BOULDER SUMMER SCHOOL LECTURE NOTES TOPOLOGICAL SEMIMETALS AND SYMMETRY PROTECTED TOPOLOGICAL PHASES

ASHVIN VISHWANATH
HARVARD UNIVERSITY, USA.

In the previous lectures you have heard about topological insulators and topological superconductors. These are phases with an energy gap in the bulk, and can be described by free fermions. Here we will attempt to remove these restrictions one at a time. First we will close the band gap, while remaining close to a noninteracting model, and ask if topological properties can survive in semimetals or metals. We will see that the answer is yes, provided we also assume translation symmetry (i.e. a perfect crystalline lattice) and we will spend most of the first part of these lectures discussing the simplest realization of this physics, Weyl semimetals. Next, we will introduce interactions, but to keep things tractable, we will revert to 'insulators' with an energy gap in the bulk. We will show that there are interacting generalizations of topological insulators, the symmetry protected topological phases. A few years back we used to jokingly refer to the interacting SPT direction as 'Beyond the band-wagon' since it was so much less popular than topological band insulators. However, it has now developed into a major field but experiments still lags significantly behind theory, and we need more experimentally feasible proposals for their realization. Therefore, to allow for a broader understanding of these topics, we will adopt an intuitive viewpoint here - that of focusing of symmetry defects (duality) and their condensation - to describe these phases. While this is not the most general approach, it provides valuable intuition and will hopefully whet your appetite for the more formal and complete treatments that will follow in coming weeks.

PART 1 - Weyl Semimetals

1. Background Material

Where we review the theory of accidental degeneracies, relativistic wave equations for fermions and topological invariants of band insulators

In this section we will review two seemingly unrelated topics that both originated in the early days of quantum mechanics - relativistic wave equations for fermions, including the Dirac and Weyl equations and the problem of level repulsion and accidental degeneracies in quantum mechanics. We will see that these provide complementary perspectives on Weyl semimetals and are unified by the identification of their topological properties. To this end we will review topological invariants of insulators in the third part.

1.1. **Accidental Degeneracies.** In the early days shortly after the birth of quantum mechanics,von-Neumann and Wigner . asked the following question - under what conditions can a pair of energy levels be degenerate. In addition to symmetry enforced degeneracies, one could also have an accidental degeneracy, and in the latter case the question of relevance is - how many parameters does one need to tune to obtain the degeneracy? Given a spectrum $E(\lambda_i)$ as a function of some parameters λ_i [**?**], let us focus on a pair of energy levels 1, 2. Now change one parameter and ask if we can bring these levels into degeneracy. Naively it may appear that we can simply tune the energy difference to zero - but actually we need to consider generic perturbations of the Hamiltonian. Then, the energy levels (upto an overall constant) are determined by the most general 2×2 Hamiltonain: $\Delta H = \epsilon_z \sigma_z + \epsilon_x \sigma_x + \epsilon_y \sigma_y$, and the energy splitting is $\Delta E = 2\sqrt{\epsilon_x^2 + \epsilon_y^2 + \epsilon_z^2}$. Therefore, a degeneracy requires tuning all three terms to zero simultaneously. In general this cannot be accomplished by tuning just one parameter. In the absence of any symmetry, by the same logic, we need to tune three independent parameters to attain a degeneracy. This will be relevant for band structures where the number of independent parameters is the number of spatial dimensions of the crystal, each of which provides a crystal momentum. This was realized by Conyers Herring who was the first to apply the theory of accidental crossings to band structure in the1930s. The connection to topology may be less clear at the moment - but these points of degeneracy in the extended three parameter space are termed diabolic points and have been discussed in the context of Berry's phase [1]which provides the link to topological properties as we discuss below. Essentially, the presence of the degeneracy at the origin of parameter space (say) can be inferred by monitoring the energy levels on the surface of a sphere surrounding the origin.

There is an interesting connection between the cast of characters involved here, which is probably no coincidence. VonNeumann and Wigner went to the same high school in Budapest before they ended up as colleagues at Princeton. The only change - their names. Jeno Pal and Janos became Eugene and John respectively. Conyers Herring's thesis was entitled'Energy Coincidences in the Theory of Brillouin Zones', and his thesis advisor was of course Eugene Wigner.

Exercise 1: Consider a system, for example an atom, with spin orbit interactions such that each energy level is doubly degenerate (Kramers degeneracy). We would like to bring a pair of these degenerate pairs together by tuning parameters. How many parameters would you need to tune to do this - assuming that the only symmetry present is time reversal symmetry.

1.2. **Dirac and Weyl Fermions.** The Dirac equation is general dimension (let d denote spatial dimension, and we set the speed of light $c = 1$) is:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{1}$$

where $\mu = 0, 1 \ldots, d$ label time and space dimensions, and the $d+1$ Gamma matrices satisfy the anticommutation relation $\{\gamma^\mu, \gamma^\nu\} = 0$ for $\mu \neq \nu$ and $(\gamma^0)^2 = -(\gamma^i)^2 = \mathbb{I}$, where $i = 1, \ldots, d$. The minimal sized matrices that satisfy this property depend of course on the dimension, and are $2^{k+1} \times 2^{k+1}$ dimensional matrices in both spatial dimensions $d = 2k+1$ and $2k+2$. You are probably more familiar with following form of the Dirac equation in 3+1D:

$$i\partial_t \psi = -i\vec{\alpha} \cdot \vec{\nabla}\psi + m\beta\psi \tag{2}$$

we then identify $\gamma^0 = \beta$ and $\gamma^i = \beta\alpha^i$.

Hermann Weyl noticed that this equation can be further simplified [**?**] in *certain* cases (massless fermions) in *odd* spatial dimensions. For simplicity let us consider $d = 1$. Then we just need two anti-commuting matrices eg. the 2×2 Pauli matrices eg. $\gamma^0 = \sigma_z$ and $\gamma^1 = i\sigma_y$. Therefore the Dirac equation in $1+1$ dimension is involves a two component spinor and can be written as: $i\partial_t\psi = (\gamma^0\gamma^1 p + m\gamma^0)\psi$, where $p = -i\partial_x$. If we were describing a *massless* particle $m = 0$, this equation can be further simplified by simply picking eigenstates of the Hermitian matrix $\gamma_5 = \gamma^0\gamma^1 = \sigma^x$. If $\gamma_5\psi_\pm = \pm\psi_\pm$. We then have the 1D Weyl equation:

$$i\partial_t\psi_\pm = \pm p\psi_\pm \tag{3}$$

The resulting dispersion is simply $E_\pm(p) = \pm p$ which denotes a right (left) moving particle, which are termed chiral, or Weyl, fermions. Analogous dispersions arise at the one dimensional edge of an Integer Quantum Hall state, but are not allowed in an isolated one dimensional system where Chiral fermions must appear in opposite pairs. The fermion mass term interconverts the opposite chiralities. We will see that an analogous situation prevails in 3+1 dimensions - indeed the analogy with 1D fermions will be a theme that we will repeatedly return to.

In any odd spatial dimension $d = 2k+1$ one can form the Hermitian matrix $\gamma_5 = i^k\gamma^0\gamma^1 \ldots \gamma^d$. This is guaranteed to commute with the 'velocity' matrices $\gamma^0\gamma^i$, which can be simultaneously diagonalized along with the massless Dirac equation. At the same time it differs from the identity matrix since it anticommutes with γ^0. In even spatial dimensions, the latter property no longer holds, since all the Gamma matrices are utilized, and their product is just the identity.

Let us now specialize to $d = 3$. The Gamma matrices are, as Dirac found, now 4×4 matrices, and can be represented as $\gamma^0 = \mathbb{I} \otimes \tau_x$, $\gamma^i = \sigma^i \otimes i\tau_y$ and $\gamma^5 = -\mathbb{I} \otimes \tau_z$. Again, if we identify chiral components: $\gamma_5\psi_\pm = \pm\psi_\pm$, where $\psi_\pm$ are effectively two component vectors, we have for the massless Dirac equation:

$$\begin{aligned} i\partial_t\psi_\pm &= H_\pm\psi_\pm \\ H_\pm &= \mp\vec{p} \cdot \vec{\sigma} \end{aligned} \tag{4}$$

Thus Weyl fermions propagate parallel (or antiparallel) to their spin, which defines their chirality. We will see that a single chirality of Weyl fermions cannot be realized in 3D, but opposite pairs can arise, which is separated by momentum are the Weyl semimetals.

Here is an excerpt from an interesting (and comical) interview with Paul Dirac, when he visited Wisconsin in 1929. Dirac's monosyllabic answers are quite in keeping with all the other stories about him. But there is an interesting connection to Hermann Weyl - he mentions him as the one person he cannot understand!

Journalist: They tell me that you and Einstein are the only two real sure-enough high-brows and the only ones who can really understand each other. I wont ask you if this is straight stuff for I know you are too modest to admit it. But I want to know this — Do you ever run across a fellow that even you can't understand?"

Dirac: 'Yes,'

Journalist: 'This well make a great reading for the boys down at the office,' says I. 'Do you mind releasing to me who he is?'

Dirac: 'Weyl,'

The complete interview is here:

https : //www.math.rutgers.edu/ greenfie/mill_courses/math421/int.html

1.2.1. *Topological Invariants for Band Insulators.* Band theory describes the electronic states within a crystal in terms of one particle Bloch wave-functions $|u_n(\vec{k})\rangle$ that are defined within the unit cell and are labelled by crystal momentum $\vec{k}$ and band index n. The Berry phase of the Bloch wavefunctions is captured by the line integral of the Berry connection $\mathcal{A}_n(\vec{k}) = -i\langle u_n(\vec{k})|\nabla_{\vec{k}}|u_m(\vec{k})\rangle$, or equivalently the surface integral of the Berry flux: $\mathcal{F}_{ab}(\vec{k}) = (\partial_{k_a}\mathcal{A}_b - \partial_{k_b}\mathcal{A}_a)$. For a two dimensional insulator, the net Berry flux, which is a single component object, is quantized to integers:

$$\sum_n \int \frac{d^2\vec{k}}{2\pi}\mathcal{F}(\vec{k}) = N \tag{5}$$

which is directly related to the quantized Hall conductance. In a three dimensional crystal, the Berry flux behaves like a dual magnetic field $\epsilon_{abc}\mathcal{B}_c(\vec{k}) = \mathcal{F}_{ab}(\vec{k})$, switching the roles of position and momentum. The semiclassical equations of motion for an electron now take the form:

$$\dot{r}_a = \partial\mathcal{E}(k)/\partial k_a + \epsilon_{abc}\dot{k}_b\mathcal{B}_c \tag{6}$$

$$\dot{k}_a = -\partial\mathcal{V}(r)/\partial r_a - \epsilon_{abc}\dot{r}_b e B_c \tag{7}$$

The presence of the Berry phase term (the so called anomalous velocity) restores a symmetry that - however one difference between B and $\mathcal{B}$ the latter has magnetic monopoles. We will see that these are precisely Weyl points in the band structure.

Exercise 2: Using these semiclassical equations, how that for a 2D Chern band with Chern number C the Hall effect is quantized.

1.3. **Topological Properties of Weyl Semimetals.** *Where we apply the concepts developed in the previous section to characterize Weyl semimetals.* In topological semimetals, the conduction and valence bands coincide in energy over some region of the Brillouin Zone. A key input in determining conditions for such band touchings is the degeneracy of bands, which in turn is determined by symmetry. If spin rotation symmetry is assumed, eg. by ignoring spin orbit coupling, the bands are doubly degenerate. Alternately, doubly degenerate bands also arise if both time reversal $\mathcal{T}$ and inversion symmetry $\mathcal{I}$ ($r \to -r$) are simultaneously present in a three dimensional crystal.

Exercise 3: Show that the combination of inversion and time reversal leads to doubly degenerate bands in a 3D crystal. What about a 2D crystal where inversion is defined again as $r \to -r$ which is just a 180 degree rotation? Do we expect doubly degenerate bands here? If not, provide a simple counterexample.

Therefore we must break at least one of these two symmetries to obtain nondegenerate bands. The conditions for a pair of such nondegenerate bands to touch can be captured by discussing just a pair of levels whose effective Hamiltonian can generically be expanded as: $H(\vec{k}) = f_0(\vec{k})\mathbb{I} + f_1(\vec{k})\sigma_x + f_2(\vec{k})\sigma_y + f_3(\vec{k})\sigma_z$. To bring the bands in coincidence we need to set all three of the coefficients $f_1 = f_2 = f_3 = 0$ simultaneously, which, by the arguments of the previous section requires that we have three independent variables, i.e. that we are in three spatial dimensions. Then, we can expect band touchings, without any special fine tuning (although for certain choices of parameters there may be no band touchings in the Brillouin Zone).

We can further discuss the generic dispersion near the band touching point, by expanding the Hamiltonian about $\vec{k} = \delta\vec{k} + \vec{k}_0$. This gives:

$$H(\vec{k}) \sim f_0(\vec{k}_0)\mathbb{I} + \mathbf{v}_0 \cdot \delta\vec{k}\ \mathbb{I} + \sum_{a=x,y,z} \mathbf{v}_a \cdot \delta\vec{k}\ \sigma^a \tag{8}$$

where $\mathbf{v}_\mu = \nabla_k f_\mu(\vec{k})\big|_{\vec{k}=\vec{k}_0}$ is an effective velocity which is typically nonvanishing in the absence of additional symmetries. Note, if we revert to the special limit where $\mathbf{v}_a = \hat{a}$, we obtain the Weyl equation (9). We therefore refer to these band touchings as *Weyl nodes*. While this makes a connection to Weyl fermions with a fixed chirality (Chirality $= \text{Sign}\,(\mathbf{v}_x \cdot \mathbf{v}_y \times \mathbf{v}_z)$) it remains unclear why Weyl nodes should come in opposite chirality pairs. To realize this we need a topological characterization of Weyl nodes which is furnished by calculating the Berry flux on a surface surrounding the Weyl point.

Weyl points are monopoles of Berry flux (Chirality is charge - opposite charges and charge 2 Weyl node etc.) That is, if we consider the sphere surrounding a Weyl point, and think of this as a 2D band structure, it has a nonvanishing Chern number $C = \pm 1$. However, if we expand this surface so that it covers the entire Brillouin one, then by periodicity, it is actually equivalent to a point, and must have net Chern number zero. Therefore, the net Chern number of all Weyl points in the Brillouin Zone must vanish

1.3.1. *Weyl Semimetals with Broken $\mathcal{T}$ symmetry.* The simplest setting to discuss Weyl semimetals is to assume broken time reversal symmetry, but to preserve inversion. This allows for the minimal number of Weyl nodes, i.e. two, with opposite chirality. Inversion symmetry guarantees they are at the same energy and furthermore provides a simple criterion to diagnose the existence of Weyl points based on the parity eigenvalues at the time reversal invariant momenta (TRIMS).

Let us discuss this in the context of the following toy model. We imagine a magnetically ordered system so the bands have no spin degeneracy, but has a pair of orbitals on each site of a simple cubic lattices. Further assume that the orbitals have opposite parity (such as s, p orbitals), so τ_z, which is diagonal in the orbital basis, is required in the definition of inversion symmetry: $H(\vec{k}) \to \tau^z H(-\vec{k})\tau_z$.

$$
\begin{aligned}
H(\vec{k}) &= t_z(2 - \cos k_x - \cos k_y + \gamma - \cos k_z)\tau_z \\
&\quad + t_x \sin k_x \tau_x + t_y \sin k_y \tau_y
\end{aligned}
\tag{9}
$$

for $-1 < \gamma < 1$ we have a pair of Weyl nodes at location $\pm\vec{k}_0 = (0, 0, \pm k_0)$ where $\cos k_0 = \gamma$. The low energy excitations are obtained by approximating $H_\pm(\vec{q}) \approx H(\pm\vec{k}_0 + \vec{q})$ where we assume small $|\vec{q}| \ll k_0$. Then, $H_\pm = \sum_a v_a^\pm q_a \tau_a$ where $v^\pm = (t_x,\, t_y,\, t_z \sin k_0)$.

Note, at the 8 TRIM momenta $(n_x,\, n_y,\, n_z)\pi$ where $n_a = 0,\, 1$, only the first term in the Hamiltonian is active, and if $\gamma > 1$ the parity eigenvalues of all the TRIMs is the same. However, if eg. $\gamma = 0$, the parity eigenvalue of the Γ point is inverted. It is readily shown that this immediately implies Weyl nodes, i.e. an odd number of inverted parity eigenvalues is a diagnostic of Weyl physics. At the same time let us compare the Chern numbers $\Omega(k_z)$ of two planes in momentum space $k_z = 0$ and $k_z = \pi$. It is readily seen that the Chern number vanishes at $\Omega(k_z = \pi) = 0$, but $\Omega(k_z = 0) = 1$. As we increase $\gamma \to 1$, the entire Brillouin Zone is filled with unit Chern number along the k_z direction, and a three dimensional version of the integer quantum Hall state is realized [**?**, **?**]. Therefore the Weyl semimetal appears as a transitional state between a trivial and a topological insulator.

When the chemical potential is at $E_F = 0$, the Fermi surface consists solely of two points $\pm\vec{k}_0$. On increasing E_F, two nearly spherical Fermi surfaces appear around the Weyl points and we have a metal. The Fermi surfaces as 2D sections of the Brillouin Zone and compute the Berry Flux penetrating them. These turn out to be quantized to ± 1, which is the special feature characteristic of a Weyl *metal*. When $E_F > E^* = t_z(1 - \gamma)$ the Fermi surfaces merge through a Lifshitz transition and the net Chern number on the Fermi surface vanishes. At this point, we would cease to call this phase a Weyl metal. This discussion highlights the importance of the Weyl nodes being sufficiently close, compared to E^*, to the chemical potential. Ideally, we would like the chemical potential to be tuned to the location of the Weyl nodes just from stoichiometry, as happens for graphene.

1.4. **Physical Consequences of Topological Properties.** We have seen that there are topological aspects of Weyl semimetals which is most simply stated in terms of them being monopoles of Berry curvature. Here we will explore some of the consequences that topology. From experience with topological insulators and quantum Hall states, we are used to two different manifestations of topology. The first is to look for nontrivial surface states, and the second is to study response to an applied electric and/or magnetic field. We will follow this general guideline in this case as well. We will not be disappointed - indeed Weyl semimetals have an unusual response to electric and magnetic fields die to the previously known chiral anomaly. And they have have special surface states called Fermi arcs. The latter was not anticipated in the high energy literature as it requires a boundary.

1.4.1. *Fermi arc surface states.* . Surface states are usually associated with band insulators. Within the bulk band gap, surface states are well defined and exponentially localized near the surface. How can we define surface states when the bulk is gapless, as in Weyl semimetals? For this we need to further assume translation invariance, so we label surface states by crystal momenta within the 2D surface Brillouin zone (sBZ). Then, we only requires that there are regions of the sBZ which are free of bulk states at the same energy. Indeed if we consider the idealized limit of a pair of Weyl nodes at the chemical potential ($E_F = 0$) at momenta $\pm\vec{k}_0^*$ in the sBZ, once can define surface states at the same energy at all momenta except the projection of the Weyl points onto the sBZ. At those two point, surface states can leak into the bulk even at $E_F = 0$ and are not well defined. If we move away from this energy, the region occupied by bulk states will grow as shown in the FIGURE. The presence of these bulk states allow for surface states that are impossible to realize, not just in 2D but also on the surface of any fully gapped bulk band structure.

Now we can discuss the nature of the surface states that arise in Weyl semimetals, which, at $E_F = 0$ are Fermi arcs that terminate at $\pm\vec{k}_0^*$. These are a direct consequence of the fact that Weyl nodes are sources and sinks of Berry flux. Hence, if we consider a pair of planes at $k_z = 0$ and $k_z = \pi$ in model (9). Since they enclose a Weyl node, or Berry monopole, there must be a difference in Berry flux piercing these two planes that accounts for this source. Indeed, in model (9) we see that $k_z = 0$ ($k_z = \pi$) has Chern number $C = 1$ ($C = 0$). In fact, any plane $-k_0 < k_z < k_0$ will have Chern number $C = 1$, so each of the of 2D Hamiltonians $H_{k_z}(k_x, k_y)$, represents a 2D Chern insulator. If we consider a surface perpendicular to the x direction, we can still label states by k_z, k_y. The 2D Chern insulators H_{k_z} will each have a Chiral edge mode that will disperse as $\epsilon \sim vk_y$ near the Fermi energy. in the simplest model, v is independent of k_z as long as it is between the Weyl nodes. The Fermi energy $E_F = 0$ crosses these states at $k_y = 0$ for all $-k_0 < k_z < k_0$, leading to a Fermi arc that ends at the Weyl node projections on the sBZ, and in this particular model, is a straight line (symmetry of reflection and time reversal).

> Exercise: Show by explicit calculation on the microscopic model that fermi arc surface states exist.

It is worth mentioning an alternate continuum derivation of the Fermi arc surface states contained in the lecture notes (see Witten's Lectures), where suitable boundary conditions are formulated to characterize scattering of Weyl electrons from the boundary of the solid.

1.4.2. *Weyl Semimetals with Broken $\mathcal{I}$ symmetry.* If $\mathcal{T}$ is preserved then inversion symmetry must be broken to realize a Weyl semimetal. A key difference from the case of the

$\mathcal{T}$ broken Weyl semimetals is that the total number of Weyl points is always a multiple of four. This arises since under time reversal a Weyl node at $\vec{k}_0$ is converted into a Weyl node at $-\vec{k}_0$ with the *same* chirality. Since the net chirality must vanish, there must be another pair with the opposite chirality.

A useful perspective on $\mathcal{T}$ symmetric Weyl semimetals is to view them as the transition phase between a 3D topological insulator and a trivial insulator. The 3D TI is defined in the presence of just time reversal symmetry, but when inversion is also present there is a simple 'parity' criterion to diagnose band topology (Fu-Kane). To achieve a transition between a topological and trivial state, Kramers doublets with opposite parities cross each other at one of the TRIMS. their opposite parity eigenvalues allowing for such a crossing while tuning just a single parameter. At the transition point, a four fold degeneracy occurs at the TRIM, leading generically to a Dirac dispersion on moving away from it. Hence the transition between a trivial and topological insulator, in the presence of inversion, proceeds via a Dirac point. However, on breaking inversion, the Kramers doublets at the TRIM points can no longer cross each other while adjusting just a single tuning parameter. How then is the transition accomplished? The key observation is that the bands are now nondegenerate away from the TRIMs, and by our previous counting, can be brought into coincidence by tuning just the crystal momenta. Tuning an additional parameter to drive the transition involves moving the Weyl nodes annihilating them after suitably modifying the topology (see article by Murakami).

2. Summary

We have shown that a topological semimetal occurs when a submanifold of the Brillouin zone that avoids the band touching (a plane in this case) has a topological index (chern number) which varies on moving through the Brillouin zone. This requires the band gap to close leading to a topological semimetal. Note, this references the classification of gapped topological insulators, but in a different dimension than the physical dimension.

> Exercise: Is graphene a topological semimetal? If so what is the relevant classification? State which symmetries you assume.

Your feedback on these notes will be very helpful. In addition to catching typos and mistakes in equations - let me know if you have a better way to arrive at the physical conclusions or a more compact derivation of the results. Or, if you know of an interesting anecdote relating to this material!

References

[1] Berry, M V, 1985, "Aspects of Degeneracy in Chaotic behavior in quantum systems" , ed.Giulio Casatil, Plenum, New York, 123-140