Axion Response in Gapless Systems

Doron L. Bergman

Physics Department, California Institute of Technology, MC 114-36, 1200 East California Boulevard, Pasadena, California 91125, USA

(Received 22 February 2011; published 17 October 2011)

The strong topological insulator in 3D is expected to realize a quantized magnetoelectric response, the so-called axion response. However, many of the materials predicted to be topological insulators have turned out to be metallic, with bulk Fermi surfaces. Following the result of Bergman and Refael [Phys. Rev. B 82, 195417 (2010)] that the surface states of the topological insulator persist even when the band structure gap is closed, we explore the fate of the magnetoelectric response in such systems. We find that a nonquantized magnetoelectric coupling remains once a bulk Fermi surface opens. More generally, we find higher-dimensional analogs of the intrinsic anomalous Hall effect for all Chern forms—quantized transport coefficients in the gapped case become nonquantized when the gap is closed. In particular, the nonquantized magnetoelectric response in 3D descends from the intrinsic anomalous Hall effect analog in 4D.

DOI: 10.1103/PhysRevLett.107.176801

PACS numbers: 73.20.At, 72.20.-i

Axion electrodynamics, long confined to the realm of high-energy physics [1], has been recognized recently as a possible effective phenomenon in solids [2]. The strong topological insulator [2–5], at its boundary, is predicted to exhibit the physics of axion electrodynamics. The physical consequences of this fact are fantastic—ranging from a quantization of magnetoelectric coupling in materials with neither inversion symmetry breaking nor time-reversal breaking all the way to effective magnetic monopole image charges appearing when an external real charge is placed outside the material.

Perhaps the greatest current issue with 3D topological insulator candidate materials, which has so far prevented putting to the test the interesting predictions mentioned above, is the fact that they have for the most part turned out to be metallic [6–10]. The metallic states in these materials do show one characteristic feature of the topological insulator—the helical surface states. Theory tells us that topological states are well defined only when the spectrum of the bulk is gapped, yet in these materials the surface states coexist with a bulk Fermi surface.

Having noticed the coexistence of the helical surface states and bulk Fermi surfaces, and the observed experimental fact that they seem at least mildly robust to disorder effects [6], Bergman and Refael [11] recently explored the physics of the surface states of a hybrid system of metallic and topological insulator bands, concluding that it is entirely possible to have helical surface states (a litmus test of the topological insulator) coexisting with a bulk Fermi surface.

The simplest way to think of these “helical metals” (bulk metals with helical surface states) is as follows. The minimal feature in the bulk band structure of a solid giving topological insulator behavior is a point in the first Brillouin zone around which the dispersion is approximately that of a massive Dirac point dispersion $E = \pm \sqrt{M^2 + q^2}$. If a Fermi surface opens around some other point in the Brillouin zone, an approximate low-energy model for this band structure would consist of multiple valleys—separate valleys for the massive Dirac point and for the Fermi surface(s) [see Fig. 1(b)]. Much like the case of graphene [see Fig. 1(a)], where the two Dirac points are to a good approximation independent because of the large momentum separation between them [12], the massive Dirac point and the Fermi surface(s) will be to a good approximation independent when they have a large separation in the Brillouin zone. From this simplistic picture, we can understand that the helical surface states occur at the surface as a consequence of the massive Dirac point, and the Fermi surface simply does not always interfere with these surface states. In particular, the surface states are the Jackiw-Rebbi bound states of the Dirac valley. The band structure of Sb, expected to be a helical metal [11] as it exhibits helical surface states, can be described in low energies as a number of massive Dirac points, in addition to quadratic-dispersing Fermi

FIG. 1 (color online). Multivalley low-energy description of graphene and of the helical metal. (b) shows the two Dirac cones of graphene, and (a) shows a (dark-shaded) massive Dirac cone at the BZ center and a (light-shaded) quadratic-dispersion Fermi pocket at the BZ corner.
surfaces—this would suggest that the simple multivalley picture for the helical metal applies to Sb directly.

Using the simplistic multivalley picture, where the Fermi surface and the massive Dirac point are for the most part independent, one would naively expect that the axion response of the 3D topological insulator would also appear in this (hybrid) system. In this Letter, we explore this question in detail. We will find that the axion response can indeed appear, but its coefficient is no longer quantized.

A less obvious, yet even simpler model of the helical metal is the Hamiltonian

$$H = -\mu + q_1 \gamma_1 + q_2 \gamma_2 + q_3 \gamma_3 + (M - b q^2) \gamma_5,$$  \hspace{1cm} (1)

where $\gamma_{1...5}$ are $4 \times 4$ matrices, satisfying a Euclidean Clifford algebra

$$\{\gamma_a, \gamma_b\} = 2 \delta_{ab},$$  \hspace{1cm} (2)

and the chemical potential is outside the bulk gap. Note that, despite our notation, the Hamiltonian (1) is nonrelativistic. In our convention the $\gamma_{1...4}$ are odd under both time reversal and inversion, and $\gamma_5$ is even under both, so that this Hamiltonian is time-reversal and inversion symmetric. This model qualitatively reproduces most features of the low-energy spectrum of Bi$_2$Se$_3$ [9]—namely, a single massive Dirac point in the low-energy bulk spectrum and a single massless Dirac surface Fermi surface. Both (1) and minimal multivalley models can be captured in a generic 4-band model of the form $H = d_0(q) + d_4(q) \cdot \gamma_5$, which we will use to show that axion response can exist in gapless systems.

The axion response action $S_{\text{axion}} = \frac{e^2}{4\pi \hbar c} \int_{x,t} \phi \bar{E} \cdot \hat{B}$ is symmetric in its space-time coordinates and $q$ denotes a momentum-frequency vector. One needs to expand in small $q$ (a gradient expansion) $K_{\rho\rho}(q) = \text{const} + (\text{linear in } q) + \Omega_{\mu\nu\lambda\rho} q_\mu q_\lambda$, in order to recover the two derivatives in $\partial_\mu A_\nu \partial_\lambda A_\rho$. For the $\bar{E} \cdot \hat{B}$ term, one would need $\Omega_{\mu\nu\lambda\rho} = \Omega_0 \epsilon_{\mu\nu\lambda\rho} + \cdots$. Unfortunately, $q_\mu q_\lambda$ is symmetric in its two indices, and such a term would vanish. Therefore, linear response cannot give a nonvanishing $\bar{E} \cdot \hat{B}$ response. Remarkably, this also implies that even in a system with inversion and time-reversal symmetries explicitly broken, where the term $\bar{E} \cdot \hat{B}$ does not violate any symmetries, no axion term can come from linear response in a translationally invariant system. Consequently, we need the system to be inhomogeneous—and this is provided by the auxiliary field $\phi$, varying in space. Response theory is easier to calculate by assuming a translationally invariant system, so one can treat the $\phi$ field as an additional external field perturbing the system and calculate nonlinear response instead of linear response, similar to the procedure in Refs. [13,14]. In analogy with a gauge field, whose derivatives are the physical observable, only the derivatives of $\phi$ appear in the modified Maxwell’s equations and have measurable physical consequences.

In the helical metal, the gapless version of the topological insulator, we will find that the coefficient in front of the axion action is modified to a nonuniversal value

$$S_{\text{axion}} = \tilde{c} \frac{1}{4\pi^2} \int_{x,t} \phi \bar{E} \cdot \hat{B},$$  \hspace{1cm} (3)

with $\tilde{c} \neq 1$ and depending on the system details. The non-quantized magnetoelectric response coefficient is analogous to the Hall conductivity in the intrinsic anomalous Hall effect (AHE) in 2D ferromagnetic metals; because there is no gap, the Hall conductance is not quantized, even though it does have a topological origin—the Berry phase—and when opening up a bulk gap, the Hall conductance is quantized. We can immediately deduce one measurable consequence of the axion response. The surface of a solid naturally breaks inversion symmetry, and, if we apply a magnetic field in that region, that would cause $\phi$ to appear. Since the inversion breaking happens only at the surface, the gradient is of the form $\partial_z \phi = \lambda \delta(z)$, where the surface is taken as the $z = 0$ plane. Using this in (3), and assuming the other fields have no $z$ dependence, we can easily arrive at $S = \lambda \frac{\tilde{c}}{8\pi^2} \int_{x,t} \epsilon_{\nu\rho\lambda\mu} A_\nu \partial_\lambda A_\rho |_{z=0}$, which is a 1st Chern form, indicating a topological Hall effect on the surface of the material. Since its coefficient is not quantized, this is an intrinsic anomalous Hall effect on the surface. More generally, with a spatial gradient in $\phi$, we can find the electric current from the effective electromagnetic action...
\( J_a = \frac{\delta S_{\text{axion}}}{\delta A_a} = \tilde{c} \left( \partial_\phi \phi \right) \frac{1}{8\pi^2} \epsilon_{abc0} (\partial_c A_0 - \partial_0 A_c) = \tilde{c} \left( \partial_\phi \phi \right) \frac{1}{8\pi^2} \epsilon_{abc} E_c, \)  

\( \text{where} \ a, \ b, \ \text{and} \ c \ \text{are spatial indices. Basically, wherever there is a gradient in} \ \phi, \ \text{a Hall response will appear.} \)

We will start by demonstrating that an analog of the AHE occurs in gapless noninteracting systems in any odd space-time dimension \( d = 2n + 1, \) and from it, we will understand how the axion response appears in the helical metal. Being the first attempt at calculating this response, we will ignore the effects of disorder and of finite temperature.

From Kubo response theory, we can find that the coefficients of the Chern-Simons forms in \( d = 2n + 1 \) space-time dimensions are [15]

\[ S_{\text{Chern}} = \chi_n \epsilon_{\mu_1 \cdots \mu_{2n+1}} \int \mathcal{L}_A \partial_{\mu_1} A_{\mu_2} \cdots \partial_{\mu_{2n}} A_{\mu_{2n+1}}. \]

The coefficient \( \chi_n \) is given by the expression

\[ \chi_n = \frac{(-1)^{n+1}}{i(n+1)(2n+1)!} \int \partial_q \left[ \prod_{j=1}^{2n+1} (G \partial_{\mu_j} G^{-1}) \right], \]

where \( G = [i\omega_n - H(q)]^{-1} \) is the single particle Green’s function and \( \partial_{\mu_j} = \frac{\partial}{\partial q^\mu_j} \) with \( q_0 \) the real frequency, \( q_{1 \cdots 2n} \) being the Brillouin zone coordinates.

We will calculate this coefficient for a Hamiltonian of the form \( H = d_0(q) + d_\gamma(q) \cdot \gamma_\mu, \) where the matrices \( \gamma_{1 \cdots 2n+1} \) satisfy a Clifford algebra \( \{ \gamma_a, \gamma_b \} = 2\delta_{ab}, \) as well as the identity

\[ \text{Tr} \left[ \gamma_{a_1} \gamma_{a_2} \cdots \gamma_{a_{2n}} \gamma_{a_{2n+1}} \right] = 2n i \epsilon_{a_1 \cdots a_{2n+1}}. \]

With \( d_0 = 0 \) this is precisely the form of a minimal model for the Chern insulator in \( d = 2 + 1 \) and for the time-reversal invariant 4D quantum Hall state in \( d = 4 + 1 \) [16,17] (which realizes the strong topological insulator when reducing to \( d = 3 + 1 \)). With \( d_0 \) closing the gap in some regions of the Brillouin zone, this describes the helical metal. First, we “normalize” the Hamiltonian by \( R = |\vec{d}|, \) by writing \( H = R h = R (r_0 + r_\gamma \gamma_\mu), \) where now \( \vec{r}^2 = 1 \) and \( g = GR = R \left[(i\omega_n - H)^{-1} \right]^{-1}. \)

It follows that \( G \partial_{\mu} G^{-1} = g \partial_{\mu} g^{-1} + \partial_{\mu} \ln(R) \), in which the last term is a scalar and so can be moved outside of the trace in (6). As a result, when expanding the formula (6), any term where more than one instance of \( G \partial_{\mu} G^{-1} \rightarrow \partial_{\mu} \ln(R) \) appears will vanish, since \( \partial_{\mu} \ln(R) \partial_{\mu'} \ln(R) \) is symmetric in the indices \( \mu \) and \( \mu', \) and we have the completely antisymmetric tensor multiplying the entire term. From the remaining cases where zero or one instances of \( G \partial_{\mu} G^{-1} \rightarrow \partial_{\mu} \ln(R) \) appear, we have

\[ \chi_n = \frac{(-1)^{n+1}}{i(n+1)(2n+1)!} \int \left[ \prod_{j=1}^{2n+1} (g \partial_{\mu_j} g^{-1}) \right] \]

\[ + \sum_{j=1}^{2n+1} \partial_{\mu_j} \ln(R) \text{Tr} \left[ \prod_{j \neq j} (g \partial_{\mu_j} g^{-1}) \right], \]

The terms \( \partial_{\mu_j} g^{-1} \) will have a scalar part \( \partial_\mu (i\omega/R - r_0), \) which can leave the trace. If two or more instances of this scalar part appear, it will be symmetric in the two derivative indices and vanish under the antisymmetrization. Therefore, at most one such scalar part can appear, but this is also the minimal number, because one of the derivatives has to be with respect to frequency \( \partial_\omega g = 1/R, \) as a result, for all the other terms, we can ignore the scalar part of the current terms and take \( g \partial_{\mu} g^{-1} = \frac{(i\omega/R - r_0)^{-1} \gamma_\mu}{(i\omega/R - r_0)^{-1} \gamma_\mu} \) (at \( T = 0 \)). The terms \( \partial_{\mu} g^{-1} \) will have a scalar part \( \partial_\mu (i\omega/R - r_0), \) which can leave the trace. If two or more instances of this scalar part appear, it will be symmetric in the two derivative indices and vanish under the antisymmetrization. Therefore, at most one such scalar part can appear, but this is also the minimal number, because one of the derivatives has to be with respect to frequency \( \partial_\omega g = 1/R. \) For every permutation of \( \mu_1 \cdots 2n+1 \) we can cyclically rotate the \( \partial_0 \) term to the left end in the trace formula, yielding

\[ \chi_n = \frac{(-1)^{n+1}}{i(n+1)(2n+1)!} \int \left[ \prod_{j=1}^{2n+1} (g \partial_{\mu_j} g^{-1}) \right] \]

where now the indices run only over momentum coordinates. From \( \vec{r}^2 = 1 \) we get \( r \cdot \vec{\partial} \vec{r} = 0, \) which we can use to find \( g \partial_{\mu_j} g^{-1} g \partial_\mu g^{-1} = \frac{\gamma_\mu \gamma_\mu' r_{\mu_j} \partial_\mu' r_\mu}{(i\omega/R - r_0)^{-1}}. \) Plugging this in (9), we find
We first turn our attention to the trace over \( \gamma \) matrices. Because of the antisymmetry in the \( \mu_j \) indices, the product of \( \gamma_{a_1 \ldots a_{2n}} \) should also be antisymmetric in its indices. In particular, those indices should be all distinct. It is then easy to show that the trace over these \( 2n \) matrices must vanish: 
\[
\text{Tr}[\gamma_{a_1 \ldots a_{2n}}] = 0.
\]
Together with (7), we then find that
\[
\chi_n = \frac{(-1)^{n+1} e_{\mu_1 \ldots \mu_{2n}} e_{a_1 \ldots a_{2n+1}}}{(n+1)(2n-1)!} \int_q \frac{\text{Tr}[\gamma_{a_1 \ldots a_{2n}} \gamma_{b_1 \ldots b_{n}}]}{[(i\omega/R - r_0)^2 - 1]^{n+1}R}.
\tag{10}
\]

Performing the \( T = 0 \) Matsubara frequency integral, we get
\[
\int_0^\infty \frac{da}{(a-M)^2} = \frac{(-1)^{n+1} \sqrt{\pi} \Gamma(n+1/2)}{\Gamma(1+n)} = (1-\text{sign}(q)) \Theta_{1-\text{sign}(q)}
\]
and the integral gives a quantized value. The quantization can be understood by realizing that in the gapped case gives the intrinsic anomalous Hall effect contribution \([18]\). For \( n = 2 \) the gapped case gives the 4D quantum Hall effect \([16,17]\), while for the gapless case we get a nonuniversal transport coefficient. The \( n = 2 \) Chern response is the precursor of the axion response in 3D, and so our results strongly suggest that such a response can be found in 3D gapless systems, to which we turn now.

The simplest way to think of the axion response in 3 dimensions is to start with the \( n = 2 \) Chern form (5) and replace the extra vector potential component \( A_4 \) with the auxiliary field \( \phi \), mentioned earlier, which couples to matter fields via the free-fermion propagator \( G = G(q, \phi) \). Note that since \( A_4 \) is odd under time reversal and inversion, \( \phi \) must be odd as well, fitting its noncompact variety. The auxiliary field \( \phi \) couples to an extra current operator
\[
J_4 = -\partial_{\phi} G^{-1}(q, \phi)|_{\phi=0},
\]
and all the other current operators are now \( J_\mu = -\partial_{\mu} G^{-1}(q, \phi)|_{\phi=0}. \) In addition, the 4th spatial dimension gets dropped, and the antisymmetric symbols now include a 5th coordinate representing the extra field \( \phi \). The derivation of the nonlinear response follows through in a similar manner as described above for the \( d = 2n + 1 \) case, using \( G(q, \phi) \), except perhaps the integration by parts arguments used when converting to the “normalized” bands. However, the integration by parts used on (8) could be done with respect to any one of the \( 2n \geq 2 \) different space-time indices \( \mu_j \), so we can easily choose one that is a frequency-momentum component rather than \( \phi \). The calculation therefore follows through exactly in the same manner as in \( d = 4 + 1 \), and we get after a bit more algebra the coefficient from (3):
\[
\tilde{c} = 6\pi^2 e_{a_1 \ldots a_5} \int_q \Theta(|\vec{d}| - |d_0|) r_{a_1} \partial_{q_1} r_{a_2} \partial_{q_2} r_{a_3} \partial_{q_3} r_{a_4} \partial_{\phi} r_{a_5}.
\tag{13}
\]

To demonstrate that \( \tilde{c} \) can be nonzero in a gapless system, we will calculate it explicitly for (1), coupled to \( \phi \) as follows:
\[
H(\phi) = -\mu + q_1 \gamma_1 + q_2 \gamma_2 + q_3 \gamma_3 + \phi M \gamma_4 + (M - bq^2) \gamma_5.
\tag{14}
\]

Using the formula (13), we can easily find
\[
\tilde{c} = 3M^2 q^2 (M + b q^2) \int_0^\infty \frac{3Mq^2(M + b q^2)^{3/2}}{[(M - b q^2)^2 + q^2]^{5/2}} \times \Theta(q^2 + (M - b q^2)^2 - \mu^2).
\tag{15}
\]

Calculating numerically for \( M = 1 \), \( b = 0.5 \), and a range of values for the chemical potential \( \mu \), we find the results of Fig. 2, which clearly indicate a quantized value \( \tilde{c} = 1 \) for the gapped case and a nonquantized value for the gapless case.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2}
\caption{(color online). Calculation of the magnetoelectric response coefficient in a model system (14) with \( M = 1 \) and \( b = 0.5 \), realizing both a gapped and gapless spectrum. The transport coefficient \( \tilde{c} \) is quantized as long as a gap remains \( |\mu| < M \). Once the gap is closed, \( \tilde{c} \) is nonquantized.}
\end{figure}
In summary, we have shown that in higher dimensions an analog of the intrinsic anomalous Hall effect exists in gapless systems closely related to Chern and topological insulators. As a consequence, the 3D descendant of the 4D effect is a nonquantized axion response in 3D gapless systems. To experimentally probe the axion response in gapless systems, one could look for an AHE at the surface of the solid, with time reversal broken either by applying a magnetic field or by proximity to a ferromagnet. More generally, the nonquantized axion response should appear in helical metal materials doped with magnetic impurities, perhaps in Bi$_{2-x}$Mn$_x$Te$_3$ [19], doped to form a bulk Fermi surface.

The future directions following what we have shown here are interesting and numerous. Borrowing from what is known about the AHE [18], we expect that including disorder will bring about corrections to the axion response—both from a finite lifetime for quasiparticles and from vertex corrections. Given that the system is gapless, these corrections may be significant, and one must proceed with caution. Still, given the robustness of the AHE in $d = 2 + 1$ to disorder, it would be surprising if its higher-dimensional analogs (the nonquantized axion response among them) were completely washed out by disorder effects. Another question that arises is how Coulomb screening will affect the electromagnetic system with this strange new response—contrary to the topological insulator, the helical metals will have free carriers in the bulk that can screen electric fields. Regardless, probing ac response (where screening is no longer effective) may reveal the axion response. We leave these questions to future work.

D.L.B. was supported by the Sherman Fairchild foundation and acknowledges discussion with G. Refael and M. P. A. Fisher. Also, I thank G. E. Volovik, for pointing me to a discussion of topological effects in gapless systems [21].

Note added.—During the completion of this work, the author was made aware of the related work of Barkeshli and Qi [20], discussing much, though not all, of the same physics.