Gapped state of a carbon monolayer in periodic magnetic and electric fields

I. Snyman*
National Institute for Theoretical Physics, Private Bag X1, 7602 Matieland, South Africa
and Department of Physics, Stellenbosch University, Private Bag X1, 7602 Matieland, South Africa
(Received 6 May 2009; revised manuscript received 29 June 2009; published 5 August 2009)

When smooth zero-on-average periodic magnetic and electric fields are applied to a carbon monolayer (graphene), a gap between the valence and conduction band is introduced. Here this gapped state is studied analytically. It is found that it does not correspond to a band insulator: a constant electric field induces a quantized Hall current even though the magnetic flux through the sample is zero. The phenomenon is of the same type as in Haldane's model for a quantum Hall effect without Landau levels, although there is the following important difference between the two models: Haldane's model requires control over external magnetic fields on length scales less than an angstrom and is, therefore, hard to realize experimentally. For the model studied here, control over external fields on length scales that are larger by several orders of magnitude is sufficient. The Hall effect is explained in terms of the topological theory of Thouless, Kohmoto, Nightingale, and den Nijs. For the system studied in this paper, a complementary explanation in terms of simple physical principles is also presented.

DOI: 10.1103/PhysRevB.80.054303 PACS number(s): 72.15.−v, 73.43.−f

I. INTRODUCTION

The filled conduction band and the empty valence band of a carbon monolayer (graphene) touch at two inequivalent points in the Brillouin zone.[1] Regions of the Brillouin zone in the vicinity of these points are called valleys. In each of the two valleys, low-energy excitations are described by a two-dimensional massless Dirac equation. As a consequence, electrons in graphene can propagate through electrostatic potential barriers. In the jargon of relativistic quantum mechanics, this is known as Klein tunneling.[2,3] When a gap is induced between the valence and conduction band, Klein tunneling is suppressed and the electronic properties of graphene change radically. Several distinct situations have been uncovered.

As one would expect, a band insulating state, in which the linear response of electrons to an external electric field is zero, can be obtained. This has been demonstrated experimentally.[4] The insulating state was realized by placing the graphene sample on an appropriate substrate that induced a potential-energy difference between the two triangular sublattices that constitute the sample's honeycomb lattice.

The band insulating state is, however, not the only gapped state that is possible. In a seminal paper by Haldane,[5] it was demonstrated that graphene's honeycomb lattice also supports a fundamentally different kind of gapped state which I will call anomalous. Even though there is a gap between the filled valence and empty conductance bands, the state shows a finite linear response to an external electric field. More precisely, there is a quantized Hall effect without Landau levels and in a zero-average magnetic field. The resulting Hall conductivity is a topological invariant.[6]

In Haldane's model, the anomalous state is realized by the combined effect of next-nearest-neighbor hopping together with a magnetic field that has the same periodicity as the honeycomb lattice and zero flux through a lattice unit cell. Because of the very short length scales (~ 1 Å) on which this magnetic field must be controlled, Haldane remarks of the model that it "is unlikely to be directly physically realizable." Recently, Kane and Mele[7] pointed out that the spin-orbit interaction can generate an anomalous gapped state in graphene without the need for any external magnetic field. In this case, since time-reversal symmetry is not broken, no Hall effect with electric charge current is produced. Rather, a spin-Hall effect occurs. However, as they also point out, the effect may be hard to observe because of the smallness of the spin-orbit interaction in graphene.

A gapped state can also occur when periodic magnetic and electric fields that are smooth on the scale of the graphene lattice constant are applied jointly to a graphene sample. In this paper I show analytically that the resulting gapped state is of the same anomalous type as in Haldane's model.

Since the fields have a periodicity $\gg 1$ Å, experimental realization is more feasible than in the case of Haldane's model. An experimental setup employing a periodic array of magnetic stripes or dots[8] is a good candidate for the future realization of the system I consider. The challenge is to reach the regime where the two magnetic length scales, namely, the periodicity of the electromagnetic fields and the magnetic length associated with the amplitude of the magnetic field, are smaller than the mean-free path of electrons in graphene. With current technology, values on the order of 1 μm have been attained for the periodicity and the mean-free path, while a value of 0.2 μm has been attained for the magnetic length (corresponding to an amplitude of 20 mT of the field). It may be necessary to refine fabrication techniques in order to decrease the periodicity somewhat for a clear effect to be observable.

Apart from being experimentally relevant, the system studied here has another interesting feature. Its Hall effect can be explained without recourse to the topological theory. Because of this, it is possible in the present instance to identify a physical mechanism responsible for the Hall effect. This hopefully takes some of the mystery out of the general but abstract topological theory.[6,9–16]

A gapped graphene state that results from smoothly varying fields was first revealed in the numerical study of Ref. 17 (see Fig. 2, last frame, of that work) and is also mentioned in Ref. 18. The effective magnetic field studied in these references is produced by lattice deformations. The single-valley physics of such a field is identical to that of a real magnetic field. The analytical results for a single valley that is presented here confirm and explain numerical results obtained in Refs. 17 and 18.

A caveat is warranted though: despite the close similarity between a graphene system with a real magnetic field and one with a deformed lattice, there is a fundamental difference. The gapped state produced by deformations is not anomalous. The reason is that for deformation-induced fields, the anomalous behavior of the two valleys cancel each other, while for real magnetic fields the behavior is the same in both valleys and there is no cancellation.

Several other publications deal with systems similar to that studied here. In Refs. 19 and 20 the two-dimensional nonrelativistic Schrödinger equation with a periodic magnetic field that averages to zero is studied. It is interesting to compare this system to the graphene system in the limit of strong magnetic fields when the magnetic length is much smaller than the periodicity of the fields. In this regime, the anomalous Hall effect[21] was found to disappear in the case of the Schrödinger equation.[20] It remains present in graphene as is shown here. Two other relevant studies[22,23] dealt with graphene in a periodic electric field alone. It was found that a strong electric field can induce additional touching points between graphene's valence and conduction bands. In the present work, only the case of weak electric fields will be considered so that additional touching points do not arise. What happens when both the electric and magnetic fields are strong, or when the electric field is strong but the magnetic field weak, is a topic for further investigation.

There are also at least two relevant works in the carbon nanotube literature. In Ref. 24 the low-energy spectrum of a carbon nanotube in a magnetic field that is transverse to the axial direction is calculated analytically. In Ref. 25 the same system is considered and shown to support a quantized Hall current in the axial direction. The nanotube system can be considered a special one-dimensional case of the general two-dimensional system considered below. In order to induce a gap in the nanotube spectrum, an electric field must be applied parallel to the magnetic field. This was not considered in the cited references. Consequently, the Hall effect found in Ref. 25 is fundamentally different from the Hall effect considered here. This will be discussed in more detail below. However, some results that hold for the system discussed in this work should also hold for the nanotube system. Where there should be agreement between results in this paper and in the nanotube studies, agreement is found.

The plan of the paper is as follows: in Sec. II the system to be studied is defined mathematically. In Sec. III its zero-energy eigenstates are found when the electric field is zero, in which case there is no gap. In Sec. IV the low-energy description of the gapped state is developed. In Sec. V the topological theory of the anomalous Hall effect is briefly reviewed and then applied to the graphene system in order to calculate the quantized Hall conductivity. In Sec. VI the anomalous Hall effect is considered again, this time without invoking the topological theory. In Sec. VII an example is presented in order to illustrate the general results of the previous sections. Conclusions are presented in Sec. VIII.

II. STATEMENT OF THE PROBLEM

The interatomic distance between carbon atoms in graphene is approximately 1.42 Å. I consider a graphene sample in the presence of static electromagnetic fields U and $\boldsymbol{A}$ that vary smoothly on this length scale. A long-wavelength description is, therefore, appropriate. This involves the Dirac Hamiltonian

$$H = v\boldsymbol{\sigma}\cdot[-i\hbar\boldsymbol{\partial}_{\boldsymbol{r}} + e\boldsymbol{A}(\boldsymbol{r})] + U(\boldsymbol{r}), \tag{2.1}$$

where $\boldsymbol{\sigma}=(\sigma_x,\sigma_y)$ are standard Pauli matrices, $\boldsymbol{r}=(x,y)$, and $\boldsymbol{\partial}_{\boldsymbol{r}}=(\partial_x,\partial_y)$. Close to the charge neutrality point, electrons in graphene obey the Dirac equation $\varepsilon\Psi=H\Psi$. In principle, two species of Dirac fermion should be distinguished because the low-energy spectrum of graphene consists of two sets of Dirac cones. However, since the fields $\boldsymbol{A}$ and U are smooth, no scattering between species is possible. Furthermore, by employing here the valley-isotropic representation,[3] the Hamiltonians for the two valleys are rendered identical not only in form but also in the actual values assumed by the parameters. An index that distinguishes between valleys is, therefore, omitted. Results obtained will apply equally to electrons in both valleys unless otherwise indicated.

I consider a magnetic field $B\hat{\boldsymbol{z}}=\boldsymbol{\partial}_{\boldsymbol{r}}\times\boldsymbol{A}$ and a scalar potential U that are periodic in space with the same periodicity and that average to zero. The periodicity of the fields define a Bravais lattice with a primitive unit cell denoted as UC and basis vectors $\boldsymbol{a}_k$, $k=1,2$, such that $B(\boldsymbol{r}+\boldsymbol{a}_k)=B(\boldsymbol{r})$ and similarly for U.

In what follows, any function that has this property will be called UC periodic. Furthermore, whenever the term "lattice" is used below, it will refer to the magnetic field lattice and not the crystal lattice of the carbon atoms in graphene. The two lattices are entirely distinct and only the former is relevant for present purposes. As a consequence of the smoothness of $\boldsymbol{A}$, UC is much larger than the unit cell of the graphene crystal lattice. Because B averages to zero, the magnetic flux through UC is zero.

The graphene sample is assumed to be much larger than UC and only bulk effects will be considered. It is, therefore, convenient to impose periodic boundary conditions

$$\Psi(\boldsymbol{r}) = \Psi(\boldsymbol{r}+\Omega\boldsymbol{a}_k), \quad k=1,2, \tag{2.2}$$

where $\Omega\gg 1$ is an integer. All functions with this property will be called sample periodic. Note that because Ω is an integer, any UC-periodic function is also sample periodic.

A reciprocal lattice is associated with the real-space lattice. The basis vectors $\boldsymbol{b}_j$, $j=1,2$, of the reciprocal lattice are defined through the equations $\boldsymbol{b}_j\cdot\boldsymbol{a}_k=2\pi\delta_{j,k}$. The Brillouin zone is a primitive unit cell of the reciprocal lattice and will be denoted as BZ. It is convenient to define the Fourier components

$$B_{mn} = l^{-2} \int_{UC} d^2 r e^{-i\boldsymbol{k}_{mn}\cdot\boldsymbol{r}} B(\boldsymbol{r}), \tag{2.3}$$

of the magnetic field. Here $\boldsymbol{k}_{mn}=m\boldsymbol{b}_1+n\boldsymbol{b}_2$ with m and n integers and $l^2=|\boldsymbol{a}_1\times\boldsymbol{a}_2|$ is the area of UC. Note that $B_{00}=0$ due to the zero-flux condition imposed on B and $B_{m,n}=B^*_{-m,-n}$ because $B(\boldsymbol{r})$ is real.

III. ZERO-ENERGY MODES

The task is now to analyze the low-energy (i.e., small $|\varepsilon|$) spectrum of H. For $U=0$, zero-energy eigenstates $\Psi=(\phi_+,\phi_-)$ of H solve

$$[\partial_x \pm i\partial_y + i(\alpha_x \pm i\alpha_y)]\phi_\pm = 0, \tag{3.1}$$

with $\boldsymbol{\alpha}=e\boldsymbol{A}/\hbar$. Following Jackiw,[26] I work in the Coulomb gauge where $\boldsymbol{\partial_r}\cdot\boldsymbol{A}=0$. Then Eq. (3.1), subject to the sample-periodic boundary conditions of Eq. (2.2), is solved by

$$\phi_\pm(\boldsymbol{r}) = c_\pm \exp[\pm F(\boldsymbol{r})], \tag{3.2}$$

where $c_\pm$ are arbitrary constants. The real function F is UC periodic and satisfies

$$\partial_{\boldsymbol{r}}^2 F = eB/\hbar. \tag{3.3}$$

In terms of the Fourier components B_{mn} of the magnetic field, F is given by

$$F(\boldsymbol{r}) = -\frac{e}{\hbar}\sum_{mn\neq 0}\frac{B_{m,n}}{|\boldsymbol{k}_{mn}|^2}e^{i\boldsymbol{k}_{mn}\cdot\boldsymbol{r}}. \tag{3.4}$$

One demonstrates uniqueness of the solutions in Eq. (3.2) as follows. Consider the equation for ϕ_-. Suppose that $\tilde{\phi}_-(\boldsymbol{r})$ solves Eq. (3.1) and is sample periodic as required by the boundary conditions. Define a function $g(\boldsymbol{r})=\tilde{\phi}(\boldsymbol{r})e^{F(\boldsymbol{r})}$. Since F is UC periodic, g is sample periodic. The function g satisfies the Cauchy-Riemann equation $(\partial_x - i\partial_y)g=0$. Thus g is an analytical function of the complex variable $z=x+iy$. As a consequence g is either constant or unbounded in some directions at large $\boldsymbol{r}$. Unboundedness is incompatible with g being sample periodic and, therefore, g must be constant so that $\tilde{\phi}_-(\boldsymbol{r})=ge^{-F(\boldsymbol{r})}$ is proportional to $\phi_-(\boldsymbol{r})=c_-e^{-F(\boldsymbol{r})}$. The uniqueness of ϕ_+ is demonstrated in the same manner.

The Dirac equation $H\Psi=E\Psi$ with $U=0$, therefore, has exactly two zero-energy solutions which I will take to be

$$\Psi_{+,0}(\boldsymbol{r}) = \frac{N_+}{\Omega}\begin{pmatrix} e^{F(\boldsymbol{r})} \\ 0 \end{pmatrix}, \quad \Psi_{-,0}(\boldsymbol{r}) = \frac{N_-}{\Omega}\begin{pmatrix} 0 \\ e^{-F(\boldsymbol{r})} \end{pmatrix}, \tag{3.5}$$

with $N_\pm=[\int_{UC} d^2r \exp\pm 2F(\boldsymbol{r})]^{-1/2}$ ensuring normalization of $\Psi_{\pm,0}$ to unity over the sample. The two eigenstates have opposite sublattice polarization: in valley K, $\Psi_{+,0}$ is A sublattice polarized while $\Psi_{-,0}$ is B sublattice polarized. Since the valley isotropic representation is employed, the roles of the sublattices are reversed in the K' valley. Hence, the state $\Psi_{+,0}$ in the K' valley is B polarized while the K' state $\Psi_{-,0}$ is A polarized. The result is consistent with the Atiyah Singer index theorem. It states that in a given valley the difference between the number of zero-energy eigenstates that are A polarized and B polarized equals the total flux through the sample.[27] Also note that because F is UC periodic, the same is true for $\Psi_{\pm,0}$ even though only sample periodicity is imposed on $\Psi_{\pm,0}$ by the boundary condition [Eq. (2.2)].

The solutions in Eq. (3.5) were found without invoking Bloch's theorem. I now briefly reexamine them from this point of view. Since the Hamiltonian H is periodic, an eigenbasis exists in which all eigenfunctions of H are of the form

$$\Psi_{n,\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\Omega}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\psi_{n,\boldsymbol{k}}(\boldsymbol{r}), \tag{3.6}$$

with $\boldsymbol{k}\in BZ$ and n a discrete index. The functions $\psi_{n,\boldsymbol{k}}(\boldsymbol{r})$ are UC periodic and will be referred to as Bloch states. They satisfy $H(\boldsymbol{k})\psi_{n,\boldsymbol{k}}=\varepsilon_{n,\boldsymbol{k}}\psi_{n,\boldsymbol{k}}$, where $\varepsilon_{n,\boldsymbol{k}}$ are the energies of the eigenstates. The effective Hamiltonian $H(\boldsymbol{k})$ is given by

$$H(\boldsymbol{k}) = v\boldsymbol{\sigma}\cdot(-i\hbar\partial_{\boldsymbol{r}} + \hbar\boldsymbol{k} + e\boldsymbol{A}) + U = H + \hbar v\boldsymbol{\sigma}\cdot\boldsymbol{k}. \tag{3.7}$$

Normalizing $\psi_{n,\boldsymbol{k}}(\boldsymbol{r})$ to unity over UC ensures that $\Psi_{n,\boldsymbol{k}}$ is normalized to unity over the whole sample. The zero eigenstates of Eq. (3.5) are UC periodic and, therefore, correspond to $\boldsymbol{k}=0$ solutions in this labeling scheme. The corresponding normalized zero-energy eigenstates of $H(\boldsymbol{k}=0)$ (with $U=0$) are $\psi_{\pm,0}=\Omega\Psi_{\pm,0}$.

IV. NEAR-ZERO EIGENSTATES AND A NONZERO BUT SMALL SCALAR POTENTIAL

The near-zero-energy eigenstates in the vicinity of $\boldsymbol{k}=0$ can be studied by treating the term $\hbar v\boldsymbol{\sigma}\cdot\boldsymbol{k}$ in $H(\boldsymbol{k})$ perturbatively. The effect of a weak periodic scalar potential U can be treated perturbatively at the same time.

To leading order in $\boldsymbol{k}$ and U, the problem is solved by Bloch states of the form

$$u_{\eta,\boldsymbol{k}}(\boldsymbol{r}) = \chi^+_{\eta,\boldsymbol{k}}\psi_{+,0}(\boldsymbol{r}) + \chi^-_{\eta,\boldsymbol{k}}\psi_{-,0}(\boldsymbol{r}), \tag{4.1}$$

where $\chi_{\eta,\boldsymbol{k}}=(\chi^+_{\eta,\boldsymbol{k}},\chi^-_{\eta,\boldsymbol{k}})$ satisfies $h(\boldsymbol{k})\chi_{\eta,\boldsymbol{k}}=\varepsilon_{\eta,\boldsymbol{k}}\chi_{\eta,\boldsymbol{k}}$. Here $\eta=\pm$ labels the bands ($-$ for the valence band and $+$ for the conduction band), $\varepsilon_{\eta,\boldsymbol{k}}$ is the energy of the state with label $\eta,\boldsymbol{k}$, and $h(\boldsymbol{k})$ is the Hamiltonian $H(\boldsymbol{k})$ projected onto the zero eigenspace. It has the form of a 2×2 Dirac Hamiltonian in k space

$$h(\boldsymbol{k})_{\eta\eta'} = \int_{UC} d^2 r \psi_{\eta,0}(\boldsymbol{r})^\dagger[\hbar v\boldsymbol{\sigma}\cdot\boldsymbol{k} + U(\boldsymbol{r})]\psi_{\eta',0}(\boldsymbol{r}),$$

$$\Rightarrow h(\boldsymbol{k}) = \hbar\tilde{v}\boldsymbol{\sigma}\cdot\boldsymbol{k} + \mu\sigma_z, \tag{4.2}$$

where a term proportional to identity is omitted because it simply leads to a redefinition of the zero energy. Here $\tilde{v}=l^2N_+N_-v$ is the renormalized Fermi velocity and l^2 is the area of UC. The effective mass μ is given by

$$\mu = \frac{1}{2}\int_{UC} d^2 r U(\boldsymbol{r})(N_+^2 e^{2F(\boldsymbol{r})} - N_-^2 e^{-2F(\boldsymbol{r})}). \tag{4.3}$$

The signs of $U(\boldsymbol{r})$ and $F(\boldsymbol{r})$ are said to be correlated when there is a large overlap between regions of positive U and positive F and anticorrelated when there is a large overlap

between regions of positive U and negative F. When the signs of U and F are either correlated or anticorrelated, the two terms in the integrand tend not to cancel and μ is nonzero.

The leading-order perturbative expressions for the Bloch state $u_{\pm,\boldsymbol{k}}(\boldsymbol{r})$ in the vicinity of $\boldsymbol{k}=0$ is

$$u_{\pm,\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{2}}\left[\sqrt{1 \pm \frac{\mu}{|\varepsilon_{\pm,\mathbf{k}}|}}\psi_+(\boldsymbol{r}) + \sqrt{1 \mp \frac{\mu}{|\varepsilon_{\pm,\mathbf{k}}|}}e^{i\theta}\psi_-(\boldsymbol{r})\right], \tag{4.4}$$

with $\theta=\arg(k_x+ik_y)$ and $\varepsilon_{\pm,\boldsymbol{k}}=\pm\sqrt{(\hbar v|\boldsymbol{k}|)^2+\mu^2}$ the energy of the state.

It is worth emphasizing here that the same effective Hamiltonian $h(\boldsymbol{k})$ and, therefore, the same mass term is induced in both valleys. This is to be contrasted with the mass term that arises when a staggered on-site potential is applied to otherwise clean graphene by, for instance, placing the sample on an appropriate substrate.[4] If the staggered potential takes on the value m on the A sublattice and $-m$ on the B sublattice, then a mass term $m\sigma_z$ is induced in valley K while a mass term $-m\sigma_z$ is induced in valley K'.[28]

V. ANOMALOUS HALL EFFECT

In the previous section, I demonstrated that there is a gap in the spectrum and that a low-energy description in terms of massive Dirac fermions is possible. The gap has several important consequences. First, low-energy electrons can be localized in selected regions of a graphene sample by inducing a gap in regions where electrons are to be excluded. The gap can be controlled by electromagnetic fields that are smooth on the scale of the graphene lattice constant and there is no need to break the sublattice inversion symmetry.[4] Second, since one has control over the sign of the mass, it is possible to produce a sample in which the sign of the mass differs in different spatial regions. At the interface between two such regions, chiral edge states are expected to appear.[7] Third, as is the case in Ref. 5, an electric field $\boldsymbol{E}=E\hat{y}$ applied to the bulk of the sample (the y direction is chosen arbitrarily) will produce a quantized current $j_x=\sigma_{xy}E$ in the direction perpendicular to it. This is remarkable since the magnetic field through the sample averages to zero.

Here, I briefly review the general theory for this phenomenon which is known as the anomalous Hall effect.[9–16] It states that the Hall conductivity σ_{xy} may be nonzero, even in the absence of a net magnetic flux through the sample, provided time-reversal invariance is still broken. Furthermore it is quantized in units of e^2/h and can be written in terms of a set of integer topological invariants as $\sigma_{xy}=(e^2/h)\Sigma_n \mathrm{Ch}_n$.[10–16] Here n is the discrete index that labels the energy bands and the sum ranges over all occupied bands. (The theory assumes that the Fermi energy is in a gap between bands so that there are no partially filled bands.) The invariant Ch_n is known as a Chern number and can be calculated if the Bloch state for band n is known throughout the whole Brillouin zone. Since only the Bloch states in the vicinity of $\boldsymbol{k}=0$ and only for the two bands that touch at $\varepsilon=0$ were found in Sec. IV, direct calculation of the Ch_n is not feasible here.

Fortunately, there is a way to work around this problem. The Chern numbers can only change when bands touch.[12] Before and after a touching of bands, they differ by an integer. Only the Chern numbers of bands involved in the touching change. Furthermore, there is a sum rule relating the Chern numbers before and after a band touching. Suppose that, by varying a parameter in the Hamiltonian, one induces bands n and $n+1$ to touch before moving apart again. Let $\mathrm{Ch}_m^<(\mathrm{Ch}_m^>)$ be the Chern numbers before (after) the bands touched. Then $\mathrm{Ch}_n^<+\mathrm{Ch}_{n+1}^<=\mathrm{Ch}_n^>+\mathrm{Ch}_{n+1}^>$, i.e., the sum of the two Chern numbers is preserved. If bands n and $n+1$ are both completely filled, σ_{xy} stays constant although Ch_n and Ch_{n+1} change. Thus σ_{xy} only changes if the Fermi energy lies in the gap between bands n and $n+1$.

In order to calculate the change in Chern numbers, one only needs information about the bands that touch and only in regions of the Brillouin zone that are close to the points where the bands touch.[15,16] This information may be obtained by means of a perturbative expansion in the wave vector $\boldsymbol{k}$ and the external parameter that controls the closing of the gap. (The general procedure is the same as was used in Sec. IV of this text.) If the first-order coefficients do not vanish, the touching point is characterized by a Dirac fermion. Its effective 2×2 Hamiltonian is brought into a standard form (essentially the valley-isotropic representation) by means of appropriate 2×2 unitary transforms. It turns out that the change in Chern numbers is completely determined by the sign of the mass term in this representation.[15] If the bands touch at several points in $\boldsymbol{k}$ space simultaneously, then each touching point gives an independent contribution.

For the graphene system I am considering, this leads to a result

$$\mathrm{Ch}_\pm^> - \mathrm{Ch}_\pm^< = \pm\{\mathrm{sign}[\mu^>] - \mathrm{sign}[\mu^<]\}. \tag{5.1}$$

Here $+(-)$ refers to the conduction (valence) band and $\mu^<(\mu^>)$ is the mass induced by the scalar potential before (after) the bands touched. There are two contributions to the result, one from each touching point. Since the same mass term appears at both touching points (in the valley-isotropic representation), the contributions from the two touching points are the same. If the sign of the mass were opposite in the two valleys, as is the case for a substrate-induced mass, the two contributions would have canceled.

There are two ways to change the sign of the mass term [cf. Equation (4.3)]. One may either change the sign of the external potential U or one may change the sign of the magnetic field and hence of F. According to the theory, both methods give the same change in the Hall conductivity, $\Delta\sigma_{xy}=\Delta\sigma_{xy}(U\to-U)=\Delta\sigma_{xy}(B\to-B)$. The transformation $B\to-B$ is equivalent to time reversal, which changes the sign of the Hall conductivity: $\sigma_{xy}=-\sigma_{xy}$. Therefore, $\Delta\sigma_{xy}=2\sigma_{xy}$. Thus, knowing the change in σ_{xy} when the mass changes sign, one also knows the actual value of σ_{xy}.

If $\sigma_{xy}^\pm$ is the Hall conductivity of a sample in which the mass term has sign $\pm$ then $\sigma_{xy}^+-\sigma_{xy}^-=-2e^2/h$. Therefore, according to the above argument

$$\sigma_{xy}^\pm = \mp\frac{e^2}{h}. \tag{5.2}$$

VI. ORIGIN OF NONZERO HALL CONDUCTIVITY

Suppose that the topological theory discussed in the previous section was not known. Would it have been possible to show that the graphene system studied here displays a Hall effect without at the same time essentially deriving the topological theory? As I show in this section, the answer is "yes." In the process, the underlying physical mechanism that here produces the anomalous Hall effect is identified.

The starting point of the argument is to consider the two zero-energy solutions $\Psi_{\pm,0} \propto e^{\pm F}$ that are obtained when $U=0$. A key feature of these wave functions is that the sign of F is anticorrelated with that of B, i.e., F is negative on average where the magnetic field points in the positive z direction and positive on average where the magnetic field points in the negative z direction. To prove this assertion I show that the number $P=(e/\hbar)\int_{UC} d^2 r B(\boldsymbol{r}) F(\boldsymbol{r})$ is negative. The proof is elementary. From Eq. (3.3) follows that $P=\int_{UC} d^2 r B(\boldsymbol{r}) \partial_{\boldsymbol{r}}^2 B(\boldsymbol{r})$. Using integration by parts and invoking Gauss's theorem, I find

$$P = \underbrace{\oint_{\partial UC} d\boldsymbol{r}_\perp \cdot B \boldsymbol{\partial}_{\boldsymbol{r}} B}_{=0} - \int_{UC} d^2 r |\boldsymbol{\partial}_{\boldsymbol{r}} B|^2 < 0. \tag{6.1}$$

The integration in the first term is around the border of UC and $d\boldsymbol{r}_\perp$ is normal to the border (in the $x-y$ plane) pointing outwards and yields zero because B is UC periodic. The result [Eq. (6.1)] implies that a particle in the state $\Psi_{-,0}(\Psi_{+,0})$ is largely confined to regions where $\boldsymbol{B}$ points in the positive (negative) z direction. In Sec. III it was noted that the eigenstates $\Psi_{\pm,0}$ have opposite sublattice polarization. The correlation between B and F then implies that sublattice polarization is correlated with the sign of the magnetic field. For instance, in valley K, regions where the magnetic field points in the negative z direction are A sublattice polarized while regions where the magnetic field points in the positive z direction are B sublattice polarized. This is consistent with the numerical results presented in Fig. 3 of Ref. 18. There it was pointed out that the state $\Psi_{-,0}$ closely resembles a state in the zero Landau level of a graphene sample with a constant magnetic field that points in the positive z direction. This would have explained the existence of a Hall effect were it not for the state $\Psi_{+,0}$ that also belongs to the zero eigenspace. This state is similar to a zero-Landau-level state in a magnetic field pointing in the negative z direction and is responsible for a Hall effect that cancels that of $\Psi_{-,0}$.

In order to have a Hall effect, a perturbation that lifts the degeneracy between $\Psi_{+,0}$ and $\Psi_{-,0}$ is needed. This is the role of the electrostatic potential U. If the sign of U is anticorrelated with that of B, it will raise the energy of the state $\Psi_{+,0}$ while lowering the energy of the state $\Psi_{-,0}$. This leads to a gap with associated positive mass. A scalar potential U whose spatial profile is correlated with that of B will do the opposite and leads to gap associated with a negative mass.

Turning to the expression [Eq. (4.4)] for the $U \neq 0$ valence-band Bloch states $u_{-,\boldsymbol{k}}$, one sees that the dominant term in the linear combination is proportional to $e^{-\mathrm{sign}(\mu)F}$. Since the signs of F and B are anticorrelated, valence-band states are confined to regions of positive B when U is anticorrelated with B and confined to regions of negative B when U is correlated with B. This is consistent with the response of $\Psi_{\pm,0}$ to U discussed above.

One thus arrives at the following fundamental result: when the mass is positive (negative), there is a finite positive (negative) magnetic flux through the region of space occupied by valence-band electrons. The valence band will then experience a "usual" Hall effect. An electric field $\boldsymbol{E}=E\hat{\boldsymbol{y}}$ will, through the Lorentz force, lead to a current in the direction $-\mathrm{sign}(\mu)\hat{\boldsymbol{x}}$. Electrons in the conduction band would have seen an opposite magnetic field and produced an opposite Hall effect. However, the Fermi energy is in the gap between the two bands so that no conduction-band states are occupied. The sign obtained here for the Hall current agrees with the result derived from the topological theory [cf. Eq. (5.2)].

So far I have said nothing about the quantization of the Hall current. Indeed, the above argument does not prove that the anomalous Hall conductivity precisely equals $\mp e^2/h$. It only indicates that σ_{xy} is finite and predicts its sign. However, the argument does relate the anomalous Hall effect in graphene with periodic electromagnetic fields to the usual integer Hall effect in a constant magnetic field. It thus shows that the quantization of σ_{xy} has the same origin in both cases.

Finally, a word on the nanotube system: the Hall effect found there[25] is not produced by states localized where the magnetic field has a well-defined direction, although such states are of course also present in that system. Rather, it is produced by higher-energy states that are localized where the magnetic field changes sign. Thus, a different mechanism is responsible for the effect reported in Ref. 25. Other features also distinguish it from the Hall effect considered in the present work: it is not accompanied by a gap between the conduction and valence bands. It is associated with two sets of states on opposite sides of the circumference of the nanotube that carry current along the length of the nanotube but in opposite directions. The quantization is approximate. This is to be contrasted with the following properties of Hall effect discussed in the present work. It appears when the Fermi energy is in the gap between the conduction and valence bands. All the current travels in one direction throughout the whole sample. The Hall quantization of the Hall conductivity is sharp.

The conclusion is that the Hall effect discussed in this paper has not yet been found in the context of carbon nanotubes. But is it there? The topological theory cannot answer this question. It only works for systems that are much larger than the periodicity of the magnetic field, whereas the length of the nanotube system in the circumferential direction is the same as the periodicity of the magnetic field. It is not clear if the intuitive argument presented in this section can be applied either. It also involves viewing the system with a spatial resolution that is larger than the periodicity of the magnetic field. The question therefore remains open.

VII. EXAMPLES

Example 1: as a simple example that illustrates the general results derived in this text, consider the magnetic field $B(\boldsymbol{r})=B_0(\sin 2\pi x/\lambda + \sin 2\pi y/\lambda)$ with periodicity λ in both

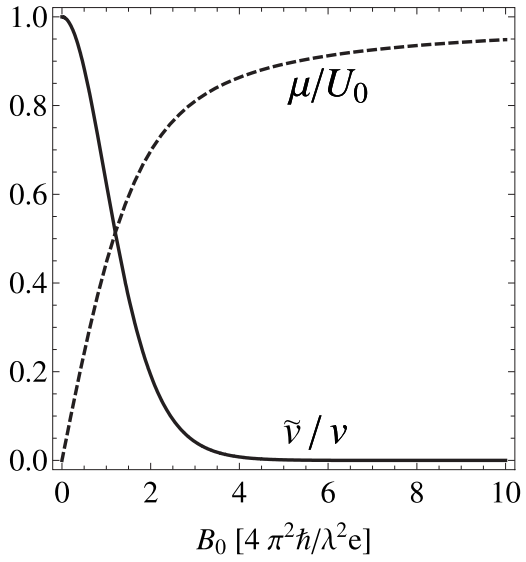

FIG. 1. The renormalized Fermi velocity $\tilde{v}$ (solid line) according to Eq. (7.2) and the mass μ (dashed line) according to Eq. (7.3) for the model with magnetic field $B(\boldsymbol{r})=B_0(\sin 2\pi x/\lambda+\sin 2\pi y/\lambda)$ perpendicular to the graphene plane.

the x and y directions. It is convenient to define the magnetic length $l_m=\sqrt{\hbar/e|B_0|}$ and a dimensionless constant $\beta=e\lambda^2|B_0|/\hbar=(\lambda/l_m)^2$. According to Eq. (3.4), F is given by

$$F(\boldsymbol{r})=-\frac{\beta}{4\pi^2}\frac{B(\boldsymbol{r})}{|B_0|}. \tag{7.1}$$

This gives the clearest possible illustration that $\Psi_{\pm,0}$ are localized to regions of positive or negative B. The degree of localization is controlled by the ratio λ/l_m.

The renormalized Fermi velocity can be calculated explicitly as

$$\tilde{v}=\frac{v}{I_0(\beta/4\pi^2)^2}, \tag{7.2}$$

where I_0 is the zeroth modified Bessel function of the first kind. When $\beta\gg 1$, or equivalently $l_m\ll\lambda$, the asymptotic form $\tilde{v}\simeq\beta e^{-\beta/2\pi^2}v/2\pi$ is valid. In this limit the Fermi velocity becomes very small and bands become nearly dispersionless in the region of $\boldsymbol{k}=0$. The result is very similar to the nanotube result.[24] This is due to the fact that the spatial profile of the magnetic field in the nanotube case is also sinusoidal. In Fig. 1 the solid line shows the renormalized Fermi velocity plotted as a function of the magnetic field strength B_0.

As was pointed out in Ref. 17 and 18, the flat bands are reminiscent of Landau levels. It should however be pointed out that flat bands do not guarantee a Hall effect: the nonrelativistic Schrödinger equation in the presence of the magnetic field of this example was studied numerically in Ref. 20. Flat bands were found in the limit $\lambda/l_m\gg 1$. For that system, the Chern numbers of flat bands were found to be zero leading to a zero Hall conductivity.

Neither are the flat bands a necessary condition for the Hall effect to occur. The effect also occurs for $l_m>\lambda$ when the renormalized Fermi velocity is on the same order of magnitude as the unrenormalized Fermi velocity. Counterintuitively, it is not the flat bands but the Dirac cones at the touching points between bands that lead to the Hall effect.

Returning to the analysis of the example, I now introduce a scalar potential $U(\boldsymbol{r})=-U_0B(\boldsymbol{r})/B_0$ with the same spatial profile as the magnetic field. The induced mass is

$$\mu=\mathrm{sign}(B_0)U_0\frac{I_1(\beta/4\pi^2)}{I_0(\beta/4\pi^2)}\simeq\mathrm{sign}(B_0)U_0, \tag{7.3}$$

where I_1 is the first modified Bessel function of the first kind. The asymptotic form is valid in the same limit as before, namely, when $l_m\ll\lambda$. This result is straightforward to interpret. In the strong magnetic field limit, the eigenstates become exponentially well confined to the maxima and minima of the magnetic field. These regions are also the maxima and minima of U so that the effective potential seen by the eigenstates has magnitude $|U_0|$. The behavior of the mass μ as a function of the magnetic field strength B_0 is shown (dashed line) in Fig. 1.

Example 2: finally, I consider an example in which the magnetic and electric fields are constant in one direction and periodic in the other. The full spectrum can be obtained analytically. The approximate formulas for the low-energy spectrum in the previous sections can, therefore, be tested against exact results.

The magnetic field and electrostatic potential are taken as

$$\boldsymbol{B}=\hat{z}B_0\lambda\sum_{n\in Z}\left\{\delta(y-n\lambda)-\delta\left[y-\left(n+\frac{1}{2}\right)\lambda\right]\right\},$$

$$U=-U_0\lambda\sum_{n\in Z}\left\{\delta(y-n\lambda)-\delta\left[y-\left(n+\frac{1}{2}\right)\lambda\right]\right\}. \tag{7.4}$$

The function F is given by

$$F=\frac{\lambda^2}{2l_m^2}\left(\frac{1}{4}-\left|\frac{y_{\mathrm{mod}\,\lambda}}{\lambda}-\frac{1}{2}\right|\right). \tag{7.5}$$

According to the results presented in Sec. IV, the renormalized Fermi velocity and mass are

$$\tilde{v}=\frac{\beta}{4\sinh(\beta/4)}v,\quad \mu=\mathrm{sign}(B_0)\frac{\beta U_0}{2}, \tag{7.6}$$

with β defined as in example 1. Again the renormalized Fermi velocity is an exponentially decreasing function of β and hence of $|B_0|$. It holds, in general, that the renormalized Fermi velocity is less than the bare Fermi velocity and that for large magnetic fields the reduction is exponential in $|B_0|$. As in a nanotube,[24,25] the effect can be measured in a tunneling experiment as an increase $\sim(v/\tilde{v})^2$ in the density of states at the Dirac point. For this experiment, the electric field must be zero so that a gap does not open. Note, furthermore, that since the electron density fluctuates spatially, the local density of states, measured by a scanning tunneling microscope, will also show oscillations. The increase in the density of states is revealed in the average of the local den-

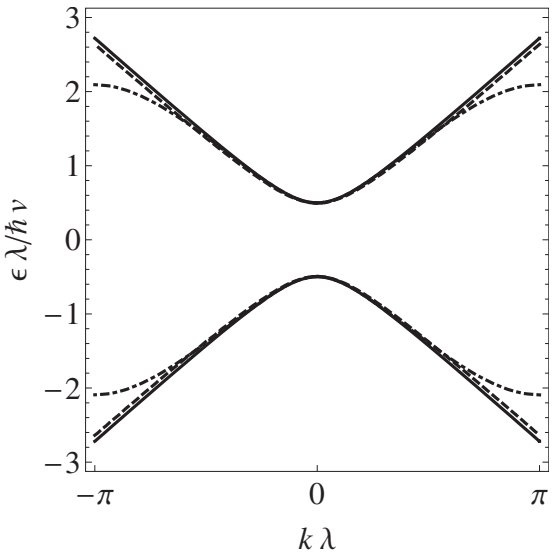

FIG. 2. The low-energy dispersion of the Dirac equation with the magnetic field and electrostatic potential of Eq. (7.4). The dashed line shows the exact result for the wave vector $\boldsymbol{k}$ chosen along the x direction. The dot-dashed line shows the exact result for the wave vector $\boldsymbol{k}$ along the y direction. The solid line shows the Dirac dispersion $\varepsilon = \pm\sqrt{(\hbar\tilde{v}|\boldsymbol{k}|)^2+\mu^2}$ with μ and $\tilde{v}$ as in Eq. (7.6). A magnetic length $l_m=\lambda/2$ and a potential strength $U_0=\hbar v/4\lambda$ was used.

sity of states over one unit cell of the magnetic field lattice.

The full spectrum of the model can be obtained from the transcendental equation

$$\cos(k_y\lambda) = \cos\left(\frac{\lambda p_+}{2}\right)\cos\left(\frac{\lambda p_-}{2}\right) - \frac{\sin\left(\frac{\lambda p_+}{2}\right)}{p_+}\frac{\sin\left(\frac{\lambda p_-}{2}\right)}{p_-}\left[\left(\frac{\varepsilon}{\hbar v}\right)^2 - k_+k_- \cos\left(\frac{2U_0\lambda}{\hbar v}\right)\right], \tag{7.7}$$

where $p_\pm=\sqrt{(\varepsilon/\hbar v)^2-k_\pm^2}$, $k_\pm=k_x\pm\lambda/2l_m^2$, ε is the energy, and the wave vector is $\boldsymbol{k}=k_x\hat{\boldsymbol{x}}+k_y\hat{\boldsymbol{y}}$. In Fig. 2 the exact dispersion of the two bands closest to $\varepsilon=0$ is plotted in the x and y directions. On the same plot is also shown the hyperbolic dispersion relation $\varepsilon=\pm\sqrt{(\hbar\tilde{v}|\boldsymbol{k}|)^2+\mu^2}$ with μ and $\tilde{v}$ as calculated in Eq. (7.6). It is seen that the low-energy dispersion relation is reproduced very well by the approximate formulas.

VIII. CONCLUSION

In this paper I showed analytically that when a graphene sample is exposed to a periodic magnetic field that (a) is smooth on the scale of the graphene lattice and (b) is zero on average, the valence and conduction bands still touch at two Dirac points but the Fermi velocity is reduced. The reduction is the same in all directions and is exponential in the magnetic field strength. This result generalizes the previously obtained numerical[17,18,25] and analytical[24] results for magnetic fields that are periodic in one direction and constant in the other. For simple models that introduce no other length scales beyond the magnetic length l_m and the periodicity λ of the magnetic field, the renormalized Fermi velocity $\tilde{v}$ behaves asymptotically like $\tilde{v}/v\propto\exp(c\lambda^2/l_m^2)$ with c of order unity. The zero eigenspace is twofold degenerate. A basis exists in which one eigenstate confines particles to regions where magnetic field points in the positive z direction and to sublattice $A(B)$ in valley K (K'). The other eigenstate does the opposite, i.e., confines particles to regions where the magnetic field points in the negative z direction and to sublattice $B(A)$ in valley K (K'). An electrostatic potential that is either correlated or anticorrelated with the magnetic field induces a gap between the valence and conduction bands. This analytical result is consistent with previously obtained numerical results.[17,18] I showed that the gapped state supports a quantized Hall effect with $\sigma_{xy}=\pm e^2/h$. The positive sign refers to the correlated case and the negative sign refers to the anticorrelated case. This is an instance of the known phenomenon of a Hall effect in zero overall flux.[5] It can be explained in terms of the topological properties of Bloch states of the occupied bands.[9–16] In the present system, I showed that the effect is also simply related to the usual integer Hall effect in a nonzero-average magnetic field: in the presence of the electrostatic potential, the spatial region in which the filled valence-band states are localized, is permeated by a finite total magnetic flux. As a result, the valence-band electrons see a magnetic field with a nonzero average even though the magnetic field averages to zero over the sample as a whole.

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation (NRF) of South Africa.

*isnyman@sun.ac.za

[1] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
[2] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim, Nat. Phys. **2**, 620 (2006).
[3] C. W. J. Beenakker, Rev. Mod. Phys. **80**, 1337 (2008).
[4] S. Y. Zhou, G.-H. Gweon, A. V. Fedorov, P. N. First, W. A. de Heer, D.-H. Lee, F. Guinea, A. H. Castro Neto, and A. Lanzara, Nature Mater. **6**, 770 (2007).
[5] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
[6] F. D. M. Haldane, Phys. Rev. Lett. **93**, 206602 (2004).
[7] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[8] P. D. Ye, D. Weiss, R. R. Gerhardts, M. Seeger, K. von Klitzing, K. Eberl, and H. Nickel, Phys. Rev. Lett. **74**, 3013 (1995).
[9] R. Karplus and J. M. Luttinger, Phys. Rev. **95**, 1154 (1954).
[10] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[11] Q. Niu, D. J. Thouless, and Y.-S. Wu, Phys. Rev. B **31**, 3372

(1985).
[12] J. E. Avron, R. Seiler, and B. Simon, Phys. Rev. Lett. **51**, 51 (1983).
[13] B. Simon, Phys. Rev. Lett. **51**, 2167 (1983).
[14] M. Kohmoto, Ann. Phys. (N.Y.) **160**, 343 (1985).
[15] M. Oshikawa, Phys. Rev. B **50**, 17357 (1994).
[16] M. Onoda and N. Nagaosa, J. Phys. Soc. Jpn. **71**, 19 (2002).
[17] F. Guinea, M. I. Katsnelson, and M. A. H. Vozmediano, Phys. Rev. B **77**, 075422 (2008).
[18] T. O. Wehling, A. V. Balatsky, A. M. Tsvelik, M. I. Katsnelson, and A. I. Lichtenstein, EPL **84**, 17003 (2008).
[19] A. Krakovsky, Phys. Rev. B **53**, 8469 (1996).
[20] M. Taillefumier, V. K. Dugaev, B. Canals, C. Lacroix, and P. Bruno, Phys. Rev. B **78**, 155330 (2008).
[21] Unfortunately, "anomalous Hall effect" is also the standard terminology for another phenomenon in graphene, namely, the unusual quantization of Landau levels. This is not what the term refers to in this text. Rather, it refers to a Hall effect in a zero-average magnetic field.
[22] C.-H. Park, Y.-W. Son, L. Yang, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **103**, 046808 (2009).
[23] L. Brey and H. A. Fertig, Phys. Rev. Lett. **103**, 046809 (2009).
[24] H.-W. Lee and D. S. Novikov, Phys. Rev. B **68**, 155402 (2003).
[25] S. Bellucci, J. González, F. Guinea, P. Onorato, and E. Perfetto, J. Phys.: Condens. Matter **19**, 395017 (2007).
[26] R. Jackiw, Phys. Rev. D **29**, 2375 (1984).
[27] J. K. Pachos and M. Stone, Int. J. Mod. Phys. B **21**, 5113 (2007).
[28] The sign of the mass term in a given valley depends on the representation. All statements made in this text hold for the valley-isotropic representation. Care should be taken when comparing with results in the literature since several other representations are also common. See footnote 2 of 3.