



This is the accepted manuscript made available via CHORUS. The article has been published as:

Topological insulator and the θ vacuum in a system without boundaries

Kuang-Ting Chen and Patrick A. Lee

Phys. Rev. B **83**, 125119 — Published 29 March 2011

DOI: [10.1103/PhysRevB.83.125119](https://doi.org/10.1103/PhysRevB.83.125119)

Topological Insulator and the θ -vacuum in a system without boundaries

Kuang-Ting Chen¹ and Patrick A. Lee¹

¹*Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139*

(Dated: February 3, 2011)

In this paper we address two questions concerning the effective action of a topological insulator in one and three dimensional space without boundaries, such as a torus. The first is whether a uniform θ -term with $\theta = \pi$ is generated for a strong topological insulator. The second is whether such a term has observable consequences in the bulk. The answers to both questions are positive, but the observability in three dimension vanishes for infinite system size.

I. INTRODUCTION

The topological insulators are characterized by insulating band structures with nontrivial topology, which cannot be smoothly deformed back to an atomic insulator, sometimes under certain discrete symmetry.¹⁻³ The most well-known example is the integer quantum Hall effect (IQHE) in two dimensions (2D). Here despite the fact that the bulk of the system is gapped, the system possesses n gapless chiral edge states where n is the number of occupied Landau levels. In addition, if we perturb the system using local electric fields, there will be local transverse current in the system. This effect is best captured by the Chern-Simons term $\int d^2x dt \epsilon^{\mu\nu\lambda} A_\mu \partial_\nu A_\lambda$ in the effective theory.

In three dimensions (3D), under time-reversal (TR) symmetry there is a Z_2 classification for band insulators,^{2,3} which distinguishes the usual insulator from the topological insulator. In contrast to the IQHE states, the two classes of insulators are often distinguished only on the edge: in the bulk both insulators do not have a current response to the applied field. However, there is an odd number of spin-filtered gapless edge states, i.e., helical Dirac cones, on the surface of the topological insulator. Another physical effect of the topological insulator can be observed when the TR symmetry is locally broken on the edge: the Dirac cones would be gapped and result in an $\frac{1}{2}$ -integer quantum Hall effect on the edge.⁴ This effect also leads to the bulk electromagneto-polarizability.⁵

The effect mentioned above can also be understood using the effective topological field theory.⁴ This theory postulates that the band topological insulator can be described by a bulk " θ -term",

$$\mathcal{L}_\theta \equiv \frac{\theta e^2}{32\pi^2} \epsilon^{\mu\nu\lambda\omega} F_{\mu\nu} F_{\lambda\omega}, \quad (1)$$

with $\theta = \pi$. The integrand in eq. (1) is a total derivative so the equation of motion remains unaltered in the bulk. The edge quantum Hall effect can be derived with a smooth change of θ from π to 0 in space. Finally, the gapless spin-filtered edge states, are understood as a limit when one preserves the TR symmetry on the boundary.

There is, however, some confusion as to whether the bulk θ -term, given by eq. (1) is indeed required to describe the topological insulator. Specifically, the " $\nabla\theta$ -

term",

$$\mathcal{L}_{\nabla\theta} \equiv -\frac{e^2}{16\pi^2} \partial_\mu \theta \epsilon^{\mu\nu\lambda\omega} A_\nu F_{\lambda\omega}, \quad (2)$$

which differs from eq. (1) by a total derivative, is what ~~actually~~ has been derived **directly** in the **literature as it** ~~previous theories and can~~ produces all the effects mentioned above as well. **We note that eq. (2) is not gauge invariant under a large gauge transform, e.g. $A_x \rightarrow A_x + 2\pi/eL_x$, and cannot be the complete low energy Lagrangian. Nevertheless, it is still desirable to give a direct derivation of eq. (1). Furthermore, we would like to address the following conceptual question: in a system without boundaries, whether there are physical quantities which distinguish between the topological insulator and the normal insulator. There is an important conceptual distinction between the two theories: with the former theory, it looks as if the bulk of the topological insulator has some special property, which manifests itself either on the edges or when the Hamiltonian is varied; with the latter theory, the bulk of the topological insulator is not intrinsically different from an ordinary insulator, and all the possible physical consequences only occur at the edge between it and a normal insulator, or with a variation of the Hamiltonian.**

In this paper, first we show that there are physical consequences which distinguish between the two theories, and then we show that it is the former theory, \mathcal{L}_θ , which describes the topological insulator. The paper is organized as follows: in the following section, we study the effective theory on a closed manifold and show that the two theories given by eq. (1) and eq. (2) behave differently. In section III, we start from the fermionic band structure and derive the effective theory \mathcal{L}_θ , without the ambiguity of a total derivative.

II. THE PHYSICAL CONSEQUENCE OF THE θ TERM

Without a boundary, $\mathcal{L}_{\nabla\theta}$ will have no effect if θ is uniform. Therefore, the distinction between the two aforementioned theories will be evident if there is any physical consequence of a θ -term with uniform θ . In the following we shall discuss the effect of an uniform θ -term in one and three dimensions, with various topologies.

θ -term in one spatial dimension (1D):

~~Here we shall follow the approach of the θ vacuum, where the θ term gives a prescription to form gauge invariant states. Our strategy consists of two steps: first we show that the θ -term in the path integral is equivalent to a prescription of forming gauge invariant states. Then we use the Hamiltonian formalism with the free Maxwell Hamiltonian with those states to calculate the partition function at finite temperature. Please see appendix A for a derivation directly using the path integral.~~

~~For completeness we first show that the θ vacuum description is equivalent to a path integral with \mathcal{L}_θ . The θ -term in 1D is defined as~~

$$\mathcal{L}_{\theta,1D} = \frac{e\theta}{2\pi} \epsilon^{\mu\nu} \partial_\mu A_\nu. \quad (3)$$

Let us first take the $A_0 = 0$ gauge. Define $\tilde{A}_1(q) = \int A_1 \exp(-iqx) dx$ as the Fourier transform of A_1 . On a circle of circumference L , configurations satisfying $\int A_1 dx \equiv \tilde{A}_1(0) = 0$ can be gauge transformed into configurations satisfying $\tilde{A}_1(0) = 2\pi n/e$, with n an integer (the winding number). Therefore, when we consider a state that is an eigenstate of the quantized operator $\tilde{A}_1(0)$, say, with eigenvalue 0, we should consider instead a linear combination of all states, each with eigenvalue $2\pi n/e$. The linear combination has to be gauge invariant, and the remaining arbitrary choice would be the phase between states with consecutive n . We call this relative phase θ and call the vacuum of this phase the θ -vacuum:

$$|\theta, \text{phys}\rangle = \sum_n \exp(-i\theta n) |n, \text{phys}\rangle. \quad (4)$$

the subscript θ denotes that we only trace over the sector whose vacuum is the θ -vacuum. $\phi = e\tilde{A}(q=0, \tau=0)$ is the initial configuration of the gauge field. Note that we have inserted $1 = \int_{-\infty}^{\infty} \frac{d\ell}{2\pi} |\ell\rangle \langle \ell|$, where ℓ is the eigenvalue of $(\tilde{E}^1(q=0)/eL)$ and $|\ell\rangle$ the eigenstate. The canonical conjugate pairs (x, p) can be determined from the Lagrangian with $p = \frac{\partial \mathcal{L}}{\partial \dot{x}}$; if we choose $(e\tilde{A}^1(q=0))$ as x it conjugates to $(\tilde{E}^1(q=0)/eL)$. Therefore we have

$$\langle \phi + 2\pi m | \ell \rangle = \exp(i(\phi + 2\pi m)\ell). \quad (9)$$

There is translational symmetry in m and n and the sum over $m+n$ just gives an overall normalization constant.

Notice that if we write down the path integral from some state with winding n to some other state with winding m by turning on $A_1(t)$, the winding number can be written as an integral:

$$m - n = \frac{e}{2\pi} \int_n^m dx dt \frac{\partial A_1}{\partial t}; \quad (5)$$

here the limits of the integral denote the winding number of the initial and final configuration. The vacuum-vacuum amplitude can thus be expressed as

$$\begin{aligned} & \sum_{m,n} \langle m, 0 | \exp(iHt) | n, 0 \rangle \exp(i\theta(m-n)) \\ &= \sum_{m,n} \int_n^{m,0} [DA_1] \exp(iS + i\frac{e\theta}{2\pi} \int dx dt \frac{\partial A_1}{\partial t}); \end{aligned} \quad (6)$$

here S in the exponent is just the ordinary action corresponding to H and the scripts of the integral specifies the initial and final boundary conditions. The θ -vacuum description is thus equivalent to adding \mathcal{L}_θ to the Lagrangian.

Now we proceed to derive the physical consequence of the term. Consider a Maxwell Lagrangian with vacuum angle θ at finite temperature $1/\beta$. Taking $A_0 = 0$, the Maxwell Hamiltonian is

$$H = \sum_q \frac{1}{2L} \left| \frac{\partial \tilde{A}_1(q)}{\partial t} \right|^2 = \sum_q \frac{1}{2L} |\tilde{E}^1(q)|^2 \equiv \sum_q H_q \quad (7)$$

Since $\tilde{A}_1(q \neq 0)$ decouples from $\tilde{A}^1(0)$ we can calculate them independently. θ only couples to the $q=0$ sector as all operators at finite q have the same eigenvalue for states which differ by arbitrary winding. Let us focus on the partition function of the $q=0$ sector:

$$Z_{q=0} = \text{Tr}_\theta(e^{-\beta H_0}) = \frac{e}{2\pi} \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} \frac{d\ell}{2\pi} \sum_m \sum_n \langle \phi + 2\pi m | \ell \rangle \langle \ell | e^{-\frac{\beta L e^2}{2} \ell^2} | \ell \rangle \langle \ell | \phi + 2\pi n \rangle e^{i(m-n)\theta}; \quad (8)$$

If we replace $(m-n)$ by n , we have

$$Z_{q=0} = \int_{-\infty}^{\infty} \frac{d\ell}{2\pi} \sum_n e^{in(\theta+2\pi\ell)} e^{-\frac{\beta L e^2}{2} \ell^2}. \quad (10)$$

If we sum over n first, we have

$$\sum_n e^{in(\theta+2\pi\ell)} \sim \sum_m \delta\left(\frac{\theta}{2\pi} + \ell + m\right). \quad (11)$$

Physically, this means the effect of the uniform θ term is to cause the average electric field to be quantized in integer units of charges, but shifted by $e\theta/2\pi$. This is a well-known result with open boundary conditions⁶, where one can imagine fractional charges at the end produce the

electric field. With periodic boundary conditions it is less intuitive.

If $\theta = \pi$, this would imply that the vacuum has two degenerate configurations characterized by $\frac{1}{L} \int dx E = \pm \frac{1}{2}e$. The matrix elements between the two states become exponentially small as $L \rightarrow \infty$, so we should think of this as a spontaneous symmetry breaking situation where the parity (P) and charge-conjugation (CC) symmetry are spontaneously broken by the electric field. The electric field would choose one direction and stay for a time period proportional to e^L .

Before we conclude, let us note that in the presence of electric permittivity ϵ , the physics becomes different from the setting with boundaries. In our formulation, we can take the electric permittivity into account by changing the coefficient of the Maxwell Hamiltonian by a factor of ϵ , with the θ -term unaltered. Following the same derivation, we see that the quantization of the electric field stays the same; the electric field threading through the sample in the ground state is still $\pm \frac{1}{2}e$. However, with boundaries, the electric field threading through would be $\pm \frac{1}{2} \frac{e}{\epsilon}$ instead. This distinction can be intuitively understood as follows: in the presence of boundaries, the electric field generated by the charge at the end would polarize the electrons and screen itself by a factor of $1/\epsilon$. With periodic boundary conditions, while the electrons would still be polarized, this polarization does not create any net charge and does not alter the electric field. In other words, the quantization of the electric field is universal in the setting without boundaries, whereas with boundaries, both the charge at the boundary and the generated electric field are nonuniversal.

In conclusion, there is indeed a real measurable difference between \mathcal{L}_θ and $\mathcal{L}_{\nabla\theta}$, where with $\theta = \pi$ in the former theory there will be huge ground state electric field at $q = 0$ and the CC symmetry is spontaneously broken, whereas in the latter theory there will be no effect and the symmetry is preserved.

Now we turn our attention to three dimensions. We consider two settings without boundaries: the first is a

3-torus, and the second is the 3-sphere. We restrain ourselves to consider only U(1) gauge fields.

Abelian gauge field on a 3-torus:

Since we need a periodic lattice to produce the topological insulator, it is natural first to consider the world as a 3-torus. Again taking the gauge choice $A_0 = 0$, the θ -term in three spatial dimensions can be written as a difference of the Chern-Simons term on the initial and the final states in the imaginary time direction:

$$\begin{aligned} \int_S d^4x \mathcal{L}_\theta &= \int_S \frac{d^4x}{8\pi^2} \epsilon^{\mu\nu\lambda\gamma} \partial_\mu A_\nu \partial_\lambda A_\gamma \\ &= \int_{\partial S} \frac{d^3x}{8\pi^2} A_i \partial_j A_k \epsilon^{ijk}, \end{aligned} \quad (12)$$

where i, j, k now run through only the spatial directions. One superficial difference to the situation in 1D is that it seems all finite- q components contribute. However, as we require the initial and final states to differ from each other only by a gauge transformation, $\vec{A}_{\text{final}} = \vec{A}_{\text{initial}} + \nabla\phi/e$, we can see the integral on the three-dimensional boundary becomes a total derivative,

$$\int d^3x \partial_i (\phi \partial_j A_k \epsilon^{ijk}/e) = \int d^3x \partial_i (\phi B_i). \quad (13)$$

Let us assume ϕ only has a winding in the z direction, i.e., $\phi(x, y, L_z) - \phi(x, y, 0) = 2\pi n$, then eq. (13) becomes $2\pi n \Phi_B$ where Φ_B is the total flux threading the torus in the z direction. Assuming $\Phi_B = m\Phi_0$ with $\Phi_0 = hc/e = 2\pi/e$, we find

$$\int_S d^4x \epsilon^{\mu\nu\lambda\gamma} \partial_\mu A_\nu \partial_\lambda A_\gamma = \frac{8\pi^2}{e^2} nm, \quad (14)$$

Thus, with m units of the fundamental flux quantum in the z direction, the " θ -vacuum" consists of linear superposition of states with configurations satisfying $\int A_z dz = 2\pi n/e$, where n is an integer. Since the Hamiltonian is still quadratic, we can calculate the relevant part of the partition function similar to the calculation in 1D. The analog of eq. (10) is

$$\begin{aligned} Z_{q=0} &\sim \sum_{m,n} \int \frac{d\ell}{2\pi} e^{imn\theta} e^{i2\pi n\ell} \exp\left(-\frac{\beta V}{2} \left(\left(\frac{e\ell}{L_x L_y}\right)^2 + \left(\frac{2\pi m}{e L_x L_y}\right)^2\right)\right) \\ &\sim \sum_{m,n'} \int \frac{d\ell}{2\pi} \delta\left(\frac{m\theta}{2\pi} + \ell + n'\right) \exp\left(-\frac{\beta V}{2} \left(\left(\frac{e\ell}{L_x L_y}\right)^2 + \left(\frac{2\pi m}{e L_x L_y}\right)^2\right)\right). \end{aligned} \quad (15)$$

$V = L_x L_y L_z$ is the world volume and we choose our conjugate variables to be $(e \int A_z d^3x / L_x L_y)$ and $(\int E_z d^3x / e L_z)$, with the eigenvalue of the latter labeled by ℓ . We find that with a fix flux $\Phi_B = m\Phi_0$ in the z direction, the electric flux in the same direction is quan-

tized:

$$E_z L_x L_y = e(n - m\theta/2\pi) = ne - \theta\Phi_B/\Phi_0^2, \quad (16)$$

with n an integer.

Let us first take the strict $T = 0$ limit. Here the thermal fluctuation of the magnetic flux is suppressed and we

find that, the θ -term only has nontrivial effect if there is a finite flux threading through. For $\theta = \pi$, when we have an odd magnetic flux, the electric flux in the same direction would be quantized in half units of e . The electric field goes to zero if the world volume goes to infinity, however.

At finite T , the thermal fluctuation of the magnetic field can generate some finite fluxes, and we would have some effect even with $B = 0$ in average. For simplicity let us again set $\theta = \pi$ and consider $L_x = L_y = L_z = L$. If $T \gg 1/L$, The correlation function of the electric field would contain an extra term comparing to the usual Maxwell theory:

$$\langle E(x)E(y) \rangle \sim \langle E(x)E(y) \rangle|_{\theta=0} + \frac{e^2}{8L^4}. \quad (17)$$

One can understand this constant correlation by imagining that half of the states in the ensemble have an odd number of magnetic fluxes. The state with an odd number of fluxes would have a ground state electric field squared to $(e/2L^2)^2$, and the average is just a half of that. This extra part of the correlation function is long ranged, and can easily be distinguished from the Maxwell part. However, the magnitude again vanishes in the large L limit. Since it is not possible to have a 3D torus without embedding it in 4D space, these effects are of academic interests only.

Before we end this subsection, we should note that from this calculation, it is clear that any local magnetic field will not produce any effect. Therefore, one would not see an electric field inside a solenoid, nor any charge at the end of it.

Abelian gauge field on a 3-sphere and magnetic monopoles:

Since we cannot have global nonzero magnetic flux in any direction in a 3-sphere, there will be no effect of the θ -term. This is in contrast to the case with a magnetic monopole, where it is predicted that there will be charge $e(n - \theta/2\pi)$ attached to it in a θ -vacuum. This effect can be understood as follows: magnetic monopole is a singularity in terms of the abelian gauge field. Suppose we have a pair of monopole-antimonopole far away in a 3-sphere so that we have one fundamental flux going from one to the other. The geometry is now a 3-sphere with two punctures. From the calculation of the previous section we can see the electric flux threading from one hole to the other must be quantized, $\Phi_E = e(n - \theta/2\pi)$, and we would attribute this as the charge of the magnetic monopole.

Franz et. al. showed that there is Witten effect inside the topological insulator.⁷ We emphasize here that this does not prove that a bulk θ -term exists, as the Witten effect can also come from the " $\nabla\theta$ " theory, provided that we characterize it by $\theta = 0$ inside the monopole. Given that a monopole can only live in a unit cell, and the band structure is absent in the unit cell, it is not unnatural to set $\theta = 0$ inside a monopole.

As a side note, if we consider nonabelian gauge fields, the θ -term in general does have effect in a 3-sphere. This effect, however, is usually associated with the physics of instantons and is quite different from what we have discussed.

III. THE EFFECTIVE θ -TERM IN THE PRESENCE OF BAND ELECTRONS

In this section, we investigate how the presence of fermions can alter the vacuum θ -angle. We briefly review the topological band invariant which characterizes the topological insulators in 1D ~~one and three spatial dimensions~~ and then we show that ~~they are related to the shift of θ .~~ **with periodic boundary conditions, the effective vacuum angle θ is shifted by π in the topological insulator. We then carry the same procedure in 3D.** For the clarity of the equations, ~~sometimes~~ we take the units $e = 1$ **when there is no ambiguity.** ~~i.e., we absorb e into the definition of the external electromagnetic gauge fields.~~

The topological insulator in 1D is characterized by the polarization⁸

$$P = \int \frac{dk}{2\pi} \sum_{\text{occ}} -i \langle u_i | \frac{\partial}{\partial k} | u_i \rangle = \int \frac{dk}{2\pi} \text{tr}(\mathcal{A}_x); \quad (18)$$

with

$$\mathcal{A}_{\mu,nn'} \equiv \langle u_{nk} | -i \frac{\partial}{\partial k^\mu} | u_{n'k} \rangle \quad (19)$$

which is the so-called Berry's phase gauge field in momentum space. $|u_{nk}\rangle$ is the periodic part of the Bloch wave function.

P is forced to either take the value $\frac{1}{2}$ or 0 when there is charge conjugation (CC) symmetry. If the CC symmetry is preserved everywhere, then on the boundary there will be an odd number of zero modes. If the CC symmetry is locally broken in some way on the boundary, then there will be a $n + 1/2$ charge, where n depends on the detail of the local symmetry breaking. A cartoon showing the effect is depicted in Fig. (1).

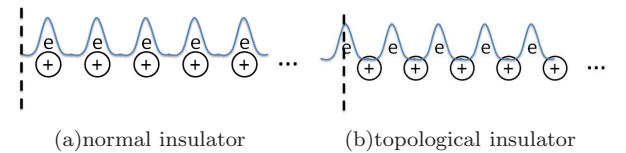


FIG. 1. Topological insulator in 1D. The dashed line shows the edge. The envelope shown is the Wannier wave function of the electrons. **There will be $\pm \frac{1}{2}e$ on the edge of the topological insulator, depending on how the lattice is terminated.**

Naively one might think this already shows that θ is shifted by $2\pi P$: after all, the θ -term in 1D is nothing but an energy term proportional to the electric field, in

which the energy of dipoles, $-\int dx P \cdot E$, fits perfectly. However, we should note that normally this dipole energy arises from separating the charges to the boundary, away from each other. On a circle with uniform polarization, therefore, one would not anticipate such energy term is present.

If we look back at how θ change the physical property of the system, it comes in by adding a phase to the amplitude between vacua with different winding numbers. Specifically, θ is precisely the additional phase of the amplitude between vacua with consecutive winding numbers. In the presence of gapped fermions, this phase can come from integrating out the fermions, in other words, the dynamical phase the fermionic system obtains under a time-dependent background field. This phase has two contributions, one is just the time-dependent energy of the fermions and the other is the Berry's phase of the process. The phaseshift from the energy depends on the time duration and is not just a function of the initial and final state; therefore it does not alter θ . Therefore, similar to the consideration in Ref.⁹, we are led to consider the accumulated geometric phase of the band electrons, when the external field is slowly turned on. See Fig. 2(a) for a cartoon of the procedure.

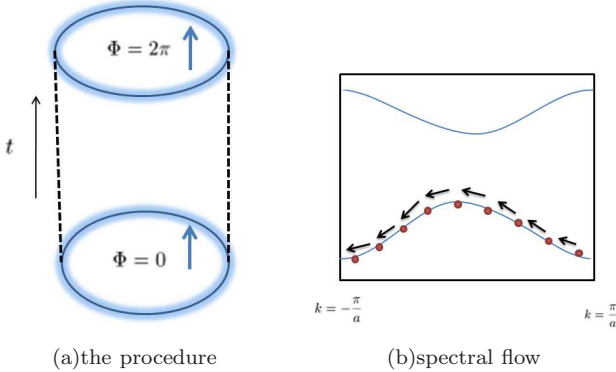


FIG. 2. (a) A flux is slowly threaded through. $\Phi = 0$ and $\Phi = 2\pi$ are the same physical state related by a gauge transform. We calculate the Berry's phase of the process. (b) During the process, at every allowed momentum by the periodic boundary conditions, the energy and the periodic part of the wave function moves slowly to the values of the state to the left, according to eq. (22). When a full flux is threaded, each one of them takes the eigenvalue and the eigenvector of the one at its left. Note that the momentum quantum number k , however, does not change. When we sum over the Berry's phase contribution from all the single particle states, it becomes an integral over the entire Brillouin zone.

First we shall consider how the single particle wave function change as we increase A_1 uniformly. We have

$$\psi_{nk}(x) = u_{nk}(x)e^{ikx} \quad (20)$$

which is the wave function in position basis, and $u_{nk}(x)$ is periodic and satisfies

$$\left((\nabla - (k + eA_1))^2 + V(x)\right) u_{nk}(x) = E_{nk} u_{nk}(x). \quad (21)$$

As we increase A_1 uniformly to $A_1 + \eta$, the momentum k cannot change as it is fixed by the finite size L and the periodic boundary condition. On the other hand, following eq. (21), $u_{nk}(x)$ changes as

$$u_{nk}(A_1 + \eta) = u_{n(k-e\eta)}(A_1), \quad (22)$$

which is just a corresponding shift of k by $-e\eta$. if $e\eta = 2\pi/L$, the system returns to its original state, but in a different gauge (i.e., with winding number different by one.) Notice that while $u_{nk}(x)$ goes to the next available value on the left, the k in the exponential stays the same. The electronic wave function is therefore different from its starting state. Nevertheless, as discussed further below, if we include the gauge field, the final state differs from the initial state by a large gauge transformation, and the Berry's phase accumulated in the process is exactly what we want to calculate.

As a side note, the situation is similar if we put electrons on a lattice which couples to the gauge field via Peierls substitution. The single particle eigenfunction can be written as $\psi_{nk} = \sum_i u_{nk,m} \exp(ikx_i) |m, x_i\rangle$, with u_{nk} now a vector in the orbital space. With an increase of A_1 , only u_{nk} changes.

Now we are ready to calculate the accumulated Berry's phase of the band electrons under the process, where the winding of the gauge field is increased by one:

$$\begin{aligned} \phi_{\text{Berry}} &= i \int_{2\pi n/e}^{2\pi(n+1)/e} d\tilde{A}_1(0) \langle \Psi_e | \frac{\partial}{\partial \tilde{A}_1(0)} | \Psi_e \rangle \\ &= i \int_{2\pi n/e}^{2\pi(n+1)/e} d\tilde{A}_1(0) \sum_{k_i, \alpha \in \text{occ}} \langle \psi_{k_i \alpha} | \frac{\partial}{\partial \tilde{A}_1(0)} | \psi_{k_i \alpha} \rangle \\ &= i \sum_{k_i, \alpha \in \text{occ}} \int_{k_i}^{k_i + 2\pi/L} dk \langle u_{k\alpha} | \frac{\partial}{\partial k} | u_{k\alpha} \rangle \\ &= i \int_{\text{BZ}} dk \sum_{\alpha \in \text{occ}} \langle u_{k\alpha} | \frac{\partial}{\partial k} | u_{k\alpha} \rangle \\ &= -2\pi P. \end{aligned} \quad (23)$$

In the second equality, we wrote the derivative acting on the Slater determinant as a sum of derivatives acting on single particle wave functions. In the third equality we then plug in the dependence of the wave functions, and change variables to k . Whenever $\tilde{A}_1(0)$ increases by $2\pi/e$, each u_{nk} reaches the next allowed eigenstate to the left by the periodic boundary condition (without actually changing the momentum eigenvalue.) As we sum over all the integral of eigenstates at different allowed k 's, the whole Brillouin zone (BZ) is covered exactly once and we reach the fourth equality.

While we calculate the Berry's phase for process where the winding number of the initial and final states differs by one, evidently the phase is proportional to the difference of the winding number in general. Therefore, it leads to a shift of θ by ϕ_{Berry} . Specifically, if the vacuum has an vacuum angle $\theta = 0$, we see that the topological insulator in 1D will be described by $\theta = \pi$.

A few remarks are in order. Firstly, this calculation is good for a finite-size system, where both the lattice spacing and the length of space are finite. Despite that the eigenstates in such a system would be discrete points in the BZ, the whole BZ is covered by the integral and there is no finite-size effect. Secondly, one might notice that the fermionic wave functions of the initial and final state are different, and it seems that our process does not form a close loop as usually required by a physical (gauge-invariant) Berry's phase. This does not invalidate our calculation, however, since the initial state and the final state are related by a large gauge transform. Once the convention of the phase of the initial state is determined, the phase convention of the final state is also determined via gauge-transforming the initial state to have the desired winding number. ~~The residual gauge degree's of freedom without altering the external gauge field (which is just an arbitrary phase of the final state) can be absorbed into the vacuum angle θ in the absence of the fermions.~~ **The uniform phase of the wave function, which is undetermined by the large gauge transform, is fixed by the prescription of the original vacuum angle θ .**

Now we turn to the 3D strong topological insulator. The 3D strong topological insulator, defined under TR symmetry, is characterized by the band structure invariant⁴

$$\frac{1}{4\pi} \int d^3k \epsilon^{abc} \text{Tr}(\mathcal{A}_a \partial_b \mathcal{A}_c - i \frac{2}{3} \mathcal{A}_a \mathcal{A}_b \mathcal{A}_c) = \pi. \quad (24)$$

Similar to the case in 1D, a θ -term is generated if we can find a Berry's phase of the electrons which is proportional to the difference of the winding number of the initial and final gauge configuration. More specifically, the procedure is as follows: we apply a constant finite magnetic field, say, in the z direction on the 3-torus. We then slowly change the gauge field in the same direction uniformly until the final state is connected to the initial state by a large gauge transform. Then we apply the magnetic field in some other direction and repeat the calculation. We can also consider procedures such as applying magnetic field in z direction and changing the gauge y direction; the phase of this process leads to a term $\propto E_y B_z$ in the effective Lagrangian. In general, we therefore expect the full effective theory to take the form

$$\mathcal{L}_{\text{eff}} = \sum_{ij} \alpha_{ij} E_i B_j. \quad (25)$$

The θ -term is the isotropic part of the effective Lagrangian:

$$\theta = \frac{4\pi^2}{3e^2} \sum_i \alpha_{ii}. \quad (26)$$

Two things are different from our calculation in 1D: firstly, the state at a given k_z is **the Slater determinant of all the wave functions at every k_x and k_y , already a many body wave function.** Secondly, we ~~have to~~ **will**

calculate everything in a finite (but maybe small) magnetic field. ~~In principle one can go over all the Landau levels at a given k_z and sum up their Berry's phases, but in practice this is not easy to do as the wave functions are not perturbative in B . It turns out that the density matrix perturbation theory introduced in Ref.¹⁰ is suitable for this calculation with an extra trick as will be described below.~~ **Due to the difficulty of solving the Landau levels even for small B for an arbitrary band structure, we will take advantage of the density matrix perturbation theory introduced in Ref.¹⁰, with an extra trick as will be described below.**

Before we dig into the calculation, let us clarify that our calculation, despite taking advantage of the same formalism, is distinct from Ref.¹⁰. There they first calculate the current flowing through the bulk as they vary the Hamiltonian with time under a small magnetic field, then they relate the time integral of the current to the polarization. While the uniform θ -term in 3D with boundaries would give rise to a magneto-polarization effect, the converse cannot be said. As mentioned in the introduction, both \mathcal{L}_θ and $\mathcal{L}_{\nabla\theta}$ can produce this effect, so a derivation of the effect does not distinguish between the two theories. To our best knowledge the following calculation is the first demonstrating **directly** that it is indeed \mathcal{L}_θ which describes the topological insulator.

Suppose we apply a small magnetic field along the z -direction, $\vec{B} = B\hat{z}$. Following the calculation in 1D, we calculate the Berry's phase of the process:

$$\phi_{\text{Berry}} = i \oint dk_z \langle \Psi_{\vec{B}} | \frac{\partial}{\partial k_z} | \Psi_{\vec{B}} \rangle; \quad (27)$$

here $\Psi_{\vec{B}}$ denotes the Slater determinant of the 2D electron wave functions, for a given k_z in the magnetic field. Analogous to the case in 1D, the derivative is understood to be taken only on the periodic part of the Bloch wave function.

Despite that the integrand can be written as a sum over single-particle wave functions, we immediately notice that it cannot be expressed as a function of the single-particle density matrix. This is due to the fact that the wave function depends on the vector potential which is gauge dependent. One easy way to realize the fact is to consider a change of phase in the wave functions. The integrand is not invariant (the integral as a whole, on the other hand, is invariant modulo 2π) whereas the density matrix remain unaltered under the transformation.

We can, however, express the integral as a whole in terms of the density matrix by the following trick. The accumulated Berry's phase can be expressed as an integral of the Berry's curvature in one extra dimension using Stokes theorem, with the region of integration bounds by the original k_z integral. The Berry's curvature can now readily be expressed in terms of the density matrix extended into the extra dimension, which is chosen continuously but otherwise arbitrarily with the constraint such that on the boundary, we have the original density

matrix. We therefore have

$$\begin{aligned} i \oint_{\partial S} dk_z \langle \Psi_{\vec{B}} | \frac{\partial}{\partial k_z} | \Psi_{\vec{B}} \rangle &= i \int_S d^2 k \epsilon^{\alpha\beta} \partial_\alpha \langle \tilde{\Psi}_{\vec{B}} | \partial_\beta | \tilde{\Psi}_{\vec{B}} \rangle \\ &= i \int_S d^2 k \epsilon^{\alpha\beta} \text{Tr} (\tilde{\rho} \partial_\alpha \tilde{\rho} \partial_\beta \tilde{\rho}); \end{aligned} \quad (28)$$

$|\tilde{\Psi}_{\vec{B}}\rangle$ is the 2D electron many-body wave function, which in addition to being a function of k_z , has been extended to some extra direction k_w . $\tilde{\rho} = \sum_i |\tilde{\psi}_i\rangle \langle \tilde{\psi}_i|$ is the extended 2D single particle density matrix, where $|\tilde{\psi}_i\rangle$ is a 2D single particle eigenstate in magnetic field \vec{B} . α, β run through two directions which is spanned by k_z and k_w . The trace sums over both the band and the (x, y) -position basis. As mentioned above, the integral is chosen to be performed on the area such that the boundary is at $(k_z, k_w = 0)$ and the density matrix has known values. It is straight forward to show the second equality, and the derivation is provided in the appendix. Different choices of density matrices inside the boundary can only alter the integral by multiples of $2\pi i$. To avoid cluttering of the equations, in the following we omit the tilde for the extended objects when there is no ambiguity.

Then following the formalism in Ref.¹⁰, we take the large size limit and expand ρ to linear order in B . As discussed there, the density matrix in real space basis can be decomposed into two parts, one of which is translationally invariant:

$$\rho_{r_1, r_2} = \bar{\rho}_{r_1, r_2} \exp(-i\vec{B} \cdot (\vec{r}_1 \times \vec{r}_2)/2), \quad (29)$$

where ρ_{r_1, r_2} denotes the density matrix in position basis, and $\bar{\rho}$ is translationally invariant. While the other part seems to be affected by the infinite range of \mathbf{r} , in our expression three ρ 's appear together and the combination is short-ranged and can be expanded in B . It is thus straight forward to expand ρ explicitly and calculate. (Please see appendix for details of the calculation.)

Up to first order in B , the result is

$$\begin{aligned} \phi_{\text{Berry}} &= \int_S d^2 k \int_{BZ} \frac{d^2 k'}{(2\pi)^2} L_x L_y \epsilon^{\alpha\beta} \\ &\quad \left[\epsilon^{\gamma\delta} B \text{Tr} \left(-\frac{1}{4} \mathcal{F}_{\alpha\beta} \mathcal{F}_{\gamma\delta} + \frac{1}{2} \mathcal{F}_{\gamma\beta} \mathcal{F}_{\alpha\delta} \right) \right. \\ &\quad \left. - i \partial_\alpha \text{Tr} (\partial_\beta \rho_{0k'} (1 - \rho_{0k'}) \bar{\rho}_{k'} \rho_{0k'} - h.c.) \right] \end{aligned} \quad (30)$$

α, β span k_z, k_w and γ, δ span k_x, k_y . The integral of k' is performed on the 2D Brillouin zone in xy -plane. $\bar{\rho}_{k'} = \langle k' | \bar{\rho} | k' \rangle$ is the translationally invariant part of the density matrix at a given (\vec{k}, k_w) and ρ_{0k} is the density matrix in zero field. $\mathcal{F}_{\mu\nu}$ is the nonabelian Berry curvature of the occupied bands,

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu - i[\mathcal{A}_\mu, \mathcal{A}_\nu]. \quad (31)$$

Notice that in eqn. (30), the tensor structure in the first and the second term is different and we can rewrite

the first term using the total-antisymmetric tensor in 4 dimensions:

$$\begin{aligned} \phi_{\text{Berry}} &= \phi_I + \phi_A \\ \phi_I &= \frac{-\Phi_B}{32\pi^2} \int_{S \times BZ} d^4 \mathbf{k} \epsilon^{abcd} \text{Tr} (\mathcal{F}_{ab} \mathcal{F}_{cd}) \end{aligned} \quad (32)$$

$$\phi_A = \frac{-i\Phi_B}{4\pi^2} \int_{S \times BZ} d^4 \mathbf{k} \epsilon^{\alpha\beta} \partial_\alpha M_{\beta z} \quad (33)$$

$$M_{\alpha\beta} = \text{Tr} \left(\partial_\alpha \rho_{0k'} (1 - \rho_{0k'}) \frac{\partial \rho_{k'}}{\partial B^\beta} \rho_{0k'} - h.c. \right). \quad (34)$$

We have explicitly expanded the second term to first order in \vec{B} . a, b, c, d runs through all directions. Both integrals are total derivatives and we can integrate back to the boundary which is the original 3D Brillouin zone:

$$\phi_I = \frac{-\Phi_B}{8\pi^2} \int d^3 k \epsilon^{abc} \text{Tr} (\mathcal{A}_a \partial_b \mathcal{A}_c - i \frac{2}{3} \mathcal{A}_a \mathcal{A}_b \mathcal{A}_c) \quad (35)$$

$$\phi_A = \frac{-i\Phi_B}{4\pi^2} \int d^3 k M_{zz}. \quad (36)$$

ϕ_I is isotropic, i.e., independent of the direction of the applied magnetic field. ϕ_A , on the other hand, is anisotropic in the sense that if we do the same calculation for the magnetic field in x or y direction, the result in general would be different. Now we consider the gradual gauge transform in the i -direction and the magnetic field in the j -direction, the same calculation still goes through, provided that we take α, β in the i -direction and the extra direction, and γ, δ in the directions perpendicular to the magnetic field. We get

$$\begin{aligned} \phi_{I,ij} &= \phi_I \delta_{ij}; \\ \phi_{A,ij} &= \frac{-i\Phi_B}{4\pi^2} \int d^3 k M_{ij}. \end{aligned} \quad (37)$$

In terms of the effective theory, this means that the effective Lagrangian not only contains $\vec{E} \cdot \vec{B}$, in general we have $\sum_{ij} \alpha_{ij} E_i B_j$, where

$$\alpha_{ij} = \int \frac{d^3 k}{(2\pi)^3} \left(\epsilon^{abc} \frac{-1}{2} \text{Tr} (\mathcal{A}_a \partial_b \mathcal{A}_c - i \frac{2}{3} \mathcal{A}_a \mathcal{A}_b \mathcal{A}_c) \delta_{ij} + M_{ij} \right). \quad (38)$$

By calculating the Berry's phase of these processes, not only do we get the coefficient of the topological term but we also get a part which is a physical response which agrees with Ref.^{10,11}. In general $\sum_i M_{ii}$ also contributes to θ . If TR symmetry is present then $M_{ij} = 0$ and we see that the vacuum angle is shifted by π in the presence of the strong topological insulator. We stress again that the calculation present here shows directly that the vacuum angle θ is shifted in the presence of the electronic band structure whereas the previous calculations only show that one can get current responses when one smoothly varies the Hamiltonian. Physically, our result predicts that there will be a half-charge electric flux if we put a strong topological insulator on a 3-torus with an odd number of magnetic flux, as described in the previous section; whereas in previous derivations, it is unclear

if one can observe anything without either a boundary or a change of the Hamiltonian.

IV. SUMMARY

In this paper, we first show that there is a measurable difference for an effective theory containing either \mathcal{L}_θ or $\mathcal{L}_{\nabla\theta}$, in one and three spatial dimensions without a boundary. Specifically, with \mathcal{L}_θ , in 1D there will be an electric field $\theta e/2\pi$ in the ground state. When $\theta = \pi$ the electric field can be in either direction and the CC symmetry and parity symmetry is spontaneously broken. In 3D, the same effect can be found on a 3-torus, but since it is the electric flux which is proportional to θe in 3D, this effect vanishes in the thermodynamic limit.

We then go on to show that the topological insulators in 1D and 3D can be described by \mathcal{L}_θ instead of $\mathcal{L}_{\nabla\theta}$. **While one may argue from the start that $\mathcal{L}_{\nabla\theta}$ is inconsistent and one would have no choice but to adapt \mathcal{L}_θ , our result here is, to our knowledge, the first derivation which shows directly that \mathcal{L}_θ exists in the bulk without the boundaries.** ~~While the expression of θ agrees with previous results, to our knowledge this is the first derivation which distinguishes between \mathcal{L}_θ and $\mathcal{L}_{\nabla\theta}$.~~

We thank N. Nagaosa, X. G. Wen and F. Wilczek for helpful discussions. We acknowledge the support of NSF under grant DMR 0804040.

V. APPENDIX

A. Path integral formulation for the θ -term

In the main text, we derive the physical consequence of the θ -term using the notion of θ -vacuum, which is similar to a Hamiltonian formalism. One may wonder why we do not directly carry out the path integral. The first reason is that the quantization is not obvious if we just do the euclidean path integral as done below. The second reason is that if we calculate the fluctuation of the electric field at finite temperature, a naive calculation would give us a sum of *negative* values, which does not make sense. It turns out that for a free theory the position space path integral is ill-behaved and a positive finite term is expanded as an infinite negative sum. A similar situation occurs when one calculates the ground state energy of the bosonic string using mode expansion. Let us start right from the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{e\theta}{2\pi}\epsilon^{\mu\nu}\partial_\mu A_\nu \quad (39)$$

using the gauge $A_0 = 0$, for the $q = 0$ sector at finite temperature $1/\beta$ we have the partition function

$$Z'_{q=0} = \left(\prod_i \left(\frac{2\pi}{\beta L \omega_i^2} \right) \right) \sum_n e^{in\theta} \exp \left(-\frac{1}{2\beta L} \left(\frac{2\pi n}{e} \right)^2 \right), \quad (40)$$

where $\omega_i = 2\pi i/\beta$ are the Matsubara frequencies. Again the finite frequency part decouples and the zero frequency part agrees with eqn.(10) if we integrate ℓ first instead:

$$Z'_{q=0} \propto Z_{q=0} = \sum_n e^{in\theta} e^{-\frac{1}{2\beta L} \left(\frac{2\pi n}{e} \right)^2} \equiv \sum_n W_n. \quad (41)$$

Nevertheless, it is hard to see from this form that θ corresponds to a quantization condition for the electric field. Without reversing the ℓ integral, another way to see the θ -dependence is to calculate the expectation value and the fluctuation of the electric field. For $\theta = \pi$, the expectation value would vanish and we can only rely on the fluctuation.

If we calculate $\langle |\tilde{E}^1(0)|^2 \rangle$, however, we would encounter a problem in the path integral as now all finite frequency part contributes and their sum seems to be infinitely negative:

$$\langle |\tilde{E}^1(0)|^2 \rangle = \frac{L^2}{Z_0} \left(\sum_n - \left(\frac{2\pi n}{\beta L e} \right)^2 W_n \right) - \sum_i \frac{2L}{\beta}. \quad (42)$$

If we compare this to what we would have got using the Hamiltonian formalism,

$$\langle |\tilde{E}^1(0)|^2 \rangle = \frac{L^2}{Z_0} \left(\sum_n W_n \left(- \left(\frac{2\pi n}{\beta L e} \right)^2 + \frac{1}{\beta L} \right) \right). \quad (43)$$

It seems we have to require

$$\sum_i (1) = -\frac{1}{2}. \quad (44)$$

for the two expressions to agree. We can understand this equality by thinking of the left hand side as the zeta function at zero, $\zeta(0)$, written in a series. While the series is divergent at zero, the zeta function is well-defined and is indeed $-\frac{1}{2}$.

The function $\sum_n n^2 W_n$ is related to the elliptic Θ -function. If one subtracts the fluctuation at $\theta = 0$ from the expression and calculate at $\beta \rightarrow \infty$, one recovers that

$$\langle |\tilde{E}^1(0)|^2 \rangle - \langle |\tilde{E}^1(0)|^2 \rangle_{\theta=0} = L^2 e^2 \left(\frac{\theta}{2\pi} \right)^2, \quad (45)$$

which implies the quantization.

B. Derivation of eq. (28) and eq. (30)

Here we show explicitly the derivation of the second equality in eq. (28). First we note $|\Psi_B\rangle = \det(\psi_i)$, and

$$\partial_\alpha \langle \Psi_B | \partial_\beta | \Psi_B \rangle = \sum_i \partial_\alpha \langle \psi_i | \partial_\beta | \psi_i \rangle. \quad (46)$$

Now we plug in $\rho = \sum_i |\psi_i\rangle \langle \psi_i|$ to the right hand side of the second equality, we have

$$\begin{aligned} & \text{Tr}(\rho \partial_\alpha \rho \partial_\beta \rho) \\ &= \sum_{ijk} |\psi_i\rangle \langle \psi_i| \left((\partial_\alpha |\psi_i\rangle) \langle \psi_j| + |\psi_j\rangle (\partial_\alpha \langle \psi_i|) \right) \end{aligned}$$

$$\begin{aligned}
& \left((\partial_\beta |\psi_i\rangle) \langle \psi_i| + |\psi_i\rangle (\partial_\beta \langle \psi_i|) \right) \\
&= \sum_{ij} \langle \psi_i | \partial_\alpha | \psi_j \rangle \langle \psi_j | \partial_\beta | \psi_i \rangle \\
&+ \sum_i (\partial_\alpha \langle \psi_i |) (\partial_\beta | \psi_i \rangle); \tag{47}
\end{aligned}$$

in the derivation we have taken advantage of the fact that $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ and thus $(\partial_\alpha \langle \psi_i |) | \psi_j \rangle = -\langle \psi_i | \partial_\alpha | \psi_j \rangle$.

Contract both eq. (46) and eq. (47) with $\epsilon^{\alpha\beta}$, we can see that they agree.

In the following we apply the magnetic field in the z -direction and take ρ as a function of k_z , and \vec{r} lies in the xy -plane. We take $\hbar = e = 1$. $\bar{\rho}$ has the following matrix elements up to first order in B :

$$\begin{aligned}
\langle \psi_{nk} | \bar{\rho} | \psi_{n'k'} \rangle &= \delta_{kk'} \left(\delta_{nn'} - \frac{1}{4} B \epsilon^{\gamma\delta} \mathcal{F}_{\gamma\delta, nn'} \right) \\
\langle \psi_{mk} | \bar{\rho} | \psi_{m'k'} \rangle &= \frac{1}{4} \delta_{kk'} B \epsilon^{\gamma\delta} \tilde{\mathcal{F}}_{\gamma\delta, mm'} \\
\langle \psi_{nk} | \bar{\rho} | \psi_{mk'} \rangle &= \delta_{kk'} \left(\frac{i}{2} B \epsilon^{\gamma\delta} \frac{\langle \psi_{nk} | \{ \partial_\gamma \rho_{0k}, \partial_\delta H_k \} | \psi_{mk} \rangle}{E_{nk} - E_{mk}} \right. \\
&\quad \left. + \frac{\langle \psi_{nk} | H'_k | \psi_{mk} \rangle}{E_{nk} - E_{mk}} \right); \tag{48}
\end{aligned}$$

note that the momentum here is two-dimensional and everything has implicit k_z, k_w dependence. n, n' are indices for occupied bands and m, m' are for empty bands. $\tilde{\mathcal{F}}$ is the nonabelian field strength for the Berry's phase gauge field defined from the unoccupied bands:

$$\begin{aligned}
\tilde{\mathcal{A}}_{\mu, mm'} &= -i \langle u_{mk} | \frac{\partial}{\partial k^\mu} | u_{m'k} \rangle \\
\tilde{\mathcal{F}}_{\mu\nu} &= \partial_\mu \tilde{\mathcal{A}}_\nu - \partial_\nu \tilde{\mathcal{A}}_\mu - i [\tilde{\mathcal{A}}_\mu, \tilde{\mathcal{A}}_\nu]. \tag{49}
\end{aligned}$$

In the following computation one would find these expressions useful:

$$\begin{aligned}
\mathcal{F}_{\mu\nu, nn'} &= -i \sum_m \langle \psi_{nk} | \partial_\mu | \psi_{mk} \rangle \langle \psi_{mk} | \partial_\nu | \psi_{n'k} \rangle \\
&\quad - (\mu \leftrightarrow \nu); \\
\tilde{\mathcal{F}}_{\mu\nu, mm'} &= -i \sum_n \langle \psi_{mk} | \partial_\mu | \psi_{nk} \rangle \langle \psi_{nk} | \partial_\nu | \psi_{m'k} \rangle \\
&\quad - (\mu \leftrightarrow \nu); \\
\epsilon^{\mu\nu\lambda\omega} \text{Tr}(\mathcal{F}_{\mu\nu} \mathcal{F}_{\lambda\omega}) &= \epsilon^{\mu\nu\lambda\omega} \text{Tr}(\tilde{\mathcal{F}}_{\omega\mu} \tilde{\mathcal{F}}_{\nu\lambda}). \tag{50}
\end{aligned}$$

Note that in the expression for the Berry's curvature \mathcal{F} , we use the whole Bloch wave function $|\psi\rangle$ instead of the periodic part $|u\rangle$ but here it makes no difference.

Now we start from eqn.(28). Write explicitly in position basis, we have

$$\begin{aligned}
\text{Tr}((\partial_\alpha \rho) \rho (\partial_\beta \rho)) &= \int dr_1 dr_2 dr_3 (\partial_\alpha \bar{\rho}_{12}) \bar{\rho}_{23} (\partial_\beta \bar{\rho}_{31}) \\
&\quad \exp \left(-\frac{i}{2} B \epsilon^{\gamma\delta} (r_2 - r_1)_\gamma (r_3 - r_1)_\delta \right) \\
&= L_x L_y \int \frac{d^2 k'}{(2\pi)^2} \left(\text{Tr}((\partial_\alpha \bar{\rho}) \bar{\rho} (\partial_\beta \bar{\rho})) \right)
\end{aligned}$$

$$+ \frac{i}{2} B \epsilon^{\gamma\delta} \text{Tr}((\partial_\alpha \partial_\gamma \rho_0) \rho_0 (\partial_\delta \partial_\beta \rho_0)) + \mathcal{O}(B^2), \tag{51}$$

where in the second equality we Taylor-expand in B , keep up to first order and go back to momentum space. We have also taken the infinite-size limit and make the sum of discrete momenta an integral. The trace on the right hand side traces over only the band indices.

The remaining task would be to plug in $\bar{\rho}$ and calculate explicitly to first order in B . One thing to notice is that when taking derivatives of $\bar{\rho}$, it acts not only on the matrix element but also on the basis. It is also useful to note that $\partial_u \rho_0$ only has non-vanishing matrix elements between the original occupied and empty states.

As we can see from eqn.(48), the inter-gap and intra-gap matrix element of $\bar{\rho}$ look pretty different. Let us denote the former as ρ' . ρ' contributes only through the first term in the right hand side of eqn.(51); since ρ' is already first order in B the remaining $\bar{\rho}$ can be replaced by ρ_0 . Let $\rho' = A + A^\dagger$ with $A = (1 - \rho_0) \rho' \rho_0$ (that is, A is the matrix element connecting occupied bands to empty bands and vice versa for A^\dagger), after explicit calculation, similar to eq. (47), we have

$$\begin{aligned}
& \epsilon^{\alpha\beta} (\text{Tr}(\partial_\alpha \rho' \rho_0 \partial_\beta \rho_0) + \text{Tr}(\partial_\alpha \rho_0 \rho' \partial_\beta \rho_0) + \text{Tr}(\partial_\alpha \rho_0 \rho_0 \partial_\beta \rho')) \\
&= -\partial_\alpha (\langle n | \partial_\beta | m \rangle A_{mn} - c.c.) \\
&= \partial_\alpha \text{Tr}(\partial_\beta \rho_0 (1 - \rho_0) \rho' \rho_0 + h.c.); \tag{52}
\end{aligned}$$

$|n\rangle$ is the short hand notation of $|\psi_{nk}\rangle$ and repeated indices are summed over.

Now that the inter-gap matrix elements are dealt with, the remaining part of the first term can also be expanded and calculated:

$$\begin{aligned}
\epsilon^{\alpha\beta} \text{Tr}(\partial_\alpha \bar{\rho} \partial_\beta \bar{\rho})|_{\text{remaining}} &= \epsilon^{\alpha\beta} \left(\frac{i}{2} \text{Tr}(\mathcal{F}_{\alpha\beta}) \right. \\
&\quad \left. - \frac{3i}{8} \epsilon^{\gamma\delta} \text{Tr}(\mathcal{F}_{\alpha\beta} \mathcal{F}_{\gamma\delta} - \tilde{\mathcal{F}}_{\alpha\beta} \tilde{\mathcal{F}}_{\gamma\delta}) \right. \\
&\quad \left. + \mathcal{O}(B^2) \right). \tag{53}
\end{aligned}$$

The first term on the right hand side is proportional to B^0 and is similar to the polarization in 1D.

The only remaining part is the second term in eqn.(51). This part proves to be somewhat tricky to calculate as one has to manually group terms into expressions of \mathcal{F} and $\tilde{\mathcal{F}}$. Nevertheless, it is otherwise straight forward and one gets

$$\begin{aligned}
& \epsilon^{\alpha\beta} \epsilon^{\gamma\delta} \text{Tr}(\partial_\gamma \partial_\alpha \rho_0 \rho_0 \partial_\delta \partial_\beta \rho_0) = \\
& \text{Tr} \left(\frac{3}{4} \mathcal{F}_{\alpha\beta} \mathcal{F}_{\gamma\delta} - \frac{1}{4} \tilde{\mathcal{F}}_{\alpha\beta} \tilde{\mathcal{F}}_{\gamma\delta} + \tilde{\mathcal{F}}_{\delta\alpha} \tilde{\mathcal{F}}_{\gamma\beta} \right). \tag{54}
\end{aligned}$$

Combining eqn.(52), eqn.(53), and eqn.(54), and with the help of eqn.(50) we get eqn. (30).

-
- ¹ C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005)
- ² Liang Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007)
- ³ J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007)
- ⁴ Xiao-Liang Qi, Taylor L. Hughes, and Shou-Cheng Zhang, Phys. Rev. B **78**, 195424 (2008)
- ⁵ Andrew M. Essin, Joel E. Moore, and David Vanderbilt, Phys. Rev. Lett. **102**, 146805 (2009)
- ⁶ S. Coleman, Ann. Phys. (N.Y.) **101**, 239 (1976).
- ⁷ G. Rosenberg and M. Franz, Phys. Rev. B. **82**, 035105 (2010)
- ⁸ R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993)
- ⁹ N. Read and S. Sachdev, Phys. Rev. Lett. **62**, 1694 (1989)
- ¹⁰ A. M. Essin, A.M. Turner, J. E. Moore, and D. Vanderbilt, Phys. Rev. B **81**, 205104 (2010)
- ¹¹ Malashevich et. al., New J. Phys. **12**:05032 (2010)