Dirac fermion quantization on graphene edges: Isospin-orbit coupling, zero modes, and spontaneous valley polarization

Grigory Tkachov

Max Planck Institute for the Physics of Complex Systems, Dresden, Germany
(Received 17 November 2008; published 30 January 2009)

The paper addresses boundary electronic properties of graphene with a complex edge structure of the armchair/zigzag/armchair type. It is shown that the finite zigzag region supports edge bound states with discrete equidistant spectrum obtained from the Green's function of the continuum Dirac equation. The energy levels exhibit the coupling between the valley degree of freedom and the orbital quantum number, analogous to a spin-orbit interaction. The characteristic feature of the spectrum is the presence of a zero mode, the bound state of vanishing energy. It resides only in one of the graphene valleys, breaking spontaneously Kramers' symmetry of the edge states. This implies the spontaneous valley polarization characterized by the valley isospin $\pm 1/2$. The polarization is manifested by a zero-magnetic-field anomaly in the local tunneling density of states, and is directly related to the local electric Hall conductivity.

PACS number(s): 73.20.At, 73.22.Gk, 73.63.Bd

I. INTRODUCTION

Due to the close connection between their topological and physical properties, two-dimensional (2D) electron systems have traditionally been in the focus of fundamental research. From the practical side, device functionalities in the 2D geometry are of great importance for applications and particularly suitable for lateral electronic architecture. The interest in these general aspects of 2D electron systems has recently revived in the light of the experimental success in isolating individual layers of graphite, preserving the honeycomb crystal structure.[1,2] Such a system—graphene—exhibits elementary excitations behaving at low energies and long distances as massless Dirac fermions.[3,4] Due to its massless quasiparticles graphene stands out among other 2D electron systems, which is probably most prominently manifested by the unconventional quantum Hall physics (e.g., Refs. [1], [2], and [5]–[8]), the phenomenon of Klein tunneling[9] and fermion bound states on extended defects such as graphene boundaries,[10]–[20] to name a few. In particular, understanding boundary effects in clean and disordered[21,22] graphene and the need for their characterization are among the outstanding current challenges in the field, arising from potentially promising electronic applications of graphene ribbons[23,24] and quantum dots.[25]

One of the reasons why the boundary effects in graphene should matter was pointed out quite a time ago by Fujita *et al.*[10] Using tight-binding calculations they predicted a new branch of quasiparticle states localized on the so-called "zigzag" edge. It is one of the most common types of the honeycomb lattice termination formed by two parallel crystal faces of the triangular sublattices of the honeycomb structure [see Fig. 1(a)]. The properties of the zigzag edge states are better understood when compared to the edge states in conventional 2D quantum Hall systems.[26] Unlike the latter, the zigzag edge states exist without any external magnetic field and any excitation gap in the 2D bulk. They are nonchiral: there is a Kramers' pair of counterpropagating modes originating from two nonequivalent nodal points of graphene's Brillouin zone [see Figs. 1(b) and 1(c)]. The zigzag edge

states have essentially the same origin as the bound states of massless fermions on domain walls.[27] Here the role of the domain wall is assumed by the out-of-plane rotation of the "sublattice" spin, which in the continuum limit corresponds to the zigzag edge.[20] Experimental evidence for the bound states on graphene edges comes from both tunneling[12,13] and angle-resolved photoemission spectroscopies.[14]

The present study is motivated by the observation that in experiments one has to deal with finite-length zigzag edges that represent a section of the graphene boundary sided usually by two armchair edges.[12,13] As the armchair sides do not support edge states,[10,16] one should generally expect quantization of the propagating modes in the finite zigzag section. This type of quantization is distinct from the size quantiza-
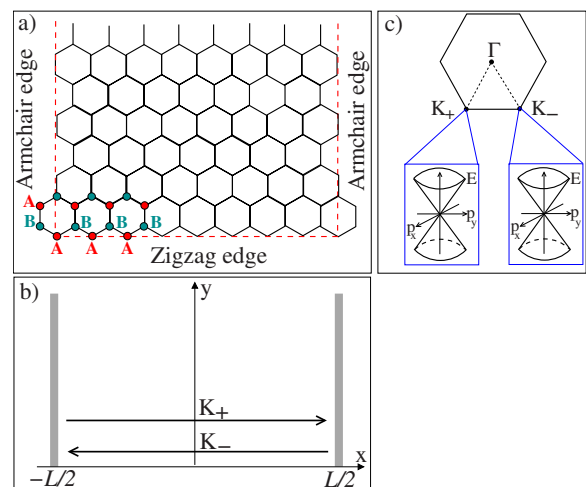

FIG. 1. (Color online) (a) Example of a finite-length zigzag edge sided by two armchair boundaries. A and B mark the sites of the two triangular sublattices. (b) Geometry of the continuum model for the system in panel (a): The zigzag edge at $y=0$ supports a Kramers' pair of counterpropagating edge states from K_+ and K_- valleys. They transform into each other via intervalley scattering caused by the armchair sides at $x=\pm L/2$. (c) Schematic view of the spectrum near the nodal points K_+ and K_- of graphene's Brillouin zone (Ref. [3]).

tion in zigzag graphene ribbons studied earlier[10,11,15,16] because it can occur on an *isolated* zigzag boundary, which is the typical situation in scanning tunneling experiments.[12,13] The consequences of such a quantization have not been studied previously. In the present work, they are addressed within the Dirac fermion confinement model derived from the lattice structure shown in Fig. 1(a).

In our approach the time-reversal symmetry and Kramers' degeneracy of the zigzag edge states comes as a result of an effective isospin-orbit coupling. The isospin is introduced as a convenient formal representation for the two nonequivalent nodal points of graphene's Brillouin zone. The rotations generated by the isospin leave the 2D Dirac equation invariant. We show that this continuous symmetry is broken by the zigzag confinement, and the edge-state spectrum explicitly depends on the confinement parameters controlling the isospin-orbit coupling. The quantization of the edge states is achieved by imposing effective boundary conditions at the ends of the zigzag edge [see Fig. 1(b)]. They cause the intervalley scattering connecting the incident and outgoing edge states, which models the armchair confinement. It turns out that the quantized spectrum contains a zero mode, i.e., the state with vanishing momentum and energy. Remarkably, it couples only to one of the isospin projections, which exists only in one of the valleys, breaking spontaneously the Kramers' symmetry of the edge states. This leads to the spontaneous isospin (valley) polarization with the total edge-state isospin $\pm 1/2$. This mechanism of the valley polarization differs from the previous proposals.[28] We demonstrate that the spontaneous symmetry breaking can be detected through the magnetic-field dependence of the tunneling density of states, and also find a direct relation between the isospin polarization and the local electric Hall conductivity.

The subsequent sections give a complete account of our approach: In Sec. II we formulate the boundary problem for a finite zigzag edge and analyze it in terms of the discrete and continuous symmetries of the 2D Dirac fermions. The Green's function solution of the boundary problem and the spectrum of the quantized Dirac fermion edge states are discussed in Sec. III. Section IV addresses the valley polarization effects, both spontaneous and induced. The latter is the analog of the quantum spin Hall polarization. Finally, Sec. V describes the signatures of the valley polarization in observables, such as the tunneling density of states and the local electric Hall conductivity, and contains concluding discussion.

II. BOUNDARY PROBLEM

A. 2D massless fermions, chiral symmetry, and isospin

The two distinct nodal points (valleys) of graphene's Brillouin zone result in a pair of massless Weyl fermions whose wave functions, ψ_+ and ψ_-, satisfy the matrix equation

$$\epsilon \Psi = H\Psi, \quad \Psi = \begin{bmatrix} \psi_+ \\ \psi_- \end{bmatrix}, \quad H = v \begin{bmatrix} \boldsymbol{\sigma}\mathbf{p} & 0 \\ 0 & U(\boldsymbol{\sigma}\mathbf{p})U^{-1} \end{bmatrix}. \tag{1}$$

It is assumed that the Hamiltonian H is diagonal in valley space $(+,-)$. The intravalley Hamiltonians are expressed in

terms of the Pauli matrices $\sigma_{x,y,z}$ acting on the functions

$$\psi_{\pm} = \begin{bmatrix} \psi_{A_{\pm}} \\ \psi_{B_{\pm}} \end{bmatrix} \tag{2}$$

that have two components due to the bipartite lattice structure of graphene, with two sublattices denoted as A and B in Fig. 1(a); v and ϵ are the Fermi velocity and energy with respect to the Fermi level, and the quasiparticle momentum $\mathbf{p}$ is confined to the plane of the system.

We further assume that the intravalley Hamiltonians are related to each other by the chiral symmetry

$$U(\boldsymbol{\sigma}\mathbf{p})U^{-1} = -\boldsymbol{\sigma}\mathbf{p}, \tag{3}$$

where U is a unitary matrix. In this way we explicitly account for the generic property of nodal lattice quasiparticles known as fermion doubling: they come in pairs of opposite-chirality (Weyl) species that together obey the Dirac equation.[29] We note that in the 2D case the unitary transformation, Eq. (3), is always achieved by one of the σ matrices. If, for instance, the system is located in the x,y plane (Fig. 1), we have

$$\boldsymbol{\sigma}\mathbf{p} = \sigma_x p_x + \sigma_y p_y, \quad U = \sigma_z. \tag{4}$$

The discrete chiral symmetry, Eq. (3), can be promoted to a continuous one. Let us introduce another set of the Pauli matrices $\tau_{1,2,3}$, acting in the valley space, and consider the vector operator

$$\mathbf{I} = \frac{1}{2}(\tau_1 \otimes \sigma_z, \tau_2 \otimes \sigma_z, \tau_3 \otimes \sigma_0), \quad [I_k, I_l] = i\varepsilon_{klm}I_m, \tag{5}$$

whose components I_k $(k=1,2,3)$ formally satisfy the commutation relations of an angular momentum. It is easy to see that the Hamiltonian H is invariant under rotations generated by I_k,

$$e^{i\theta_k I_k} H e^{-i\theta_k I_k} = H, \quad H = v\tau_3 \otimes (\boldsymbol{\sigma}\mathbf{p}), \tag{6}$$

where θ_k is the rotation angle. This means that the original choice of the upper and lower components of Ψ [Eq. (1)] as being the "+" and "−" valley functions, respectively, is not physically distinguished. One can rather treat them as the "up" and "down" states of the effective spin (isospin) 1/2. We will nevertheless keep the original notations $\psi_{\pm}$ for the upper and lower components of Ψ, interpreting them as the projections

$$\psi_{\pm} = \left(\frac{I_0}{2} \pm I_3\right)\Psi, \tag{7}$$

where $I_0 = \tau_0 \otimes \sigma_0$ is the 4×4 unit matrix (the direct product of the 2×2 unit matrices τ_0 and σ_0).

B. Boundary condition for the zigzag edge and broken isospin rotation symmetry

The zigzag edge is a type of the honeycomb lattice termination where the outermost lattice sites all belong to one of the sublattices [Fig. 1(a)]. It does not couple the states from the K_+ and K_- valleys,[10] due to which the continuum bound-

ary condition for the zigzag edge can be obtained by rather simple reasoning.[16] To be concrete let us assume that the outermost sites are all of the A type and the next (missing) atomic row would be of the B type, as in Fig. 1(a). On the missing B row one can impose the hard-wall condition $\psi_{B\pm}(x,0)=0$, while keeping $\psi_{A\pm}(x,0)$ arbitrary. In spinor notations [Eq. (2)], this reads

$$\psi_{\pm}(x,0) = \boldsymbol{\sigma}\mathbf{l}_{\perp}\psi_{\pm}(x,0), \quad \mathbf{l}_{\perp} = (0,0,1). \tag{8}$$

This boundary condition admits the generalization beyond the hard-wall approximation. It is achieved by rotating the unit vector $\mathbf{l}_{\perp}$ about the normal $\mathbf{n}_B$ to the boundary [in Fig. 1(b), $\mathbf{n}_B \| \hat{\mathbf{y}}$], which is consistent with the requirement for the normal component of the current to vanish at the edge.[18,30] Moreover, the rotation can be made valley dependent: $\mathbf{l}_{\perp} \rightarrow \mathbf{l}_{\perp \pm}$. Using the four component spinors, we can therefore write

$$\Psi(x,0) = M\Psi(x,0), \tag{9}$$

$$M = \frac{\tau_0 + \tau_3}{2} \otimes \boldsymbol{\sigma}\mathbf{l}_+ + \frac{\tau_0 - \tau_3}{2} \otimes \boldsymbol{\sigma}\mathbf{l}_-, \tag{10}$$

$$\mathbf{l}_{\pm}^2 = 1, \quad (\mathbf{l}_{\pm}\mathbf{n}_B) = 0. \tag{11}$$

Further restrictions on $\mathbf{l}_{\pm}$ are imposed by the discrete symmetries of the problem. As the lattice prototype of our system has two identical sides [Fig. 1(a)], our continuum model should inherit spatial parity with respect to coordinate reflection along the edge, i.e., $x \rightarrow -x$. It is the symmetry of the Dirac equation (1) since the coordinate reflection can be compensated by the spinor transformation

$$\Psi^P(x,y) = \Lambda\Psi(-x,y), \quad \Lambda = \tau_1 \otimes \sigma_x, \tag{12}$$

simultaneously swapping both the valley and sublattice spinor components. However, the boundary condition, Eq. (9), does not share this symmetry because M and Λ do not commute

$$\Lambda M \Lambda^{-1} = \frac{\tau_0 - \tau_3}{2} \otimes (\sigma_x l_{x+} - \sigma_z l_{z+})$$

$$+ \frac{\tau_0 + \tau_3}{2} \otimes (\sigma_x l_{x-} - \sigma_z l_{z-}), \tag{13}$$

unless there is a relation between $\mathbf{l}_+$ and $\mathbf{l}_-$ such that

$$l_{x+} = l_{x-} \equiv l_x, \quad l_{z+} = -l_{z-} \equiv l_z, \quad \mathbf{l} = (l_x, 0, l_z). \tag{14}$$

These restrictions also make the zigzag boundary invariant under time-reversal operation $\Psi^T(x,y) = \Lambda\Psi^*(x,y)$.

We are now prepared to prove that the zigzag boundary condition, Eq. (9) violates the isospin rotation symmetry. More specifically, we are talking about the nontrivial rotations generated by the I_1 and I_2 components of the isospin, Eq. (5). Indeed, the matrix M (10) does not commute with $I_{1,2}$:

$$I_{1,2}M\Gamma_{1,2}^{-1} = \frac{\tau_0 - \tau_3}{2} \otimes (-\sigma_x l_{x+} + \sigma_z l_{z+})$$

$$+ \frac{\tau_0 + \tau_3}{2} \otimes (-\sigma_x l_{x-} + \sigma_z l_{z-}), \tag{15}$$

unless $\mathbf{l}_+$ and $\mathbf{l}_-$ satisfy the conditions:

$$l_{x+} = -l_{x-}, \quad l_{z+} = l_{z-}. \tag{16}$$

These are incompatible with the requirements for the x-parity and time-reversal symmetry, Eq. (14). *In Sec. III B we demonstrate that the broken isospin rotation symmetry implies an analog of the spin-orbit coupling controlled by the components of the vector* $\mathbf{l}$ *in Eq. (14).*

C. Parity-symmetric armchair edges

We now turn to the boundary conditions at the armchair sides $x = \pm L/2$. They should account for the valley and sublattice mixing specific to the armchair lattice termination[10] and, at the same time, possess both the x-parity and time-reversal symmetry. The suitable boundary conditions can be written as[31]

$$\Psi\left(\pm\frac{L}{2},y\right) = \Lambda\Psi\left(\pm\frac{L}{2},y\right), \tag{17}$$

with the same off-diagonal matrix Λ as in Eq. (12). They meet the requirement of the vanishing of the normal component of the Dirac current:

$$j_x\left(\pm\frac{L}{2},y\right) = \Psi^{\dagger}\left(\pm\frac{L}{2},y\right)\tau_3 \otimes \sigma_x\Psi\left(\pm\frac{L}{2},y\right)$$

$$= \Psi^{\dagger}\left(\pm\frac{L}{2},y\right)\tau_1 \otimes \sigma_x(\tau_3 \otimes \sigma_x)\tau_1 \otimes \sigma_x\Psi\left(\pm\frac{L}{2},y\right)$$

$$= -\Psi^{\dagger}\left(\pm\frac{L}{2},y\right)\tau_3 \otimes \sigma_x\Psi\left(\pm\frac{L}{2},y\right)$$

$$= -j_x\left(\pm\frac{L}{2},y\right) = 0, \tag{18}$$

where we have switched to the creation $\Psi^{\dagger}(x,y)$ and annihilation $\Psi(x,y)$ operators.

Importantly, the x parity of the problem allows us to reduce the boundary conditions, Eq. (17), to the usual symmetric boundary conditions

$$\Psi\left(\frac{L}{2},y\right) = \Psi\left(-\frac{L}{2},y\right). \tag{19}$$

To prove this we first notice that the original function $\Psi(x,y)$ and the transformed one $\Psi^P(x,y)$ [Eq. (12)] correspond to the same solution of Eqs. (1), (9), and (17), and, therefore, must coincide: $\Psi(x,y) = \Lambda\Psi(-x,y)$. In particular, at $x = \pm L/2$ we have

$$\Psi\left(\pm\frac{L}{2},y\right) = \Lambda\Psi\left(\mp\frac{L}{2},y\right). \tag{20}$$

Comparison with Eq. (17) yields Eq. (19). For the actual calculations, we will use the symmetric boundary conditions modulated by a magnetic phase ϕ

$$\Psi\left(\frac{L}{2},y\right) = \Psi\left(-\frac{L}{2},y\right)\exp(2\pi i\phi), \qquad (21)$$

$$E_x = -(h/eL)\dot{\phi}. \qquad (22)$$

In this way we account for a weak magnetic field perpendicular to the plane x,y. If its vector potential is chosen to be parallel to the zigzag edge, $\mathbf{A}(y)\|\hat{\mathbf{x}}$ and to vanish at $y\to\infty$, then at $y=0$ the phase ϕ exactly equals to the flux through the strip in units of ch/e. For weak magnetic fields, the spatial variation in ϕ with the coordinate y can be neglected, while its adiabatic variation with time implies an electric field along $\hat{\mathbf{x}}$ given by Eq. (22).

III. DIRAC FERMION EDGE STATES ON FINITE ZIGZAG EDGES

A. Green's function of the system

To study the spectral properties of zigzag graphene edges it is convenient to use the Green's function approach. The specifics of its implementation to boundary problems in graphene is still scarcely covered in literature (e.g., Refs. 17 and 32). Below we describe in some detail the main calculation steps leading to the final result given by Eqs. (36)–(38).

We begin by introducing the retarded Green's function $G^R(\mathbf{r}t,\mathbf{r}'t')$ as a 4×4 matrix in space of the valley (isospin) and the sublattice degrees of freedom whose matrix elements are given by

$$G^R_{\alpha\beta}(\mathbf{r}t,\mathbf{r}'t') = \frac{\Theta(t-t')}{i\hbar}\times\langle\psi_\alpha(\mathbf{r}t)\psi^\dagger_\beta(\mathbf{r}'t')+\psi^\dagger_\beta(\mathbf{r}'t')\psi_\alpha(\mathbf{r}t)\rangle, \qquad (23)$$

where the brackets $\langle\cdots\rangle$ denote averaging with the equilibrium statistical operator and the indices α and β independently run through all possible combinations of the isospin and sublattice indices: $\alpha,\beta=A_+,A_-,B_+,B_-$. As the zigzag edge [Eq. (9)] possesses the isospin rotation symmetry generated by I_3 (i.e., does not couple the valleys), G^R can be decomposed into the direct product

$$G^R(\mathbf{r}t,\mathbf{r}'t') = \frac{1}{2}\sum_{\tau=\pm1}(\tau_0+\tau\tau_3)\otimes G^R_\tau(\mathbf{r}t,\mathbf{r}'t'), \qquad (24)$$

where $\tau=\pm1$ labels the valleys (i.e., the two isospin projections) and

$$G^R_\tau(\mathbf{r}t,\mathbf{r}'t') = \begin{pmatrix} G_{AA|\tau}(\mathbf{r}t,\mathbf{r}'t') & G_{AB|\tau}(\mathbf{r}t,\mathbf{r}'t') \\ G_{BA|\tau}(\mathbf{r}t,\mathbf{r}'t') & G_{BB|\tau}(\mathbf{r}t,\mathbf{r}'t') \end{pmatrix} \qquad (25)$$

is the matrix Green's function in sublattice space. Its time Fourier transform satisfies the equation

$$(\epsilon\sigma_0 - v\tau\boldsymbol{\sigma}\mathbf{p})G^R_\tau(\mathbf{r},\mathbf{r}') = \sigma_0\delta(\mathbf{r}-\mathbf{r}'). \qquad (26)$$

In terms of $G^R_\tau(\mathbf{r},\mathbf{r}')$ the boundary conditions, Eqs. (9) and (21), of the previous section read

$$G^R_\tau = (\boldsymbol{\sigma}\mathbf{l}_\tau)G^R_\tau|_{y=0}, \quad \mathbf{l}_\tau \equiv \mathbf{l}_\pm, \qquad (27)$$

$$G^R_\tau|_{x=L/2} = G^R_\tau|_{x=-L/2}\exp(2\pi i\phi). \qquad (28)$$

The solution to Eq. (26) can be sought in the form

$$G^R_\tau(\mathbf{r},\mathbf{r}') = \left(\sigma_0 + \frac{v\tau}{\epsilon}\boldsymbol{\sigma}\mathbf{p}\right)$$
$$\times\sum_{n\in\mathbb{Z}}\begin{pmatrix} G_{AA|\tau k_n}(y,y') & 0 \\ 0 & G_{BB|\tau k_n}(y,y') \end{pmatrix}\frac{e^{ik_n(x-x')}}{L}, \qquad (29)$$

where the diagonal matrix elements are the Green's functions on sublattices A,B. They are expanded in plane waves $e^{ik_n x}$ with the wave number

$$k_n = (2\pi/L)(n+\phi), \quad n\in\mathbb{Z}(0,\pm1,\ldots), \qquad (30)$$

given by the boundary condition, Eq. (28). For $G_{AA,BB|\tau k_n}(y,y')$ one has the ordinary differential equation,

$$(\partial_y^2 - q_n^2)G_{AA,BB|\tau k_n}(y,y') = \frac{\epsilon}{\hbar^2 v^2}\delta(y-y'), \qquad (31)$$

and the boundary conditions following from Eq. (27),

$$\partial_y G_{AA|\tau k_n} = \left[\frac{\tau\epsilon(1-l_{z\tau})}{\hbar v l_{x\tau}} - k_n\right]G_{AA|\tau k_n}\Bigg|_{y=0}, \qquad (32)$$

$$\partial_y G_{BB|\tau k_n} = \left[\frac{-\tau\epsilon(1+l_{z\tau})}{\hbar v l_{x\tau}} + k_n\right]G_{BB|\tau k_n}\Bigg|_{y=0}, \qquad (33)$$

where $q_n=\sqrt{k_n^2-\epsilon^2/\hbar^2 v^2}$. We seek the solution (finite at $y\to\infty$) in the form

$$G_{AA,BB|\tau k_n}(y,y') = C_{A,B}(y')e^{-q_n y} - \frac{\epsilon}{2\hbar^2 v^2 q_n}e^{-q_n|y-y'|},$$

where the first term is the solution of the homogeneous Eq. (31) and the second one is the Green's function of the unbounded system. The coefficients $C_{A,B}$ are obtained from Eqs. (32) and (33) with the following results:

$$G_{AA|\tau k_n}(y,y') = \frac{\epsilon}{2\hbar^2 v^2 q_n}(e^{-q_n(y+y')} - e^{-q_n|y-y'|})$$
$$+ \frac{(1+l_{z\tau})(q_n+k_n) - \tau\epsilon l_{x\tau}/\hbar v}{2(\epsilon - \hbar v\tau k_n l_{x\tau} + i0)}e^{-q_n(y+y')}, \qquad (34)$$

$$G_{BB|\tau k_n}(y,y') = \frac{\epsilon}{2\hbar^2 v^2 q_n}(e^{-q_n(y+y')} - e^{-q_n|y-y'|})$$
$$+ \frac{(1-l_{z\tau})(q_n-k_n) + \tau\epsilon l_{x\tau}/\hbar v}{2(\epsilon - \hbar v\tau k_n l_{x\tau} + i0)}e^{-q_n(y+y')}. \qquad (35)$$

The results of the above calculations can be summarized in the expression for the full matrix Green's function,

$$G^R(\mathbf{r}, \mathbf{r}') = \sum_{\tau=\pm 1, n \in \mathbb{Z}} \left(\frac{\tau_0 + \tau \tau_3}{2} \right) \otimes \left(\sigma_0 + \frac{\upsilon \tau}{\epsilon} \boldsymbol{\sigma} \mathbf{p} \right)$$

$$\times [G^s_{\tau k_n}(y, y') \sigma_0 + G^a_{\tau k_n}(y, y') \sigma_z] \frac{e^{ik_n(x-x')}}{L}. \tag{36}$$

We introduce the symmetric $G^s_{\tau k_n}(y, y') = (G_{AA|\tau k_n} + G_{BB|\tau k_n})/2$ and asymmetric $G^a_{\tau k_n}(y, y') = (G_{AA|\tau k_n} - G_{BB|\tau k_n})/2$ sublattice functions given explicitly by

$$G^s_{\tau k_n}(y, y') = \frac{\epsilon}{2\hbar^2 \upsilon^2 q_n} (e^{-q_n(y+y')} - e^{-q_n|y-y'|})$$

$$+ \frac{q_n + k_n l_{z\tau}}{2(\epsilon - \hbar \upsilon \tau k_n l_{x\tau} + i0)} e^{-q_n(y+y')}, \tag{37}$$

$$G^a_{\tau k_n}(y, y') = \frac{k_n + q_n l_{z\tau} - \tau \epsilon l_{x\tau}/\hbar \upsilon}{2(\epsilon - \hbar \upsilon \tau k_n l_{x\tau} + i0)} e^{-q_n(y+y')}. \tag{38}$$

In the equations above the denominators vanish at $\epsilon = \hbar \upsilon \tau k_n l_{x\tau}$. To identify this as a pole, we should make sure that the nominators remain finite as $\epsilon \to \hbar \upsilon \tau k_n l_{x\tau}$. In this limit, Green's function (36) behaves as

$$G^R(\mathbf{r}, \mathbf{r}') \approx -\frac{1}{2L} \sum_{\tau=\pm 1, n \in \mathbb{Z}} (\tau_0 + \tau \tau_3) \otimes (\sigma_0 + \boldsymbol{\sigma} \mathbf{l}_\tau)$$

$$\times \frac{\Theta(k_n l_{z\tau})}{\epsilon - \hbar \upsilon \tau k_n l_{x\tau} + i0} \partial_y \, e^{-|k_n l_{z\tau}|(y+y')+ik_n(x-x')}, \tag{39}$$

and we can see that the pole exists only if the unit step function $\Theta(k_n l_{z\tau})$ is not zero,

$$\epsilon_{\tau,n} = \hbar \upsilon \tau k_n l_{x\tau}, \quad k_n l_{z\tau} > 0. \tag{40}$$

This is the spectrum of the states, decaying exponentially from the edge $y=0$ and propagating along x.

B. Edge-state spectrum, isospin-orbit coupling, and zero modes

Let us analyze the edge-state spectrum, Eq. (40), in some more detail. With the requirements of the x parity and time-reversal symmetry [see Eq. (14)] and for k_n given by Eq. (30), we have

$$\epsilon_{\tau,n} = \text{sgn}(l_x) \Delta \tau(n + \phi), \quad \tau(n + \phi)l_z > 0, \tag{41}$$

$$\Delta = h\upsilon|l_x|/L, \quad n = 0, \pm 1, \dots. \tag{42}$$

It is equidistant with the level spacing Δ and particle-hole asymmetric because of the restriction $(n+\phi)\tau l_z > 0$ [see Fig. 2]. The phase ϕ results in the homogeneous shift of the levels. Let us consider $|\phi| \ll 1$ and neglect the shift in all of the states except the zero mode $n=0$,
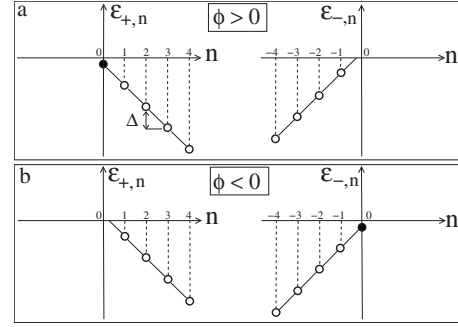

FIG. 2. Edge states in $+$ and $-$ valleys, Eq. (41), for (a) positive and (b) negative magnetic flux ϕ. We assume the zigzag confinement parameters, $l_x < 0$ and $l_z > 0$, so that the edge states exist below the Fermi level as inferred from tunneling spectroscopic measurements (Refs. 12 and 13). The external magnetic flux shifts the levels in $+$ and $-$ valleys in the opposite directions such that the zero mode $n=0$ (filled circle) occurs only in one of the valleys: $+$ one for $\phi > 0$ and $-$ one for $\phi < 0$. The valley-dependent zero mode violates Kramers' symmetry of the edge-state spectrum for arbitrary small ϕ.

$$\epsilon_{\tau,n} = \text{sgn}(l_x)\Delta \tau n, \quad \tau n l_z > 0, \quad n = \pm 1, \dots, \tag{43}$$

$$\epsilon_{\tau,0} = \text{sgn}(l_x)\Delta \tau \phi, \quad \tau \phi l_z > 0, \quad n = 0. \tag{44}$$

We see that the states with $n = \pm 1, \dots$ exhibit Kramers' symmetry under $\tau, n \to -\tau, -n$, resulting from the coupling between the valley (isospin) degree of freedom τ and the orbital quantum number n. The isospin-orbit coupling originates from the broken isospin rotation symmetry discussed in Sec. II B. The coupling constants are given by the parameters l_x and l_z of the zigzag confinement. For the hard-wall zigzag edge ($l_x=0$), we find the degenerate zero-energy state $\epsilon_{\tau,n}=0$. This is in agreement with the tight-binding calculations for zigzag graphene ribbons (e.g., Refs. 10 and 11) if their results are extrapolated to the case of the infinite width when the edges become isolated.

The zero mode, Eq. (44), stands out because it is due to the coupling between the isospin and the electromagnetically induced momentum $k_0 = (2\pi/L)\phi$. This mode breaks the Kramers' symmetry of the edge-state spectrum since it exists only for one of the isospin projections $\tau = \text{sgn}(\phi l_z)$, i.e., only in one of the valleys. In other words, there is a valley polarization effect. It is studied quantitatively in Sec. IV.

IV. VALLEY POLARIZATION

A. Spontaneous polarization

To quantitatively characterize the valley polarization effect we introduce the local isospin polarization,

$$p(\epsilon, \mathbf{r}) = -\frac{1}{\pi} \text{Im Tr } I_3 G^R(\mathbf{r}, \mathbf{r}) = -\frac{2}{\pi L} \sum_{\tau=\pm 1, n \in \mathbb{Z}} \frac{\tau}{2} \text{Im } G^s_{\tau k_n}(y, y), \tag{45}$$

where Im denotes the imaginary part, the trace Tr of Green's function (36) is taken in $\tau \otimes \sigma$ space, and the function

$G_{\tau k_n}^s(y,y)$ is given by Eq. (37) of Sec. III. As we are interested in the edge isospin polarization, the relevant contribution to Im $G_{\tau k_n}^s(y,y)$ comes from the pole in Eq. (37),

$$p_e(\epsilon,y) = -\frac{1}{L}\partial_y \sum_{\tau,n\in\mathbb{Z}} \frac{\tau}{2} e^{-2|k_n l_z|y}\Theta(k_n\tau l_z)\delta(\epsilon-\epsilon_{\tau,n}). \quad (46)$$

Note that for the zero mode the step function $\Theta(\phi\tau l_z)$ indicates the breaking of the Kramers' symmetry.

Next we calculate the zero-temperature ground-state isospin density localized at the edge as

$$i_e(y) = \int_{-\infty}^0 d\epsilon\, p_e(\epsilon,y), \quad (47)$$

$$= -\frac{1}{L}\partial_y \sum_{\tau,n\in\mathbb{Z}} \frac{\tau}{2} e^{-2|k_n l_z|y}\Theta(k_n\tau l_z)\Theta(-k_n\tau l_x), \quad (48)$$

$$= -\frac{\Theta(-l_x l_z)}{L}\partial_y \sum_{\tau,n\in\mathbb{Z}} \frac{\tau}{2} e^{-2|k_n l_z|y}\Theta(k_n\tau l_z). \quad (49)$$

To obtain the last formula we used the identity $\Theta(x)\Theta(y) = \Theta(x)\Theta(xy)$. The summations over $\tau=\pm 1$ and n can be done exactly,

$$i_e(y) = -\frac{N}{2L}\partial_y \sum_{n=-\infty}^{\infty} \operatorname{sgn}(n+\phi)e^{-|n+\phi||y|/\lambda}, \quad (50)$$

$$= -\frac{N}{2L}\partial_y\left[\operatorname{sgn}\phi\, e^{-|\phi||y|/\lambda} - \frac{2\sinh(\phi y/\lambda)}{e^{y/\lambda}-1}\right], \quad (51)$$

$$N = \Theta(-l_x l_z)\operatorname{sgn} l_z, \quad \lambda = \frac{L}{4\pi|l_z|}. \quad (52)$$

In Eq. (51) the first term, nonanalytic in ϕ, is due to the zero edge mode $n=0$. Its penetration length depends on the flux ϕ and diverges at $\phi\to 0$. The second term accounts for the rest of the edge states $n=\pm 1,\ldots$. It is an analytic function of ϕ. The edge-state penetration length is measured in units of λ given in Eq. (52), and the flux is confined to a one-period interval chosen as $-1/2 < \phi < 1/2$.

We note that depending on the boundary parameters the factor N [Eq. (52)] takes integer values 0 or ± 1. The case $N=0$ corresponds to edge states above the Fermi level $\epsilon=0$, for which the zero-temperature occupation number $\Theta(-l_x l_z) = 0$. In what follows, we focus on the opposite situation, i.e., the edge states below the Fermi level and

$$N = \operatorname{sgn} l_z = \pm 1, \quad l_x l_z < 0, \quad (53)$$

which is supported by the tunneling spectroscopy.[12,13]

Finally, we obtain the total isospin carried by the edge states as

$$I_e = L\int_0^\infty dy\, i_e(y) = \left(\frac{\operatorname{sgn}\phi}{2} - \phi\right)\operatorname{sgn} l_z. \quad (54)$$

The zero mode results in the discontinuity at $\phi=0$, due to which in the limit $\phi\to 0$ the total isospin remains finite (half integer),

$$I_e = \frac{1}{2}\operatorname{sgn}(\phi l_z), \quad \phi\to 0. \quad (55)$$

This implies that the ground state does not share the time-reversal symmetry of the original Eqs. (26)–(28) in the limit $\phi\to 0$. In this sense, the zero mode violates the time-reversal and Kramers' symmetries spontaneously, with the resulting spontaneous valley polarization.

B. Valley Hall polarization

The accumulated isospin, Eq. (54), contains a linear term $\propto\phi$. It comes from the Kramers' degenerate edge states with $n=\pm 1,\ldots$ [see Eq. (43)]. Such a property of Kramers' degenerate edge states was first noticed in the theory of quantum spin Hall systems (e.g., Refs. 33–36). The recent interest in these systems is motivated by the principal possibility to realize a time-reversal invariant integer quantum Hall state in which the spin Hall conductance is quantized. From Eq. (54) it is possible to derive the analog of the quantum spin Hall conductance. Let us calculate the isospin current as the rate of adiabatic change in the isospin: $\dot{I}_e = \dot{\phi}\partial I_e/\partial\phi = -(eE_x L/h)\partial I_e/\partial\phi$, which assumes $\phi\neq 0$ and Eq. (22). The derivative $\dot{I}_e$ gives the transverse isospin flow in response to the voltage drop $E_x L$ along the edge,

$$\dot{I}_e = G_{iH}E_x L, \quad G_{iH} = \frac{e}{h}\operatorname{sgn} l_z, \quad (56)$$

where the quantum isospin Hall conductance G_{iH} takes the universal values $\pm e/h$. As the zigzag-terminated graphene supports the edge states without any excitation gap in the bulk, conductance (56) is hardly the signature of any bulk topological order.[37] We rather interpret it as the measure of the valley polarization rate at the edges.

V. SIGNATURES OF THE VALLEY POLARIZATION IN OBSERVABLES

A. Tunneling density of states

The presence of the valley polarization can be inferred from the magnetic-field dependence of the zero mode. One of the possibilities is to measure the local tunneling conductance in the presence of a weak magnetic. At zero bias and the finite temperature T, the tunneling conductance is proportional to the tunneling density of states

$$\bar{\nu}(T,\phi) = \int_{-\infty}^{\infty} d\epsilon\left(-\frac{\partial f(\epsilon,T)}{\partial\epsilon}\right)\nu(\epsilon,\phi), \quad (57)$$

which is the convolution of the local spectral density of states, $\nu(\epsilon,\phi)$, and the energy derivative of the Fermi
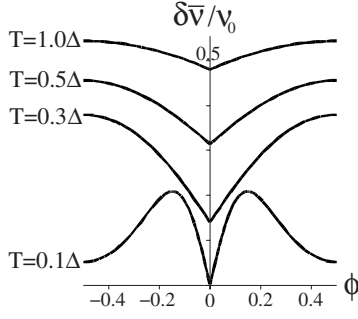
FIG. 3. Tunneling density of states, Eq. (60), vs magnetic flux (in units of ch/e) at various temperatures; $y=0.5\lambda$, $\nu_0=1/\Delta L\lambda$.

FIG. 4. (Color online) Tunneling density of states, Eq. (60), vs temperature (in units of level spacing Δ) for small ($\phi=0.01$) and large ($\phi=0.5$) flux values; $y=0.5\lambda$, $\nu_0=1/\Delta L\lambda$.

distribution function, $f(\epsilon,T)$. The local spectral density of states is obtained from the Green' function, Eq. (36), as

$$\nu(\epsilon,\mathbf{r}) = -\frac{1}{\pi}\text{Im Tr } G^R(\mathbf{r},\mathbf{r}) = -\frac{2}{\pi L}\sum_{\tau=\pm 1, n\in\mathbb{Z}}\text{Im } G^s_{\tau k_n}(y,y). \tag{58}$$

The edge-state contribution to Im $G^s_{\tau k_n}(y,y)$ comes from the pole in Eq. (37),

$$\nu_e(\epsilon,y) = -\frac{1}{L}\partial_y\sum_{\tau,n\in\mathbb{Z}}e^{-2|k_n l_z||y|}\Theta(k_n\tau l_z)\delta(\epsilon-\epsilon_{\tau,n}). \tag{59}$$

From Eqs. (57) and (59) one can obtain the edge-state contribution to the tunneling density of states as

$$\delta\bar{\nu}(T,\phi) = \frac{1}{4LT\lambda}\left\{\frac{|\phi|e^{-|\phi|y/\lambda}}{\cosh^2\left(\frac{\phi\Delta}{2T}\right)} + \sum_{n=1}^{\infty}\left[\frac{(n+\phi)e^{-\phi y/\lambda}}{\cosh^2\left(\frac{(n+\phi)\Delta}{2T}\right)}\right.\right.$$
$$\left.\left. + \frac{(n-\phi)e^{\phi y/\lambda}}{\cosh^2\left(\frac{(n-\phi)\Delta}{2T}\right)}\right]e^{-ny/\lambda}\right\}, \tag{60}$$

where Δ is the level spacing given by Eq. (42).

Figure 3 shows that the flux dependence of $\delta\bar{\nu}$ is nonanalytic, indicating the spontaneous valley polarization at $|\phi|\rightarrow 0$. The nonanalyticity is present in a wide range of temperatures. The reason is that the zero-mode term always dominates the flux dependence near $\phi=0$ because it is linear in $|\phi|$, while the rest of the sum varies as ϕ^2. As demonstrated in Fig. 4, for small $\phi=0.01$ (red curve) the zero-mode also dominates the low-temperature behavior of $\delta\bar{\nu}$, showing a $1/T$ increase when T becomes much smaller than the level spacing Δ. This feature is due to the fact that for $\phi\ll 1$ the energy of the zero mode $\propto\phi\Delta\ll\Delta$. In contrast, for the rest of the sum in Eq. (60) the relevant energy scale is set by the level spacing Δ.

B. Local electric Hall conductivity

Although the zero-mode behavior in the tunneling density of states signals the valley polarization effect, this observable does not provide the direct access to the accumulated isospin. Here we intend to show that the accumulated isospin is
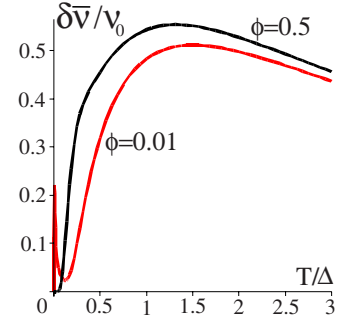
directly related to the local electric Hall conductivity.

As the first step, we use Eq. (59) to calculate the zero-temperature ground-state charge density localized at the edge,

$$\rho_e(y) = e\int_{-\infty}^{0}d\epsilon\,\nu_e(\epsilon,y), \tag{61}$$

$$= -\frac{e}{L}\partial_y\sum_{\tau,n\in\mathbb{Z}}e^{-2|k_n l_z||y|}\Theta(k_n\tau l_z)\Theta(-k_n\tau l_x), \tag{62}$$

$$= -\frac{e\Theta(-l_x l_z)}{L}\partial_y\sum_{\tau,n\in\mathbb{Z}}e^{-2|k_n l_z||y|}\Theta(k_n\tau l_z). \tag{63}$$

Again the summations over $\tau=\pm 1$ and n can be done explicitly,

$$\rho_e(y) = -\frac{e\Theta(-l_x l_z)}{L}\partial_y\sum_{n=-\infty}^{\infty}e^{-|n+\phi|y/\lambda}, \tag{64}$$

$$= -\frac{e\Theta(-l_x l_z)}{L}\partial_y\left[e^{-|\phi|y/\lambda} + \frac{2\cosh(\phi y/\lambda)}{e^{y/\lambda}-1}\right], \tag{65}$$

with $|\phi|\leq 1/2$. The ϕ dependence of Eq. (65) allows us to take the adiabatic time derivative, $\dot{\rho}_e=\dot{\phi}\partial_\phi\rho_e = -(eE_x L/h)\partial_\phi\rho_e$, and obtain the following continuity equation:

$$\dot{\rho}_e = -\partial_y j_y, \quad j_y = \mathrm{s}_{yx}E_x, \tag{66}$$

where j_y is the Hall current density induced by the transverse electric field E_x, and s_{yx} is the local *position-dependent* Hall conductivity given by

$$\mathrm{s}_{yx}(y,\phi) = \frac{4e^2 l_z}{\hbar}y\int_{y}^{\infty}dy'\,i_e(y',\phi). \tag{67}$$

It is expressed in terms of the edge isospin density given by Eq. (51). The existence of the electric current density, j_y, normal to the system's boundary is consistent with the charge conservation because at the edge $y=0$ the conductivity $\mathrm{s}_{yx}(0,\phi)$ is zero [see also Fig. 5(a)]. It also vanishes far away from the edge: $\mathrm{s}_{yx}(y\rightarrow\infty,\phi)\rightarrow 0$, so that the total edge charge is conserved: $\int_0^\infty dy\dot{\rho}_e=0$.
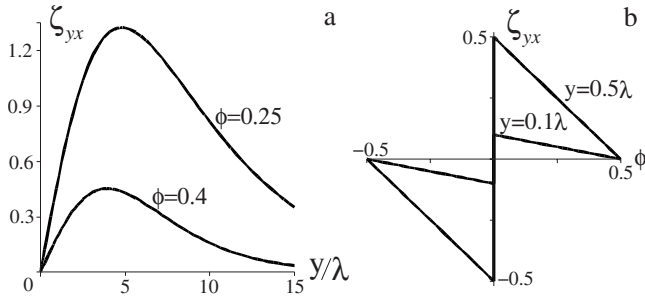
FIG. 5. Hall conductivity in units of e^2/h, Eq. (67), vs (a) position and (b) dimensionless magnetic flux.

At distances smaller than the characteristic penetration length, $y \ll \lambda$, the Hall conductivity is simply proportional to the total isospin carried by the edge states,

$$s_{yx}(y, \phi) \approx \frac{2e^2 \, \mathrm{sgn} \, l_z}{h} \frac{y}{\lambda} I_e(\phi) = \frac{2e^2}{h} \frac{y}{\lambda} \left(\frac{\mathrm{sgn} \, \phi}{2} - \phi \right),$$

(68)

showing the same nonanalytic flux dependence as $I_e(\phi)$ in Eq. (54) [see Fig. 5(b)].

In conclusion, we discuss the applicability of the results of this paper. First of all, the lattice prototype of our continuous model [Fig. 1(a)] is only one of many possible realizations of a finite-length zigzag boundary. It is nevertheless clear that for the honeycomb lattice an armchair/zigzag/armchair edge structure is rather typical [see, e.g., Fig. 6]. Independently of its concrete realization, the edge states must experience multiple intervalley backscattering from the two armchair regions, resulting in the bound states. The experimental estimate[13] of the typical length of zigzag edges is of order of 10 nm. This is large enough for the applicability of our continuum model and, on the other hand, is shorter than the typical mean free path in graphene, which is required for the ballistic quantization. For samples with longer
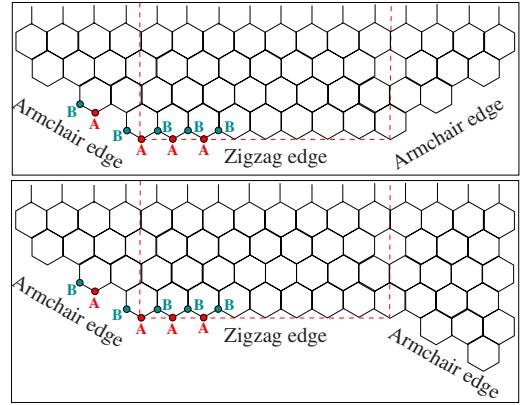

FIG. 6. (Color online) Other realizations of a finite-length zigzag boundary between two armchair edges.

edges, multiple electron scattering due to boundary and bulk disorder may come into play, as revealed by recent numerical studies.[21,22] The quantization effects studied in this paper are characteristic to isolated zigzag edges as opposed to the size quantization in zigzag graphene ribbons. Because the control over graphene edges is still a serious experimental issue, it seems easier to obtain graphene samples with isolated zigzag edges rather than to produce zigzag-terminated ribbons. In this sense, local tunneling spectroscopy is currently the most adequate tool for investigating the edge states in graphene. As for the results on local electric Hall conductivity (67), their verification may present a challenging experimental task. It should however be achievable with increasing control over the boundary effects in graphene.

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, Nature (London) **438**, 197 (2005).

[2] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, Nature (London) **438**, 201 (2005).

[3] P. R. Wallace, Phys. Rev. **71**, 622 (1947).

[4] G. W. Semenoff, Phys. Rev. Lett. **53**, 2449 (1984).

[5] V. P. Gusynin and S. G. Sharapov, Phys. Rev. Lett. **95**, 146801 (2005).

[6] E. McCann and V. I. Falko, Phys. Rev. Lett. **96**, 086805 (2006).

[7] K. Nomura and A. H. MacDonald, Phys. Rev. Lett. **96**, 256602 (2006).

[8] D. A. Abanin, K. S. Novoselov, U. Zeitler, P. A. Lee, A. K. Geim, and L. S. Levitov, Phys. Rev. Lett. **98**, 196806 (2007).

[9] For recent reviews, see, e.g., M. I. Katsnelson, K. S. Novoselov, and A. K. Geim, Nat. Phys. **2**, 620 (2006); C. W. J. Beenakker, Rev. Mod. Phys. **80**, 1337 (2008).

[10] M. Fujita, K. Wakabayashi, K. Nakada, and K. Kusakabe, J. Phys. Soc. Jpn. **65**, 1920 (1996); K. Nakada, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **54**, 17954 (1996).

[11] K. Wakabayashi and M. Sigrist, Phys. Rev. Lett. **84**, 3390 (2000).

[12] Y. Kobayashi, K. I. Fukui, T. Enoki, K. Kusakabe, and Y. Kaburagi, Phys. Rev. B **71**, 193406 (2005).

[13] Y. Niimi, T. Matsui, H. Kambara, K. Tagami, M. Tsukada, and H. Fukuyama, Phys. Rev. B **73**, 085421 (2006).

[14] S. Y. Zhou, G.-H. Gweon, J. Graf, A. V. Federov, C. D. Spataru, R. D. Diehl, Y. Kopelevich, D.-H. Lee, S. G. Louie, and A. Lanzara, Nat. Phys. **2**, 595 (2006).

[15] N. M. R. Peres, F. Guinea, and A. H. Castro Neto, Phys. Rev. B **73**, 125411 (2006).

[16] L. Brey and H. A. Fertig, Phys. Rev. B **73**, 235411 (2006).

[17] G. Tkachov, Phys. Rev. B **76**, 235409 (2007).

[18] A. R. Akhmerov and C. W. J. Beenakker, Phys. Rev. B **77**, 085423 (2008).

[19] E. V. Castro, N. M. R. Peres, J. M. B. Lopes dos Santos, A. H. Castro Neto, and F. Guinea, Phys. Rev. Lett. **100**, 026802 (2008); E. V. Castro, N. M. R. Peres, and J. M. B. Lopes dos Santos, Europhys. Lett. **84**, 17001 (2008).

[20] G. Tkachov and M. Hentschel, arXiv:0810.0632 (unpublished); arXiv:0803.0713 (unpublished).

[21] M. Evaldsson, I. V. Zozoulenko, H. Xu and T. Heinzel, Phys. Rev. B **78**, 161407(R) (2008).

[22] E. R. Mucciolo, A. H. Castro Neto, and C. H. Lewenkopf, arXiv:0806.3777 (unpublished).

[23] M. Y. Han, B. Özyilmaz, Y. Zhang, and P. Kim, Phys. Rev. Lett. **98**, 206805 (2007).

[24] Z. Chen, Y.-M. Lin, M. J. Rooks, and P. Avouris, Physica E **40**, 228 (2007).

[25] L. A. Ponomarenko, F. Schedin, M. I. Katsnelson, R. Yang, E. H. Hill, K. S. Novoselov, and A. K. Geim, Science **320**, 356 (2008).

[26] B. I. Halperin, Phys. Rev. B **25**, 2185 (1982).

[27] R. Jackiw and C. Rebbi, Phys. Rev. D **13**, 3398 (1976).

[28] A. Rycerz, J. Tworzydlo, and C. W. J. Beenakker, Nat. Phys. **3**, 172 (2007).

[29] H. B. Nielsen and M. Ninomiya, Nucl. Phys. B **185**, 20 (1981).

[30] E. McCann and V. I. Fal'ko, J. Phys.: Condens. Matter **16**, 2371 (2004).

[31] For more general boundary conditions, see Refs. 18 and 30.

[32] P. Burset, A. L. Yeyati, and A. Martin-Rodero, Phys. Rev. B **77**, 205425 (2008).

[33] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).

[34] L. Sheng, D. N. Sheng, C. S. Ting, and F. D. M. Haldane, Phys. Rev. Lett. **95**, 136602 (2005).

[35] B. A. Bernevig and S. C. Zhang, Phys. Rev. Lett. **96**, 106802 (2006).

[36] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).

[37] This issue requires a separate discussion that will be presented elsewhere.