

# Unified Topological Response Theory for Gapped and Gapless Free Fermions

Daniel Bulmash, Pavan Hosur, Shou-Cheng Zhang, Xiao-Liang Qi

*Department of Physics, Stanford University, Stanford, California 94305-4045, USA*

(Dated: May 29, 2015)

We derive a scheme for systematically characterizing the responses of gapped as well as gapless systems of free fermions to electromagnetic and strain fields starting from a common parent theory. Using the fact that position operators in the lowest Landau level of a quantum Hall state are canonically conjugate, we consider a massive Dirac fermion in  $2n$  spatial dimensions under  $n$  mutually orthogonal magnetic fields and reinterpret physical space in the resulting zeroth Landau level as phase space in  $n$  spatial dimensions. The bulk topological responses of the parent Dirac fermion, given by a Chern-Simons theory, translate into quantized insulator responses, while its edge anomalies characterize the response of gapless systems. Moreover, various physically different responses are seen to be related by the interchange of position and momentum variables. We derive many well-known responses, and demonstrate the utility of our theory by predicting spectral flow along dislocations in Weyl semimetals.

## I. INTRODUCTION

Spurred by the discovery of topological insulators, topological phases have become a vital part of condensed matter physics over the last decade[1–4]. Even in the absence of interactions, a wide variety of gapped topological phases of fermions are now known, ranging from the quantum Hall[5, 6] and the quantum spin Hall[7–13] insulators among insulators to the chiral  $p$ -wave superconductor[14, 15] and the  $B$  phase of Helium-3[16, 17] among superconducting phases. All these phases share some common features: as long as certain symmetry conditions are upheld, they have a bulk band structure that cannot be deformed into that of an atomic insulator – a *trivial* insulator by definition – without closing the band gap along the way. Moreover, they all have robust surface states that mediate unusual transport immune to symmetry-respecting disorder. This leads one to wonder whether all gapped phases of free fermions can be unified within a common mathematical framework.

Two different approaches have been developed to provide unified characterization of gapped phases of free fermions. In the topological band theory approach [18–20], homotopy theory and  $K$ -theory are applied to classify free fermion Hamiltonians in a given spatial dimension and symmetry class. The topological band theory provides a complete topological classification of free fermion gapped states in all dimensions and all the 10 Altland-Zirnbauer symmetry classes[21]. However, it does not directly describe physical properties of the topological states. In comparison, the topological response theory approach[22–26] describes topological phases by topological terms in their response to external gauge fields and gravitational fields. The advantage of this approach is that the topological phases are characterized by physically observable topological effects, so that the robustness of the topological phase is explicit and more general than in the topological band theory. Since it is insensitive to details of the microscopic Hamiltonian, a

response theory based classification scheme can be further extended to strongly interacting systems[27].

Recently, the advent of Weyl semimetals (WSMs) has triggered interest in gapless topological phases of free fermions[28–34]. These phases are topological in the sense that they cannot be gapped out perturbatively as long as momentum and charge are conserved. In this regard, ordinary metals are also topological since their Fermi surfaces are robust in the absence of instabilities towards density waves or superconductivity. Additionally, gapless topological phases may have non-trivial surface states such as Fermi arcs[28, 35–37] and flat bands[38]. Teo and Kane[39] applied homotopy arguments to classify topological defects such as vortices and dislocations in gapped phases; Matsuura *et. al.*[40] used an analogous prescription to classify gapless phases by observing that gapless regions in momentum space such as Fermi surfaces and Dirac nodes can be viewed as topological defects in momentum space in a gapped system. Thus, a common mathematical formalism to describe the Bloch Hamiltonians of gapless phases was derived.

Unlike their gapped counterparts, however, it is not clear whether the response theories of gapless phases are amenable to a unified description. For gapped systems, the path from the Hamiltonian to the response theory is conceptually straightforward: the fermions are coupled to gauge fields and integrated out to get the low energy effective field theory, which describes the topological response properties. In contrast, the low energy theory of gapless phases contains fermions as well as gauge fields, and is distinct from the response theory which contains only gauge fields. Thus, it is not obvious how the topological properties of the Hamiltonian affect the response. The response depends on system details in general and therefore, recognizing its universal features and then unifying the responses of various gapless phases is a non-trivial task. A few cases of topological response properties of gapless fermions have been studied. One example is the intrinsic anomalous Hall effect of a two-

dimensional Fermi gas[41–45]. A generalization of this effect in three-dimensional doped topological insulators has been discussed[46]. Another example is the topological response of Weyl semimetals, which has been described in the form of the axial anomaly[34, 47–60]. This refers to the apparent charge conservation violation that occurs for each Weyl fermion branch in the presence of parallel electric and magnetic fields, although the net charge of the system must still be conserved. Recently, these ideas were generalized to find the topological responses of point Fermi surfaces in arbitrary dimensions[61]. The DC conductivity of metals has also been proposed to be related to a phase space topological quantity [62]. However, a general theory that describes the topological properties of gapless fermions in a unified framework has not been developed yet.

In this work, we achieve the above goals for free fermions: we show that gapless systems have universal features, independent of system details, and derive a unified description of their response. Remarkably, this description also captures the response of gapped systems. In particular, the response of gapped phases arises from the bulk response of a certain parent topological phase, while the universal features of the gapless phases correspond to its edge anomalies. We elucidate this idea below.

The backbone of our construction is a mapping from  $n$ -dimensional gapped *or* gapless systems to a *gapped* quantum Hall (QH) system which lives in  $2n$ -dimensional phase space. Such a phase space system has both bulk responses, given by a  $2n$ -dimensional Chern-Simons (CS) theory, and boundary (axial) anomalies. We identify the bulk responses with topological responses of insulators. However, the key insight that allows us to include gapless systems is to identify a Fermi surface in real space with a phase space boundary in the momentum directions. Likewise, real space excitations near the Fermi surface are identified with the gapless edge excitations in phase space. The universal features of the response of gapless systems are thus the anomalies associated with these phase space edges.

There is an important technical point required in order to bestow the  $2n$ -dimensional QH system with the interpretation of phase space. Specifically, we must choose the QH system to consist of a massive Dirac fermion under  $n$  uniform magnetic fields of strength  $B_0$  in  $n$  orthogonal planes, then project to the zeroth Landau level (ZLL) of the total field. In this case, the projected operators for pairs of dimensions acquire the usual canonical commutation relations that relate ordinary real and momentum space (up to an overall factor of  $B_0$ ). This allows us to interpret the ZLL of the  $2n$ -dimensional QH system as phase space for the  $n$ -dimensional physical space. We interpret additional perturbations in the phase space gauge fields as physical quantities such as the  $n$ -dimensional system’s electromagnetic (EM) field, strain field, Berry

curvatures, and Hamiltonian. Topological defects, such as monopoles, in the phase space gauge fields allow us to generalize to systems with dislocations and with point Fermi surfaces such as graphene and Weyl semimetals. These ideas are summarized in Table I.

This construction enables us to systematically enumerate all possible intrinsic, topological responses to electromagnetic and strain fields in the DC limit in any given dimension. One simply has to write down the Chern-Simons action in phase space, vary it with respect to each gauge field, and consider each boundary to obtain all the bulk, boundary and gapless responses in real space. Following this procedure, we show carefully that screw dislocations in Weyl semimetals trap chiral modes which are well-localized around the dislocation at momentum values away from the Weyl nodes. A related but different effect has been studied previously [63]. However, our framework provides a unified and natural description of this effect and other topological effects.

It is crucial that the  $2n$ -dimensional system be gapped even in the absence of the background magnetic fields of strength  $B_0$ . This ensures that its response theory contains terms depending on  $B_0$  in addition to fluctuations in the gauge fields. In  $n$  dimensions, we will see that the  $B_0$ -dependent terms translate into quasi-lower-dimensional responses, such as the polarization of a system of coupled chains along the chains. If the  $2n$ -dimensional system is gapless in the absence of the background fields, such responses will be missed by the unified theory.

A caveat is that our construction does not capture responses to spatial, momental and temporal variations in the field strengths, such as the gyrotropic effect which is an electric response to a spatially varying electric field. Note that the regular Maxwell response, given by  $j^\mu = \partial_\nu F^{\nu\mu}$  is a response to a variation in the field strength. Another caveat is that in phase space dimensions equal to 4 and above, the Maxwell term in the action is equally or more relevant than the Chern-Simons term and hence will, in general, dominate the DC response. However, our central objective is to demonstrate that there exists a theory which unifies the responses of gapped and gapless systems, namely, the phase space Chern-Simons theory.

The rest of this paper is structured as follows. In Section II, we review the key property of the ZLL which provides the physical justification for our construction. In Section III we explain the interpretation of the gauge fields in our mapping and give an example illustrating the validity of the CS theory. In Section IV we write down an explicit model with a CS response theory and show the precise way in which it behaves as the phase space response theory of a lower-dimensional model. In Sections V and VI we explain the responses and anomalies (respectively) that come from the CS theory in various dimensions, applying our framework to describe spectral flow in Weyl semimetals with dislocations. Finally, in

Phase space	Real space
Bulk responses	Quantized insulator responses
Anomalies from momentum direction edges	Gapless response
Anomalies from real direction edges	Real edge anomalies
Gauge field strength	EM field strength/ $k$ -space Berry curvature/Strain
Monopole in gauge field	Magnetic monopole/Weyl node/dislocation

TABLE I. Dictionary for interpreting phase space quantities in real space.

Section VII we summarize our work and suggest extensions of our theory to more nontrivial systems.

## II. A REVIEW OF THE ALGEBRA OF THE ZEROth LANDAU LEVEL

One of the key features that we use in the intuition for our approach is the fact that projecting position operators to the ZLL yields nonzero commutators between those operators. We now review this fact, for concreteness as well as for later convenience, for the case of Dirac electrons in a uniform magnetic field in two spatial dimensions in Landau gauge. Although we consider the ZLL of Dirac electrons here, the non-commutativity of position operators is simply a consequence of minimal coupling and Landau quantization of cyclotron orbits and hence is true for other dispersions as well as for other Landau levels for a Dirac dispersion.

Consider a 2D massive Dirac Hamiltonian in a uniform magnetic field

$$H = (p_x - eBy)\sigma_x + p_y\sigma_y + m\sigma_z \quad (1)$$

Here  $\sigma_i$  are the Pauli matrices. We have set the Fermi velocity to unity, written the electron charge as  $-e$ , and chosen the Landau gauge  $\mathbf{A} = -By\hat{\mathbf{x}}$  with  $B > 0$  for definiteness. Note that  $p_x$  commutes with the Hamiltonian, so we may replace it by its eigenvalue. We can define an annihilation operator  $a = (p_x - eBy - ip_y)/\sqrt{eB}$ , which has  $[a, a^\dagger] = 1$ , and the Hamiltonian becomes

$$H = \begin{pmatrix} m & \sqrt{2eB}a \\ \sqrt{2eB}a^\dagger & -m \end{pmatrix} \quad (2)$$

It is straightforward to show that the eigenstates are labeled by an eigenvalue  $n \geq 0$  of the number operator  $a^\dagger a$ , with dispersion  $\pm\sqrt{2eBn + m^2}$  for  $n \neq 0$ . For  $n = 0$ , the eigenvalue is  $-m$ , the spin state is  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , and the state is annihilated by  $a$ . This is the expected result that the kinetic energy is quenched and the spectrum becomes discrete, highly degenerate Landau levels.

Let  $|k_x\rangle$  be the state in the ZLL with  $p_x$  eigenvalue  $k_x$ . Then the projection operator to the ZLL is

$$P = \int dk_x \frac{L}{2\pi} |k_x\rangle\langle k_x| \quad (3)$$

with  $L$  the system length in the  $x$  direction. Writing  $y = ((p_x/\sqrt{eB}) - (a + a^\dagger))/\sqrt{eB}$ , the projected  $y$  operator becomes

$$PyP = \int dk_x \frac{L}{2\pi} \frac{k_x}{eB} |k_x\rangle\langle k_x| \quad (4)$$

where we have used the fact the  $a$  and  $a^\dagger$  describe inter-Landau level processes and thus, vanish under projection onto the ZLL. Next, using the fact that  $|k_x\rangle$  is an eigenstate of  $p_x$ , we find

$$PxP = \int dk_x \frac{L}{2\pi} i\partial_{k_x} |k_x\rangle\langle k_x| \quad (5)$$

The commutator can then easily be computed to be

$$[PxP, PyP] = \frac{i}{eB} \quad (6)$$

Hence if we absorb the factor of  $eB$  into  $y$ , then  $PxP$  and  $PyP$  have the correct commutator structure for us to imbue them with the interpretation of the position and momentum operators, respectively, of a 1D system. This interpretation is the primary physical motivation for the construction which follows. As mentioned earlier, other dispersions will also result in commutation relations similar to Eq. (6) and thus imbue  $x$  and  $y$  with interpretations of position and momentum of a 1D system. However, a massive Dirac dispersion is ideal for deriving the unified response theory because it does not miss any quasi-lower-dimensional responses, as mentioned earlier and detailed later.

In higher dimensions, the Dirac model is  $H = \sum_i (p_i + eA_i)\Gamma_i$ , where the  $\Gamma_i$  are anticommuting elements of the Clifford algebra of  $2n$  by  $2n$  matrices. If we apply constant magnetic fields  $F_{ij}$  for disjoint pairs  $(i, j)$  of coordinates, we can form an annihilation operator for each such pair. Annihilation operators from different pairs commute, and the analysis above carries through so that the position operators within each pair no longer commute after projection.

## III. PHASE SPACE CHERN-SIMONS THEORY

The key idea of our construction is to represent a (possibly gapless)  $n$ -dimensional system by a gapped  $2n$ -dimensional phase space system, specifically a massive

Dirac model coupled to a gauge field. As we just showed, we can interpret a  $2n$ -dimensional system as living in phase space by adding background magnetic fields between disjoint pairs of spatial directions and projecting to the ZLL. Moreover, since the phase space system is gapped, we can immediately write down a response theory for it, the topological part of which can be proved to be a CS theory[22, 64–66]. Note that in a CS theory, real and momentum space gauge fields enter the action in similar ways, analogous to our idea of treating position and momentum on the same footing in phase space.

Before proceeding, we fix some notation and conventions. We will always use the Einstein summation convention where repeated indices are summed. Phase space coordinates will be labelled by  $x, y, z$  and  $\bar{x}, \bar{y}, \bar{z}$ . After projection,  $x, y, z$  will be interpreted as the corresponding real space coordinates, while  $\bar{x}, \bar{y}, \bar{z}$  will be interpreted as momentum space coordinates  $k_x, k_y, k_z$  respectively. In phase space, we will refer to the  $U(1)$  background gauge field which generates the real space commutator structure as  $A_\mu$  with its nonzero field strengths being  $F_{\bar{i}\bar{j}} = B_0$  for  $i = (x, y, z)$ . We denote all other contributions to the gauge field by  $a_\mu$  and the total gauge field by  $\mathcal{A}_\mu = A_\mu + a_\mu$ . Likewise, we write  $f_{\mu\nu}$  and  $\mathcal{F}_{\mu\nu} = F_{\mu\nu} + f_{\mu\nu}$  for the non-background and total field strengths respectively. The (non-Abelian) field strengths are as usual defined by  $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu - [a_\mu, a_\nu]$ .

We will also abbreviate the CS Lagrangian by  $\epsilon a \partial a \equiv \epsilon^{\mu\nu\sigma} a_\mu \partial_\nu a_\sigma$ , where  $\epsilon$  is the totally antisymmetric Levi-Civita tensor, with an analogous abbreviation for higher-dimensional CS terms. Finally, we set  $e = \hbar = 1$ , and also assume that  $\sqrt{B_0} \sim 1/l_B$  is very large compared to all other wavenumbers in the problem.

### A. Interpretation of Phase Space Gauge Fields

Our prescription is that the non-background contributions  $a_\mu$  to the phase space gauge field should be interpreted as the Berry connection for the lower-dimensional system:

$$a_\mu^{\alpha\beta} = i \langle u_{k\alpha} | \partial_\mu | u_{k\beta} \rangle \quad (7)$$

where  $|u_{k\alpha}\rangle$  is the (local) Bloch wavefunction at momentum  $k$  for the  $\alpha$  band. Here  $\partial_{\bar{x}, \bar{y}, \bar{z}}$  should be interpreted as  $B_0 \partial_{k_x, k_y, k_z}$ . We can think of the physical EM vector potential as a Berry connection, which means that it is included in the real-space components of  $a$ .

As such, we will often use the following heuristic interpretations in order to more clearly see the physics:  $-a_t$  is the lower-dimensional band Hamiltonian plus the physical EM scalar potential,  $a_{x,y,z}$  is the physical EM vector potential, and  $a_{\bar{x}, \bar{y}, \bar{z}}$  is the momentum-space Berry connection. The field strengths which do not mix real and momentum space then have natural interpretations as the physical EM field strengths and Berry curvatures.

The physical interpretation of “mixed” field strengths such as  $f_{x\bar{y}}$  (in 4 or higher dimensional phase space) is less obvious. Here we present two ways to think about them. First consider a gauge where  $\partial_{\bar{y}} a_x = 0$ . We find that

$$\int d\bar{y} f_{x\bar{y}} = \partial_x \int d\bar{y} a_{\bar{y}} = 2\pi \partial_x P_y \quad (8)$$

where  $P_y$  is the one-dimensional polarization of the system[67]. A spatially varying polarization can be thought of as strain of the electron wavefunction, which can come from either mechanical strain or some other spatial variation of the parameters entering the band structure.

To make the connection of  $f_{x\bar{y}}$  to mechanical strain more explicit, we change the gauge to set  $\partial_x a_{\bar{y}} = 0$ . An intuitive way to think about a nonzero  $f_{x\bar{y}}$  in this gauge is in terms of dislocations. In particular, adiabatically moving a particle around a real space dislocation leads to a translation, but if the particle can locally be treated as a Bloch wave, then that translation is equivalent to the accumulation of a phase. This (Berry) phase is equal to  $\mathbf{k} \cdot \mathbf{b}$ , with  $\mathbf{b}$  the Burgers vector of the dislocation. In particular, this is a *momentum-dependent* Berry phase resulting from adiabatic motion in real space. Hence  $f_{x\bar{y}}$  is nonzero. It can be shown explicitly[44] in the perturbative regime that strain typically leads to such a Berry phase.

### B. Example: 2D Phase Space

We first consider the case where our real space theory consists of a single filled band living in 1D and that a 2D CS response term in phase space with a background field describes the expected responses. We will, for simplicity, only consider Abelian physics in this example. Consider the CS action

$$S_{CS} = \frac{1}{4\pi} \int dt d^2 x C(\bar{x}, x) \epsilon \mathcal{A} \partial \mathcal{A} \quad (9)$$

(In Section IV, we will show in an explicit model how Eq. 9, with this (quantized) coefficient, appears, but for now we simply assume that it is the relevant response theory.) Here  $C(\bar{x}, x)$  accounts for the filling at different points; for example, if the system occupies  $x > 0$ , then  $C(\bar{x}, x)$  will be proportional to  $\Theta(x)$  with  $\Theta$  the Heaviside step function, as shown in Fig. 1(a). Likewise, if the system has a Fermi momentum  $k_F$ , then  $C(\bar{x}, x)$  will be proportional to  $(\Theta(\bar{x} + k_F/B_0) - \Theta(\bar{x} - k_F/B_0))$ , as shown in Fig. 1(b).

Let us assume that there are no edges so that  $C = 1$  everywhere. Then, the responses for this action, given by  $j^\mu = \delta S / \delta \mathcal{A}_\mu$ , are

$$j_{2D}^\mu = \frac{1}{2\pi} \epsilon^{\mu\nu\sigma} \mathcal{F}_{\nu\sigma} \quad (10)$$



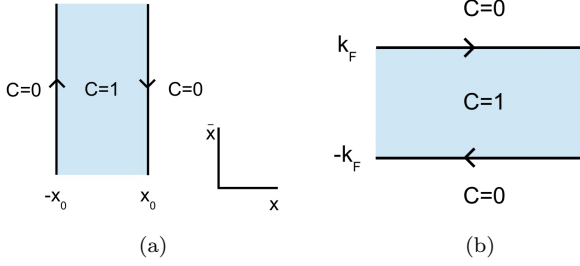


FIG. 1. Phase space realization of (a) real-space edges (b) Fermi points for a 1D real space system. Arrows indicate the direction of the edge modes.

where  $\mathcal{F}_{\nu\sigma} = \partial_\nu \mathcal{A}_\sigma - \partial_\sigma \mathcal{A}_\nu$  is the field strength tensor corresponding to  $\mathcal{A}$ . Let us consider each component, assuming for conciseness that the background field is in Landau gauge  $A_{\bar{x}} = B_0 x$ .

First we examine the real-space response  $j_{2D}^x = \mathcal{F}_{\bar{x}t}/2\pi$ . This current in general depends on  $\bar{x}$ , which we interpret as  $k_x/B_0$ ; the observable current should then be given by integrating the 2D current with respect to  $\bar{x}$ , as the real-space current has contributions from all occupied momenta. The resulting 1D current is

$$j_{1D}^x = \frac{1}{2\pi} \left( \int d\bar{x} \partial_{\bar{x}} a_t - \partial_t \int d\bar{x} a_{\bar{x}} \right) \quad (11)$$

Interpreting the  $\bar{x}$ -dependent part of  $a_t$  as the dispersion, the first integral generically gives zero. The second integral is, for a gapped system, exactly the time derivative of the polarization  $P_x = \frac{1}{2\pi} \int a_{\bar{x}} d\bar{x}$ , which is the expected 1D real-space current response  $j_{1D}^x = -\partial_t P_x$ . Similarly, the  $k$ -space response is  $j_{2D}^{\bar{x}} = \mathcal{F}_{tx}/2\pi$ . Interpreting  $j_{2D}^{\bar{x}}$  as  $dk/dt$ , we recover the real space semiclassical equation of motion  $dk/dt = E/2\pi$  with  $E$  the electric field.

Finally, the charge response is given by

$$\rho_{1D} = \frac{1}{2\pi} \int d\bar{x} (F_{x\bar{x}} + \partial_x a_{\bar{x}} - \partial_{\bar{x}} a_x) \quad (12)$$

In units  $B_0 = 1$ , the first term simply gives the total charge in the occupied band, which can be thought of a quasi-0D response.

The second term of Eq. (12), in a gauge where  $a_x = 0$ , becomes  $\partial_x P_x$  for a gapped system. This is again intuitive; if, say, the system is strained, then the polarization and hence the charge density will change accordingly.

If we now impose a pair of edges at  $\bar{x} = \pm k_F$ , two things happen. Firstly, the 1D system lies between a pair of momentum points  $\pm k_F$ , so the integrals in (11) and (12) run from  $-k_F$  to  $+k_F$  instead of the full Brillouin zone. Consequently, the background charge becomes  $\rho_{1D}^{bg} = \frac{1}{2\pi} \int_{-k_F}^{k_F} F_{x\bar{x}} = k_F/\pi$ , as expected, while the terms proportional to  $f_{x\bar{x}}$  cease to have a simple interpretation as the polarization but can be non-zero

nonetheless. Secondly, the 2D system develops a chiral anomaly at the edge, given by

$$\partial_t \rho_{2D} + \partial_x j_{2D}^x = \frac{1}{2\pi} \mathcal{F}_{tx} \partial_{\bar{x}} f(\bar{x}) = \frac{E + \partial_x \varepsilon}{2\pi} \partial_{\bar{x}} C(\bar{x}) \quad (13)$$

where  $C(\bar{x}) = \Theta(\bar{x} + k_F) - \Theta(\bar{x} - k_F)$ . Integrating over 1D real space under a constant electric field and translational invariance yields

$$\partial_t \int dx \rho_{2D} = \frac{L}{2\pi} (\delta(\bar{x} + k_F) - \delta(\bar{x} - k_F)) E \quad (14)$$

where  $L$  is the length of the system and  $\delta$  is the Dirac delta function. This is precisely the chiral anomaly in the 1D system: the electric field tilts the 1D Fermi surface, effectively converting right-moving charge in the vicinity of one Fermi point into left-moving charge near the other. Thus, we have derived a property of a gapless 1D band structure from the edge anomaly of the parent 2D QH system.

Notice also that integration of (13) over momentum space leads to

$$\partial_t \rho_{1D} + \partial_x j_{1D}^x = 0 \quad (15)$$

which correctly tells us that there is no anomaly in the total charge. The precise value of  $\rho_{1D}$  and  $j_{1D}^x$  depends on system details; therefore, calculating them in our formalism would require knowledge of non-universal properties of the 2D QH edge, such as the velocity of the chiral modes. However, we have shown here that they still have universal properties that reflect the universal properties of a higher dimensional topological state.

A different type of anomaly occurs when the system has real-space edges and a filled band. In this case, the anomaly equation in 2D is

$$\partial_t \rho_{2D} + \partial_{\bar{x}} j_{2D}^{\bar{x}} = -\frac{1}{2\pi} \mathcal{F}_{t\bar{x}} \partial_x C(x) \quad (16)$$

Integrating the above in  $\bar{x}$  yields

$$\partial_t \rho_{1d} = ((\delta(x - x_0) - \delta(x + x_0)) \partial_t P(x) \quad (17)$$

This is the known result[68] that charge can be adiabatically pumped from one edge of the system to the other via a time-dependent local polarization.

We thus see that the standard responses, including anomalies, that we expect in a 1D theory are retrieved from the 2D CS theory. However, detail-dependent edge responses are described in our theory only by the anomaly (or lack thereof) that they create. We expect the same procedure to generalize to higher dimensions, and we will show that the expected topological responses appear in Section V.

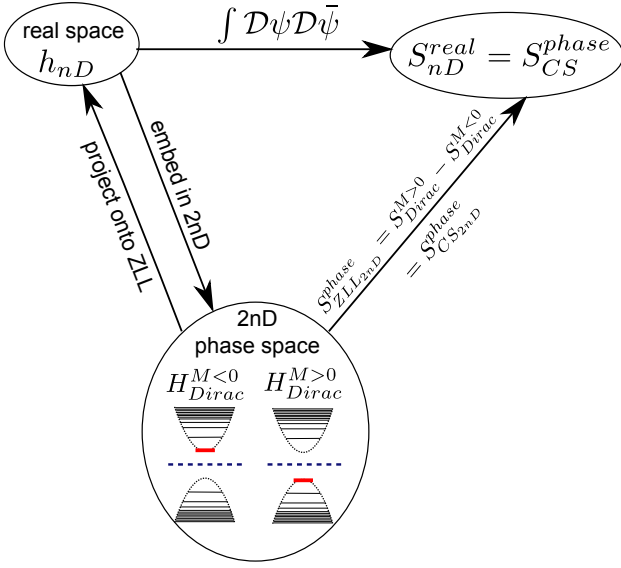


FIG. 2. Logical flow of the derivation in Section IV. An  $nD$  Hamiltonian  $h_{nD}$  can be obtained by projecting a  $2nD$  massive Dirac Hamiltonian  $H^M_{Dirac}$  in a magnetic field onto the ZLL, denoted by the thick red bar. For a Fermi level in the Dirac mass gap, the  $M > 0$  and  $M < 0$  ground states differ only in the occupation of the ZLL, while their response theories  $S^M_{Dirac}$  differ by the CS-term in  $2nD$ . Thus, the response theory of the ZLL  $S_{ZLL2nD}$ , which is the real space response theory  $S^{real}_{nD}$  of  $h_{nD}$ , is the phase space CS-theory  $S^{phase}_{CS}$ .

#### IV. EXPLICIT MODEL

In this section, we elucidate the precise way in which a ZLL behaves as the phase space of a system in half the number of spatial dimensions. In particular, we explain why the response theory of the lower dimensional system should be of CS form in phase space and describe the physical meaning of projecting onto the ZLL. We also answer the question of when a CS theory in  $2nD$  can be interpreted as a phase space response theory in  $nD$ .

To begin, consider a massive Dirac Hamiltonian in  $2n$  dimensions coupled to the gauge field  $\mathcal{A}$  defined in Sec. III:

$$H_{2nD} = \sum_{i=1}^n [\Gamma_i(p_i - \mathcal{A}_i) + \Gamma_{\bar{i}}(p_{\bar{i}} - \mathcal{A}_{\bar{i}})] + \Gamma_0 M \quad (18)$$

$\mathcal{A}$  corresponds to large constant background fields  $B_0$  in  $n$  orthogonal planes plus small fluctuations; thus,  $F_{i\bar{i}} = B_0 \gg f_{ij}, f_{i\bar{j}}, f_{\bar{i}\bar{j}}$  for all  $i, j \in 0 \dots n$ . The  $\Gamma$ 's are  $2n \times 2n$  anticommuting matrices with eigenvalues  $\pm 1$  and satisfy  $\Gamma_0 = \prod_{i=1}^n \Gamma_i \Gamma_{\bar{i}}$ . To zeroth order in  $f$ , the spectrum of  $H_{2nD}$  can be easily derived by generalizing the calculation of Sec II; it consists of Landau levels with energies  $\pm \sqrt{2kB + M^2}$  for positive integers  $k$  together with a ZLL state which has energy  $-M$  and a spinor wavefunction that has a  $\Gamma_0$  eigenvalue of  $-1$ .

We have two tasks. First, we must isolate the topological response theory of the ZLL of this system, which we expect to be a Chern-Simons theory. Second, we must relate this Hamiltonian, projected onto the ZLL of the total field, to the Hamiltonian of the real space system.

For the first task, note that the ZLL is occupied (unoccupied) in the ground state if  $M > 0$  ( $M < 0$ ), while the occupation of all the other Landau levels is independent of the sign of  $M$ . This should hold for non-zero  $f$  as well if  $M \gg \sqrt{B_0}$ . As a result, the response of the ZLL to  $\mathcal{A}$  is given by the terms in the total response that are *odd* in  $M$ . Moreover, it is known that the two signs of  $M$  correspond to a topological and a trivial insulator (which sign corresponds to which phase is determined by the regularization far away from the Dirac point). Therefore, the difference between their response theories, which by definition is the topological part of the effective action, equals the response of the ZLL. In the absence of vertex corrections, this is known to be the  $2n$ -dimensional CS action with coefficient 1 to lowest order in the coupling constant  $e$ . In short, the response of the ZLL is precisely the CS action with coefficient 1 in appropriate dimensions under suitable well-controlled perturbative approximations. We emphasize that this statement is true even if  $H_{2nD}$  is modified at high energies to change the total ( $n^{\text{th}}$ ) Chern numbers of the occupied and unoccupied bands. The only requirement is that the Chern numbers of the  $M > 0$  and  $M < 0$  cases differ by unity; their actual values are irrelevant for determining the ZLL response.

Next, we recall that  $[x_i, \bar{x}_i] = i$  in the ZLL as shown in Sec II, so  $x_i$  and  $\bar{x}_i$  can be thought of as a pair of canonically conjugate position and momentum variables. Therefore, projecting  $H_{2nD}$  onto the ZLL gives an  $n$ -dimensional system whose response theory is guaranteed to be of CS form in phase space. In this response theory, the gauge fields in the “momentum” directions are to be reinterpreted as momentum space Berry connections. This flow of logic is depicted in Fig 2 (where we have renamed  $H_{2nD}$  as  $H^M_{Dirac}$  to make the figure self-contained).

Having shown that the response of the  $n$ -dimensional system is given by the phase space CS theory, we turn to our second task and show in detail how the Hamiltonian in  $n$ -dimensions is related to  $H_{2nD}$ . For clarity, we choose  $n = 1$ , i.e., we demonstrate this in 1D real space with a  $U(1)$  gauge field; the procedure generalizes straightforwardly to more dimensions and to larger gauge groups. To construct the 2D phase space model, let  $\Gamma_x$ ,  $\Gamma_{\bar{x}}$ , and  $\Gamma_0$  be the Pauli matrices  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$ . (The  $\Gamma$  notation is for consistency with the higher-dimensional generalization in Eq. (18).) The appropriate 2D Hamiltonian is

$$H_{2D} = (p_x - \mathcal{A}_x)\Gamma_x + (p_{\bar{x}} - \mathcal{A}_{\bar{x}})\Gamma_{\bar{x}} + M\Gamma_0 + \mathcal{A}_0 \quad (19)$$

We are projecting onto the ZLL of the total field  $\mathcal{A}$ , so

we need to make some approximations to make progress. We assume that the field fluctuations  $f$  are much smaller than  $B_0$ . That is, we identify  $1/\sqrt{B_0}$  with some microscopic length scale like a lattice constant for the underlying real space system, and assume that all the gauge field fluctuations are small over that length scale. If this is true, then we can make the gauge choice that  $\partial_\mu a_\nu \ll B_0$  for all  $\mu, \nu$ . In this case, the Hamiltonian of the ZLL of  $\mathcal{A}$  can be computed by considering  $a$  to be a perturbation on the Hamiltonian with  $\mathcal{A} = A$ . We implement degenerate perturbation theory as follows.

Let us write

$$H_{2D} = H_0 + H' \quad (20)$$

with

$$H_0 = (p_x - A_x)\Gamma_x + (p_{\bar{x}} - A_{\bar{x}})\Gamma_{\bar{x}} + M\Gamma_0 - A_0 \quad (21)$$

$$H' = -a_x\Gamma_x - a_{\bar{x}}\Gamma_{\bar{x}} - a_0 \quad (22)$$

Since  $A$  is a constant background field of strength  $B_0$ , we know how to diagonalize  $H_0$ ; let  $|n, k\rangle$  be the eigen-

states where  $n$  labels the LL and  $k$  labels a momentum (in Landau gauge). Using  $\langle \bar{x}|0, k\rangle \propto e^{-B_0(\bar{x}-k/B_0)^2/2}(1, 0)$ , where the spinor indicates that the ZLL states are polarized in the basis of  $\Gamma_0$  eigenstates, and denoting  $k_\pm = k \pm q/2$ , first order degenerate perturbation theory gives an effective 1D Hamiltonian as

$$\begin{aligned} \langle 0, k_-|H'|0, k_+\rangle &\propto -\int dx d\bar{x} e^{iqx} a_0(x, \bar{x}) e^{-B_0(\bar{x}-k/B_0)^2} e^{-q^2/4B_0} \\ &\propto -a_0(-i\delta'(q), k/B_0) \\ \implies h_{1d}(k) &\equiv -a_0(i\partial_k, k/B_0) \end{aligned} \quad (23)$$

Thus, the desired 1D Hamiltonian  $h_{1d}(x, k)$  can easily be obtained by choosing  $a_0(x, \bar{x}) = -h_{1d}(x, B_0\bar{x})$ . Since the ZLL is spin-polarized, the dependence on  $a_x$  and  $a_{\bar{x}}$  disappears from (23); these fields only appear at second order in perturbation theory. Degenerate perturbation theory tells us that, if  $P$  is the projector onto the degenerate subspace, then the second order correction to the energy is given by the eigenvalues of

$$\langle \psi_i|H_{2D}|\psi_j\rangle = \langle 0, k_i|(H_0 + H' + (H' - PH'P)(H_0 - E_0)^{-1}(H' - PH'P))|0, k_j\rangle \equiv \langle 0, k_i|H_{2D} + H^{(2)}|0, k_j\rangle \quad (24)$$

where

$$|\psi_i\rangle = |0, n_i\rangle + \sum_{n>0, l} |n, k_l\rangle \langle n, k_l|(H_0 - E_0)^{-1}(H' - PHP)|n=0, k_i\rangle \quad (25)$$

is a basis for the perturbed ZLL wavefunctions up to first order in  $H'$ . In particular, the unitary transformation  $U$  which takes  $H_{2D}$  to  $H_{2D} + H^{(2)}$ , to second order in  $H'$ , is the one which takes  $|0, k_i\rangle$  to a state living in the ZLL of the full Hamiltonian, to first order in  $H'$ .

Therefore, if we find this unitary transformation and then perform the projection in the ZLL of the background field, we still get our desired projected Hamiltonian. We write  $U = \exp(iS)$  with  $S$  Hermitian, and expand  $S = S_1 + S_2 + \dots$  where the subscripts indicate an expansion in orders of  $H'$  (by inspection  $S$  can be chosen to have no zeroth order term). Then we can match, order by order, terms in  $e^{iS}H_{2D}e^{-iS}$  with those in  $H_{2D} + H^{(2)}$  to find the conditions

$$[H_0, S_1] = 0 \quad (26)$$

$$[H_0, S_2] = iH^{(2)} \quad (27)$$

We do not claim to be able to demonstrate explicitly a

unitary transformation which obeys the second of these conditions, as computing  $H^{(2)}$  is highly nontrivial. However, we will proceed first by exhibiting an ansatz for  $U$ , then showing that the projection onto the ZLL of  $A$  yields the correct Hamiltonian in real space, and finally giving the physical motivation for the ansatz.

Let us start in the gauge  $A_x = -B_0\bar{x}$ ,  $A_{\bar{x}} = A_t = 0$ . Then let

$$U = \left(e^{ia_x p_x ds/B_0}\right)^N e^{-iB_0 x \bar{x}} \left(e^{-ia_x p_x ds/B_0}\right)^N e^{iB_0 x \bar{x}} \quad (28)$$

where  $ds$  is an infinitesimal parameter and  $N \rightarrow \infty$  such that  $Nds = 1$ . This transformation is a gauge transformation, followed by a translation of  $\bar{x}$  by  $-a_x$ , followed by the reverse gauge transformation, followed by a translation of  $x$  by  $a_{\bar{x}}$ .

By inspection  $U$  commutes with  $H_0$ , so Eq. (26) is satisfied. To second order in  $H'$ , we now have

$$UH_{2D}U^\dagger \approx H_0 - a_x \left(x + \frac{a_{\bar{x}}}{B_0}, \bar{x} - \frac{a_x}{B_0}\right) \Gamma_x - a_{\bar{x}} \left(x + \frac{a_{\bar{x}}}{B_0}, \bar{x} - \frac{a_x}{B_0}\right) \Gamma_{\bar{x}} - a_0 \left(x + \frac{a_{\bar{x}}}{B_0}, \bar{x} - \frac{a_x}{B_0}\right) \quad (29)$$

where  $a_i$  appearing without explicit functional dependence means  $a_i(x, \bar{x})$ . The terms that we have neglected are “double-nestings” of  $a/B_0$ ; our aforementioned approximation that  $a$  is slowly varying (which was a gauge choice possible when the corresponding field strengths were weak) allows us to write

$$a_0 \left( x + \frac{a_{\bar{x}}}{B_0}, \bar{x} - \frac{a_x \left( x + \frac{a_{\bar{x}}}{B_0}, \bar{x} \right)}{B_0} \right) \approx a_0 \left( x + \frac{a_{\bar{x}}}{B_0}, \bar{x} - \frac{a_x}{B_0} \right) \quad (30)$$

Let us now perform the projection on the ZLL of the background field. As before, everything projects to zero except for the  $a_0$  term and the mass term of  $H_0$ . The latter just projects to a constant which we can absorb by a shift of  $a_0$ . However, we now obtain a different 1D Hamiltonian  $h'_{1d}(x, \bar{x}) = -a_0(x + a_{\bar{x}}, \bar{x} - a_x)$ , which is simply  $h_{1d}(x, \bar{x})$  with minimal coupling to the gauge fields  $a_{\bar{x}}$  and  $a_x$ , respectively ( $B_0$  set to unity for convenience). We therefore have correctly retrieved the full 1D Hamiltonian from a projection to the ZLL, as the functional form of the projected Hamiltonian is correct if we imbue  $x$  and  $\bar{x}$  with the interpretations of a parameter tracking a locally periodic Hamiltonian in space and Bloch momentum respectively.

A major question remains: why, physically, should this choice of  $U$  be the correct one? First of all, the projected Hamiltonian, if it is to describe a real system, must be gauge invariant. Hence the gauge fields should be minimally coupled, and  $U$  indeed accomplishes this goal.

A more fundamental reason, though, is the following. Consider  $H_{2D}$  in some local region over which  $a$  is approximately constant, and for convenience choose a gauge in which  $a_{\bar{x}}$  is zero. In this region,  $a_x$  functions as a constant shift of the momentum  $p_x$  which dictates, in the ZLL of  $A$ , the wavefunction center in  $\bar{x}$ . Hence we should, roughly speaking, identify the (local) eigenvalue of  $p_x - a_x$  with  $\bar{x}$ . In the original basis, then, the variable canonically conjugate to  $x$  is identified in the ZLL with  $\bar{x} + a_x$ . If we are to interpret the commutator of the projected  $x$  and  $\bar{x}$  operators in phase space as being the canonical commutation relation of  $x$  and  $p$  in real space, then we need to shift  $\bar{x}$  by  $-a_x$  in order to do so. By a similar argument in the gauge where  $a_x = 0$ , we should shift  $x$  by  $a_{\bar{x}}$  to identify  $x$  with  $p_{\bar{x}}$  in the ZLL.

Having derived the real space Hamiltonian from an ansatz for the solution to the phase space one, we now comment on a few details.

First notice that this derivation generalizes easily to higher dimensions, as the background field only couples  $x$  to  $\bar{x}$ ,  $y$  to  $\bar{y}$ , etc. The primary difference is that in  $2n$ -dimensional phase space, the  $\Gamma$  matrices must be anticommuting elements of the Clifford algebra of  $2n$  by  $2n$  matrices with  $\text{diag}(\Gamma_i) = 0$  for  $i \neq 0$ .

We next comment on gauge invariance. It may appear that there is extra gauge invariance in the phase

space theory; in particular, it may seem strange that the Berry connections  $a_{\bar{x}}$  can be gauge transformed into real space gauge fields  $a_x$  and vice-versa. We claim that this is simply a reflection of the usual gauge invariance in the lower-dimensional Hamiltonian. To see this, consider a unitary operator  $U = \exp(if(x, \bar{x}))$  which implements the gauge transformation  $a_\mu \rightarrow a_\mu + \partial_\mu f$ , and let the ZLL wavefunctions be  $|n\rangle$  for some set of labels  $n$ . Since  $U$  is a gauge transformation in the phase-space system, it must commute with the projection operator  $P$  (as  $U$  must take states in a given LL to the same LL). Hence we can project  $U$  to get its action on the projected Hamiltonian; by the same argument we used for projecting the Hamiltonian, we must have  $PUP = \exp(if(x, k))$ . To understand the meaning of this operator, recall that locally, any state can be labeled as a Bloch wavefunction  $|k; x\rangle$  at momentum  $k$  for a local Hamiltonian at  $x$ . Therefore, a gauge transformation in the higher-dimensional system is equivalent to a spatially dependent  $U(1)$  gauge transformation on the eigenstates  $|k; x\rangle$  of the local Hamiltonian parametrized by  $x$ .

Finally, after seeing the derivation, we may answer the following question: when can a CS theory in  $2nD$  be interpreted as the phase space response theory of a system in  $nD$ ? The key physical requirement in our derivation was that the total field in the  $2nD$  system could be separated into two parts: a uniform background field, which sets some length scale, and another portion which varies slowly on that length scale. When this condition holds, the CS theory may be interpreted as a phase space theory for some lower-dimensional system.

## V. ENUMERATION OF BULK RESPONSES

Having shown that the phase space CS theory is the correct unified theory, we now systematically enumerate all the bulk responses of the CS theory for each possible dimensionality of phase space, and interpret them in real space. To avoid cluttering the notation, we set  $B_0 = 1$ .

The 2D responses were discussed in Section III B. There we showed that the real space current density is the rate of change of polarization, while the  $k$ -space current density reflects the expected relation  $dk/dt \sim E$  with  $E$  the electric field. The charge density response is just the band filling corrected for strain-induced changes in the lattice constant.

We summarize the 2D responses in Table II.

### A. 4D Phase Space

The action is given by

$$S = \frac{1}{24\pi^2} \int dt d^4x \text{tr} [C(x, \bar{x}) \epsilon A \partial A \partial A + \dots] \quad (31)$$



Current component	Response
Real space	Change in polarization
$k$ -space	Electric force
Charge density	Band filling

TABLE II. Summary of 2D phase space responses.

where  $+...$  indicates the non-Abelian terms. We set  $C = 1$  uniformly to look at bulk responses.

#### Spatial components:

A priori there is no difference between the spatial directions  $x$  and  $y$ , so we focus on the  $x$  responses.

$$j_{2D}^x = \frac{1}{4\pi^2} \int d^2\bar{x} \text{tr} [\mathcal{F}_{y\bar{y}} \mathcal{F}_{t\bar{x}} + \mathcal{F}_{\bar{x}\bar{y}} \mathcal{F}_{t\bar{y}} + \mathcal{F}_{y\bar{x}} \mathcal{F}_{\bar{y}t}] \quad (32)$$

The first term includes the background field; setting  $\mathcal{F}_{y\bar{y}} = F_{y\bar{y}} = B_0$  turns this into  $j_{2D}^x = \frac{1}{2\pi} \int d\bar{x} \text{tr} [\mathcal{F}_{t\bar{x}}] = \partial_t P_x$  with  $P_x$  the polarization in the  $x$  direction. This is the same response that appears in 1D, and is illustrated in Fig. 3(a). The second term is the anomalous Hall response; in the simple case where  $\mathcal{F}_{t\bar{y}}$  is simply an electric field, this term gives a current  $j^x = E_y \int d^2\bar{x} \text{tr} [\mathcal{F}_{\bar{x}\bar{y}}] / 4\pi^2 = E_y C_1 / 2\pi$  with  $C_1$  the first Chern number of the occupied bands. This formula also applies to systems with open boundary in the  $\bar{x}, \bar{y}$  directions, in which case  $C_1$  is not quantized but still determines the intrinsic Hall conductivity of the two-dimensional Fermi liquid[41–45].

The third term, illustrated in Fig. 3(b), says the following. Suppose that there is a change in time of the polarization in the  $y$ -direction, i.e.  $\mathcal{F}_{\bar{y}t} \neq 0$ , without any strain in the system. If we now add shear in the system, i.e. have  $\mathcal{F}_{y\bar{x}} \neq 0$ , then some of that polarization change becomes a current along the  $x$  direction as defined before adding strain.

#### $k$ -space components:

The  $\bar{x}$ -direction responses are:

$$j_{4D}^{\bar{x}} = \frac{1}{4\pi^2} \text{tr} [\mathcal{F}_{t\bar{x}} \mathcal{F}_{y\bar{y}} + \mathcal{F}_{t\bar{y}} \mathcal{F}_{y\bar{x}} + \mathcal{F}_{t\bar{y}} \mathcal{F}_{x\bar{y}}] \quad (33)$$

The first term is quasi-1D, saying that  $dk_x/dt$  is proportional to the electric field  $E_x$ . The second term says that an electric field  $E_y$  leads to a change in  $k_x$  if there is shear in the system. The third term is, semiclassically, the Lorentz force - changing the polarization in the  $y$  direction ( $\mathcal{F}_{t\bar{y}} \neq 0$ ) leads to a 1D current in the  $y$  direction, which then feels the Lorentz force of the magnetic field ( $\mathcal{F}_{x\bar{y}} \neq 0$ ), causing  $k_x$  to change.

#### Charge component:

The charge responses are

$$j_{2D}^t = \frac{1}{4\pi^2} \int d^2\bar{x} \text{tr} [\mathcal{F}_{x\bar{y}} \mathcal{F}_{y\bar{x}} + \mathcal{F}_{x\bar{x}} \mathcal{F}_{y\bar{y}} - \mathcal{F}_{x\bar{y}} \mathcal{F}_{y\bar{x}}] \quad (34)$$

The first term is the Hall response for a Chern insulator. Specifically, if  $F_{xy}$  is just the magnetic field,

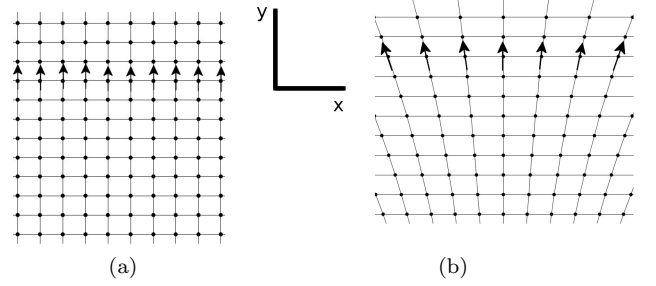


FIG. 3. Cartoon of the polarization response of an (a) unstrained square lattice (b) sheared square lattice. In (b),  $F_{y\bar{x}} \neq 0$  because motion along the  $y$  lattice vector leads to translation in the unstrained  $x$  direction. The directed lines are the flow of charge due to a positive  $\partial_t P_y$ ; the polarization is measured along the lattice vector, which is deformed in (b) due to strain. Only (b) has a nonzero current in the  $x$  direction in (b), which is due to this deformation.

this term becomes  $j^t = (B_z/4\pi^2) \int d^2\bar{x} F_{y\bar{x}} = C_1 B_z / 2\pi$ , where  $C_1$  is the first Chern number.

Consider the remaining terms

$$j_{2D}^t = \frac{1}{4\pi^2} \int d^2\bar{x} \text{tr} [\mathcal{F}_{x\bar{x}} \mathcal{F}_{y\bar{y}} - \mathcal{F}_{x\bar{y}} \mathcal{F}_{y\bar{x}}] \quad (35)$$

In the simplest case of a single featureless, flat band,  $f_{i\bar{j}} \propto \partial_i u_j$  where  $\mathbf{u}$  is the displacement vector. This is because infinitesimal motion  $dx_i$  in the  $i$  direction leads to a translation  $dx_j = \partial_i u_j dx_i$  in the  $j$  direction, which is, for Bloch wavefunctions, the same as accumulating a  $\bar{j}$ -dependent phase  $B_0 \bar{j} dx_j$ . Hence  $a_i = \bar{j} \partial_i u_j$  with our convention of  $B_0 = 1$ .

In this simple case, then, Eq. (35) becomes

$$j_{2D}^t = \frac{1}{4\pi^2} \int d^2\bar{x} ((1 + \partial_x u_x)(1 + \partial_y u_y) - \partial_x u_y \partial_y u_x) \quad (36)$$

The expression inside the parentheses is the determinant of the deformation gradient, that is, the area of the strained unit cell in units of the original unit cell area. Hence the non-background terms are just due to the change in the area of the unit cell. Adding features to the bands will lead to corrections due to, for example, strain changing the local density of states.

We summarize the 4D responses in Table III.

Current density component	Response
Real space	Change in polarization Hall response Change in polarization with strain
$k$ -space	Electric force Sheared response to electric field Lorentz force
Charge density	Hall response Change in unit cell area due to strain

TABLE III. Summary of responses from 4D phase space.

### B. 6D Phase Space

The action is given by

$$S = \frac{1}{192\pi^3} \int dt d^6x \text{tr} [C(x, \bar{x}) \epsilon \mathcal{A} \partial \mathcal{A} \partial \mathcal{A} \partial \mathcal{A} + \dots] \quad (37)$$

with  $+\dots$  again representing the terms for a non-Abelian gauge field. For simplicity of exposition and interpretation, we will assume that the momentum-space and time components of  $\mathcal{A}$  are  $U(N)$  and the real-space components are  $U(1)$ , that is, the latter couple to all the bands in the same way. This assumption is not necessary for our theory, however. We again assume that  $C = 1$  uniformly to look at bulk responses.

There are 15 different responses in each component. If we separate  $\mathcal{F}_{x\bar{x}}$  into its background and non-background components, for the spatial and momentum components we get an extra 7 terms for a total of 22. For the charge component, there are 28. We sort them, neglecting relative minus signs between the groups.

#### Spatial components:

Quasi-1D response (5 terms):

$$j_{3D}^x = \frac{1}{8\pi^3} \int d^3\bar{x} \text{tr} [\mathcal{F}_{t\bar{x}} ((F_{y\bar{y}} + f_{y\bar{y}})(F_{y\bar{y}} + f_{y\bar{y}}) - \mathcal{F}_{y\bar{z}} \mathcal{F}_{z\bar{y}})] \quad (38)$$

By the same computation that was done for the charge response in 4D,  $\mathcal{F}_{y\bar{y}} \mathcal{F}_{z\bar{z}} - \mathcal{F}_{y\bar{z}} \mathcal{F}_{z\bar{y}}$  is the change in area perpendicular to the current. This response thus has the form of the 1D real space response (time-varying polarization) times the change in area perpendicular to the current.

Layered Chern insulator response (2 terms):

$$j_{3D}^x = \frac{1}{8\pi^3} \int d^3\bar{x} \text{tr} [\mathcal{F}_{t\bar{y}} \mathcal{F}_{\bar{x}\bar{y}} F_{z\bar{z}} + (y \leftrightarrow z)] \quad (39)$$

where  $(y \leftrightarrow z)$  means to switch  $y$  and  $z$  as well as  $\bar{y}$  and  $\bar{z}$ . This is the Hall response corresponding to thinking of the 3D system as 2D systems layered in momentum space. Note that this includes the Hall response of a Weyl semimetal (WSM) [34, 52, 53, 55, 69] appearing from its monopoles of  $\mathcal{F}_{\bar{x}\bar{y}}$ . This can be seen by thinking of the (2-node) WSM as stacks of 2D insulators parametrized by

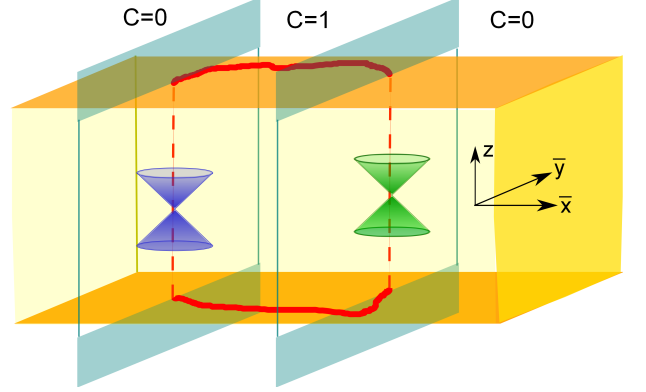


FIG. 4. Slab of WSM with Weyl nodes separated along  $\bar{x}$ . Each slice in momentum space with fixed  $\bar{x}$  can be characterized by a Chern number  $C_1(\bar{x})$ , which changes by unity across the Weyl nodes. Thus, the region between the nodes in the above figure is a series of Chern insulators. The edge states of these Chern insulators constitute the Fermi arcs, marked as thick red lines with an irregular shape. Note that the cones are only present as a cartoon to depict the position of the Weyl nodes; the vertical direction in the figure is  $z$ , and should not be confused with energy. Figure adapted from Ref. [34].

the momentum direction along which the Weyl nodes are split; as shown in Fig. 4, each insulator lying between the nodes is a Chern insulator and thus contributes to  $\mathcal{F}_{\bar{x}\bar{y}}$  (for  $k_z$ -direction Weyl node splitting). In this special case, integration yields

$$j_{3D}^x = \frac{1}{4\pi^2} (E_y \Delta k_z - E_z \Delta k_y) \quad (40)$$

where  $\Delta k_i$  is the splitting of the Weyl points in the  $k_i$  direction.

Topological magnetoelectric effect (3 terms):

$$j_{3D}^x = \frac{1}{8\pi^3} \int d^3\bar{x} \text{tr} [(\mathcal{F}_{\bar{x}t} \mathcal{F}_{\bar{y}z} - \mathcal{F}_{\bar{x}\bar{y}} \mathcal{F}_{t\bar{z}} + \mathcal{F}_{\bar{x}\bar{z}} \mathcal{F}_{t\bar{y}}) \mathcal{F}_{yz}] \quad (41)$$

Assuming that  $\mathcal{F}_{yz}$  does not depend on momentum for simplicity, this term is an  $x$ -direction current proportional to  $B_x$ . Indeed, if we assume that the real space system is gapped so that there are no monopoles of Berry

curvature, simple but tedious manipulations (see Appendix A) turn Eq. (41) into

$$j_{3D}^x = -\frac{1}{16\pi^3} \partial_t \int d^3 \bar{x} \epsilon^{IJK} \text{tr} \left[ a_I \partial_J a_K + \frac{2}{3} a_I a_J a_K \right] B_x$$

$$\equiv -\frac{1}{2\pi} (\partial_t P_3) B_x \quad (42)$$

where  $I, J, K$  run over  $\bar{x}, \bar{y}, \bar{z}$  and  $P_3$  is the 3-dimensional analog of charge polarization. Eq. (42) is precisely the contribution of topological magnetoelectric effect to  $j^x$ . [22]

Topological insulator (TI)-type anomalous Hall response (6 terms):

$$j_{3D}^x = \frac{1}{8\pi^3} \int d^3 \bar{x} \text{tr} [\mathcal{F}_{tz} (\mathcal{F}_{\bar{x}\bar{y}} \mathcal{F}_{y\bar{z}} - \mathcal{F}_{\bar{y}\bar{z}} \mathcal{F}_{\bar{x}y})$$

$$+ \mathcal{F}_{ty} \mathcal{F}_{\bar{x}\bar{y}} f_{z\bar{z}} + (y \leftrightarrow z)] \quad (43)$$

If we choose a gauge such that the real space Berry connections do not depend on momentum, then by similar logic to the topological magnetoelectric effect terms (and with similar assumptions), we can manipulate this contribution into the form

$$j_{3D}^x = -\frac{1}{2\pi} (E_y \partial_y - E_z \partial_z) P_3 \quad (44)$$

Here  $E$  is the real space electric field. This is the anomalous Hall effect that appears in a 3D TI. [22] It differs from the Hall effect that appears as a quasi-2D response in that it does not originate from having a nonzero total Chern number at each 2D slice of momentum space.

Sheared polarization responses (6 terms):

$$j_{3D}^x = \frac{1}{8\pi^3} \int d^3 \bar{x} \text{tr} [\mathcal{F}_{t\bar{y}} (\mathcal{F}_{\bar{x}z} \mathcal{F}_{z\bar{y}} + \mathcal{F}_{\bar{x}y} (F_{z\bar{z}} + f_{z\bar{z}}))$$

$$+ (y \leftrightarrow z)] \quad (45)$$

The first term here corresponds to a current flowing in  $y$  due to a changing polarization, but that current being redirected into the  $z$  and then the  $x$  direction by shear. The second term is the same current in  $y$  being redirected into the  $x$  direction together with a change in the perpendicular area due to uniaxial strain. These are the 3D real space analogues of the 2D real space response illustrated in Fig. 3(b).

#### **k-space components:**

Quasi-1D response (5 terms):

$$j^{\bar{x}} = \frac{1}{8\pi^3} \text{tr} [\mathcal{F}_{tx} ((F_{y\bar{y}} + f_{y\bar{y}})(F_{z\bar{z}} + f_{z\bar{z}}) - \mathcal{F}_{y\bar{z}} \mathcal{F}_{z\bar{y}})] \quad (46)$$

As in the real space response, this is the 1D response accounting for changes in the area perpendicular to the current.

Strained electric forces (6 terms):

$$j^{\bar{x}} = \frac{1}{8\pi^3} \text{tr} [\mathcal{F}_{ty} (\mathcal{F}_{x\bar{z}} \mathcal{F}_{\bar{y}z} + \mathcal{F}_{x\bar{y}} (F_{z\bar{z}} + f_{z\bar{z}})) + (y \leftrightarrow z)] \quad (47)$$

These terms correspond to a typical electrical force in the  $y$  direction, which is then redirected to the  $x$  direction by shears and correcting for change in the area perpendicular to the current.

Strained polarization/Lorentz force responses (8 terms):

$$j^{\bar{x}} = \frac{1}{8\pi^3} \text{tr} [\mathcal{F}_{t\bar{y}} (\mathcal{F}_{zx} \mathcal{F}_{y\bar{z}} + \mathcal{F}_{x\bar{z}} \mathcal{F}_{yz} + \mathcal{F}_{xy} (F_{z\bar{z}} + f_{z\bar{z}}))$$

$$+ (y \leftrightarrow z)] \quad (48)$$

The first two terms say that if the polarization changes in the  $y$  direction, then either shear or the Lorentz force can change this into a current in the  $z$  direction. That current can then be redirected by the Lorentz force or shear (respectively) to the  $x$  direction. The third term is the same current due to polarization, leading to a current in the  $x$  direction by the Lorentz force, correcting for change in the area perpendicular to the new current.

WSM-type  $\mathbf{E} \cdot \mathbf{B}$  charge pumping (3 terms):

$$j^{\bar{x}} = \frac{1}{8\pi^3} \text{tr} [(\mathcal{F}_{tx} \mathcal{F}_{yz} + \mathcal{F}_{ty} \mathcal{F}_{zx} + \mathcal{F}_{tz} \mathcal{F}_{xy}) \mathcal{F}_{\bar{y}\bar{z}}] \quad (49)$$

Assuming the real-space field strengths are  $k$ -independent, then this term is  $\mathbf{E} \cdot \mathbf{B}$  times the Berry curvature. If we integrate over  $\bar{y}$  and  $\bar{z}$ , we find

$$\int d\bar{y} d\bar{z} j^{\bar{x}} = \frac{1}{4\pi^2} C_1(\bar{x}) \mathbf{E} \cdot \mathbf{B} \quad (50)$$

where  $C_1(\bar{x})$  is the Chern number of the slice of the Brillouin zone at fixed  $\bar{x}$ . For the case of a WSM with Weyl points split along the  $\bar{x}$  direction,  $C_1(\bar{x})$  is nonzero between the Weyl points (see Fig. 4), so this response is a current from one Weyl point to the other. This is exactly the chiral anomaly [34, 51–53, 55, 69, 70], which says that an  $\mathbf{E} \cdot \mathbf{B}$  field in a WSM pumps charge from one Weyl point to the other.

#### **Charge component:**

All 16 terms which only contain mixed field strengths  $\mathcal{F}_{i\bar{j}}$  combine to form the unit cell volume, corrected for strain. The other two types of response are:

Layered Chern insulator Hall response (3 terms):

$$j^0 = \frac{1}{8\pi^3} \int d^3 \bar{x} \text{tr} [\mathcal{F}_{xz} \mathcal{F}_{\bar{z}\bar{x}} F_{y\bar{y}} + \text{perm.}] \quad (51)$$

where “perm.” indicates terms created by cyclically permuting  $(x, y, z)$  and  $(\bar{x}, \bar{y}, \bar{z})$ . This is the charge density counterpart to the spatial layered Chern insulator Hall response; it analogously comes from adding the Hall response of each subsystem at fixed  $k_i$  to a magnetic field  $B_i$ .

TI-type Hall response (9 terms):

$$j^0 = \frac{1}{8\pi^3} \int d^3\bar{x} \text{tr} [\mathcal{F}_{xz} (\mathcal{F}_{\bar{y}\bar{z}} \mathcal{F}_{y\bar{x}} + \mathcal{F}_{\bar{x}\bar{y}} \mathcal{F}_{y\bar{z}} + \mathcal{F}_{\bar{z}\bar{x}} \mathcal{F}_{y\bar{y}}) + \text{perm.}] \quad (52)$$

By the same methods used to derive Eq. (42), these terms can be massaged into the form

$$j^0 = \frac{1}{2\pi} \mathbf{B} \cdot \nabla P_3 \quad (53)$$

This is exactly the charge component of the Hall response that appears in a TI[22].

We summarize the 6D responses in Table IV.

## VI. ANOMALIES

In the previous section, we enumerated the phase space bulk responses, which, as we have seen, correspond to the topological responses of filled states in real space. This includes the responses of insulators, semimetals and the responses of metals that involve all the occupied states, such as the anomalous Hall effect. We now wish to describe the topological features of Fermi surfaces and real space system edges. We will also approach the response of semimetals from another perspective. All these features take the form of anomalies in phase space.

Given the CS theory for a phase space system, such as that which appears in Eq. (9) or its higher-dimensional and/or non-Abelian generalization, suppose that  $\partial_i C \neq 0$  for some coordinate  $i$ . This means that the phase space system contains an edge, that is, the real space system has a Fermi surface or an edge. Then in general, the responses of the system will depend on details; for example, edge currents of quantum Hall systems depend on the non-universal edge mode velocity. However, there will be a universal anomaly (or lack thereof) along such edges. We see this from the anomaly resulting from the phase-space CS term in  $2nD$ :

$$\sum_{\mu \neq i} \partial_\mu j^\mu = \frac{1}{n! 2^{2n} \pi^n} \partial_i C \epsilon^{i\alpha_1 \alpha_2 \dots \alpha_{2n}} \text{tr} [\mathcal{F}_{\alpha_1 \alpha_2} \dots \mathcal{F}_{\alpha_{2n-1} \alpha_{2n}}] \quad (54)$$

From here, integration over the appropriate phase space directions determines the anomalies in the real system. We list some physically interesting effects below.

### A. Fermi Surface Anomalies

In Section IIIB, we computed the chiral anomaly in 1D momentum space by imposing edges at  $\bar{x} = \pm k_F$  in 2D phase space. Integrating over  $x$  yielded the fact that an electric field pumps electrons in states near  $+k_F$  into states near  $-k_F$ , or vice versa. At a down-to-earth

level, this simply corresponds to tilting of the 1D Fermi surface in an electric field due to semiclassical motion of the electrons.

This idea is straightforward to generalize to higher dimensions. For instance, an open Fermi surface in 2D is obtained by confining the 4D phase space system between  $\bar{x} = \pm k_F$  while leaving the other three directions infinite, while a 3D spherical Fermi surface results from making the  $\bar{x}$ ,  $\bar{y}$  and  $\bar{z}$  directions in 6D phase space finite under the constraint  $\bar{x}^2 + \bar{y}^2 + \bar{z}^2 = k_F^2$  and leaving the  $x$ ,  $y$  and  $z$  directions unconstrained. For each phase space geometry, the corresponding anomaly characterizes properties of the resultant Fermi surface. Importantly, if a special object such as a Dirac or a Weyl point is buried under the Fermi surface, its observable effects in local transport phenomenon should emerge from the anomaly equation.

We demonstrate this first for a 3D spherical Fermi surface, which carries a Chern number in general and exhibits a chiral anomaly proportional to the Chern number and the electromagnetic field  $\mathbf{E} \cdot \mathbf{B}$ . The most well-known occurrence of this phenomenon is in Weyl semimetals. We begin with the anomaly equation in 6D phase space. In the absence of any strains and ignoring quasi-lower-dimensional terms (i.e., terms such as  $F_{x\bar{x}}$  that contain the background field), it reads

$$\sum_{\mu} \partial_\mu j^\mu - \partial_{\bar{r}} j^{\bar{r}} = \frac{1}{8\pi^3} \delta(\bar{r} - k_F) \mathcal{F}_{\bar{\theta}\bar{\phi}} (\mathcal{F}_{t\bar{x}} \mathcal{F}_{y\bar{z}} + \mathcal{F}_{t\bar{y}} \mathcal{F}_{z\bar{x}} + \mathcal{F}_{t\bar{z}} \mathcal{F}_{x\bar{y}}) \quad (55)$$

where  $(\bar{r}, \bar{\theta}, \bar{\phi})$  are the spherical co-ordinates corresponding to  $(\bar{x}, \bar{y}, \bar{z})$ . Integrating over the barred co-ordinates immediately yields the chiral anomaly in Weyl semimetals:

$$\sum_{\mu=t,x,y,z} \partial_\mu j_{3D}^\mu = \frac{1}{4\pi^2} C_{FS} \mathbf{E} \cdot \mathbf{B} \quad (56)$$

where  $C_{FS} = \frac{1}{2\pi} \oint d\bar{\theta} d\bar{\phi} \sin \bar{\theta} \mathcal{F}_{\bar{\theta}\bar{\phi}} \in \mathbb{Z}$  is the Chern number of the Fermi surface and equals the total chirality of all Weyl points enclosed by it.

Unlike in 3D, in 2D systems the one-dimensional Fermi surface carries a non-quantized Berry phase instead of a Chern number. An analogous analysis, i.e., starting with the anomaly equation in 4D phase space with a  $(\bar{x}, \bar{y})$  boundary that satisfies  $\bar{x}^2 + \bar{y}^2 = k_F^2$  and integrating over  $(\bar{x}, \bar{y})$  gives

$$\sum_{\mu=t,x,y} \partial_\mu j_{2D}^\mu = \frac{1}{4\pi^2} B_z \partial_t \gamma \quad (57)$$

ignoring strain and quasi-lower-dimensional terms, where  $\gamma = \oint_{FS} \mathbf{a}_{\mathbf{F}} \cdot d\bar{\mathbf{r}}$  is the Berry phase on the Fermi surface. Eq (57) is the statement that adiabatically changing the Hall conductivity of an anomalous Hall metal in a magnetic field creates charged excitations bound to the field.

Current component	Response
Real space	Quasi-1D responses Quasi-2D (layered Chern insulator) response Topological magnetoelectric effect TI-like anomalous Hall response Change in polarization with strain
$k$ -space	Quasi-1D response Electric fields with strain Change in polarization plus Lorentz force with strain WSM-like $\mathbf{E} \cdot \mathbf{B}$ charge pumping
Charge density	Density response to change in unit cell volume Layered Chern insulator Hall response TI-like anomalous Hall response

TABLE IV. Summary of 6D phase space responses.

In the presence of strains, both (56) and (57) contain more terms on their right hand sides. We will encounter these terms in the next subsection when we discuss the effects of dislocations. Before moving on, however, we wish to stress that the physical anomaly in a given dimension is independent of the topology of the Fermi surface. However, certain topologies are more convenient for studying a given physical anomaly. For instance, the chiral anomaly in Weyl metals is easier to see for a spherical Fermi surface, but it can be equally well be derived for open Fermi surfaces that span one or two directions in the Brillouin zone.

## B. Anomalies in Real Space

### 1. Real space edge

The simplest example of a real space edge anomaly was derived in Sec III B, where we imposed  $x$ -direction edges in 2D phase space and showed that a time-dependent polarization in 1D can be used to pump charge across the length of the chain. As a more non-trivial example, consider a real, single-band 2D Chern insulator which occupies  $x > 0$ . Then in 4D phase space, the anomaly equation (54) with  $C(x) = \Theta(x)$  reads

$$\partial_t \rho + \partial_y j^y + \partial_{\bar{x}} j^{\bar{x}} + \partial_{\bar{y}} j^{\bar{y}} = \frac{1}{4\pi^2} \delta(x) (\mathcal{F}_{ty} \mathcal{F}_{\bar{x}\bar{y}} - \mathcal{F}_{t\bar{x}} \mathcal{F}_{y\bar{y}} + \mathcal{F}_{t\bar{y}} \mathcal{F}_{y\bar{x}}) \quad (58)$$

Integrating  $\partial_{\bar{x}} j^{\bar{x}} + \partial_{\bar{y}} j^{\bar{y}}$  over momentum space gives zero since there is no boundary in those directions. Hence, integrating the previous equation over momentum space gives

$$\partial_t \rho_{2D} + \partial_y j_{2D}^y = \frac{1}{4\pi^2} \delta(x) \int d^2 k \mathcal{F}_{\bar{x}\bar{y}} E_y = \frac{1}{2\pi} \delta(x) C_1 E_y \quad (59)$$

with  $C_1$  the first Chern number of the occupied band of the 2D Hamiltonian. We have ignored quasi-1D terms and terms containing strain. Eq. (59) is recognizable as the usual anomaly for a 2D Chern insulator where an electric field parallel to the edge builds up a charge density along that edge.

### 2. Dislocations

**4D phase space:** The simplest example of a dislocation is an edge dislocation in 2D real space. The key feature of the dislocation, as we discussed in Section III A, is that, far from the dislocation line itself, electrons accumulate a Berry's phase of  $\mathbf{k} \cdot \mathbf{b}$  upon encircling the dislocation. We can thus model the dislocation by a Berry connection  $(a_r, a_\theta) = (0, \frac{\mathbf{b} \cdot \mathbf{k}}{2\pi})$ , leading to a  $\mathbf{k}$ -independent Berry curvature  $\mathcal{F}_{\theta\bar{i}} = b_i/2\pi r$ . Our theory breaks down at the dislocation itself because the system changes quickly on the scale of a lattice constant. We can avoid this problem by keeping the Berry connection but surrounding the dislocation by a finite size puncture in the system of radius  $r_0$ , i.e. choose  $C = \Theta(r - r_0)$  with  $r$  the radial coordinate in the  $xy$ -plane. The resulting anomaly equation reads

$$\sum_\mu \partial_\mu j^\mu - \partial_r j^r = \frac{1}{8\pi^3} (\mathcal{F}_{t\bar{x}} b_y - \mathcal{F}_{t\bar{y}} b_x) \quad (60)$$

plus quasi-1D terms on the right-hand-side, which we ignore. Integrating over  $\bar{x}, \bar{y}$  and  $\theta$  gives the charge radiating from the core of an edge dislocation in the presence of a time-dependent polarization:

$$\partial_t \rho_{2D} = \hat{\mathbf{z}} \cdot (\mathbf{b} \times \partial_t \mathbf{P}) \quad (61)$$

This is shown in Fig 5, which makes the physical picture of the anomaly clear in the limit of weakly coupled chains perpendicular to  $\mathbf{b}$ ; the core of the dislocation is the end



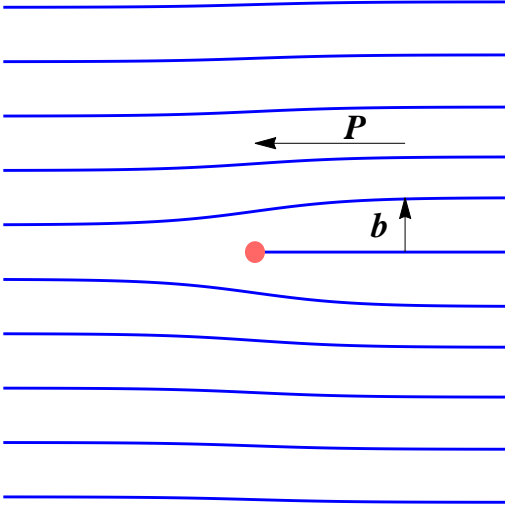


FIG. 5. An edge dislocation in 2D with Burger's vector  $\mathbf{b}$ . In the presence of a polarization  $\mathbf{P} \perp \mathbf{b}$ , charge gets accumulated at the core of the dislocation, shown by the red dot.

of such a chain, so polarizing that chain adds charge to its end. The non-trivial result is that the extra charge remains bound to the dislocation core and does not leak into other chains even when they are strongly coupled.

**6D phase space:** A similar analysis for a dislocation in 3D real space running along  $\hat{\mathbf{z}}$  and with Burger's vector  $\mathbf{b}$  gives

$$\begin{aligned} \partial_t \rho + \frac{1}{r} \partial_\theta j^\theta + \partial_z j^z \\ = \frac{\delta(r - r_0)}{8\pi^3} \int d^3 \bar{x} (\mathcal{F}_{\bar{x}\bar{y}}(r\mathcal{F}_{\theta\bar{z}}) + \mathcal{F}_{\bar{y}\bar{z}}(r\mathcal{F}_{\theta\bar{x}}) + \mathcal{F}_{\bar{z}\bar{x}}(r\mathcal{F}_{\theta\bar{y}})) \mathcal{F}_{t\bar{z}} \\ = -\frac{\delta(r - r_0)}{8\pi^3} E_z \int d^3 \bar{x} \Omega \cdot \mathbf{b} \end{aligned} \quad (62)$$

where  $\Omega_i = \frac{1}{2} \epsilon_{ijk} \mathcal{F}_{j\bar{k}}$  is the Berry curvature of the bands in the plane perpendicular to  $i$ .

To understand (62), let us first consider a layered Chern insulator, that is, a system composed of layers of Chern insulators stacked along a certain direction. The integral in (62) then gives the Chern number of the layers in each direction, so

$$\partial_t \rho + \frac{1}{r} \partial_\theta j^\theta + \partial_z j^z = -\frac{\delta(r - r_0)}{\pi} E_z \mathbf{C} \cdot \mathbf{b} \quad (63)$$

where  $C_i = \frac{1}{8\pi^2} \int d^3 \bar{x} \Omega_i$ . Now, add a dislocation running along  $\hat{\mathbf{z}}$  with Burger's vector  $\mathbf{b}$ . Edge dislocations are defined by  $\mathbf{b} \perp \hat{\mathbf{z}}$  whereas screw dislocations have  $\mathbf{b} \parallel \hat{\mathbf{z}}$ . The two scenarios are shown in Fig 6. Eq (63) says that in either case, there exists a chiral mode along the dislocation that participates in an anomaly in response to  $E_z$ . We can understand this as follows.

An edge dislocation can be thought of as a semi-infinite sheet perpendicular to  $\mathbf{b}$  and unbounded along  $\hat{\mathbf{z}}$  inserted

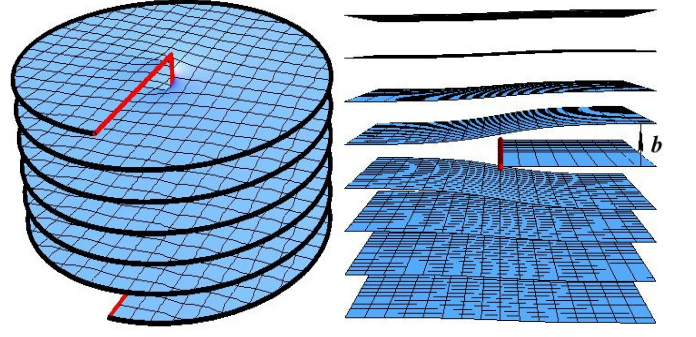


FIG. 6. Screw (left) and edge (right) dislocations in 3D. Dislocations in an axion insulator harbor chiral modes, denoted by red lines in both figures. In the screw dislocation, thick black lines represent the standard chiral edge mode. The screw dislocation geometry with Weyl nodes split along the screw axis was used for the numerical results presented in Fig 7.

into the 3D lattice. If the sheet has a Chern number, we expect it to have a chiral mode along  $\hat{\mathbf{z}}$ . For weakly coupled sheets, this is precisely the chiral mode along the edge dislocation. For a screw dislocation, the existence of a chiral dislocation mode follows from an argument adapted from one that predicts helical dislocation modes in weak topological insulators[71, 72]. Suppose that our system is of finite size in the  $z$  direction. Then on each surface, there is a semi-infinite edge emerging from the dislocation. However, this edge must carry a chiral mode since the surface layer is a Chern insulator. By charge conservation, this chiral mode cannot terminate at the dislocation, so the chiral mode must proceed along the dislocation to the other surface. Moreover, in each case, the chiral mode is expected to survive for strongly coupled layers as well, where the system is better thought of as stacked sheets in momentum space and is typically termed an axion insulator. This is because layered Chern insulators and axion insulators are actually the same phase – there is no phase transition as the interlayer coupling is strengthened – so their topological defects such as dislocations have qualitatively similar behavior. Indeed, the presence of a chiral mode was shown explicitly for an axion insulator created from a charge density wave instability of a WSM in Ref. 73.

**Spectral flow due to dislocations in Weyl semimetals:** Having seen examples of anomalies being the universal feature of gapless systems, we use our theory's anomaly machinery to derive a new result: prediction of an anomaly at dislocation lines in a WSM. This is closely related to the case of a layered Chern insulator (axion insulator) just discussed.

Consider a WSM with two nodes split by  $\mathbf{K}$  (and thus having broken time reversal symmetry) with a dislocation along  $z$ . In contrast to the layered Chern insulators or axion insulators, WSMs have a gapless bulk and thus, cannot support localized modes the same way that the

former do. However, our theory allows us to confirm that there is indeed an anomaly at the dislocation in a WSM. The anomaly calculation is identical to the axion insulator case, except that  $\frac{1}{4\pi} \int d^3x \Omega_i = K_i$  instead of is  $2\pi$ . The result is that the anomaly is  $\delta(r-r_0)E_z \mathbf{K} \cdot \mathbf{b}/2\pi$ , which reflects the fact that chiral modes appear only in the region of momentum space between the Weyl nodes, where the Chern number of the layers is  $\pm 1$ . From now on, we assume  $\mathbf{K} = K\hat{\mathbf{z}}$  for concreteness.

The physical interpretation of this anomaly is more subtle for the WSM than the axion insulator. In the latter case, due to the bulk gap, the anomaly means that there is a chiral zero mode on the dislocation. In the WSM case, there is no bulk gap. Furthermore, if the region carrying a nonzero Chern number is near  $k_z = 0$ , then that region sees only small perturbations from the dislocation because the dislocation acts like a flux proportional to  $k_z$ . Hence we should not necessarily expect a zero mode on the dislocation. On the other hand, if this region is located near  $k_z = \pi$ , then there may be such a zero mode. In general, however, the existence of a localized zero mode is not guaranteed.

Since the anomaly need not imply a localized zero mode, we numerically solved a simple  $\mathbf{k} \cdot \mathbf{p}$  model for a WSM in the presence of a dislocation in order to directly verify the anomaly. The Hamiltonian we used is

$$H = [M_0 + M_1 k_z^2 + M_2 (k_x^2 + k_y^2)] \Gamma_5 + L_1 k_z \Gamma_4 + L_2 (k_y \Gamma_1 - k_x \Gamma_2) + U_0 \Gamma_{12} \quad (64)$$

Here the anticommuting  $\Gamma$  matrices are defined by  $\Gamma_{1,2,3} = \sigma_{x,y,z} \tau_x$ ,  $\Gamma_4 = \tau_y$ ,  $\Gamma_5 = \tau_z$ , and  $\Gamma_{ij} = [\Gamma_i, \Gamma_j]/2i$  where  $\sigma$  is a spin index and  $\tau$  is an orbital index. This model leads to Weyl points at  $\mathbf{k} = \pm |U_0|/L_1 \hat{\mathbf{z}}$  when the quadratic term is neglected. This model has been previously investigated in a radial geometry[55] with no dislocation. The only effect of a screw dislocation at  $r = 0$  with Burgers vector  $b\hat{\mathbf{z}}$  is that the dependence of the components of the wavefunction on the in-plane angle  $\theta$  changes from  $e^{in\theta}$  to  $e^{i(l+bk_z/2\pi)\theta}$ , where the half-integer  $l$  is the eigenvalue of  $L_z$  in the absence of the dislocation.

We solved the discretized version of this model at fixed angular momentum for a cylinder of size  $R = 120$  sites at fixed angular momentum  $l = 1/2 + k_z/2\pi$ . For comparison, we show the band structure in the WSM phase with no dislocation in Fig. 7(a). The mode localized near  $r = 0$  (in blue) is always at higher energy than the Fermi arc, and is not topological. Adding the dislocation, we see in Fig. 7(b) that now the  $r = 0$  mode changes from unoccupied to occupied after crossing the Weyl points; an electron has been pumped from the Fermi arc (in red) to the dislocation. This is the anomaly that we discussed above, even though there is no zero energy mode localized on the dislocation. This system can smoothly evolve, by bringing the Weyl points together and annihilating them, into the axion insulator in Fig. 7(c). That state has a sin-

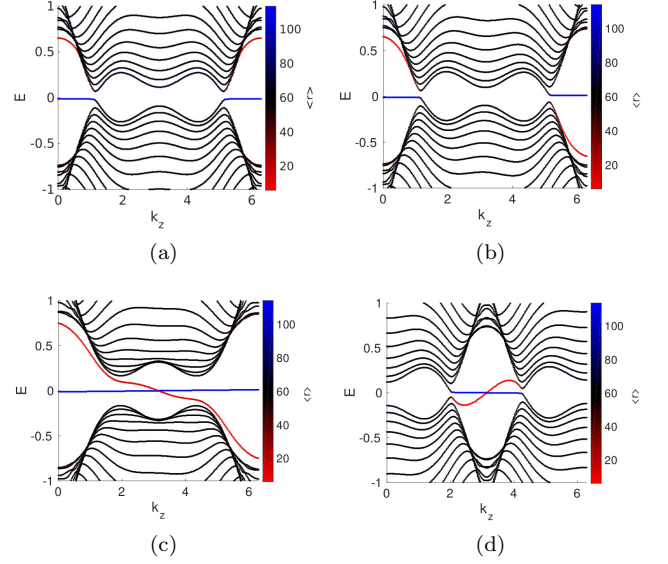


FIG. 7. Band structure of the lattice regularized version of Eqn. (64) in a cylindrical geometry. Color corresponds to  $\langle r \rangle$  with the dislocation at  $r = 0$ ; red indicates localization on the dislocation and blue is localization on the outer boundary. Parameters are  $M_0 = 0$ ,  $M_1 = 0.342 \text{ eV } \text{\AA}^2$ ,  $M_2 = 18.25 \text{ eV } \text{\AA}^2$ ,  $L_1 = 1.33 \text{ eV } \text{\AA}$ ,  $L_2 = 2.82 \text{ eV } \text{\AA}$ ,  $R = 120$  radial sites,  $l = 1/2$  angular momentum unless otherwise stated. Note that due to the dislocation the system is not periodic in  $k_z$  at fixed angular momentum. (a) WSM phase ( $U_0 = 1.3 \text{ eV}$ ), no dislocation. (b) Same as (a), but with dislocation. (c) Axion insulator phase ( $U_0 = 1.7 \text{ eV}$ ) with dislocation. (d) WSM phase ( $U_0 = -1.3 \text{ eV}$ ,  $M_1 = -0.342 \text{ eV } \text{\AA}$ ,  $M_0 = 1.4 \text{ eV}$ ) with topologically nontrivial BZ slices centered at  $k_z = \pi$  and a dislocation.

gle chiral mode localized on the dislocation which crosses the bandgap without mixing with the outer edge mode, as expected. In Fig. 7(d), we have a WSM with a topologically nontrivial region centered about  $k_z = \pi$ ; here there is a zero mode localized on the dislocation, and the charge pumping is more obvious than in Fig. 7(b).

The result of charge pumping due to disclinations has been previously predicted[63]. However, our picture is different from than the one considered there. The claim in Ref. [63] is that a chiral magnetic field, which in our case is created by the dislocation, causes a net spontaneous current to flow. As can be seen from our picture, this is not true; an electric field is necessary to have an anomaly and thus a net current. Fundamentally, the total current must vanish in the absence of an electric field. If the current did not vanish, adding an electric field parallel to the current would cause dissipation and lower the system energy, but this is impossible for a system already in its ground state. The difference in Ref. [63] stems from neglecting momentum space regions away from the Weyl nodes and the real space boundary in determining the total current. Thus, while the general expression for the

current density derived by Ref. [63] is correct, the total current vanishes when these contributions are included. For the case which we show in Fig. 7(b), the net current due to the dislocation is cancelled by the current along the Fermi arcs. In the case of Fig. 7(d), the dislocation mode near one Weyl point connects directly to the mode on the other side through a zero mode which cancels the net current.

To summarize, dislocations in a WSM indeed cause pumping of charge to (or from) the dislocation line when an electric field is applied along the dislocation. Such a charge pumping is smoothly connected to that which occurs in the axion insulator, but it may or may not, depending on details, result in a zero mode localized on the dislocation. Although we presented numerics for a screw dislocation that runs along the same direction as the Weyl node splitting, (62) and hence the qualitative result is valid for edge dislocations as well as for other directions of the Weyl node splitting.

## VII. DISCUSSION AND CONCLUSIONS

We have shown that the responses and anomalies of a gapped or gapless system living in  $n$  spatial dimensions can be described by a single response theory of a gapped system living in  $2n$  spatial dimensions. Conceptually, this is because adding magnetic fields in the  $2n$  dimensional system and projecting onto the zeroth Landau level allows us to interpret that system as living in phase space. We have used this theory to reproduce well-understood responses and anomalies in systems with non-interacting electrons and Abelian real space gauge fields, as well as to demonstrate the existence of spectral flow due to dislocations in Weyl semimetals.

There are several interesting fundamental questions about our theory which are at present open. It would be interesting to see how our theory connects to the use of phase space in statistical mechanics; how might the Landau-Boltzmann transport equation, that describes transport in Fermi liquids via Wigner functions, or Liouville's theorem, that describes the time-evolution of general classical systems in phase space via a density matrix, arise in our context? Both the Wigner function and the phase space density matrix treat real and momentum space on an equal footing; thus, our theory holds promise in capturing these phenomena.

In addition to these fundamental questions, we envision a number of extensions of our theory to more complicated systems. In particular, the responses that we have explicitly discussed have so far been only those of non-interacting systems which only feel a  $U(1)$  real space gauge field, though the  $k$ -space Berry connection has been allowed to be non-Abelian. The latter constraint is not an inherent limitation of the theory; perhaps there are interesting responses to a larger real-space gauge

group.  $SU(2)$  groups in 4D and 3D have been studied and shown to give topological insulator- and WSM-like responses, respectively[74]. It is thus conceivable that general gauge groups can lead to other topological responses, possibly of phases with emergent fermions such as partons[75] or composite fermions[76].

As for interactions, it is not immediately clear if there are sensible real space systems that are well-described by a phase space theory with only local interactions. However, if there are such real space systems, then working in phase space could be very useful because, for example, in the absence of a magnetic field mean field theory is more accurate due to the higher dimensionality. This advantage may be mitigated by the fact that our construction requires gauge fields, however. Alternatively, it is possible that there is a simple way to directly incorporate the interactions of the real space system into the phase space theory.

Another interesting question is if there is an extension of our theory which describes nodal superconductors. Our theory as written requires  $U(1)$  charge conservation; perhaps there is some way to incorporate spontaneous breaking of this symmetry. Finally, it could also be interesting to explicitly incorporate other symmetries of the lower-dimensional system; this could allow a better understanding of gapless symmetry-protected phases like Dirac semimetals.

DB is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-114747. PH is supported by the David and Lucile Packard Foundation and the U.S. DOE, Office of Basic Energy Sciences, contract DEAC02-76SF00515. SCZ is supported by the National Science Foundation under grant No. DMR-1305677. XLQ is supported by the National Science Foundation through the grant No. DMR-1151786.

- 
- [1] X.-L. Qi and S.-C. Zhang, "The quantum spin Hall effect and topological insulators," *Physics Today* **63**, 33–38 (2010).
  - [2] M. Z. Hasan and C. L. Kane, "Colloquium: Topological insulators," *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
  - [3] Xiao-Liang Qi and Shou-Cheng Zhang, "Topological insulators and superconductors," *Rev. Mod. Phys.* **83**, 1057–1110 (2011).
  - [4] M. Z. Hasan and J. E. Moore, "Three-dimensional topological insulators," *Annual Review of Condensed Matter Physics, Annual Review of Condensed Matter Physics* **2**, 55–78 (2011).
  - [5] F. D. M. Haldane, "Model for a quantum Hall effect without Landau levels: Condensed-matter realization of the "parity anomaly"," *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
  - [6] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, "Quantized Hall Conductance in a Two-



- Dimensional Periodic Potential,” *Phys. Rev. Lett.* **49**, 405–408 (1982).
- [7] C. L. Kane and E. J. Mele, “Quantum Spin Hall Effect in Graphene,” *Phys. Rev. Lett.* **95**, 226801 (2005).
- [8] C. L. Kane and E. J. Mele, “ $\mathbb{Z}_2$  Topological Order and the Quantum Spin Hall Effect,” *Phys. Rev. Lett.* **95**, 146802 (2005).
- [9] R. Roy, “Topological phases and the quantum spin Hall effect in three dimensions,” *Phys. Rev. B* **79**, 195322 (2009).
- [10] L. Fu, C. L. Kane, and E. J. Mele, “Topological Insulators in Three Dimensions,” *Phys. Rev. Lett.* **98**, 106803 (2007).
- [11] B. A. Bernevig and S.-C. Zhang, “Quantum spin hall effect,” *Phys. Rev. Lett.* **96**, 106802 (2006).
- [12] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, “Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells,” *Science* **314**, 1757–1761 (2006).
- [13] Rahul Roy, “ $\mathbb{Z}_2$ ,” *Phys. Rev. B* **79**, 195321 (2009).
- [14] N. Read and D. Green, “Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum hall effect,” *Phys. Rev. B* **61**, 10267–10297 (2000).
- [15] C. Kallin and A. J. Berlinsky, “Is  $\text{Sr}_2\text{RuO}_4$  a chiral  $p$ -wave superconductor?” *Journal of Physics: Condensed Matter* **21**, 164210 (2009).
- [16] D. D. Osheroff, W. J. Gully, R. C. Richardson, and D. M. Lee, “New Magnetic Phenomena in Liquid  $\text{He}^3$  below 3 mK,” *Physical Review Letters* **29**, 920–923 (1972).
- [17] A. J. Leggett, “Interpretation of Recent Results on  $\text{He}^3$  below 3 mK: A New Liquid Phase?” *Physical Review Letters* **29**, 1227–1230 (1972).
- [18] A. Kitaev, “Periodic table for topological insulators and superconductors,” L. D. Landau Memorial Conference “Advances in Theoretical Physics” (AIP, 2009) pp. 22–30.
- [19] Shinsei Ryu, Andreas P Schnyder, Akira Furusaki, and Andreas WW Ludwig, “Topological insulators and superconductors: tenfold way and dimensional hierarchy,” *New J. Phys.* **12**, 065010 (2010).
- [20] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, “Classification of topological insulators and superconductors in three spatial dimensions,” *Phys. Rev. B* **78**, 195125 (2008).
- [21] Alexander Altland and Martin R Zirnbauer, “Non-standard symmetry classes in mesoscopic normal-superconducting hybrid structures,” *Physical Review B* **55**, 1142 (1997).
- [22] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, “Topological field theory of time-reversal invariant insulators,” *Phys. Rev. B* **78**, 195424 (2008).
- [23] A. M. Essin, J. E. Moore, and D. Vanderbilt, “Magnetoelectric Polarizability and Axion Electrodynamics in Crystalline Insulators,” *Phys. Rev. Lett.* **102**, 146805 (2009).
- [24] Zhong Wang, Xiao-Liang Qi, and Shou-Cheng Zhang, “Topological field theory and thermal responses of interacting topological superconductors,” *Phys. Rev. B* **84**, 014527 (2011).
- [25] Shinsei Ryu, Joel E. Moore, and Andreas W. W. Ludwig, “Electromagnetic and gravitational responses and anomalies in topological insulators and superconductors,” *Phys. Rev. B* **85**, 045104 (2012).
- [26] Xiao-Liang Qi, Edward Witten, and Shou-Cheng Zhang, “Axion topological field theory of topological superconductors,” *Phys. Rev. B* **87**, 134519 (2013).
- [27] Chong Wang, Andrew C. Potter, and T. Senthil, “Classification of interacting electronic topological insulators in three dimensions,” *Science* **343**, 629–631 (2014).
- [28] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov, “Topological semimetal and fermi-arc surface states in the electronic structure of pyrochlore iridates,” *Phys. Rev. B* **83**, 205101 (2011).
- [29] A. A. Burkov and L. Balents, “Weyl semimetal in a topological insulator multilayer,” *Phys. Rev. Lett.* **107**, 127205 (2011).
- [30] A. A. Burkov, M. D. Hook, and Leon Balents, “Topological nodal semimetals,” *Phys. Rev. B* **84**, 235126 (2011).
- [31] A. M. Turner and A. Vishwanath, “Beyond Band Insulators: Topology of Semi-metals and Interacting Phases,” ArXiv e-prints (2013), [arXiv:1301.0330 \[cond-mat.str-el\]](#).
- [32] O. Vafek and A. Vishwanath, “Dirac Fermions in Solids - from High Tc cuprates and Graphene to Topological Insulators and Weyl Semimetals,” ArXiv e-prints (2013), [arXiv:1306.2272 \[cond-mat.mes-hall\]](#).
- [33] W. Witczak-Krempa and Y.-B. Kim, “Topological and magnetic phases of interacting electrons in the pyrochlore iridates,” *Phys. Rev. B* **85**, 045124 (2012).
- [34] Pavan Hosur and Xiaoliang Qi, “Recent developments in transport phenomena in weyl semimetals,” *Topological insulators / Isolants topologiques, Comptes Rendus Physique* **14**, 857–870 (2013).
- [35] Pavan Hosur, “Friedel oscillations due to fermi arcs in weyl semimetals,” *Phys. Rev. B* **86**, 195102 (2012).
- [36] A. C. Potter, I. Kimchi, and A. Vishwanath, “Quantum Oscillations from Surface Fermi-Arcs in Weyl and Dirac Semi-Metals,” ArXiv e-prints (2014), [arXiv:1402.6342 \[cond-mat.mes-hall\]](#).
- [37] F. D. M. Haldane, “Attachment of Surface ”Fermi Arcs” to the Bulk Fermi Surface: ”Fermi-Level Plumbing” in Topological Metals,” ArXiv e-prints (2014), [arXiv:1401.0529 \[cond-mat.str-el\]](#).
- [38] T.T. Heikkila, N.B. Kopnin, and G.E. Volovik, “Flat bands in topological media,” *JETP Letters* **94**, 233–239 (2011).
- [39] J. C. Y. Teo and C. L. Kane, “Topological defects and gapless modes in insulators and superconductors,” *Phys. Rev. B* **82**, 115120 (2010).
- [40] Shunji Matsuura, Po-Yao Chang, Andreas P Schnyder, and Shinsei Ryu, “Protected boundary states in gapless topological phases,” *New J. Phys.* **15**, 065001 (2013).
- [41] Robert Karplus and JM Luttinger, “Hall effect in ferromagnetics,” *Physical Review* **95**, 1154 (1954).
- [42] F. D. M. Haldane, “Berry curvature on the fermi surface: Anomalous hall effect as a topological fermi-liquid property,” *Phys. Rev. Lett.* **93**, 206602 (2004).
- [43] T. Jungwirth, Qian Niu, and A. H. MacDonald, “Anomalous hall effect in ferromagnetic semiconductors,” *Phys. Rev. Lett.* **88**, 207208 (2002).
- [44] Di Xiao, Ming-Che Chang, and Qian Niu, “Berry phase effects on electronic properties,” *Rev. Mod. Phys.* **82**, 1959–2007 (2010).
- [45] Naoto Nagaosa, Jairo Sinova, Shigeki Onoda, A. H. MacDonald, and N. P. Ong, “Anomalous hall effect,” *Rev. Mod. Phys.* **82**, 1539–1592 (2010).
- [46] Maissam Barkeshli and Xiao-Liang Qi, “Topological response theory of doped topological insulators,” *Phys.*

- Rev. Lett. **107**, 206602 (2011).
- [47] Stephen L Adler, “Axial-vector vertex in spinor electrodynamics,” *Physical Review* **177**, 2426 (1969).
- [48] John S Bell and Roman Jackiw, “A pcac puzzle:  $\pi^0 \rightarrow \gamma\gamma$  in the  $\sigma$ -model,” *Il Nuovo Cimento A* **60**, 47–61 (1969).
- [49] H.B. Nielsen and M. Ninomiya, “Absence of neutrinos on a lattice: (i). proof by homotopy theory,” *Nuclear Physics B* **185**, 20 – 40 (1981).
- [50] H. B. Nielsen and M. Ninomiya, “Absence of neutrinos on a lattice: (II). Intuitive topological proof,” *Nuclear Physics B* **193**, 173–194 (1981).
- [51] H. B. Nielsen and M. Ninomiya, “The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal,” *Physics Letters B* **130**, 389 – 396 (1983).
- [52] A. A. Zyuzin and A. A. Burkov, “Topological response in weyl semimetals and the chiral anomaly,” *Phys. Rev. B* **86**, 115133 (2012).
- [53] Y. Chen, Si Wu, and A. A. Burkov, “Axion response in weyl semimetals,” *Phys. Rev. B* **88**, 125105 (2013).
- [54] Vivek Aji, “Adler-bell-jackiw anomaly in weyl semimetals: Application to pyrochlore iridates,” *Phys. Rev. B* **85**, 241101 (2012).
- [55] Chao-Xing Liu, Peng Ye, and Xiao-Liang Qi, “Chiral gauge field and axial anomaly in a weyl semimetal,” *Phys. Rev. B* **87**, 235306 (2013).
- [56] G. Basar, D. E. Kharzeev, and H.-U. Yee, “Triangle anomaly in Weyl semi-metals,” ArXiv e-prints (2013), [arXiv:1305.6338 \[hep-th\]](#).
- [57] D. T. Son and B. Z. Spivak, “Chiral anomaly and classical negative magnetoresistance of weyl metals,” *Phys. Rev. B* **88**, 104412 (2013).
- [58] K. Landsteiner, “Anomaly related transport of Weyl fermions for Weyl semi-metals,” ArXiv e-prints (2013), [arXiv:1306.4932 \[hep-th\]](#).
- [59] P. Goswami and S. Tewari, “Axion field theory and anomalous non-dissipative transport properties of (3+1)-dimensional Weyl semi-metals and Lorentz violating spinor electrodynamics,” ArXiv e-prints (2012), [arXiv:1210.6352 \[cond-mat.mes-hall\]](#).
- [60] Adolfo G. Grushin, “Consequences of a condensed matter realization of lorentz-violating qed in weyl semi-metals,” *Phys. Rev. D* **86**, 045001 (2012).
- [61] Srinidhi T Ramamurthy and Taylor L Hughes, “Patterns of Electro-magnetic Response in Topological Semi-metals,” ArXiv e-prints (2014), [arXiv:1405.7377v1](#).
- [62] Balázs Hetényi, “dc conductivity as a geometric phase,” *Phys. Rev. B* **87**, 235123 (2013).
- [63] Zhou Jian-Hui, Jiang Hua, Niu Qian, and Shi Jun-Ren, “Topological invariants of metals and the related physical effects,” *Chinese Physics Letters* **30**, 027101 (2013).
- [64] Hong Yao and Dung-Hai Lee, “Topological insulators and topological nonlinear  $\sigma$  models,” *Phys. Rev. B* **82**, 245117 (2010).
- [65] Shou-Cheng Zhang and Jiangping Hu, “A four-dimensional generalization of the quantum hall effect,” *Science* **294**, 823–828 (2001).
- [66] Bogdan Andrei Bernevig, Chyh-Hong Chern, Jiang-Ping Hu, Nicolaos Toumbas, and Shou-Cheng Zhang, “Effective field theory description of the higher dimensional quantum hall liquid,” *Annals of Physics* **300**, 185 – 207 (2002).
- [67] R. D. King-Smith and D. Vanderbilt, “Theory of polarization of crystalline solids,” *Phys. Rev. B* **47**, 1651–1654 (1993).
- [68] D. J. Thouless, “Quantization of particle transport,” *Phys. Rev. B* **27**, 6083–6087 (1983).
- [69] K.-Y. Yang, Y.-M. Lu, and Y. Ran, “Quantum hall effects in a weyl semimetal: Possible application in pyrochlore iridates,” *Phys. Rev. B* **84**, 075129 (2011).
- [70] G.E. Volovik, *The Universe in a Helium Droplet*, International Series of Monographs on Physics (Oxford University Press, 2009).
- [71] Y. Ran, Y. Zhang, and A. Vishwanath, “One-dimensional topologically protected modes in topological insulators with lattice dislocations,” *Nature Physics* **5**, 298–303 (2009).
- [72] T. L. Hughes, H. Yao, and X.-L. Qi, “Majorana zero modes in dislocations of  $\text{Sr}_2\text{RuO}_4$ ,” ArXiv e-prints (2013), [arXiv:1303.1539 \[cond-mat.supr-con\]](#).
- [73] Zhong Wang and Shou-Cheng Zhang, “Chiral anomaly, charge density waves, and axion strings from weyl semimetals,” *Phys. Rev. B* **87**, 161107 (2013).
- [74] Yi Li, Shou-Cheng Zhang, and Congjun Wu, “Topological insulators with  $\text{su}(2)$  landau levels,” *Phys. Rev. Lett.* **111**, 186803 (2013).
- [75] Xiao-Gang Wen, “Projective construction of non-abelian quantum hall liquids,” *Phys. Rev. B* **60**, 8827–8838 (1999); B. Swingle, M. Barkeshli, J. McGreevy, and T. Senthil, “Correlated topological insulators and the fractional magnetoelectric effect,” *Phys. Rev. B* **83**, 195139 (2011); Joseph Maciejko, Xiao-Liang Qi, Andreas Karch, and Shou-Cheng Zhang, “Fractional topological insulators in three dimensions,” *Phys. Rev. Lett.* **105**, 246809 (2010).
- [76] J. K. Jain, “Composite-fermion approach for the fractional quantum hall effect,” *Phys. Rev. Lett.* **63**, 199–202 (1989).

## Appendix A: Topological Magnetoelectric Effect

Here we derive explicitly the topological magnetoelectric effect from our response theory. The topological magnetoelectric effect is only quantized in gapped systems, so we assume the system is gapped.

The relevant terms are in Eqn. (41), which we rewrite here as

$$j_{3D}^x = \frac{-1}{16\pi^3} \int d^3\bar{x} \text{tr} [\epsilon^{IJK} \mathcal{F}_{tI} \mathcal{F}_{JK} \mathcal{F}_{yz}] \quad (\text{A1})$$

where  $I, J, K$  run over  $\bar{x}, \bar{y}, \bar{z}$ . We assume that  $\mathcal{F}_{yz}$  is the real-space magnetic field  $B_x$ , so we can pull it out. For simplicity choose a gauge such that  $A_i = 0$  for  $i = x, y, z$ .

Expanding Eq. (A1) and manipulating some indices yields

$$j_{3D}^x = -\frac{B_x}{8\pi^3} \int d^3\bar{x} \epsilon^{IJK} \text{tr} [(\partial_t a_I - \partial_I a_t + [a_t, a_I])(\partial_J a_K + a_J a_K)] \quad (\text{A2})$$



We first show that several sets of terms in this expansion are zero. First, notice that

$$\int d^3\bar{x}\epsilon^{IJK}\partial_I a_t\partial_J a_K = - \int d^3\bar{x}\epsilon^{IJK}\partial_I\partial_J a_t a_K \quad (\text{A3})$$

$$= \int d^3\bar{x}\epsilon^{IJK}\partial_J a_t\partial_I a_K \quad (\text{A4})$$

$$= - \int d^3\bar{x}\epsilon^{IJK}\partial_I a_t\partial_J a_K \quad (\text{A5})$$

where we integrated by parts twice and then switched the indices  $I$  and  $J$ . Hence

$$\int d^3\bar{x}\epsilon^{IJK}\partial_I a_t\partial_J a_K = 0 \quad (\text{A6})$$

Next consider the terms

$$\int d^3\bar{x}\epsilon^{IJK}\text{tr}[[a_t, a_I]\partial_J a_K - \partial_I a_t a_J a_K] = \int d^3\bar{x}\epsilon^{IJK}\text{tr}[-\partial_J(a_t a_I - a_I a_t)a_K - \partial_I a_t a_J a_K] \quad (\text{A7})$$

$$= \int d^3\bar{x}\epsilon^{IJK}\text{tr}[\partial_I(a_t a_J - a_J a_t)a_K - \partial_I a_t a_J a_K] \quad (\text{A8})$$

$$= \int d^3\bar{x}\epsilon^{IJK}\text{tr}[(a_t\partial_I a_J - \partial_I a_J a_t - a_J\partial_I a_t)a_K] \quad (\text{A9})$$

$$= - \int d^3\bar{x}\epsilon^{IJK}\text{tr}[[a_t, a_I]\partial_J a_K - \partial_I a_t a_J a_K] \quad (\text{A10})$$

We have integrated by parts, manipulated indices, and used the cyclic property of the trace. Hence the left-hand side here is also zero.

Finally, trivial manipulations show that  $\int d^3\bar{x}\epsilon^{IJK}\text{tr}[[a_t, a_I]a_J a_K] = 0$  as well.

The remaining terms in the expansion are those that do not involve  $a_t$ :

$$j_{3D}^x = \frac{-B_x}{8\pi^3} \int d^3\bar{x}\epsilon^{IJK}\text{tr}[\partial_t a_I(\partial_J a_K + a_J a_K)] \quad (\text{A11})$$

$$= \frac{-B_x}{8\pi^3} \int d^3\bar{x}\epsilon^{IJK}\text{tr}[\partial_t(a_I\partial_J a_K + a_I a_J a_K) - a_I(\partial_t\partial_J a_K + 2\partial_t a_J a_K)] \quad (\text{A12})$$

$$= \frac{-B_x}{8\pi^3} \int d^3\bar{x}\epsilon^{IJK}\text{tr}[\partial_t(a_I\partial_J a_K + a_I a_J a_K) + \partial_J a_I\partial_t a_K - 2\partial_t a_I a_J a_K] \quad (\text{A13})$$

$$= \frac{-B_x}{8\pi^3} \int d^3\bar{x}\epsilon^{IJK}\text{tr}\left[\partial_t(a_I\partial_J a_K + a_I a_J a_K) - \partial_t a_I\partial_J a_K - \partial_t a_I a_J a_K - \frac{1}{3}\partial_t(a_I a_J a_K)\right] \quad (\text{A14})$$

$$= \frac{-B_x}{8\pi^3} \int d^3\bar{x}\epsilon^{IJK}\text{tr}\left[\partial_t\left(a_I\partial_J a_K + \frac{2}{3}a_I a_J a_K\right)\right] - j_{3D}^x \quad (\text{A15})$$

This immediately gives the desired relation (42).