 came down to π_1 of the diagonal unitary group.) We can understand the second Chern and Chern-Simons form similarly, using the homotopy invariants π_3 (gauge transformation in $d = 3$) and π_4 (quantized state in $d = 4$). The Chern-Simons integral for θ given above, in the non-Abelian case, has a $2\pi n$ ambiguity under gauge transformations, and this ambiguity counts the integer-valued homotopy invariant

$$\pi_3(SU(N)) = \mathbb{Z}, \quad N \geq 2. \quad (7)$$

In other words, there are “large” (non-null-homotopic) gauge transformations. Note that the Abelian Chern-Simons integral is completely gauge-invariant, consistent with $\pi_3(U(1)) = 0$.

The quantized state in $d = 4$ was originally discussed in the context of time-reversal-symmetric systems. The set \mathcal{Q} has one integer-valued π_4 invariant for each band pair, with a zero sum rule. These invariants survive even once T is broken, but realizing the nonzero value requires that two bands touch somewhere in the four-dimensional Brillouin zone. In this sense, the “four-dimensional quantum Hall effect” is a property of how pairs of bands interact with each other, rather than of individual bands. Even if this 4D QHE is not directly measurable, it is mathematically connected to the 3D magnetoelectric polarizability in the same way as 1D polarization and the 2D IQHE are connected.

The above Chern-Simons formula for θ works, in general, only for a noninteracting electron system. This is not true for the first Chern formula for the IQHE, or the polarization formula, so what is different here? The key is to remember that the 3D Chern formula behaves very differently in the Abelian and non-Abelian cases; for example, in the Abelian case, θ is no longer periodic as the integral is fully gauge-invariant. Taking the ground state many-body

wavefunction and inserting it into the Chern-Simons formula is not guaranteed to give the same result as using the multiple one-particle wavefunctions.

However, we can give a many-body understanding of θ that clarifies the geometric reason for its periodicity even in a many-particle system. Consider evaluating dP/dB by applying the 3D polarization formula

$$P_i = e \int_{BZ} \frac{d^3k}{(2\pi)^3} \text{Tr} \mathcal{A}_i. \quad (8)$$

to a rectangular-prism unit cell. The minimum magnetic field normal to one of the faces that can be applied to the cell without destroying the periodicity is one flux quantum per unit cell, or a field strength $h/(e\Omega)$, where Ω is the area of that face. The ambiguity of polarization (8) in this direction is one charge per transverse unit cell area, i.e., e/Ω . Then the ambiguity in dP/dB is

$$\Delta \frac{P_x}{B_x} = \frac{e/\Omega}{h/(e\Omega)} = \frac{e^2}{h} = 2\pi \frac{e^2}{2\pi h}. \quad (9)$$

So the periodicity of 2π in θ is really a consequence of the geometry of polarization, and is independent of the single-electron assumption that leads to the microscopic Chern-Simons formula.

C. Anomalous Hall effect and Karplus-Luttinger anomalous velocity

Our previous examples of Berry phases in solids have concentrated on insulators, but one of the most direct probes of the Berry phase of Bloch electrons is found in metals that break time-reversal symmetry. The breaking of T allows a nonzero transverse conductivity σ_{xy} to exist along with the metallic diagonal conductivity σ_{xx} . This ‘‘anomalous Hall effect’’ (AHE) can originate from several different microscopic processes. Here we will concentrate on the intrinsic AHE that results from Berry phases of a time-reversal-breaking band structure when the Fermi level is in the middle of a band.

Remarkably, the AHE originates from a term in the semiclassical equations of motion that is neglected in almost all textbooks. This term was first obtained by Karplus and Luttinger, but as this took place well before the modern idea of Berry phases, their results were not universally accepted. We will present a modern derivation of the Karplus-Luttinger term using the same idea as in our polarization calculation: trying to ‘‘gauge away’’ the Berry phase leads to a gauge-invariant physical effect. We will derive in this process the zero- B -field limit of the standard semiclassical equations of motion in, e.g., Ashcroft and Mermin,

$$\begin{aligned} \hbar \dot{\mathbf{k}} &= e\mathbf{E} + e\mathbf{v} \times \mathbf{B} \\ \hbar \mathbf{v} &= \nabla_k \epsilon_n(\mathbf{k}) + \dots \end{aligned} \quad (10)$$

where \dots indicate the Karplus-Luttinger term that we seek.

In class, we derived this term for an applied electric field in the form

$$\hbar \mathbf{v} = \nabla_k \epsilon_n \mathbf{k} - e\mathbf{E} \times (\nabla_k \times \mathcal{A}^{(n)}), \quad (11)$$

where $\mathcal{A}^{(n)}$ is the Berry vector potential of band n . The physical interpretation is fairly straightforward once we recall that our polarization calculation already showed that \mathcal{A} can be connected to the spatial distribution of the electron. As an electron wavepacket moves in k -space under the influence of an applied field, there are two contributions to its spatial velocity. The Karplus-Luttinger contribution describes how a change in \mathbf{k} induces a change in the real-space location because the Bloch states are changing; the first term describes how a fixed wavepacket of Bloch states still describes a moving particle.

(Our treatment followed closely chapter 4 of a tutorial by Ong and Lee that is on the course web page, so I will not reiterate it here. One note on their presentation: I believe that deriving their equation (17) requires assuming that the potential is weakly varying, because otherwise higher-order terms are not guaranteed to vanish because of some nonzero commutators.)

II. TOPOLOGICAL EFFECTS IN PATH INTEGRALS: THE HALDANE GAP

Finding the zero-temperature properties of an infinite number of interacting quantum particles is generally quite difficult. The only nontrivial successes are found in one dimension, and we will start by studying the Heisenberg spin-half antiferromagnetic chain as an example. Even in one dimension, not all problems of physical interest are exactly

solvable, and we will present the approach developed by Haldane to show that spin-one chains have qualitatively different behavior from spin-half chains. His argument involves topological properties of configurations in the path integral describing the spin chain. An alternate approach suggesting the same conclusion for the spin-one chain is also presented, based on a solvable model found by Affleck, Kennedy, Lieb, and Tasaki that is “near” the Heisenberg model.

(In class I quickly discussed the physical origin of the Heisenberg model from the Hubbard model; this is standard material and can be found, e.g., in the book of Auerbach, which is also a good reference for our treatment of the Haldane gap.)

The Heisenberg model we will concentrate on is

$$H = J \sum_i \mathbf{S}_i \cdot \mathbf{S}_{i+1}. \quad (12)$$

For $J > 0$ the ground state is quite complicated, while for $J < 0$ the ground state is simply that all spins align. More precisely, any of the states in the multiplet of maximum total angular momentum is a ferromagnetic ground state. For the antiferromagnet, we can get some insight by first considering the “Ising” case, where only one of the three spin components is retained in (12), say $S_i^z S_{i+1}^z$. The ground states in this case is easy to find: they are the Néel states $|\uparrow\uparrow\uparrow\dots\rangle$ and $|\downarrow\downarrow\downarrow\dots\rangle$. That these are ground states follows from noting that they minimize the energy of every bond independently; their energy is $-J/4$ per bond. We now turn to the question of what happens when we restore one more of the three spin components.

A. The XX chain via Jordan-Wigner mapping

In one dimension, there is a deep connection between the physics of fermions, bosons, and spins that does not hold in higher dimensions. There are several different hand-waving ways of expressing what is special in one dimension: one way to put it is that particle statistics are defined in terms of exchanges of particles, and in one dimension any exchange requires that the particles pass through each other (collide), which is not true in higher dimensions. One-dimensional systems are realized in carbon nanotubes, some organic compounds, and artificially fabricated “quantum wires.” One-dimensional quantum systems are also important because they sometimes describe experimentally important systems in higher dimensions (such as the Kondo effect: the low-energy physics is dominated by the spherically symmetric s channel, so that the problem is effectively radial), and because they provide solvable examples of interacting quantum problems.

The simplest such statistics-changing transformation in 1D is the Jordan-Wigner transformation. The z part is local (here $\hbar = 1$),

$$S_i^z = \frac{1 - 2c_i^\dagger c_i}{2}, \quad (13)$$

but we need a nonlocal “string” for raising and lowering operators because spins on different sites commute, while fermions anticommute. The representation we use is (with $S^\pm = S^x \pm iS^y$)

$$\begin{aligned} S_i^+ &= c_i^\dagger \prod_{j<i} (1 - 2c_j^\dagger c_j), \\ S_i^- &= \prod_{j<i} (1 - 2c_j^\dagger c_j) c_i. \end{aligned} \quad (14)$$

You can check that these preserve the spin commutation relations: in terms of Pauli matrices σ ,

$$[\sigma_i^+, \sigma_j^-] = \delta_{ij} \sigma_i^z, \quad [\sigma_i^z, \sigma_j^\pm] = \pm 2\delta_{ij} \sigma_i^\pm \quad (15)$$

and ordinary fermionic commutation relations. The reason the Jordan-Wigner transform works is very simple: the string is cooked up so that it changes sign from +1 to -1 depending on whether the number of fermions to the left of site i is even or odd.

Now we use the above Jordan-Wigner transformation to solve the so-called XX chain, which is like the Heisenberg model but with no z coupling:

$$H_{XX} = J \sum_i (S_i^x S_{i+1}^x + S_i^y S_{i+1}^y) = \frac{J}{4} \sum_i (\sigma_i^+ \sigma_{i+1}^- + \sigma_i^- \sigma_{i+1}^+). \quad (16)$$

Using the Jordan-Wigner transformation gives

$$\begin{aligned} H_{XX} &= \frac{J}{4} \sum_i \left[\prod_{j<i} (1 - 2c_j^\dagger c_j)^2 \left(c_i(1 - 2c_i^\dagger c_i)c_{i+1}^\dagger + c_i^\dagger(1 - 2c_i^\dagger c_i)c_{i+1} \right) \right] \\ &= \frac{J}{4} \sum_i \left(c_i(1 - 2c_i^\dagger c_i)c_{i+1}^\dagger + c_i^\dagger(1 - 2c_i^\dagger c_i)c_{i+1} \right). \end{aligned} \quad (17)$$

To simplify the terms in parentheses, note that the first $(1 - 2c_i^\dagger c_i)$ term gives a $-$ sign. So we are left with

$$H_{XX} = \frac{J}{4} \sum_i (c_{i+1}^\dagger c_i + c_i^\dagger c_{i+1}). \quad (18)$$

This is just a tight-binding model for a single fermion, which we know how to solve: the solution consists of a band of free fermions (in the continuum limit) with energies $-J/2 \leq E \leq J/2$.

So this very simple mapping tells us that the entire spectrum of the apparently nontrivial XX chain is given just by free fermions! There are two obvious $E = 0$ states: the state of all fermion states occupied (minimum S_z) and all fermion states empty (maximum S_z). The mapping to fermions also shows that not every bond can be optimized for this spin chain in the antiferromagnetic case: if every bond were optimized, the ground state energy would be $-\frac{J}{2}$ per bond, while the spinless fermions have worse energy than this since the average energy of the $N/2$ occupied states (N the number of sites) is greater than $-J/2$. We can compute the average energy using

$$\int_{-J/2}^0 dE ED(E) = \int_{-J/2}^0 dE E \frac{1}{2\pi\sqrt{(J/2)^2 - E^2}} = (J/2) \int_{-1}^0 dx \frac{x}{2\pi\sqrt{1-x^2}} = \frac{J}{4\pi} \sqrt{1-x^2} \Big|_{-1}^0 = \frac{-J}{4\pi}, \quad (19)$$

where $D(E)$ is the density of states $(1/2\pi)dk/dE$. However, we have to normalize this by dividing by the integral of $D(E)$,

$$\int_{-J/2}^0 dE D(E) = \int_{-1}^0 dx \frac{1}{2\pi\sqrt{1-x^2}} = \frac{1}{2\pi} \int_0^{\pi/2} \frac{\sin\theta}{\sin\theta} = 1/4, \quad (20)$$

so the average energy is $-J/\pi$.

Suppose now we consider the XXZ chain, which is obtained by adding a term $\lambda J s_i^z s_{i+1}^z$ to the XX Hamiltonian. Then for $\lambda = 0$ we obtain the XX chain and for $\lambda = 1$ the ordinary Heisenberg model. The $s_i^z s_{i+1}^z$ interaction becomes a four-fermion interaction in the fermionic representation. This interaction can still be handled but makes the problem much more complicated. We will study Bethe's solution of the Heisenberg point ($\lambda = 1$) as an example of how even strongly interacting problems can sometimes be solved in one dimension.

B. The Bethe solution at the isotropic point

So as not to be carrying around factors of J everywhere, we take the Hamiltonian in this section to be

$$H = \frac{1}{2} \sum_{j=1}^N (\sigma_j \cdot \sigma_{j+1} - 1). \quad (21)$$

This corresponds to $J = 2$ in (12), plus an energy offset. The energy offset means that the ferromagnetic state (all spins up, for example), will have energy 0. A lower bound for the ground state energy can be obtained by imagining that every pair of interacting spins could form a singlet simultaneously, which would then give energy -2 per bond. Based on the XX model above, it is useful to divide this into one term that hops a spin-flip, and another term that is an interaction between spin-flips, with a total number M of spin-flips:

$$H = \sum_{j=1}^N [-\sigma_j^+ \sigma_{j+1}^- - \sigma_j^- \sigma_{j+1}^+ + 2n_j n_{j+1}] - 2M. \quad (22)$$

(The $-$ sign appears because, following convention, we have performed the canonical transformation $s^x \rightarrow -s^x$, $s^y \rightarrow -s^y$, $s^z \rightarrow s^z$ on every other site. This would change k to $k + \pi$ in our solution of the XX model above but leave

all energies and other physical quantities invariant.) Here we have introduced n_j as the “number of spin flips on site j ”, so that the last term in the sum looks like an interaction between particles. The $-2M$ term is required because of how we have rewritten the σ_z term:

$$\frac{1}{2} \sum_j \sigma_j^z \sigma_{j+1}^z = \frac{1}{2} \sum_j (2n_j - 1)(2n_{j+1} - 1) = -2M + N/2 + \sum_j 4n_j n_{j+1} \quad (23)$$

and the constant term eliminates the constant term in (21).

As in the XX model above, we will build up the antiferromagnetic ground state as a system of $M = N/2$ delocalized spin-flips relative to the ferromagnetic state, but now these spin-flips will interact through interactions rather than just through their Fermi statistics. Our trial wavefunction, the “Bethe ansatz”, is

$$\Psi(x_1, x_2, \dots, x_M) = \sum_P A(P) \exp \left[i \sum_{j=1}^M k_{P_j} x_j \right]. \quad (24)$$

Here the x_j are sites on the one-dimensional lattice. The meaning of this wavefunction in terms of particles is that it gives us the location of the M spin-flips relative to the ferromagnet. It remains to determine the amplitude $A(P)$ for a given permutation P and also the momenta k_j . The amplitude will reflect the nature of collisions in one dimension: since elastic collisions can not change the momenta of the two particles, the collision’s effects are entirely captured by a phase shift. We define this as the phase $\theta(k, k')$ that appears in the solution of the two-particle problem

$$\Psi(x_1, x_2) = e^{i(kx_1 + k'x_2)} - e^{i(k'x_1 + kx_2) - i\theta(k, k')}. \quad (25)$$

Based on this, we define $A(P)$ so that, if the only change between permutations P and P' is that the two momenta k and k' are interchanged,

$$A(P')/A(P) = -e^{-i\theta(k, k')}. \quad (26)$$

We will need to compute this phase shift θ later. But first note that the energy is determined completely by the momentum distribution as long as there is some configuration where no adjacent sites are occupied, as is the case for $M \leq N/2$. Then all the energy is kinetic, and

$$E = 2 \sum_{j=1}^M (-1 - \cos k_j). \quad (27)$$

This is similar in form to the XX result (with the additional offset, and noting that here M is the number of particles, which is half the number of bonds), but the momentum distribution will be different. However, it remains to prove that there is an eigenstate with this energy and to determine the k distribution.

The phase shift θ describes the scattering of two particles that cannot sit on the same site (since any spin can be flipped at most once) and, if on adjacent sites, have a repulsive interaction energy 2. We will go a step farther and show that our many-particle Bethe ansatz wavefunction is an energy eigenstate, whose energy must then be given by (27); this is clear at least for $M \leq N/2$, where it is possible to have an initial state with no particles on adjacent sites. Consider the wavefunction $\Psi(x_1, \dots, x_M)$ defined above, and allow it to be defined even if any two x_i are identical (call this Ψ_g). The next step is somewhat subtle: if we did not require the wavefunction to vanish when any two x_i were identical, then the above kinetic energy (27) would apply and

$$H_{nonint} \Psi_g = E \Psi_g = -2M \Psi - \sum_{j=1}^M [\Psi_g(\dots, x_j + 1, \dots) + \Psi_g(\dots, x_j - 1, \dots)], \quad (28)$$

and the state would be an eigenstate of the Hamiltonian without the “interaction energy” term. We will use the interaction energy term to balance the fact that the wavefunction must in reality vanish when two particles are on the same site. In other words, we need that for a wavefunction with one nearest-neighbor pair,

$$H \Psi = E \Psi \Rightarrow H_{nonint} \Psi + H_{int} \Psi = E \Psi \Rightarrow 2\Psi + \Psi_g(\dots, x, x, \dots) + \Psi_g(\dots, x + 1, x + 1, \dots) = 0. \quad (29)$$

where the second two terms on the right side result because in reality there is no term contributing to the nearest-neighbor Ψ that came from having two particles on the same site. A similar equation, just with more terms, results if there is more than one nearest-neighbor pair.

If we look at our form for the wavefunction above, there are $M!$ terms. Given two particles x_a, x_b , these terms can be grouped into pairs where each pair assigns the same two momenta to the same two particles; the terms within a pair differ by an interchange of which particle gets which momentum. So each pair has the two-particle eigenstate $\Psi(x_a, x_b)$ as a common factor; in this sense the Bethe ansatz wavefunction is built up from two-particle wavefunctions. From the above, the many-particle wavefunction will be an eigenstate if the two-particle wavefunction satisfies

$$\Psi(x, x) + \Psi(x+1, x+1) + 2\Psi(x, x+1) = 0. \quad (30)$$

Inserting the two-body wavefunction from (25), we get

$$1 + e^{ik+ik'} + 2e^{ik'} = e^{-i\theta(k, k')} \left[1 + e^{ik+ik'} + 2e^{ik} \right] \quad (31)$$

so

$$-i\theta(k, k') = \log \left[\frac{1 + e^{ik+ik'} + 2e^{ik'}}{1 + e^{ik+ik'} + 2e^{ik}} \right] \quad (32)$$

Since $\arctan z = \frac{i}{2} \log[(1 - iz)/(1 + iz)]$, we can rewrite this if desired as

$$\theta(k, k') = i \log \left[\frac{\cos((k+k')/2) + 2e^{i(k'-k)/2}}{\cos((k+k')/2) + 2e^{i(k-k')/2}} \right] = -2 \arctan \left[\frac{\sin((k'-k)/2)}{\cos((k+k')/2) + \cos((k'-k)/2)} \right] \quad (33)$$

The phase shift takes a simpler form if we make the change of variables

$$e^{ik} = \frac{1 + 2i\alpha}{1 - 2i\alpha} \Leftrightarrow k(\alpha) = 2 \arctan(2\alpha), \quad (34)$$

so

$$\theta(k, k') = i \log \left[\frac{(1 - 2i\alpha)(1 - 2i\alpha') + (1 + 2i\alpha)(1 + 2i\alpha') + 2(1 + 2i\alpha')(1 - 2i\alpha)}{(1 - 2i\alpha)(1 - 2i\alpha') + (1 + 2i\alpha)(1 + 2i\alpha') + 2(1 + 2i\alpha)(1 - 2i\alpha')} \right] = i \log \left[\frac{4 - 4i(\alpha - \alpha')}{4 + 4i(\alpha - \alpha')} \right], \quad (35)$$

or simply

$$\theta(k, k') = 2 \arctan(\alpha - \alpha'). \quad (36)$$

The ‘‘fundamental equation’’ of the Bethe ansatz is derived from the rule the total phase shift in moving the particle around the system, a closed loop of N sites, must be a multiple of 2π . In other words,

$$e^{ikN} = \prod_{k' \neq k} \left[-e^{i\theta(k, k')} \right], \quad (37)$$

where the negative sign appears because of how we defined the phase shift above. Taking the logarithm of both sides, we obtain

$$kN = 2\pi I(k) + \sum_{k' \neq k} \theta(k, k'). \quad (38)$$

This is consistent that for free particles as in the XX model, we have momenta $k_j = 2\pi I_j/N$, $I_j = j - (M+1)/2$ ¹.

Now we come to the key mathematical step: we make the change of variables above to make a difference kernel for the integral equation that results from the continuum limit of (38). Being able to make such a change of variables is a large part of what separates solvable 1D models from others. For future reference, we note the inverse transformation

$$k(\alpha) = 2 \arctan(2\alpha), \quad k'(\alpha) = \frac{4}{1 + 4\alpha^2}, \quad (39)$$

¹ There is a subtle difference between bosons and fermions here; for an odd number M of bosons, or for fermions, we get $I(k)$ integer, while for an even number of bosons, $I(k)$ becomes half-integer. See Sutherland, *Beautiful Models*, chapter 6, for more details. We are constructing a wavefunction for bosons in this section, because the coordinates x_j label spin flips that should commute, but describing ‘‘hard-core bosons’’ or ‘‘fermions’’ in one dimension is a difference of only an overall factor $(-1)^P$.

and that the energy for a given momentum is

$$E_k = 2(-1 - \cos k) = -2\theta'(2\alpha) = -k'(\alpha). \quad (40)$$

The density of momenta $\rho(k)$ in the continuum limit then becomes a different function $R(\alpha)$ in terms of α :

$$R(\alpha) d\alpha = \rho(k) dk \Rightarrow R(\alpha) = k'(\alpha)\rho(\alpha). \quad (41)$$

The fundamental equation in terms of R , and dividing by N , is

$$k(\alpha) = 2\pi f(\alpha) + \int_{-\infty}^{\infty} \theta(\alpha - \alpha')R(\alpha') d\alpha', \quad (42)$$

where $f(\alpha)$ is the number of momenta less than α , as the integers $I(k)$ are assumed increase by one for each momenta, as in the free case. Here the boundary conditions run to $\pm\infty$ after converting the original boundary condition $k = \pm\pi$ to the α variables.

Taking the derivative of both sides with respect to α , we obtain

$$k'(\alpha) = 2\pi R(\alpha) + \int_{-\infty}^{\infty} \theta'(\alpha - \alpha')R(\alpha') d\alpha'. \quad (43)$$

We can solve this for $R(\alpha)$ by noting the convolution form of the integral equation. Dividing through by 2π , we have an equation of the form

$$(\mathbf{1} + K)R = \frac{k'}{2\pi}, \quad (44)$$

where $\mathbf{1}$ is the identity operator and K is the integral operator with kernel $\theta'/(2\pi)$. The solution of this equation is easily expressed in terms of Fourier components, since that converts the convolution integral to a product:

$$\tilde{R}_0 = \frac{1}{2\pi} \frac{\tilde{k}'}{1 + \tilde{K}}. \quad (45)$$

Now we have to do a few integrals. Since K is the integral operator with kernel $(1/\pi)(1 + (\alpha - \alpha')^2)^{-1}$, its Fourier transform is a simple exponential:

$$\tilde{K}(s) = \int_{-\infty}^{\infty} \frac{1}{\pi(1 + \alpha^2)} e^{is\alpha} d\alpha = \exp(-|s|), \quad \tilde{R}_0(s) = \frac{\exp(-|s|/2)}{1 + \tilde{K}(s)} = \frac{1}{2 \cosh(s/2)}. \quad (46)$$

where we have used also that the Fourier transform of $k'/(2\pi) = (2/\pi)(1 + 4\alpha^2)^{-1}$ is $\exp(-|s|/2)$.

From this expression for \tilde{R}_0 , we can Fourier transform back to obtain the ground state distribution of α 's,

$$R_0(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} ds e^{-is\alpha} \frac{1}{2 \cosh(s/2)} = \frac{1}{2 \cosh(\pi\alpha)} \quad (47)$$

and from that the average energy per site:

$$E_0 = \int E_k \rho(k) dk = - \int_{-\infty}^{\infty} \theta'(2\alpha) R_0(\alpha) d\alpha = - \int_{-\infty}^{\infty} k'(\alpha) R_0(\alpha) d\alpha = -2 \int_{-\infty}^{\infty} \frac{1}{\cosh(\pi\alpha)(1 + 4\alpha^2)} = -2 \log 2. \quad (48)$$

As expected, this is significantly higher than -2 , reflecting the fact that a spin cannot simultaneously form singlets with both its neighbors. To compare it to the XX and Ising values, we restore J , remove the offset, and change to the convention in the XX model above. We then get

$$E_{\text{Ising}} = -\frac{J}{4}, \quad E_{\text{XX}} = -\frac{J}{\pi} \approx -0.318J, \quad E_{\text{Heisenberg}} = -J \frac{4 \log 2 - 1}{4} \approx -0.443J. \quad (49)$$

With more work, one can confirm that the Heisenberg chain, like the XX one, is gapless and critical. Does this apply for a spin-1 chain as well?

C. Topological effects in spin chains