(New twist to an 80-year-old subject!)

1 Historical Sequence

- In the integral (or fractional) quantum Hall effect, there exists an energy gap in the bulk, but there are arbitrarily low energy states at the edge. However, these have the unusual property of being chiral (a result, inter alia, of breaking of time-reversal symmetry by an external magnetic field).

- Quantum spin Hall (QSH) effect: Although overall time-reversal symmetry may not be broken, could be broken for electrons of definite spin by the effects of spin-orbit interaction ($\mathbf{l} \cdot \mathbf{H}_{\text{ext}} \rightarrow \mathbf{l} \cdot \mathbf{\sigma}$). Then we might get states of unique chirality for each spin state separately. But spin-nonconserving processes will spoil the result.

- In the case of the integral quantum Hall effect, there exists a topological invariant characterizing the bulk states. Does anything similar characterize the quantum spin-Hall states? A calculation for graphene with spin-orbit interaction (Kane-Mele, 2005) shows indeed that there exists such a topological invariant. (“$Z_2$ topological index”, i.e., this quantity can take values 0 or 1). Is this stable against a small spin-nonconserving part?

- Bernevig et al.: A similar situation may hold for heavy-metal compounds (much more promising for experiment)

- Generalization to 3D: “strong” and “weak” topological insulators. Strong topological insulators have topologically protected 2D surface states ⇒ realistic possibility of experiments.

- Experiments in HgTe quantum wells, $\text{Bi}_x\text{Sb}_{1-x}$, $\text{Bi}_2\text{Te}_3$...

Note: In the historical development of the subject, time-reversal invariance has played a crucial role. But it is possible that this is something of a red herring...
2 Preliminaries

1. Graphene Introduce as in lecture 5 a “pseudospin” operator $\hat{\sigma}$ such that $\sigma_z = +1(-1)$ corresponds to the A(B) sublattice (and nearest-neighbor tunnelling matrix element $\propto \hat{\sigma}_x$). In the following, $\sigma \equiv$ pseudospin. $s \equiv$ real spin (so, e.g., $\hat{1}_s$ is the identity operator in real-spin space)

2. Symmetry operations:

(a) Inversion $\hat{P} : \hat{P}(r, p, s) \rightarrow (-r, -p, s) \Rightarrow$ in graphene, $\hat{P}$ exchanges $A$ and $B$ sublattices, thus $\hat{P} = \hat{1}_s \otimes \hat{\sigma}_x$

(b) Time reversal $\hat{T} : \hat{T}(r, p, s) = (r, -p, -s)$ (note that this must be true for all components of $s$) $\Rightarrow$

$$\hat{T} = i \hat{1}_s \otimes \hat{s}_y \hat{K}, \quad \hat{K} \equiv \text{complex conjugation} \quad (1)$$

Note that since $\hat{K}^2 = \hat{s}_y^2 = 1$, this means

$$\hat{T}^2 = -1. \quad (2)$$

(c) In principle we can also introduce “pseudo-time-reversal” $\hat{T}_{\sigma}$ such that $\hat{T}_{\sigma}(r, p, s, \sigma) = (r, -p, s, -\sigma)$. The explicit form is then

$$\hat{T}_{\sigma} = i \hat{1}_s \otimes \hat{s}_y \hat{K} \quad (3)$$

and $\hat{T}_{\sigma}^2 = -1$.

3. Kramers’ theorem:

For any single electron (or an odd number of electrons) with time-reversal-invariant Hamiltonian, every energy eigenstate must be (at least) doubly degenerate (this follows simply because for a single electron (etc.), $\hat{T}$ inter alia reverses the spin, and the resulting state is clearly not identical to the original one (For an even number of electrons, $S = 0$ is possible $\Rightarrow$ the time-reversed state is identical to the original).

4. Possible terms in the single-particle Hamiltonian of graphene:

<table>
<thead>
<tr>
<th></th>
<th>$P$</th>
<th>$T$</th>
<th>$T_{\sigma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) n.n. tunnelling</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>(b) lattice bias ($\propto \hat{1}_s \otimes \hat{\sigma}_x$)</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>(c) bulk spin-orbit</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>(d) Rashba spin-orbit$^1$</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

$^1$in an electric field perpendicular to the sheet.
5. The explicit form of the spinor corresponding to a spin pointing in direction \( \mathbf{n} \) which makes the components single-valued is, up to an overall phase factor,

\[
| \mathbf{n} \rangle = | \theta, \phi \rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{-i\phi} \end{pmatrix}
\]  

(4)

with \( \theta, \phi \) the azimuthal and polar angles of \( \mathbf{n} \). Thus the scalar product of \( | \mathbf{n} \rangle \) and \( | \mathbf{n}' \rangle \) with \( \phi = \phi' \) is

\[
\langle \mathbf{n} | \mathbf{n}' \rangle = \cos \frac{1}{2} (\theta - \theta'),
\]

(5)

while the scalar product of \( | \theta, \phi \rangle \) and \( | \pi - \theta, \phi \rangle \), which we shall need below is

\[
\langle \theta, \phi | \pi - \theta, \phi \rangle = \frac{1}{2} \sin \theta \left( 1 + e^{-i\phi} \right).
\]

(6)

6. Berry phase:

Imagine transporting (real or imaginary) spin \textit{adiabatically} around a closed loop (e.g. a spin in its ground state in a magnetic field \( \mathbf{B} \) which is slowly rotated around \( \hat{z} \)). Define:

\[
\phi_B = \text{Im} \oint (\psi, \nabla \psi) \cdot d\mathbf{l}.
\]

(7)

For this example, \( \phi_B = \pi (1 - \cos \theta) \) and in particular, a rotation through \( 2\pi \) (i.e. \( \cos \theta = 0 \)) gives

\[
\phi_B = \pi
\]

(standard \(-1\) of spinor under a \( 2\pi \) rotation)

3 Bulk topology and edge states: a simple illustration

Consider the simple model of graphene discussed in lecture 5 (where the only term in the Hamiltonian is the inter-sublattice tunnelling). Then \( \hat{H} \) is equivalent to the Hamiltonian of a (pseudo) spin in a magnetic field which for all \( \mathbf{k} \) lies in the \( xy \)-plane:

\[
\hat{H} = - \begin{pmatrix} 0 & \mathcal{H}_x - i\mathcal{H}_y \\ \mathcal{H}_x + i\mathcal{H}_y & 0 \end{pmatrix} \equiv -\sigma \cdot \mathcal{H}(\mathbf{k})
\]

(9)
and the ground state for given \( \mathbf{k} \) has \( \sigma \) oriented along \( \mathcal{H}(\mathbf{k}) \). The crucial point is that \( \mathcal{H}(\mathbf{k}) \) rotates through \( 2\pi \) in the \( xy \)-plane clockwise (antclockwise) as we trace out a path encircling \( K \) (\( K' \)). Suppose now we have a circular “disk” of graphene (for the moment we do not need to be too concerned about the exact boundary conditions at the edge).

Imagine constructing a semiclassical wave packet from the ground-state manifold localized near a value of \( r \) close to (but not at) the edge, and near an angle \( \theta \), and transporting adiabatically around the disk (i.e. adiabatically varying \( \theta \) while keeping \( r \) constant). As we do so, the wave vectors \( \mathbf{k} \) composing the packet will also rotate, and after a complete turn around the ring will have turned through \( 2\pi \). Since the packet is made up from the ground-state manifold (for Hamiltonian \(-\sigma \cdot \mathcal{H}(\mathbf{k})\)) this means that \( \sigma \) will also have rotated through \( 2\pi \) (recall \( \mathcal{H}(\mathbf{k}) \) and thus \( \sigma \) is always in the \( xy \)-plane). But according to the Berry-phase result (point 6 above), this implies that after one trip around the edge we come back to the same state, but with a factor \(-1\).

Now suppose that the disk is embedded in a “normal” material, that is, one in which the (matrix elements of the) Hamiltonian \( \mathcal{H}(\mathbf{k}) \) have no interesting topology. We bring the wave packet closer and closer to the edge, and eventually allow it to overlap with the normal material. What happens? It seems everything depends crucially on the energy scales; we will assume explicitly that for all the \( \mathbf{k}' \)'s of interest in our wave packet

\[
|\mathcal{H}(t)| \ll E_c
\]  

(10)

where \( E_c \) is the coupling (e.g. a tunnelling matrix element) between the disk and the normal material. Under these conditions we need to “match” the part of the wave function in the disk to the part in the normal material. However, if we make up a wave packet entirely in the normal material and transport it once around the disk, it picks up no Berry phase at all! And it then follows, provided \( E_c \) is “large”, that also the part of the wave function within the disk must have this property. In other words, we must make up a linear combination of states from the ground and other excited graphene bands in such a way that as we go around the ring, the pseudospin rotates through \( 2\pi \) around a “local” axis perpendicular to \( \mathcal{H}(\mathbf{k}) \). This gives the required extra factor of \(-1\). This means that at some point it is antiparallel to \( \mathcal{H} \), i.e., entirely in the excited band. Thus we make the conclusion that such an edge state must have a spectrum \( E(\mathbf{k}) \) which passes right through the bulk energy gap.
The above argument is heuristic at best: (a) we have not actually shown that “edge states” must exist, only that if they exist they must span the gap. (b) In the simple model of lecture 5, the “band gap” (i.e., the minimum value of $2E(k)$ as a function of arbitrary $k$) is actually zero, since $E$ is zero at the Dirac points $K$ and $K'$. This difficulty is actually resolved rather easily, by adding a nonzero lattice-bias term to the Hamiltonian (cf. below): then there is a nonzero gap of $2|\lambda_v|$, and we can still use the above argument by making up our wave packet out of states with $|E(k)| \gg |\lambda_v|$. (c) Most seriously, the topological feature of the model which we relied on, namely the “circulation” of $\mathcal{H}$ through $2\pi$ around the Dirac point $K$ is not stable, in the sense that the “vortex” centered at $K$ can annihilate with the “antivortex” around $K'$, yielding a topologically trivial state. For those and other reasons it is not thought that the edge states of graphene as described by this minimal model are particularly interesting. However, the argument may give some intuition which will be useful in the more complicated examples below.

4 $Z_2$ topological index in graphene

Let’s for the moment set the Rashba term equal to zero. Then all terms conserve $\hat{s}_z$, so we can simply consider the two $2 \times 2$ subspaces corresponding to $s_z = \pm 1$ separately. Then for the “Bloch Hamiltonian” matrix for the Bloch state whose spatial dependence is $\exp(i k \cdot R)$ we have

$$
tunneling \ term : \quad -t \begin{pmatrix} 0 & f(k) \\ f^*(k) & 0 \end{pmatrix}, \quad (11)$$

$$
lattice \ bias \ term : \quad \lambda_v \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (12)$$

$$
bulk \ spin - orbit \ term : \quad \lambda_{s_0} s_z g(k) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (13)$$

where

$$
f(k) \equiv \exp(-ik_xa) \left\{ 1 + 2\exp\left(i \frac{3}{2} k_x a \right) \cdot \cos\left(\frac{\sqrt{3}}{2} k_y a \right) \right\}, \quad (14)$$

$$
g(k) \equiv \sin\left(\sqrt{3} k_y a / 2 \right) \left\{ \cos\left(\frac{\sqrt{3}}{2} k_y a / 2 \right) - \cos\left(3 k_x a / 2 \right) \right\}. \quad (15)$$

\footnote{In the sense that we can modify the Hamiltonian in a continuous manner (while preserving the time-reversal symmetry) to remove the singularities.}

\footnote{For notational convenience I have defined $\lambda_{s_0}$ to be four times what KM denotes by that quantity.
Note that $\text{Re}f(k)$ is an even function of $k$, while $\text{Im}f(k)$ and $g(k)$ are both odd functions. The zeroes of $f(k)$ are points, at the Dirac points $K$ and $K'$, while the real function $g(k)$ has line nodes along the six perpendicular bisectors of the first Brillouin zone faces:

Note that $g(k)$ does not vanish at the Dirac points, in fact it takes the value $(3/4)\sqrt{3}$ there.

The problem is equivalent to that of a (pseudo) spin $1/2$ in a fictitious 3D magnetic field whose components are

$$
\mathcal{H}_x = -t\text{Re}f(k), \quad \mathcal{H}_y = -t\text{Im}f(k), \quad \mathcal{H}_z = -(\lambda_v + \lambda_{so}s_z g(k)).
$$

(16)

In the ground state the “spin” $\sigma$ lines up parallel to the field $\mathcal{H}(k)$.

Let’s consider first the case (a) $\lambda_{so} = 0$, $\lambda_v \neq 0$. Then the $z$-component of the field $\mathcal{H}$ dominates near the Dirac points and removes the node: in the notation of lecture 5

$$
E(k) = \left( (\Delta\epsilon)^2 + \hbar^2 v_F^2 q^2 \right)^{1/2} = \left( \lambda_v^2 + \hbar^2 v_F^2 q^2 \right)^{1/2}
$$

(17)

(Since our $\lambda_v$ is just the $\Delta\epsilon$ of Eq. (5.24)). However, this does not change the topological structure of the phase: we can still annihilate the “vortex” and “antivortex” while preserving time-reversal symmetry and keeping a nonzero value of $E(k)$ everywhere (i.e. a nonzero gap).

A much more interesting case (b) turns out to be $\lambda_v = 0, \lambda_{so} \neq 0$. Let us suppose (as is likely to be the case in real life) that $\lambda_{so} \ll t$, so that the spin-orbit coupling plays a major role only near the Dirac points. Then for a given value of $s_z$, say $+1$, the $z$-component of pseudospin $\sigma_z$ will be opposite at $K$ and $K'$ (For $s_z = -1$, the $z$-component of $\sigma$ is reversed, but the $xy$-plane components are unchanged).

The crucial question is: Are cases (a) and (b) topologically equivalent? That is, is it possible, by an arbitrary variation of “fields” which preserves a nonzero gap$^4$ (i.e. such that $|\mathcal{H}(k)| > 0$ for all $k$) to transform (b) into (a)? And if it is not, can we find an explicit “topological index” which differentiates them?

The approach taken by KM is to study the quantity

$$
P(k) \equiv \left( u_+(k), \mathcal{T}u_-(k) \right)
$$

(18)

where $u_+(k)$ and $u_-(k)$ are the two pseudospace spinors which are the ground state eigenfunctions of the Bloch Hamiltonian corresponding to (real spin) projections $s_z = \pm 1$. Some

$^4$It is fairly obvious that we can deform (b) into (a) by simply turning up the value of $\lambda_v$ through $(3/4)\sqrt{3}\lambda_{so}$. However, at the crossing point the gap tends to zero at one of the Dirac points.
care is needed in analyzing the properties of $P(k)$. In general, although the original Hamiltonian is by assumption invariant under time reversal $T$, the “Bloch Hamiltonian” $H(k)$ is not, since $T$ *inter alia* changes $k$ to $-k$; in general all we can say is that

$$\hat{T}H(k)\hat{T}^{-1} = H(-k). \quad (19)$$

However, there are some special points ($k = 0$, and the centers of the zone edges) where it follows from the symmetry and periodicity of the hexagonal lattice that $H(-k) = H(k)$. For these points we can assert that $H(k)$ commutes with $\hat{T}$, and hence $\hat{T}u_-(k)$ must be an eigenfunction of $H(k)$; since the ground state is only two-fold degenerate, this means that $\hat{\psi}T u_-(k)$ must be equal to $u_+(k)$ up to a phase factor, and hence for these points $|P(k)| = 1$. (These points are said by KM to belong to the “even subspace”). On the other hand, at the special points where the transverse component of “field” is zero, i.e., the Dirac points $K$ and $K'$, it is clear that the pseudospins corresponding to $u_+(k)$ and $\hat{T}u_-(k)$ are oppositely oriented and thus $P(k)$ must be zero. It is important to realize that the converse of this argument is not valid. The fact that the spinors corresponding to $u_+(k)$ and $u_-(k)$ are identical (as they must be on the lines where $g(k)$ vanishes, that is, the peaked lines of the figure) does not guarantee that $|P(k)| = 1$, since (unless $H(-k) = H(k)$) one cannot infer that $T u_-(k)$ is an eigenfunction of $H(k)$ (and indeed the time reversal operation will in general change the direction of $\sigma$, cf below). However, one thing we can be sure of is that over most of the first Brillouin zone $|P(k)|$ does not vanish; thus it makes sense to study the behavior of its argument (phase) over closed contours, for example one encircling the Dirac point $K$ by travelling along the lines on which $g(k) = 0$ (see figure).

Let us denote by $\theta_+$ and $\phi_+$ the azimuthal and polar angles of the pseudospin for say $s_z$ (real spin) = +1 and by $\theta_-$ and $\phi_-$ those for $s_z = -1$. Evidently in both cases (a) and (b) we have $\phi_+ = \phi_-$, so we denote $\phi_+(k) = \phi_-(k) = \phi_k$. In case (a) we also have $\theta_+(k) = \theta_-(k) = \theta$. Since the only effect of the time-reversal operator in the pseudospin space is complex conjugation, i.e. $\phi \rightarrow -\phi$, we have, according to point (5) above, with the choice of phases which makes the given components single-valued

$$P_\alpha(k) = \cos^2 (\theta/2) + \sin^2 (\theta/2) e^{-i2\phi_k} \quad (20)$$

Since in this case $u_-$ is real, $\hat{T}u_-(k) \equiv u_+(k)$
In case (6), if $\lambda_v$ is strictly equal to zero, we have $\theta_- = \pi - \theta_+$, and hence

$$P_b(k) = \sin \theta_k (1 + e^{-i\phi_k})$$  \hfill (21)$$

It is actually helpful to imagine $\lambda_v$ to have a very small negative value, in which case $P_b(k)$ is modified to

$$P_b(k) = A(\theta_k) + B(\theta_k)e^{-i\phi_k}. \quad B(\theta_k) > A(\theta_k)$$  \hfill (22)$$

It is now clear that the topological properties of $P_a(k)$ and $P_b(k)$ are different. Consider the integral

$$I \equiv \frac{1}{2\pi} \oint_C \nabla \cdot (\text{arg}P(k)) \cdot dl$$  \hfill (23)$$

This is just the number of turns around the origin in the Argand plane executed by the complex function $P(k)$ as we trace out the contour $C$. Since in case (a) $\theta_k$ is always less than $\pi/2$, so that $\cos^2 \theta_k > \sin^2 \theta_k$, it is clear that in this case $I = 0$. On the other hand, for the (modified) case (b), since $\phi_k$ rotates through $2\pi$ as we trace out the contour, $I$ is 1. Note that had we made a different choice of the spinor wave functions which maintains single-valuedness, the effect is to multiply $P(k)$ by $\exp(in\phi_k)$ with $n$ even; thus $I$ is unchanged at 0 but $I_b \rightarrow I_b + n$, $n$ even. Thus $I$ is only meaningful mod 2, and hence has acquired the name of a “$Z_2$ topological invariant”.

There is one thing which is a bit annoying about $P_b(k)$ as defined above: it makes the value of $I$ depend on the sign of $\lambda_v$! To get around this, it may be helpful to consider the sum of the $I_B$ defined as above for the Dirac point $K$, and the corresponding one for $K'$. Since the sign of the spin-orbit term is opposite at $K'$ from that at $K$, the effect of $\lambda_v$ is opposite, so the sum of $I_B(K)$ and $I_B(K')$ is always 1 (or, more generally, odd). Equivalently, one can define $I$ by a contour which traverses the edge of the complex reciprocal unit cell.

There is actually a simpler way to characterize the $Z_2$ topological invariant. Consider the four points at which $-k = k + G$ ($G$ is a reciprocal lattice vector), namely, the origin and the points $M_1, M_{-1},$ and $M_0$ (all of which lie on the original contour $C$. Note that $M_0$ is not
“equivalent” to $M_1$ or $M_{-1}$, as it is half a nonprimitive reciprocal lattice vector). At these points we must have from symmetry $\Delta(k)$ real, and thus the pseudospin $\sigma$ lies along $\hat{x}$: it is easily checked from the explicit formula for $\Delta_k$ (Eq. (8) of lecture 5)\(^6\) that $\sigma_x$ is positive for the points $\Gamma, M_1$, and $M_{-1}$, but negative for $M_0$, and thus has rotated through $\pi$ from the states at $\Gamma$. Thus, each of $u_+(k)$ and $u_-(k)$ must have acquired a phase $\pm \pi/2$. The crucial question is: Is it the same or different? Depending on the answer to this question, the value of $P(k)$ at $M_0$ will be either $+1$ or $-1$ (note that $P(k)$ must be real at $M_0$, since $\sigma$ is along the $x$-axis and hence $|\sigma\rangle^* = |\sigma\rangle$). To investigate which it is, note that by time-reversal invariance we can replace $Tu_-(k)$, up to a phase factor, by $u_+(-k) \equiv u_+(-k + G)$; thus we consider the phase accumulated by a $+$ electron from $\Gamma$ to $M_0$ on the paths $C$ and $C'$.

Recalling that the sense of the $xy$-plane precession around $K$ and $K'$ is opposite, we see that for case (a) $u_+(k)$ and $u_-(k)$ acquire the same phase ($\pi/2$), thus $P_a(M_0) = +1$. On the other hand for case (b), where the $z$-component of pseudospin is opposite, they acquire opposite phases $\pm \pi/2$, and hence $P_b(M_0) = -1$.

Actually, the “absolute” phase of $P(k)$ at $M_0$ is not a particularly meaningful quantity, given by changing the definition of the spinor basis (cf. Eq. (6) of lecture 5) one changes the quantity $\Delta_k$ by a factor $e^{i\mathbf{k} \cdot \mathbf{R}}$ where $\mathbf{R}$ is some lattice vector. Since the $\mathbf{k}$ corresponding to $M_0$ (also to $M_1$, $M_{-1}$) is half a reciprocal lattice vector, this can result in a change of sign of $\sigma(M_0)$. However, if we consider the product of $P(k)$ at all four time-reversal invariant points: $(\Gamma, M_1, M_{-1}, M_0)$, the relevant factor is $e^{i\sum_{i=1}^4 \mathbf{k} \cdot \mathbf{R}}$, and since $M_0$ is half the sum of the primitive reciprocal lattice vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, this factor is always unity. Thus the physically significant quantity is the product

$$\eta \equiv \prod_{i=1}^4 \delta_i$$

where $\delta_i$ is the value of $P(k)$ at $M_i$, and the sum goes over the four independent time-reversal-invariant points. The “$Z_2$ topological invariant” is thus 0 for $\eta = +1$ and 1 for $\eta = -1$. This result was attained by Fu and Kane by a different argument based on adiabatic transport.

It is clear that we can generalize the definition (Eq. (24)) to an arbitrary 2D lattice, with the $M_i$ again defined as the four independent points in the first Brillouin zone (including the origin) for which $\mathbf{k} = \mathbf{G}/2$ (“time-reversal-invariant points”). Moreover, once this is realized, it is clear that a generalization is possible to 3D; in this case there are 8 time-reversal-invariant points. For notational simplicity, let us consider only a simple cubic lattice

\(^6\)Where, however, the factor $e^{-ikx}$ must be removed (it is an artifact of the artificial convention used in Eq. (5.6))
(so that the reciprocal lattice is also simple cubic). Then the eight points in question are the vertices of a cube with sides 1/2 of the reciprocal lattice vectors (and a volume 1/8 of the first Brillouin zone), and we can consider not only the quantity analogous to Eq. (24), namely

\[ \eta_{3D} \equiv \prod_{i=1}^{\infty} \delta_i, \]  

but also the “partial” \( \eta_{\nu} \)’s (\( \nu = 1, 2, 3, \ldots \)) obtained by taking the product only over the four vertices in a single face of the cube\(^7\). We may distinguish three cases\(^8\)

1. “Band insulator” (or “topologically trivial case”), all \( \eta \)’s \(+1\)
2. “Weak topological insulator”, one or more \( \eta_{\nu} = -1 \) and \( \eta_{3D} = +1 \)
3. “Strong topological insulator”, \( \eta_{3D} = -1 \)

How do we find the values of the quantities \( \delta_i \) and thus of \( \eta \) for a particular physical system? Fu and Kane note that if the lattice has inversion symmetry (note that this is not the case for graphene) then \( \delta_i \) is just the parity of the occupied Bloch state at \( k_i \), something which band-structure theorists routinely calculate. On this basis several heavy-metal (Bi and Pb) compounds were suggested to be strong 3D topological insulators.

5  Stability

It is rather clear that so long as \( s_z \) remains a good quantum number and the energy gap is nonzero everywhere in the first Brillouin zone, the \( Z_2 \) topological index is well-defined; to change it one needs to suppress the gap to zero at some point(s). (For example, we know that case (a) is realized for \( \lambda_v \gg \lambda_{so} \) and case (b) for the opposite limit \( \lambda_{so} \gg \lambda_v \); however, to go between these two limits one has to pass through the point \( \lambda_v = \frac{3}{4} \sqrt{3} \lambda_{so} \), and at this point the field \( \mathcal{H}(k) \) (hence the gap) vanishes for the \( s_z = +1/-1 \) states at the \( K'(K) \) point). Moreover, a small spin (i.e. \( s_z \)-) nonconserving term in the Hamiltonian, such as the Rashba spin-orbit coupling term (\( \propto s_x \) and \( s_y \)) should not affect the situation qualitatively; while the two occupied states are no longer exactly eigenstates of \( s_z \), they should by continuity maintain their topological properties.

The essential difference, with regard to stability, between case (a) and case (b) is that in case (a) one can, by a suitable (gap-preserving) modification of the Hamiltonian bring the

---

\(^7\)The product over the opposite face is then constrained by the value of \( \eta_{3D} \).

\(^8\)In the following it is assumed that the gap is nonzero throughout the whole of the first Brillouin zone.
vortex (originally at $K$) and the antivortex (originally at $K'$) together and annihilate them, while in case (b) this is impossible without allowing the gap to tend to zero somewhere in the first Brillouin zone.

6 Edge states in the 2D case

The argument for edge states which disperse across the energy gap is similar to the one given above: in the present case we can consider, for given $s_z$, a path which runs right around the edge of the reciprocal unit cell, thereby circling both $K$ and $K'$. As we have seen, an electron traversing this circuit must pick up a phase factor $e^{i\pi}(-1)$, and by the above argument the matching to an exterior medium with the “normal” behavior ($e^{i\eta}$) must imply that any state localized on the boundary disperses across the gap. In the case of graphene, Kane and Mele verified this by an explicit calculation.

In the case of a 3D system, one needs to consider the behavior in a plane in $k$-space parallel to the physical ($r$-space) surface.