Draft Lecture III notes for Les Houches 2014

Joel E. Moore, UC Berkeley and LBNL

(Dated: August 7, 2014)

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_{+1}, y_0, y_{+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 $x_0 - x_{+1}$ is equal to $y_0 - y_{+1}$.) 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

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

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

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

(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 $\theta \to -\theta$ 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:

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

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,

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

	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 $K = \mathrm{Tr}[\mathcal{A}_i\partial_j\mathcal{A}_k - i\frac{2}{3}\mathcal{A}_i\mathcal{A}_j\mathcal{A}_k]$ above starting from the second Chern form $\mathrm{Tr}[\mathcal{F}\wedge\mathcal{F}]$; the relationship between the two is that

$$dK = \mathrm{Tr}[\mathcal{F}\wedge\mathcal{F}], \tag{6}$$

just as $\mathcal{A}$ is related to the first Chern form: $d(\mathrm{Tr}\mathcal{A}) = \mathrm{Tr}\mathcal{F}$. These relationships hold locally (this is known as Poincare'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

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

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

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

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

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

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,

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

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

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

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

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. A minor note is that the present author finds it highly non-trivial to explain how exactly the above term leads to a Hall effect once disorder is present; there are additional contributions that are not very easy to disentangle. A recent Reviews of Modern Physics by Nagaosa *et al.* explains the current state of knowledge in the field.