Topological insulators: preliminaries[1],[2]

At least as far as currently known, a good qualitative understanding of the properties of the class of materials now known as topological insulators (TI's) can be achieved on the basis of a picture of noninteracting electrons subject to a particular kind of band structure, which generally speaking involves nontrivial effects of the spin-orbit interaction (SOI). (As a result, the experimentally observed TI's generally contain heavy elements where the SOI is important). Thus, with the virtue of (considerable!) hindsight, the theory of TI's may be regarded as simply traditional undergraduate solid-state physics seen in a new light; the "new light" focuses particularly on states of the surface of the system, which have not traditionally received much attention. The theory of TI's was originally developed in the context of 2D (or strongly layered) systems, and while some of the ideas can be generalized to 3D, most of the experiments have been done on 2D or quasi-2D systems, so I shall concentrate on those in these lectures. Let's start by reviewing some relevant elementary considerations.

First, let's review some basic properties of a system which is confined to a 2-dimensional Hilbert space. Any such system is formally equivalent to a particle of spin $1/2$, and we will say that it is described by a "pseudospin" vector $\hat{\sigma}$, such that any pure state is an eigenstate of the operator $\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}$ with eigenvalue ± 1, where $\mathbf{n}$ is some unit vector:

$$\mathbf{n} \cdot \hat{\boldsymbol{\sigma}} |\psi\rangle = \pm |\psi\rangle \tag{1}$$

We will usually choose to work in the basis of eigenstates of $\hat{\sigma}_z$; depending on the context, these may be actual eigenstates of the true spin (intrinsic angular momentum) of the electron in question, but in the context of the theory of TI's are at least as as likely to correspond to different Bloch bands (cf. below).One point to note is that since any operator in a 2D Hilbert space can be written as a linear combination of terms proportional to the unit matrix and to the three Pauli matrices $\sigma_i (i = x, y, z)$, and the term proportional to the unit matrix has no effect, then the Hamiltonian of any pseudospin $-1/2$ system can always be written in the form

$$\hat{H} = -\boldsymbol{\sigma} \cdot \mathcal{H} \tag{2}$$

where $\mathcal{H}$ is a "pseudo-magnetic field" (which, if the Hamiltonian is to be hermitian, must be a real "vector", i.e. have real components $\mathcal{H}_i, i = x, y, z$).

Potted review of elementary band theory: band inversion

Consider a crystalline lattice such that the potential felt by a single electron satisfies the condition

$$V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}) \tag{3}$$

[1] Hasan and Kane, Revs. Mod. Phys. 82, 3045(2010): Qi and Zhang, ibid. 83, 1057 (2011)

[2] B.A. Bernevig and T.L. Hughes. Topological Insulators and Topological Superconductors, Princeton Univ. Press (2013)

where $\mathbf{R} \equiv l\mathbf{a} + m\mathbf{b} + n\mathbf{c}$, with l, m, n integers and $\mathbf{a}, \mathbf{b}, \mathbf{c}$ the primitive lattice translation vectors. Then Bloch's theorem assures us that the single-electron eigenstates can be written in the form (ignoring spin for the moment)

$$\psi_{\mathbf{k},n} = \exp\ (i\mathbf{k} \cdot \mathbf{r}) \cdot u_{\mathbf{k},n}(\mathbf{r}) \tag{4}$$

where $u_{\mathbf{k},n}(\mathbf{r})$ is periodic with crystal periodicity, i.e.

$$u_{\mathbf{k},n}(\mathbf{r} + \mathbf{R}) = u_{\mathbf{k},n}(\mathbf{r}) \tag{5}$$

We call $\mathbf{k}$ the quasimomentum and n the band index. The energy $\epsilon_{\mathbf{k}n}$ associate with $\psi_{\mathbf{k}n}$ is in general a function of both $\mathbf{k}$ and n.

If $\mathbf{K}$ is a reciprocal lattice vector, i.e. a vector such that

$$\exp\ i\mathbf{K} \cdot \mathbf{R} = 1. \quad \forall \mathbf{R} \tag{6}$$

then we can choose the $u_{\mathbf{k}n}$ to be periodic in $\boldsymbol{K}$, i.e.

$$u_{\mathbf{k}+\mathbf{K},n}(\mathbf{r}) \equiv u_{\mathbf{k}n}(\mathbf{r}) \quad \forall \mathbf{r} \tag{7}$$

The three minimal values of $\boldsymbol{K}$ (the "primitive lattice vectors") define the first Brillouin zone (FBZ); it is possible (and we shall do it) to use the so-called "reduced zone" scheme, in which $\mathbf{k}$ by convention is confined to the FBZ and the energies plotted as $\epsilon_n(\mathbf{k})$ within this zone.

Let us first consider the kind of simple situation indicated (in a 1D cross-section) in fig. 1,

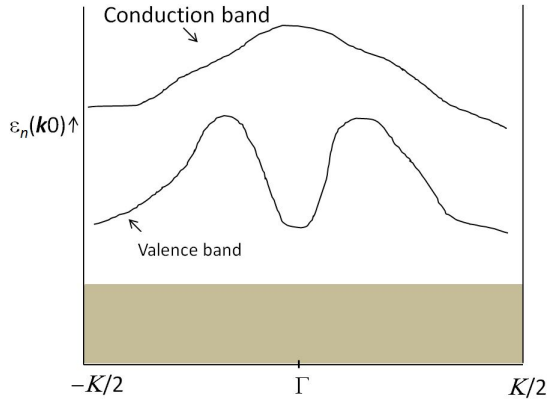


Fig. 1

i.e. a case in which for any given value of $\mathbf{k}$ there are two relevant bands, with the conduction-band state always above the valence-band state:

$$\epsilon_{\text{cond}}(\mathbf{k}) > \epsilon_{\text{val}(\mathbf{k})}, \quad \forall \mathbf{k} \tag{8}$$

Then according to the standard textbook analysis there are three possibilities, depending on the structure of the bands and the number of electrons filling them:

1. if there is no band overlap, i.e. if the quantity Δ defined by

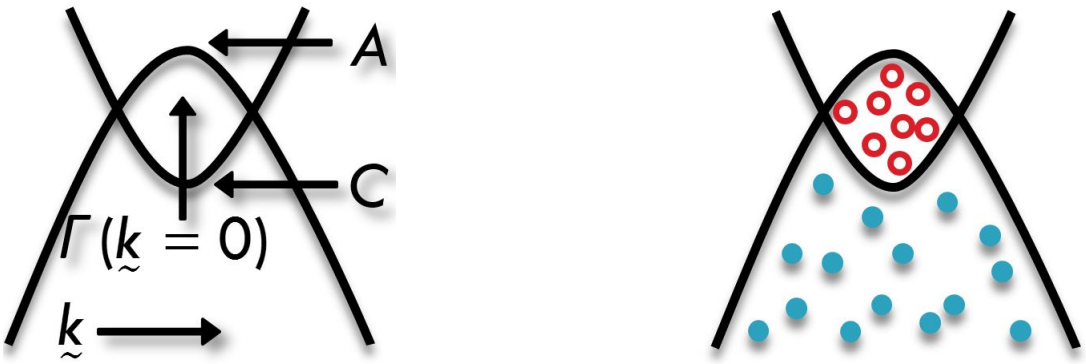
$$\Delta \equiv min_{\mathbf{k}}(\epsilon_{\rm cond}(\mathbf{k})) - max_{\mathbf{k}}(\epsilon_{\rm val}(\mathbf{k})) \tag{9}$$

   is positive, then

   (a) if the total number of electrons in the crystal is first enough to fill the valence band (i.e. $N_{el}/2N_{site} =$ integer[3]), then the system is a simple "band" insulator (or for small enough Δ a semiconductor) with band gap Δ.

   (b) if the number of electrons does not satisfy this condition, (and in particular, if $N_{el}/2N_{site} = n + 1/2$ as is a common state of affairs), then either the valence or the conduction band is partly filled, and the system is a metal.

2. if there is band overlap, i.e. if $\Delta < 0$, then both the conduction and the valence bands are partially filled; the usual name for this case is a "semimetal", and it will basically show metal-like conduction.



(a) (b)

Fig. 2

More interesting in the present context is the case where condition (8) is violated for some value $\mathbf{k}$, and moreover the condition $N_{el}/2N_{site} =$ integer is satisfied. What happens then? Lets us suppose for simplicity that in the absence of hybridization between bands, we have

$$\varepsilon_{\rm val(\mathbf{k})} = A - Bk^2, \epsilon_{cond}(\mathbf{k}^2) = C + Dk^2, \quad A > C, B + D > 0 \tag{10}$$

[3] At this point we need of course to invoke the (real) spin degree of freedom to allow population of each Bloch state by two electrons.

so that the situation is roughly as shown in fig.(2a).

Then the filling will look like the picture in fig. (2b), where the black dots represent the valence band and the light dots represent the conduction band. We can represent this situation in the language of a 2D "pseudospin" space by assigning pseudospin $|\uparrow\rangle$ to the valence band and pseudospin $|\downarrow\rangle$ to the conduction band; then, subtracting out a term which is proportional to the unit matrix, we find that for given $\mathbf{k}$ the effective Hamiltonian can be written

$$\hat{H}_k = -\boldsymbol{\sigma}_k \cdot \mathcal{H}_k \tag{11}$$

$$\mathcal{H}_{\mathbf{k}} \equiv -\hat{\mathbf{z}}(\varepsilon_{\rm cond}(\mathbf{k}) - \varepsilon_{val}(\mathbf{k})) = (A - C) - (B + D)k^2 \tag{12}$$

and in the groundstate $\boldsymbol{\sigma}_k$ will point along the direction of $\mathcal{H}_k$, i.e. along $\pm\hat{\mathbf{z}}$. The crucial point is that $\mathcal{H}_k$ changes sign at the points k_c given by

$$k_c^2 = \frac{A - C}{B + D} \tag{13}$$

(which is the full 3D case form the surface of a sphere). Hence in the pseudospin language a cross-section passing through the origin will look like

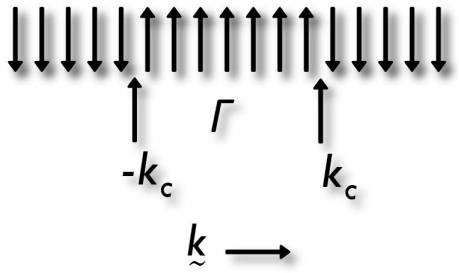


Fig. 3

So far, the energy "gap", which is twice the magnitude of $\mathcal{H}_k$, goes to zero at k_c, so the system is a semimetal just as in case(2) above! (cf. fig (2b)) ...

However, in real life there may be some term $\hat{H}'$ in the Hamiltonian which "hybridizes" the two bands, i.e. such that

$$M_k \equiv \langle u_{cond}(\mathbf{k})|\hat{H}'|u_{val}(\mathbf{k})\rangle \neq 0 \tag{14}$$

Since M_k is in general complex, it will correspond in the pseudospin model to some direction in the xy-plane. Let's suppose for definitions that $M_k \propto (k_x + ik_y)$, as is in fact the case for the SOI (see next lecture), then the transverse component of the effective "field" is

$$\mathcal{H}_{k,l} = \alpha_k \mathbf{k} \tag{15}$$

(where α_k is real and must behave appropriately at the Γ-point and at the edge of the FBZ: we put this issue aside for the moment). The pseudospin Hamiltonian is now

$$\hat{H}_k = -\boldsymbol{\sigma}_k \cdot \mathcal{H}_k \tag{16}$$

$$\mathcal{H}_k = (P - Qk^2)\hat{\mathbf{z}} + \alpha_k \hat{\mathbf{k}} \tag{17}$$

and in the groundstate the pseudospin will orient itself along $\mathcal{H}_k$. Now the crucial point is that the energy gap E_k is given as a fraction of $\mathbf{k}$ by the expression

$$E_k = 2|\mathcal{H}| = 2\{(P - Qk^2)^2 + \alpha_k^2\}^{1/2} \tag{18}$$

and in general is nowhere zero. Thus, in the groundstate, the electrons just fill up the (hybridized) lower band, and the system behaves as an insulator. Despite this, we shall see, its properties are interestingly different from those of the band insulators (case (1a) above); it is in fact a topological insulator.

Rotation of a (pseudo-)spin 1/2 "particle"

The general expression for the change of a quantum state $|\Psi\rangle$ (whether scalar, spinor, tensor ...) under rotation $\hat{R}(\theta, \hat{\omega})$ through an angle θ around an axis[4] $\boldsymbol{\omega}$ is given by

$$|\Psi'\rangle \equiv \hat{R}(\theta, \boldsymbol{\omega})|\Psi\rangle = \exp\, i\theta\boldsymbol{\omega} \cdot \hat{\mathbf{J}}|\Psi\rangle \tag{19}$$

where $\hat{\mathbf{J}}$ is the operator of total angular momentum. If for a particle of spin 1/2 we ignore the orbital motion (so that $|\Psi\rangle$ is just a 2-dimensional spinor) then $\mathbf{J}$ is just $\frac{1}{2}\boldsymbol{\sigma}$, so

$$\hat{R}(\theta, \boldsymbol{\omega}) = \exp\, \frac{i\theta}{2}\boldsymbol{\omega} \cdot \hat{\boldsymbol{\sigma}} = \left(\cos\frac{\theta}{2}\right)\hat{1} + i\left(\sin\frac{\theta}{2}\right)\boldsymbol{\omega} \cdot \hat{\boldsymbol{\sigma}} \tag{20}$$

(where the second equality follows from the commutation relation of the components of $\boldsymbol{\sigma}$). We note in particular that for a 2π rotation around any axis we have

$$\hat{R}(2\pi, \boldsymbol{\omega}) = -1 \tag{21}$$

– the celebrated "4π periodicity" of spinors, which we have always met in the context of weak localization and which has been spectacularly verified in experiments on neutron interferometry. Note also that if the rotation is through π around an axis laying in the xy-plane and making an angle χ with the x-axis, then its effect on the state $|\uparrow\rangle$ is

$$\hat{R}(\pi : \chi)|\uparrow\rangle = (\exp\, i\chi)|\downarrow\rangle \tag{22}$$

[4] $\boldsymbol{\omega}$ is assumed to be a unit vector; I omit the "hat" so as to avoid confusion with the notation for an operator.

– a result which will be useful below.

Time reversal

While it is perfectly possible to define the operation of (pseudo)-time reversal for any 2D Hilbert space, in the pseudospin case it is usually not a symmetry of the Hamiltonian, so the ensuing discussion refers primarily to real spin. For a scalar (spinless) particle the operation $\hat{T}$ of time reversal must transform the operator $\hat{\mathbf{x}}$ to itself, but transform the momentum operator $\hat{p} \equiv -i\hbar\partial/\partial x$ to $-\hat{p}$; thus an adequate prescription is that $\hat{T}$ is simply the operation of complex conjugation

$$\hat{T}_{(\text{scal})}|\psi(r)\rangle = K|\psi(r)\rangle \equiv |\psi^*(\mathbf{r})\rangle \tag{23}$$

Evidently, if $\hat{\mathbf{L}}$ is the operator of orbital angular momentum $\hat{\mathbf{r}} \times \hat{\mathbf{p}}$, we have for scalar particles

$$\hat{T}\hat{\mathbf{L}}\hat{T}^{-1} = -\hat{\mathbf{L}}, \quad \hat{T}^2 = 1 \tag{24}$$

For spinor particles, the prescription (23) is not enough, since we also need $\hat{\mathbf{S}} \to -\hat{\mathbf{S}}$, or since for spin $1/2$ $\hat{\mathbf{S}} \equiv \frac{1}{2}\hat{\boldsymbol{\sigma}}$

$$\hat{T}\hat{\boldsymbol{\sigma}}\hat{T}^{-1} = -\hat{\boldsymbol{\sigma}} \tag{25}$$

It is sufficient to take[5]

$$\hat{T} = -i\hat{\sigma}_y\hat{K} \tag{26}$$

where explicitly

$$-i\hat{\sigma}_y \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{27}$$

and $\hat{K}$ as before complex-conjugates the (components of the) wave function. Note that since $-i\hat{\sigma}_y$ is a real matrix, it commutes with $\hat{K}$, and we have

$$\hat{T}^2 \equiv (-i\hat{\sigma}_y\hat{K})^2 = (-i\hat{\sigma}_y)^2\hat{K}^2 = (-i\hat{\sigma}_y)^2 = -1 \tag{28}$$

– a statement which actually valid more generally for half-odd-integer spin.

Associated with the operation of time reversal for a spin $-1/2$ particle[6] is an important theorem: Kramers' theorem. Consider a spin $-1/2$ particle described by a spinor $|\Psi\rangle \equiv \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ (where α and β might in general be function of $\mathbf{r}$). By definition

$$\hat{T}|\Psi\rangle = (-i\hat{\sigma}_y\hat{K})|\Psi\rangle \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \begin{pmatrix} -\beta^* \\ \alpha^* \end{pmatrix} \tag{29}$$

[5]The choice of the y-axis is obviously arbitrary, but it is the one conventionally used in the literature. Note incidentally that since the operator $-i\hat{\sigma}_y$ simply rotates by π around the y-axis, the result $\hat{T}^2 = -1$ is just equivalent to the "4π symmetry."

[6]And more generally for half-odd-integral spin.

and so

$$\langle \Psi | \hat{T} \Psi \rangle = -\alpha^* \beta^* + \alpha^* \beta^* = 0 \tag{30}$$

Hence, $\hat{T}\Psi\rangle$ is orthogonal to $|\Psi\rangle$ and must be a different state. (Note that this doesn't (necessarily) work for integer spin; e.g. the state $S = 1, S_z = 0$ is $2^{-1/2}(|\uparrow\downarrow + \downarrow\uparrow\rangle)$ and thus transforms into itself (up to an overall phase) under $\hat{T}$).

Moreover, suppose the Hamiltonian is invariant under time reversal:

$$\left[\hat{H}, \hat{T}\right] = 0 \tag{31}$$

In a 2D Hilbert space, the most general form of the Hamiltonian is, as we have seen,

$$\hat{H} = \varepsilon \hat{1} - \hat{\boldsymbol{\sigma}} \cdot \mathcal{H} \tag{32}$$

where the (pseudo-) field $\mathcal{H}$ must be real. Then since $\hat{T} \equiv -i\hat{\sigma}_y \hat{K}$ anti commutes with $\hat{\boldsymbol{\sigma}}$ (rather than commuting), (32) implies that $\hat{\mathcal{H}} = 0$. Thus $\hat{H}$ must be proportional to the unit matrix; and since $|\Psi\rangle$ and $\hat{T}|\Psi\rangle$ are mutually orthogonal, $\langle \hat{T}\Psi | \hat{\mathcal{H}} | \Psi \rangle = 0$. The conclusion is therefore Kramers' theorem:

For a spin $-1/2$ particle governed by a time reversal invariant Hamiltonian, the states are always (at least) doubly degenerate

"Merons" in coordinate space: Chern number

Above we anticipated the fact that a topological insulator differs from a band insulator in having a nontrivial configuration, in **k**-space, of the pseudospin operator describing the two relevant Bloch bands. However, to get some intuitive feeling for what is going on it may be helpful to consider first a somewhat analogous situation which originally arose in a quite different context (that of superfluid ^{3}He-A) and which involves a variation in coordinate space rather than **k**-space.

The superfluid A phase of liquid ^{3}He is believed to be characterized by the formation of Cooper pairs whose relative wave function is a product of a spin function (corresponding to $S = 1$, $S_z = 0$ along some axis) which can be ignored in the present context, and a function of the relative momentum **k** where angular dependence is, if the coordinate system is appropriately chosen (see below), of the form exp $i\varphi_k$, or equivalently $k_x + ik_\varphi$, Transcribing this dependence into coordinate space, we can say that for a spatially uniform case with an appropriate choice of coordinates axis x, y, z, the pair wave function (order parameter) $F(\mathbf{r})$ has the following dependence on the relative coordinate **r**:

$$F(\mathbf{r}) = f(|\mathbf{r}|)(x + iy)/r \equiv f(|r|) \text{ exp } i\varphi \tag{33}$$

where φ is the polar angle with respect to the z-axis. We can generalize this expression to be independent of the choice of axes by introducing an orthogonal triad of unit vectors $\boldsymbol{n_1}, \boldsymbol{n_2}, \boldsymbol{l}$ such that $n_1 \times n_2 = l$ (see fig. 4) and writing

$$F(\mathbf{r}) = f(|\mathbf{r}|)(\mathbf{n_1} + i\mathbf{n_2}) \cdot \mathbf{r}/r \equiv f(|r|) \text{ exp } i\varphi_r \tag{34}$$

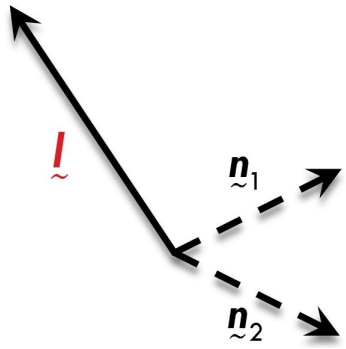


Fig. 4

(a) (b)

Fig. 5

where φ_r is the angle of rotation of $\mathbf{r}$ around $\boldsymbol{l}$. When the pair wave function is written in this way, it is clear that multiplication by an overall phase factor $e^{i\chi}$ is equivalent to rotation by χ around the axis of $\boldsymbol{l}$. The mathematical object formed by such a triad of unit vectors, with their orientation allowed to vary in space (cf. below) is technically known as a *Cosserat continuum*, and had been studied in geomechanics and elsewhere prior to the discovery of ^{3}He-A; but the incorporation of the quantity $\varphi_r \equiv \arg(\boldsymbol{n}_1 + i\boldsymbol{n}_2 + \boldsymbol{r})$ as the phase of a quantum-mechanical wave function, with the single-valuedness constraint that implies, gives it a new level of interest.

Consider now a situation in which the order parameter (pair wave-function) is spatially varying: $F = F(\mathbf{r}, \mathbf{R})$ where $\mathbf{R}$ is the coordinate of the COM of the pair. We first briefly consider the "trivial" case where the vector $\boldsymbol{l}$ is constant (fig. 5a) and the only spatial dependence is the multiplication of $F(r)$ by a phase factor $\exp i\chi(\boldsymbol{R})$, or equivalently the rotation of $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ around $\boldsymbol{l}$ by an angle $\chi(\boldsymbol{R})$:

$$F(\mathbf{r} : \mathbf{R}) = (\exp i\chi(\boldsymbol{R}))f(\mathbf{r}) \tag{35}$$

In this case we can define a "superfluid velocity" $\mathbf{v}_s(\mathbf{R})$ just as for a simple s-wave Fermi superfluid

$$\mathbf{v}_s(\mathbf{R}) \equiv \frac{h}{2m}\nabla_R\chi(\boldsymbol{R}) \tag{36}$$

and derive the standard "irrotationality" and "Onsager-Feynman" results

$$\text{curl } \mathbf{v}_s = 0 \quad (\text{curl } \equiv \boldsymbol{\nabla}_{\boldsymbol{R}}\times) \tag{37}$$

$$\oint \mathbf{v}_s \cdot d\boldsymbol{l} = nh/2m \tag{38}$$

The rotational properties in this special case are therefore identical to those of a simple s-wave superfluid.

Much more interesting is the case where the direction of the vector $\boldsymbol{l}$ is varying in space (fig. 5b). The variable χ now becomes *nonholonomic*; that is, while its $differential$ $\mathrm{d}\chi$ (or equivalently a differential rotation of the triad around the axis of $\boldsymbol{l}$) can still be defined, and hence so can $\boldsymbol{\nabla}_R\chi$, the quantity $\chi(\mathbf{R})$ itself (or equivalently the "total angle of rotation around $\boldsymbol{l}$") is not defined, since the integral of $\boldsymbol{\nabla}_R\chi$ from (say) the origin to $\mathbf{R}$ will in general depend on the path followed. Consequently, while under these conditions we can still define a "superfluid velocity" $\mathbf{v}_s$ by eqn. (36), this velocity in general no longer satisfies eqns. (37) or (38). In fact, (36) is replaced by the celebrated[7] Mermin-Ho relation: for subsequent convenience we write this in two different notations:

$$A. \quad (\text{curl } \mathbf{v}_s)_i = \frac{\hbar}{2m}\epsilon_{ijk}\epsilon_{\alpha\beta\gamma}l_\alpha\partial_j l_\beta \partial_k l_\gamma \tag{39 a}$$

$$B. \quad \text{curl } \mathbf{v}_s = \frac{\hbar}{2m}\epsilon_{\alpha\beta\gamma}l_\alpha \boldsymbol{\nabla} l_\beta \times \boldsymbol{\nabla} l_\gamma \tag{39 b}$$

While (39 a) and (39 b) are of course trivially equivalent, (39 b) is convenient for application to a 2D situation (where curl $\mathbf{v}_s$ is automatically up the axis perpendicular to the plane) whereas (39 a) has the advantage that it makes it explicit that there is no necessary relation between the "spaces" described by the two sets of axes (α, β, γ) and (i, j, k), so that it is straightforward to make the generalization to the case where the latter describe $\mathbf{k}$-space and the former pseudospin space.

The Mermin-Ho relation is sufficiently important that it is worth giving a brief and informal proof of it. It is actually a consequence of the following general theorem, which we have already exploited in the discussion of spin effects in weak localization:

Suppose $\hat{R}_1$ and $\hat{R}_2$ represent two infinitesimal rotations, characterized by vectors $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ respectively, where the direction of $\boldsymbol{\omega}_i$ specifies the axis of rotation and its (infinitesimal) magnitude the angle of rotation. Then the sequence $\hat{R}_1\hat{R}_2\hat{R}_1^{-1}\hat{R}_2^{-1} \cong 1 + [\hat{R}_1, \hat{R}_2]$ is equivalent (up to order ω^2) to a single rotation characterized by the vector $\boldsymbol{\omega}_1 \times \boldsymbol{\omega}_2$.

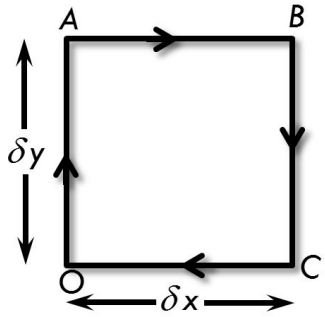


Fig. 6

We will give the proof for the 2D spatial geometry which is the case of interest for (2D) topological insulators; that is, we assume that the spatial coordinate $\mathbf{r}$ lies in the xy-plane, so that the quantity

[7] That is, to aficionados of superfluid ^{3}He. It seems to be unknown to many workers in the field of topological insulators

curl $\mathbf{v}_s$ is automatically along the z-axis. Consider an infinitesimal circuit in which we start from and return to the point O. Imagine, first, that we make infinitesimal rotations of the vector $\mathbf{l}$ on each leg of the circuit, but do not rotate around $\boldsymbol{l}$ (i.e. set $\mathbf{v}_s = 0$). Since the change $\delta\boldsymbol{l}$ in $\boldsymbol{l}$ in $\boldsymbol{l}$ must be perpendicular to $\boldsymbol{l}$ ($\boldsymbol{l}$ is a unit vector!) the rotation $\hat{R}_1 \equiv \hat{R}_{O\to A}$ is characterized by the perpendicular vector

$$\boldsymbol{\omega}_1 = \delta\boldsymbol{l}^{(1)} \times \boldsymbol{l} \equiv \delta x \cdot \left(\frac{\partial \boldsymbol{l}}{\partial x} \times \boldsymbol{l}\right) \tag{39}$$

Similarly, the rotation $\hat{R}_2 \equiv \hat{R}_{A\to B}$ is characterized by

$$\boldsymbol{\omega}_2 = \delta\boldsymbol{l}^{(2)} \times \boldsymbol{l} \equiv \delta y \left(\frac{\partial \boldsymbol{l}}{\partial y} \times \boldsymbol{l}\right) \tag{40}$$

In order to obtain results valid up to order ω^2, it is adequate to take $\hat{R}_{B\to C}$ to be equal to $\hat{R}_1^{-1}$ and $\hat{R}_{C\to O}$ to be equal to $\hat{R}_2^{-1}$. Hence, applying the above general theorem, we find that on completing the circuit we have achieved a total rotation specified by a vector $\boldsymbol{\omega}_{circ}$ given by

$$\boldsymbol{\omega}_{circ} = \delta x \delta y \left(\frac{\partial \boldsymbol{l}}{\partial x} \times \boldsymbol{l}\right) \times \left(\frac{\partial \boldsymbol{l}}{\partial y} \times \boldsymbol{l}\right) \tag{41}$$

which since $\partial\boldsymbol{l}/\partial x$ and $\partial\boldsymbol{l}/\partial y$ are both perpendicular to $\boldsymbol{l}$ can be simplified to

$$\boldsymbol{\omega}_{circ} = \delta x \delta y \left(\frac{\partial \boldsymbol{l}}{\partial x} \times \frac{\partial \boldsymbol{l}}{\partial y}\right) \tag{42}$$

and is in direction $\boldsymbol{l}$. Thus, remembering that the vector $\boldsymbol{l}$ lies in the (α, β, γ) space, we can write

$$\Delta\chi = \boldsymbol{\omega}_{circ} \cdot \boldsymbol{l} = \delta x \delta y \epsilon_{\alpha\beta\gamma} l_\alpha \frac{\partial l_\beta}{\partial x}\frac{\partial l_\gamma}{\partial y} \equiv \frac{1}{2}\partial x \partial y \epsilon_{\alpha\beta\gamma} l_\alpha \boldsymbol{\nabla} l_\beta \times \boldsymbol{\nabla} l_\gamma \quad (\times\hat{\mathbf{z}}) \tag{43}$$

(with the cross-product now in the coordinate (i, j, k) space). i.e. we have affected a rotation around $\mathbf{l}$ whose magnitude is given by the RHS of eqn. (42) . If the wave function (order parameter) is to return to its original value on completing the circuit, as it must, we need to compensate (42) by an explicit rotation around $\mathbf{l}$: again, to obtain results valid to the relevant order, it is sufficient to take $\mathbf{l}$ to be fixed at its direction at O. Then using the definition of $d\chi$ as the differential rotation around $\mathbf{l}$ and eqn. (30), we have for the resulting total rotation angle $\Delta\chi$

$$\Delta\chi = \frac{2m}{\hbar}\oint \mathbf{v}_s \cdot d\boldsymbol{l} \tag{44}$$

Since $\mathbf{v}_s$ is defined everywhere within the circuit, we can rewrite (44) as

$$\Delta\chi = \frac{2m}{\hbar}\int \text{curl } \mathbf{v}_s dA \cong \frac{2m}{\hbar}(\text{curl } \mathbf{v}_s)dxdy \tag{45}$$

and equating this to the (negative of) the RHS of (43) gives the MH relation (39), QED.

It may be easier to visualize the content of the MH theorem in integral form. Consider for example the following configuration (sometimes called a "meron"): Start at the origin with $\hat{\boldsymbol{l}} = \hat{\boldsymbol{z}}, \hat{\boldsymbol{n}}_1 = \boldsymbol{x}, \hat{\boldsymbol{n}}_2 = \boldsymbol{y}$, and rotate the triad by an angle π through an axis lying in the xy-plane at an angle θ to the x-axis. The result is

$$\hat{\boldsymbol{l}} \rightarrow -\hat{\boldsymbol{z}}, x \rightarrow x\cos 2\theta + y \sin 2\theta, y \rightarrow -x \sin 2\theta + y \cos 2\theta \tag{46}$$

i.e.

$$\boldsymbol{n}_1 + i\boldsymbol{n}_2 \rightarrow \exp\,(2i\theta)(\boldsymbol{n}_1 + i\boldsymbol{n}_2) \tag{47}$$

In other words, the effect of rotation through π around an axis in the xy-plane making an angle θ with the x-axis is

$$\chi \rightarrow \chi + 2\theta \tag{48}$$

Now imagine a circular disk of radius R, and rotate $\boldsymbol{l}$ everywhere by a position-dependent rotation $\boldsymbol{\omega}(\mathbf{r})$ given by

$$\boldsymbol{\omega}(\mathbf{r}) = \pi \mathbf{r}/R \quad \text{(or more generally } \pi \mathbf{r} f(|r|) \text{ with } f(R) = {}^1\!/_R) \tag{49}$$

The resulting texture of $\boldsymbol{l}(\mathbf{r})$ has various names ("hedgehog/ willow/ boojum"): the salient point is that at the outer edge of the disk ($|\mathbf{r}| = R$), $\boldsymbol{l}(\mathbf{r})$ is everywhere downwards. The "global" phase $\chi(\mathbf{r})$ *when taken around the circumference of the disk* is thus well-defined, and from the above argument is given by 2θ and this the quantity $\Delta\chi \equiv \oint \boldsymbol{\nabla}\chi \cdot d\boldsymbol{l}$ is given by

$$\oint \boldsymbol{\nabla}\chi \cdot d\boldsymbol{l} = 4\pi \tag{50}$$

(so that the circulation $\oint \mathbf{v}_s \cdot d\boldsymbol{l}$ is h/m (not $h/2m$!) and we have a so-called "coreless 4π vortex").

We see that this result can be generalized: Whenever we can find a contour on which the direction of $\boldsymbol{l}$ is fixed, and we require the wave function (order parameter) to be single-valued, we can define $\chi(\mathbf{r})$ on the contour and thus require

$$\oint \boldsymbol{\nabla}\chi \cdot d\boldsymbol{l} = 4n\pi \quad (n \text{ integral}) \tag{51}$$

and so can conclude by the Stokes and MH relations

$$\frac{1}{4\pi} \int dS\; \epsilon_{ijk} l_i (\boldsymbol{\nabla} l_j \times \boldsymbol{\nabla} l_k) \cdot \hat{\mathbf{z}} = n \tag{52}$$

Note that this theorem is actually purely geometric in nature; the use of quantum-mechanical considerations in requiring single-valuedness of the wave function is not essential, in fact were we to consider a purely classical Cosserat continuum and require the orientation of

the triad to be single-value we could still derive the result (52). The integer defined by the RHS of eqn. (52) is sometimes called the "Chern number" (or, in the case where the orientation is a function of $\mathbf{k}$ across a band, the "TKNN number").

While the discussion above was explicitly in terms of a vector ($J = 1$) order parameter, it is easy to see how to generalize it: for general J the only difference is that a rotation through $\delta\theta$ changes the phase of the wave function not by $\delta\theta$ but by $J\delta\theta$, and thus the n on the RHS of eqn. (52) is replaced by nJ. In particular, for $J = 1/2$ (the case of a "pseudospin") we have

$$\int dS\ \epsilon_{ijk}\sigma_i \cdot (\boldsymbol{\nabla}\sigma_j \times \boldsymbol{\nabla}\sigma_k) \cdot \hat{\mathbf{z}} = 2\pi n \tag{53}$$

Appendix: a useful special case of the MH theorem.

Consider a vector $\boldsymbol{l}$ (or $\boldsymbol{\sigma}$) which is oriented along (say) the positive z-direction at points 1 and 2, but may deviate (continuously) from this direction at other points. We can define the differential rotation $\delta\chi$ around $\boldsymbol{l}$ at any point, but in general its integral from 1 to 2 will be path-dependent. However, if we confine ourselves to paths on which the deviation lies **entirely** in a **single (specified) plane** (e.g. the xz-plane), then the quantity

$$\Delta\chi_{12} \equiv \int_1^2 (\boldsymbol{\nabla}\chi \cdot d\mathbf{r} \tag{54}$$

is uniquely defined (Moreover, by noting that the trivial case of constant $\boldsymbol{l} \parallel \hat{\mathbf{z}}$ is common to all planes, we see that $\Delta\chi$ does not depend on the choice of plane).

The proof of this theorem follows straightfowardly by considering the special case of the situation depicted in fig. 6 for which the variation of $\boldsymbol{l}$ is (say) entirely in the xz-plane. Then from the fact that $\boldsymbol{l}, \partial\boldsymbol{l}/\partial x$ and $\partial\boldsymbol{l}/\partial y$ all lie in a single plane, we see using eqn. (43) that $\Delta\chi$ must be zero, so that the quantity $\Delta\chi_{AB}$ as defined in (54) is independent of the path taken between A and B ($A \to B$ vs $A \to O \to C \to B$).

Indeed, we can actually draw from this argument the strange conclusion that under the specified conditions $\Delta\chi_{AB}$ can be defined for arbitrary points A and B, whether or not $\boldsymbol{l}_A = l_{\mathbf{B}}$. Thus, in conclusion, provided we confine ourselves to paths such that $\boldsymbol{l}(\mathbf{r})$ lies in the plane containing $\boldsymbol{l}(1)$ and $\boldsymbol{l}(2)$[8], then **the phase $\boldsymbol{\chi}(\mathbf{r})$ can be "globally" defined**. This result will be useful in lecture 23.

[8] or in the case that to $\boldsymbol{l}(1) = \boldsymbol{l}(2)$, to a single plane.