Electric polarization in a Chern insulator

Sinisa Coh* and David Vanderbilt†
Department of Physics & Astronomy, Rutgers University, Piscataway, NJ 08854-8019, USA
(Dated: October 24, 2008)

We extend the Berry-phase concept of polarization to insulators having a non-zero value of the Chern invariant. The generalization to such Chern insulators requires special care because of the partial occupation of chiral edge states. We show how the integrated bulk current arising from an adiabatic evolution can be related to a difference of bulk polarizations. We also show how the surface charge can be related to the bulk polarization, but only with a knowledge of the wavevector at which the occupancy of the edge state is discontinuous. We also present numerical calculations on a model Hamiltonian to provide additional support for our analytic arguments.

PACS numbers: 77.22.Ej, 73.43.-f, 73.20.At

In 1988 Haldane pointed out that an insulating crystal with broken time-reversal symmetry may exhibit a quantized Hall conductance even in the absence of a macroscopic magnetic field [1]. We shall refer to such a material as a "Chern insulator" (CI) because it necessarily would have a non-zero Chern invariant associated with its manifold of occupied Bloch states [2, 3]. While no CI has yet been discovered experimentally, there appears to be no reason why one could not exist, and theoretical models that behave as CIs are not difficult to construct. It seems plausible that the current blossoming of interest in exotic non-collinear magnets and multiferroics could yield an experimental example before long.

CIs occupy a middle ground between metals and ordinary insulators. Like metals, their conductivity tensor $\sigma_{\alpha\beta}$ is non-zero, their surfaces are metallic (as they are topologically required to have chiral conductance channels or "edge states" at the Fermi energy), and it is impossible to construct exponentially localized Wannier functions (WFs) for them [4]. On the other hand, only the off-diagonal (dissipationless) elements of $\sigma_{\alpha\beta}$ can be non-zero, the chiral edge states decay exponentially into the bulk (so that the deep-bulk region has a well-defined gap), the one-particle density matrix decays exponentially in the interior [5], and the localization measure $\Omega_{\rm I}$ introduced in Refs. [6, 7] is finite [5] as in other insulators. Overall it appears natural to regard a CI as an unusual species of insulator, but many aspects of its behavior remain open to investigation.

As is well known, the electric polarization $\mathbf{P}$ is not well-defined in a metal. For an ordinary insulator, its definition alternatively in terms of Berry phases or WFs is by now well established [8, 9, 10]. For a CI, the absence of a Wannier representation removes the possibility of using it to define the polarization, and we shall show below that there is a fundamental difficulty with the Berry-phase definition as well. In view of the presence of dissipationless currents and metallic edge states, one might be tempted to conclude that $\mathbf{P}$ is not well-defined at all in a CI. On the other hand, Souza et al. [11] have shown that the localization measure $\Omega_{\rm I}$ is related to the fluctuations of $\mathbf{P}$, and the finiteness of this quantity [5] suggests that the polarization might be well-defined after all.

The purpose of this Letter is to discuss whether, and in what sense, a definition of electric polarization is possible in a CI. We demonstrate that the usual Berry-phase definition does remain viable if it is interpreted with care when connecting it to observables such as the internal current that flows in response to an adiabatic change of the crystal Hamiltonian, or to the surface charge at the edge of a bounded sample.

For the remainder of this Letter we restrict ourselves to the case of a two-dimensional crystalline insulator having a single isolated occupied band. The generalization to the case of a three-dimensional multiband insulator is not difficult, but would complicate the presentation. We also restrict ourselves to a single-particle Hamiltonian; while this restriction is more difficult to remove, we note that all the principal difficulties associated with understanding CIs occur already at the one-particle level. The lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are related to the reciprocal lattice vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ in the usual way ($\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}$) and the cell area is $S = |\mathbf{a}_1 \times \mathbf{a}_2|$.

The Berry-phase expression for the electric polarization can be written as

$$\mathbf{P}_{[\mathbf{k}_0]} = \frac{e}{(2\pi)^2}\,\mathrm{Im}\int_{[\mathbf{k}_0]} d\mathbf{k}\,\langle u_{\mathbf{k}}|\nabla_{\mathbf{k}}|u_{\mathbf{k}}\rangle \qquad (1)$$

where e is the charge quantum ($e>0$), $|u_{\mathbf{k}}\rangle$ are the cell-periodic Bloch functions, and $[\mathbf{k}_0]$ indicates the parallelogram reciprocal-space unit cell with origin at $\mathbf{k}_0$ (that is, with vertices $\mathbf{k}_0$, $\mathbf{k}_0+\mathbf{b}_1$, $\mathbf{k}_0+\mathbf{b}_1+\mathbf{b}_2$, and $\mathbf{k}_0+\mathbf{b}_2$). In an ordinary insulator one insists on a smooth and periodic choice of gauge (relative phases of the $|u_{\mathbf{k}}\rangle$) in Eq. (1), and $\mathbf{P}$ is well defined (modulo $e\mathbf{R}/S$, where $\mathbf{R}$ is a lattice vector [8]) independent of $\mathbf{k}_0$. However, in a CI such a gauge choice is no longer possible. To see this, we decompose $\mathbf{P}_{[\mathbf{k}_0]} = P_1\mathbf{a}_1 + P_2\mathbf{a}_2$, $\mathbf{k} = k_1\mathbf{b}_1 + k_2\mathbf{b}_2$, and $\mathbf{k}_0 = \kappa_1\mathbf{b}_1 + \kappa_2\mathbf{b}_2$, and rewrite Eq. (1) as

$$P_1^{[\kappa_2]} = \frac{-e}{S}\int_{\kappa_2}^{\kappa_2+1} dk_2\,\frac{\theta_1(k_2)}{2\pi} \qquad (2)$$

where

$$\theta_1(k_2) = -\mathrm{Im}\int_{\kappa_1}^{\kappa_1+1} dk_1 \langle u_{k_1,k_2}|\partial_{k_1}|u_{k_1,k_2}\rangle\,. \qquad (3)$$

Eq. (3) is a Berry phase and is gauge independent modulo 2π (independent of κ_1). This allows us to make an arbitrary choice of branch for $\theta_1(k_2 = \kappa_2)$ and to insist, as part of the definition of $P_1^{[\kappa_2]}$, that $\theta_1(k_2)$ should remain continuous as k_2 is increased from κ_2 to $\kappa_2 + 1$. Since states at (k_1, κ_2) and $(k_1, \kappa_2 + 1)$ are equivalent, it follows that

$$\theta_1\Big|_{\kappa_2}^{\kappa_2+1} = -2\pi C \qquad (4)$$

where C is an integer. In fact C just defines the Chern number, and the insulator is a CI if $C \neq 0$.

Using Eqs. (2-3) and similar equations for P_2, we have arrived at a definition $\mathbf{P}_{[\mathbf{k}_0]}$ that is well-defined, modulo $e\mathbf{R}/S$ as usual, even for a CI. However, as illustrated in Fig. 1(a),

$$\mathbf{P}_{[\mathbf{k}_0+\Delta\mathbf{k}]} = \mathbf{P}_{[\mathbf{k}_0]} - \frac{eC}{2\pi}\,\hat{\mathbf{z}}\times\Delta\mathbf{k} \qquad (5)$$

where $\hat{\mathbf{z}}$ is the unit vector along $\mathbf{a}_1\times\mathbf{a}_2$. This dependence on $\mathbf{k}_0$ clearly presents a problem for the interpretation of Eq. (2) as a "physical" polarization in the case of a CI.

However, let us recall how the concept of polarization is *used*. For a normal insulator at least [8], the change of polarization during an adiabatic change of some internal parameter of the system from time t_i to t_f is given by

$$\int_{t_\mathrm{i}}^{t_\mathrm{f}} dt\,\mathbf{J}(t) = \mathbf{P}^{(\mathrm{f})}_{[\mathbf{k}_0]} - \mathbf{P}^{(\mathrm{i})}_{[\mathbf{k}_0]} \quad (\text{modulo } e\mathbf{R}/S)\,, \qquad (6)$$

where $\mathbf{J}(t)$ is the cell-averaged adiabatic current flowing in the bulk. A related statement, connected with the requirement that the charge pumped to the surface must be consistent with Eq. (6), is that the charge on an insulating surface normal to reciprocal vector $\mathbf{b}_1$ is [9]

$$\sigma = \mathbf{P}\cdot\hat{\mathbf{b}}_1 \quad (\text{modulo } e/a_2)\,. \qquad (7)$$

In the remainder of this Letter, we demonstrate that Eq. (6) remains correct for the case of a CI, provided that the *same* $\mathbf{k}_0$ (i.e., the same reciprocal-space cell) is used for $\mathbf{P}^{(\mathrm{i})}$ and $\mathbf{P}^{(\mathrm{f})}$ in Eq. (6). We also show that Eq. (7) must be modified in a CI, and show how to do so. We provide numerical tests as well as analytic arguments for both claims.

We begin by giving two arguments for the correctness of Eq. (6) in the CI case. First, it is straightforward to see that the the contribution to $J_1(t)$ can be computed independently for each k_2 [9], with the problem in (k_1, t) space effectively corresponding to that of an ordinary one-dimensional crystal. Thus, the derivation of Eq. (6) given in Ref. [8] goes through unchanged for the CI case. Second, we note that the expected result is obtained for the special case that the parameter of interest is a spatially uniform but time-dependent vector potential $\mathbf{A}(t)$. Since a slow turning on of $\mathbf{A}(t)$ causes state $u_\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}}$ to evolve into $u_{\mathbf{k}+(e/\hbar c)\mathbf{A}}e^{i\mathbf{k}\cdot\mathbf{r}}$, it follows that

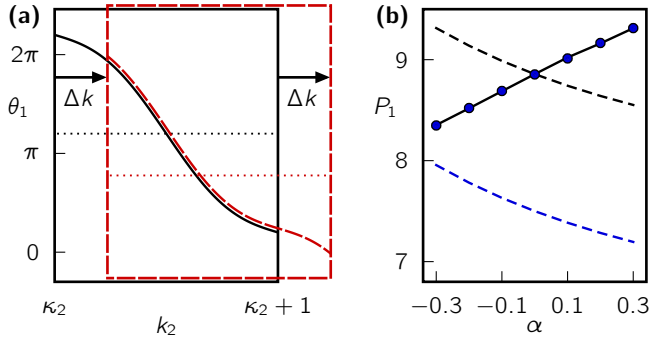

FIG. 1: (Color online) (a) Sketch of $\theta_1(k_2)$ in a Chern insulator ($C = +1$). Solid black and dashed red frames indicate reciprocal-cell origin chosen at κ_2 and $\kappa_2 + \Delta k$ respectively. Dotted lines indicate corresponding averages, proportional to P_1. (b) Computed $P_1(\alpha)$ for the modified Haldane model, in units of $-0.01e/S$, for adiabatic (dashed lines) and thermal (solid line and symbols) filling respectively. See text.

$$\mathbf{P}^{[\mathbf{A}]}_{[\mathbf{k}_0]} = \mathbf{P}^{[\mathbf{A}=0]}_{[\mathbf{k}_0]} - \frac{e^2C}{hc}\,\hat{\mathbf{z}}\times\mathbf{A}\,. \qquad (8)$$

But a time varying vector potential generates an electric field $\mathbf{E} = (-1/c)d\mathbf{A}/dt$, so that

$$\mathbf{J} = \frac{Ce^2}{h}\hat{\mathbf{z}}\times\mathbf{E}\,. \qquad (9)$$

The transverse conductivity σ_{xy} is thus quantized in units of e^2/h, expressing the fact that a CI is a realization of the integer quantum Hall effect [1].

We further confirm the validity of Eq. (6) by numerically testing our prediction on the Haldane model [1], which behaves as a CI in a certain region of its parameter space. This is a tight-binding model for spinless electrons on a honeycomb lattice (see inset of Fig. 2) at half filling (one occupied band), with staggered site energies and complex second-neighbor hoppings. Using the notation of Ref. [1], we adopt parameters $t_1 = 1$, $t_2 = 1/3$, $\phi = \pi/4$, $\Delta = 2/3$ and the lattice vectors $\mathbf{a}_1 = a_0(\sqrt{3}\hat{x}+\hat{y})/2$ and $\mathbf{a}_2 = a_0\hat{y}$ (so that $a_1 = a_2 = a_0$). Furthermore, we modify the first-neighbor hopping $t_1 \to t_1(1+\alpha)$ on the bonds parallel to $\mathbf{a}_1 + \mathbf{a}_2$ so as to break the threefold rotational symmetry and allow an adiabatic current to flow as α is varied. The compensating ionic charge is assumed to sit on the site with lower site energy.

We consider an infinite strip of the Haldane model N_1 cells wide and extending to $\pm\infty$ along y, as sketched in the inset of Fig. 2. States $\psi_{nk_2}(\mathbf{r})$ are labeled by k_2, which remains a good quantum number, and an addi-

tional index $n = 1, ..., 2N_1$. In the limit of large N_1

$$P_1 = \frac{-e}{N_1 S} \int_0^1 dk_2 \sum_{n \in \mathcal{N}(k_2)} \langle \psi_{nk_2} | r_1 | \psi_{nk_2} \rangle \,, \qquad (10)$$

where position vector $\mathbf{r}$ is decomposed as $\mathbf{r} = r_1 \mathbf{a}_1 + r_2 \mathbf{a}_2$ and $\mathcal{N}(k_2)$ is the set of occupied states to be discussed shortly. We associate the integrated current that flows along $\hat{x}$ in the interior of the strip during an adiabatic evolution from $\alpha = \alpha_\mathrm{i}$ to $\alpha = \alpha_\mathrm{f}$ by the corresponding change in Eq. (10), since by continuity the charge must arrive at the surfaces. We then compare this with the change of P_1 evaluated using a single bulk unit cell via Eqs. (2-3) to validate the theory.

There is a subtlety, however. Neutrality implies that $\mathcal{N}(k_2)$ contains N_1 states, but which ones? The problem arises because a CI is topologically required to have chiral metallic edge states. Our ribbon of CI therefore has one band of edge states along its left (L) edge and one along its right (R) edge (see inset of Fig. 2). For any given α, let $k_2^\times(\alpha)$ be the value of k_2 at which L-edge and R-edge bands cross. A *thermalized* filling of the edge states would correspond to the thick black curve for case α_i in Fig. 2, where the N_1 lowest-energy states are occupied at each k_2 and $\epsilon_\mathrm{F} = \epsilon(k_2^\times)$. Defining k_2^* to be the point at which the occupation switches between L and R edge states, we have $k_2^* = k_2^\times$ for the thermalized case.

In general $k_2^\times(\alpha)$ varies with α. However, k_2^* cannot change during an *adiabatic* evolution. Because we want to "measure" the polarization by the charge that accumulates at the surface, we specify that the adiabatic evolution is fast compared to the tunneling time between edge states but slow compared to all other processes, so that electrons cannot scatter between edges. Thus if we thermalize the system at α_i and then adiabatically carry the system from α_i to α_f, we arrive at the *adiabatic* filling illustrated by the thick red curve for case α_f in Fig. 2.

We thus expect that the change in polarization calculated from the right-hand side of Eq. (6) from the bulk bandstructure using Eqs. (1-3) should match that given by the change of Eq. (10) *only* if the adiabatic filling is maintained. We have confirmed this numerically for our modified Haldane model. The polarization as a function of α calculated using Eq. (10) and using the right-hand side of Eq. (6) are indicated in Fig. 1(b) with black and blue dashed lines respectively [12]. Eqs. (2-3) were evaluated on a 300×300 k-point mesh. Eq. (10) was calculated using five values of $N_1 \in [25, 70]$ and then extrapolating to infinity, while the k_2 integral was discretized with 5000 k-points. While there is a vertical offset between these curves that depends on the choice of $\mathbf{k}_0$ in Eq. (6), the *differences* ΔP_1 between different α are correct at the level of 10^{-5}. On the other hand, the results obtained with the thermalized filling in Eq. (10), shown by the solid line in Fig. 1(b), are drastically different. These results confirm that the appropriate comparison is with

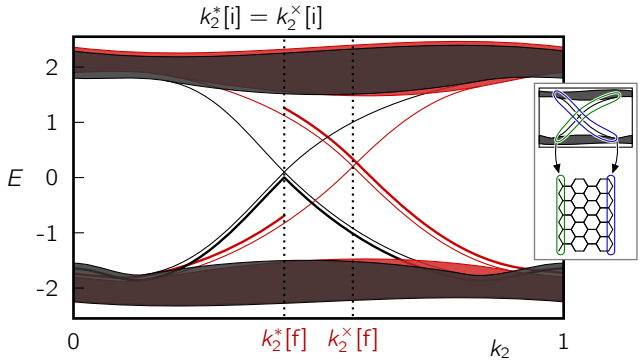

FIG. 2: (Color online) Sketch of a band structure of a finite ribbon of a Chern insulator. Solid regions indicate projected bulk bands; thin solid lines are edge states. Black and red correspond to $\alpha = \alpha_\mathrm{i}$ and $\alpha = \alpha_\mathrm{f}$ respectively; corresponding values of $k_2^\times$ are indicated. Thick lines indicate filling of edge states as dictated by k_2^*, chosen to illustrate system thermalized at α_i and then carried adiabatically to α_f. Inset: edge states associated with L (green) and R (blue) surfaces.

the adiabatic filling, and provide numerical confirmation that Eq. (6) is indeed satisfied even in a CI.

We now turn our attention to Eq. (7). A naive generalization to the CI case might be that $\sigma = \mathbf{P}_{[\mathbf{k}_0]} \cdot \hat{\mathbf{b}}_1$ (modulo e/a_2), but this cannot be correct. First, the left-hand side should be independent of $\mathbf{k}_0$, but the right-hand side is not. Second, the usual proof for ordinary insulators of the connection between surface charge and bulk polarization assumes that the surface is insulating, with the Fermi level lying in a gap common to both the bulk and surface [9]. When chiral edge states are present, the surfaces cannot be insulating, so the usual conditions are violated.

To show how Eq. (7) can be corrected for the case of a CI, let us again consider our Haldane-model ribbon at some fixed value of parameter α. Its surface charge σ can be calculated from $\sigma = \mathbf{P} \cdot \hat{\mathbf{b}}_1 = (S/a_2) P_1$ with P_1 evaluated using Eq. (10), but its value will depend on the the choice of the k_2^* at which the occupation of the edge state has its discontinuity, so that

$$\sigma^{[k_2^*]} = \frac{-e}{N_1 a_2} \int_0^1 dk_2 \sum_{n \in \mathcal{N}} \langle \psi_{nk_2} | r_1 | \psi_{nk_2} \rangle \,, \qquad (11)$$

where $\mathcal{N}$ is the set of N_1 occupied states at k_2 given the specified k_2^* (i.e., the choice whether the L or R edge state is included in $\mathcal{N}$ flips as k_2 passes through k_2^*).

Since the surface charge theorem of Eq. (7) for ordinary insulators was demonstrated via the Wannier representation [9], we take the same approach here. However, well-localized bulk WFs do not exist in a CI [4], so we focus instead on "hybrid Wannier functions" (HWFs) [13] in which the Fourier transform from Bloch functions is carried out in the r_1 direction only. Thus k_2 remains a

good quantum number and the HWF

$$W_{k_2}(r_1, r_2) = \sqrt{N_1} \int_0^1 dk_1 \, \Psi_{k_1 k_2}(r_1, r_2) \qquad (12)$$

is well localized only in the $\mathbf{a}_1$ direction. Using these we can represent the polarization

$$P_1^{[\kappa_2]} = \frac{-e}{S} \int_{\kappa_2}^{\kappa_2+1} dk_2 \, \rho_{k_2}^{[\kappa_2]} \qquad (13)$$

in terms of the HWF center $\rho_{k_2}^{[\kappa_2]} = \langle W_{k_2}|r_1|W_{k_2}\rangle$. We require that ρ be a continuous function of $k_2 \in [\kappa_2, \kappa_2+1]$ so as to guarantee a result that is equivalent to Eqs. (1-3).

To make the connection between this expression and Eq, (11), we construct Wannier-like functions along the $\mathbf{a}_1$ direction also for the finite-width strip. Constructing the $N_1 \times N_1$ matrix

$$\mathcal{R}^{[k_2^*]}_{mn,k_2} = \langle \psi_{mk_2}|r_1|\psi_{nk_2}\rangle \,, \qquad (14)$$

where $m, n \in \mathcal{N}$ as specified by k_2^*. The N_1 eigenvectors of $\mathcal{R}^{[k_2^*]}_{k_2}$ correspond to states that are Bloch-like along r_2 but localized along r_1, and we refer to them as the HWFs of the finite-width ribbon. The corresponding eigenvalue $\varrho_{jk_2}^{[k_2^*]}$ locates the center of charge of the j-th ribbon HWF. Using the basis-independence of the trace, Eq. (11) can now be rewritten as

$$\sigma^{[k_2^*]} = \frac{-e}{N_1 a_2} \int_0^1 dk_2 \sum_j \varrho_{jk_2}^{[k_2^*]} \,. \qquad (15)$$

The similarity between Eqs. (13) and (15) suggests that these can be connected. Since k_2 is a good quantum number, each k_2 can be treated independently. For each k_2 we can compare the infinite (bulk) 1D system described by Eq. (13) with the finite (ribbon) 1D system described by Eq. (15). The essential observation is that, in the limit of large N_1, the HWF centers ϱ_{jk_2} in the central part of the ribbon converge to the bulk ones given by ρ_{k_2} modulo an integer [9]. This is also illustrated in Fig. (3), where both sets of HWF centers are plotted as a function of k_2 for a ribbon of width $N_1 = 6$. Furthermore, the fact that the occupation of edge states switches between L and R edge at k_2^* is reflected in the discontinuity of ribbon HWF centers ϱ_{jk_2} at k_2^*. On the other hand, the bulk HWF centers ρ_{k_2} are chosen to be continuous across k_2^*. We can account for this discrepancy either by including a correction term proportional to $(k_2^* - \kappa_2)$,

$$\sigma^{[k_2^*]} = \frac{1}{a_2}\left[S P_1^{[\kappa_2]} + eC(k_2^* - \kappa_2)\right] \quad (\text{modulo } e/a_2) \,, \qquad (16)$$

or by realizing that by the virtue of Eq. (5) this is equivalent to shifting the reciprocal space origin to k_2^*,

$$\sigma^{[k_2^*]} = \frac{S}{a_2} P_1^{[k_2^*]} \quad (\text{modulo } e/a_2) \,, \qquad (17)$$

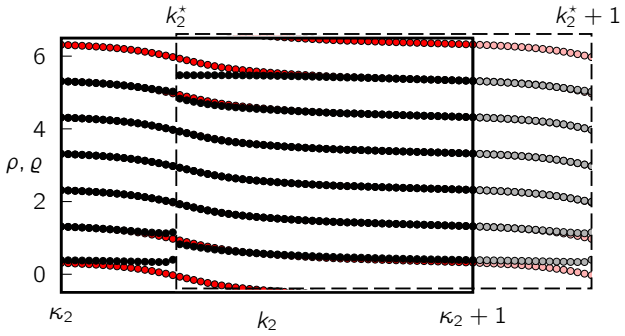

FIG. 3: (Color online) Black dots show ribbon HWF centers $\varrho_{jk_2}^{[k_2^*]}$ and red dots bulk HWF centers $\rho_{k_2}^{[\kappa_2]}$ and its periodic images as a function of k_2. Dashed frame corresponds to choice of origin at discontinuity in $\varrho_{jk_2}^{[k_2^*]}$, k_2^*.

as can be seen from the dashed frame in Fig. 3. Eq. (16) or (17) is the appropriate generalization of the surface charge theorem, Eq. (7), to the case of a CI, and should be correct in large N_1 limit for both thermalized and adiabatic fillings as long as the appropriate $k_2^\star$ is used.

We have also tested the correctness of this formula using our numerical calculations on the modified Haldane model. Recall that the solid curve in Fig. 1(b) represents the surface charge as computed from Eq. (10) for the thermalized case. Next, for each α, we locate $k_2^* = k_2^\times$ using 1000 k-points on a ribbon width $N_1 = 70$ and evaluate Eq. (17) using Eqs. (2-3) on a 250×250 k-point mesh. The resulting values are plotted as blue dots in Fig. 1(b). The agreement is clearly excellent.

In summary, we have generalized the Berry-phase concept of polarization to the case of a Chern insulator. The integrated current flow during adiabatic evolution is given by Eq. (6), where the reciprocal-space cell must be the same in both terms on the right-hand side. The surface charge at an edge of a bounded sample is given by Eq. (17), where k_2^* specifies the wavevector at which the occupation discontinuity occurs in the chiral edge state. These results may be of use in understanding the physical properties of these topological insulators, and perhaps in searching for experimental realizations.

We acknowledge useful discussions with P. Chandra. This work was supported by NSF Grant DMR-0549198.

* Electronic address: sinisa@physics.rutgers.edu
† Electronic address: dhv@physics.rutgers.edu

[1] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
[2] D. J. Thouless, *Topological Quantum Numbers in Nonrelativistic Physics* (World Scientific, Singapore, 1998).
[3] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[4] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. **98**, 046402 (2007).

[5] T. Thonhauser and D. Vanderbilt, Phys. Rev. B **74**, 235111 (2006).
[6] R. Resta and S. Sorella, Phys. Rev. Lett. **82**, 370 (1999).
[7] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
[8] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).
[9] D. Vanderbilt and R. D. King-Smith, Phys. Rev. B **48**, 4442 (1993).
[10] R. Resta and D. Vanderbilt, in *Modern Ferroelectrics.*, edited by C. Ahn and K. Rabe (Springer-Verlag, 2007), pp. 31–68.
[11] I. Souza, T. Wilkens, and R. M. Martin, Phys. Rev. B **62**, 1666 (2000).
[12] Note that $P_1 \neq 0$ even for $\alpha = 0$, since $\Delta \neq 0$ allows for an asymmetric population of the edge states.
[13] C. Sgiarovello, M. Peressi, and R. Resta, Phys. Rev. B **64**, 115202 (2001).