PHYSICAL REVIEW B **75**, 041401(R) (2007)

Spin-orbit gap of graphene: First-principles calculations

Yugui Yao,[1] Fei Ye,[2] Xiao-Liang Qi,[2] Shou-Cheng Zhang,[3] and Zhong Fang[1,4]
[1]*Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100080, China*
[2]*Center for Advanced Study, Tsinghua University, Beijing 100084, China*
[3]*Department of Physics, McCullough Building, Stanford University, Stanford, California 94305-4045, USA*
[4]*International Center for Quantum Structure, Chinese Academy of Sciences, Beijing 100080, China*
(Received 5 December 2006; published 2 January 2007)

Even though graphene is a low-energy system consisting of a two-dimensional honeycomb lattice of carbon atoms, its quasiparticle excitations are fully described by the (2+1)-dimensional relativistic Dirac equation. In this paper we show that, while the spin-orbit interaction in graphene is of the order of 4 meV, it opens up a gap of the order of 10^{-3} meV at the Dirac points. We present a first-principles calculation of the spin-orbit gap, and explain the behavior in terms of a simple tight-binding model. Our result also shows that the recently predicted quantum spin Hall effect in graphene can occur only at unrealistically low temperature.

DOI: 10.1103/PhysRevB.75.041401 PACS number(s): 73.43.−f, 71.70.Ej, 73.21.−b

Recently, the electronic properties of graphene, a single-layer graphite sheet, have attracted great interest both theoretically and experimentally. The key difference of graphene compared with most other two-dimensional materials is the linear energy spectrum around two nodal points in the Brillouin zone, which makes the low-energy dynamics of electrons in this system equivalent to that of relativistic fermions, as described by the massless Dirac equation.[1] The two sublattices in the graphene honeycomb lattice play the role of pseudospin degrees of freedom. In Refs. 2 and 3, the quantum Hall effect in graphene is observed; it shows the nonconventional quantization rule $\sigma_H=(2e^2/h)(2n+1)$, $n\in\mathbb{Z}$. Such an "abnormal" quantum Hall effect agrees with theoretical calculations based on the massless Dirac equation under external magnetic field,[4–6] and can be considered as a consequence of the chiral anomaly in two-dimensional massless fermions. Moreover, a recent experiment on the low-field magnetoresistance[7] shows that graphene remains metallic under temperatures as low as $T=4$ K, which confirms that any possible gap opened at the Dirac cones cannot be larger than $k_BT\sim 0.34$ meV.

Nevertheless, it has been proposed that a small gap can open on the two Dirac points of graphene due to spin-orbital coupling (SOC),[8] which at the same time makes the system a spin Hall insulator[9] with quantized spin Hall conductance. Physically, this proposal is a spinful version of Haldane's model for the quantum Hall effect without magnetic field,[10] in which a spin-dependent next-nearest-neighbor hopping term is introduced to induce opposite mass terms for the two Dirac cones. Reference 8 estimates the spin-orbit gap in graphene to be 2.4 K. In this paper we provide systematic calculations of the spin-orbital gap in graphene by both first-principles calculation and the tight-binding model, and show that the actual gap is much smaller compared to the crude estimate given in Ref. 8, this explains the (near) gaplessness observed in experiments and also defines a much more narrow temperature range for the quantum spin Hall effect to be observed.

The sp^2 hybridization of the $2s$ orbital and two $2p$ orbitals of the carbon atom creates σ bonds to form the honeycomb lattice of graphene, which is bipartite with two carbon atoms in one unit cell. The π band consisting of the remaining $2p$ orbitals controls the low-energy physics of graphene and makes it a semimetal. One can describe the π and σ electrons by two tight-binding (TB) Hamiltonians separately, which in momentum space is a 2×2 matrix $\mathbf{H}_\pi(\vec{k})$ for the π band, and a 6×6 matrix $\mathbf{H}_\sigma(\vec{k})$ for the σ band.[11] If the spin degeneracy of electrons is taken into account, the dimensions of these two matrices are doubled. The diagonal entities of the matrices are the on-site energies of different orbitals and the off-diagonal entities are the possible hopping between different sublattices.

The SOC is a relativistic effect described by a Hamiltonian with the form $\hbar\vec{\sigma}\cdot(\vec{\nabla}V\times\vec{p})/(4m^2c^2)\sim\vec{L}\cdot\vec{\sigma}$. $\vec{\sigma}$ is the Pauli matrix. For a single carbon atom, there is no SOC between $2s$ and $2p$ orbitals due to their different azimuthal quantum numbers, and SOC exists only among the $2p$ orbitals. Its magnitude ξ_0 can be estimated to be of order 4 meV by directly computing the overlap integral of SOC between $2p_z$ and $2p_x$ orbitals. Note that the SOC changes the magnetic quantum number accompanied with a spin flip of electrons; hence no SOC exists within the same atomic orbital.

For graphene, only the SOC in the normal direction with the form $L_z\sigma_z$ has a nonzero contribution due to the reflection symmetry with respect to the lattice plane. Even this term vanishes for the π orbitals between nearest neighbors (NNs), since there is an additional vertical reflection plane along the nearest-neighbor bond. Under mirror reflection at this plane, the $2p_z$ wave functions of the adjacent atoms are unchanged; however, the angular momentum L_z changes its sign, and hence the matrix element of $L_z\sigma_z$ between NNs vanishes. This is different from the carbon nanotube[12] where the curvature effect can provide a SOC between $2p$ orbitals of NNs although it is still vanishing for large tube radius. Thus, to realize the SOC effect of the π band within the NN approximation we need the aid of the σ band. This process turns out to be a second-order one, which is three orders of magnitude smaller than ξ_0. On the other hand, the SOC can act directly within the σ band; it will open a gap at some degenerate points with the same order of magnitude as $\xi_0\sim 4$ meV.

The SOC mixes the π and σ bands and the total Hamiltonian reads

$$H = \begin{pmatrix} \mathbf{H}_\pi & \mathbf{T} \\ \mathbf{T}^\dagger & \mathbf{H}_\sigma \end{pmatrix}. \tag{1}$$

Here, $\mathbf{H}_\pi$ and $\mathbf{H}_\sigma$ should be enlarged to be 4×4 and 12×12 matrices, respectively, by the spin indices. The SOC term $\mathbf{T}$ bridging the π and σ bands is a 4×12 matrix of order ξ_0; its explicit form is not important at present, and will be given later. The wave vector $\vec{k}$ is omitted for simplicity hereinafter because it is always a good quantum number.

Since we are concerned with the low-energy physics, an effective π-band model with SOC derived from the original Hamiltonian Eq. (1) is more advantageous. For this purpose, one can perform a canonical transformation

$$H \to H_S = e^{-S} H e^S,$$

$$S = \begin{pmatrix} 0 & \mathbf{M} \\ -\mathbf{M}^\dagger & 0 \end{pmatrix}, \tag{2}$$

where $\mathbf{M}$ should satisfy

$$\mathbf{M}\mathbf{H}_\sigma - \mathbf{H}_\pi\mathbf{M} = \mathbf{T}, \tag{3}$$

so that H_S is block diagonal up to order ξ_0^2. Clearly $\mathbf{M}$ is also a 4×12 matrix. The effective Hamiltonian H_{eff} is then extracted from the diagonal part of H_S as

$$H_{eff} \approx \mathbf{H}_\pi - \frac{1}{2}(\mathbf{T}\mathbf{M}^\dagger + \mathbf{M}\mathbf{T}^\dagger). \tag{4}$$

The second term is just the effective SOC for the π-band electrons.

The matrix $\mathbf{M}$ can be calculated iteratively through Eq. (3),

$$\mathbf{M} = \mathbf{T}\mathbf{H}_\sigma^{-1} + \mathbf{H}_\pi\mathbf{T}\mathbf{H}_\sigma^{-2} + \cdots. \tag{5}$$

Around the Dirac points, the spectrum of $\mathbf{H}_\pi$ is close to zero measured from the on-site potential of the $2p$ orbital, while that of $\mathbf{H}_\sigma$ is of order several eV; hence we can take $\mathbf{M} \approx \mathbf{T}\mathbf{H}_\sigma^{-1}$ approximately. The effective SOC of the π band then reads

$$-\mathbf{T}\mathbf{H}_\sigma^{-1}\mathbf{T}^\dagger, \tag{6}$$

whose magnitude is roughly estimated as $\xi_1 \sim |\xi_0|^2/\Delta$ with Δ being of the order of the energy difference at the Dirac points between the π and σ bands. ξ_1 is of the order 10^{-3} meV, since Δ is of order eV.

So far we have not used the explicit forms of $\mathbf{H}_\sigma$ and $\mathbf{T}$ in the above discussions. To derive ξ_1 and the SOC analytically, we need more details of $\mathbf{H}_\sigma$ and $\mathbf{T}$. $\mathbf{H}_\sigma$ can be written as

$$\mathbf{H}_\sigma = \begin{pmatrix} \mathbf{E} & \mathbf{\Sigma} \\ \mathbf{\Sigma}^\dagger & \mathbf{E} \end{pmatrix} \otimes \mathbf{I}, \tag{7}$$

where $\mathbf{I}$ is the identity matrix for the spin degrees of freedom. The matrix $\mathbf{E}$ represents the on-site energy of different atomic orbitals, which can be written as

$$\mathbf{E} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta_\varepsilon \end{pmatrix} \tag{8}$$

if we arrange the three sp^2 hybridized orbitals in the sequence of $\{2p_y, 2p_x, 2s\}$. Here, Δ_ε is the energy difference $\varepsilon_{2s} - \varepsilon_{2p}$ between the $2s$ and $2p$ orbitals. $\mathbf{\Sigma}$ describes the hopping between the two sublattices in the momentum space. To give its exact form, we first consider the hopping between the two adjacent atoms in real space, which can also be described by a 3×3 matrix. Suppose the two adjacent atoms are placed on the horizontal x axis, i.e., the bond angle is zero; then this hopping matrix can be written as follows:

$$\Sigma_0 = \begin{pmatrix} V_{pp\pi} & 0 & 0 \\ 0 & V_{pp\sigma} & V_{sp\sigma} \\ 0 & V_{sp\sigma} & V_{ss\sigma} \end{pmatrix}. \tag{9}$$

One can obtain the hopping matrix $\Sigma(\theta)$ for arbitrary bond angle θ by a rotation $\mathbf{R}(\theta)$ in the xy plane as

$$\Sigma(\theta) = \mathbf{R}^\dagger(\theta)\Sigma_0\mathbf{R}(\theta),$$

$$\mathbf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{10}$$

The parameters $V_{pp\pi}$, $V_{pp\sigma}$, $V_{sp\sigma}$, and $V_{ss\sigma}$ correspond to the σ or π bonds formed by the $2s$ and $2p$ orbitals, whose empirical value can be found in textbooks, for example, Ref. 11. Note that we do not consider the wavefunction overlap matrix in our TB approximation scheme for the sake of simplicity. Then the hopping matrix in the momentum space reads

$$\mathbf{\Sigma}(\vec{k}) = \sum_\alpha \Sigma(\theta_\alpha) e^{i\vec{k}\cdot\vec{d}_\alpha}, \tag{11}$$

where $\vec{d}_\alpha$ with $\alpha = 1,2,3$ are the bond vectors connecting the carbon atom and its three nearest neighbors and θ_α is the angle between $\vec{d}_\alpha$ and the x axis.

For $\mathbf{T}$, as we have described above, the spin flip on the same atom takes place only between the $2p_z$ and two in-plane $2p_{x,y}$ orbitals. A straightforward calculation leads to the on-site spin flip

$$\mathbf{T}_0 = \xi_0(-\sigma_x, \sigma_y, 0), \tag{12}$$

with $\sigma_{x,y}$ the Pauli matrices. Then $\mathbf{T}$ can be written as

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_0 & 0 \\ 0 & \mathbf{T}_0 \end{pmatrix}. \tag{13}$$

Notice that there are two $\mathbf{T}_0$ terms in the above matrix corresponding to different sublattices.

Since $\mathbf{H}_\sigma$ has a large gap near the Dirac points K and K^*, we can expect that $\mathbf{H}_\sigma(\vec{k}+\vec{K}) = \mathbf{H}_\sigma(\vec{K}) + o(k)$, which means we can substitute $\mathbf{H}_\sigma^{-1}(\vec{K})$ into Eq. (6) as a good approximation. Finally we get the effective Hamiltonian with SOC at the low energy scale,

$$H_{eff}^{[K]} \approx \xi_1 + \begin{pmatrix} \xi_1\sigma_z & v_F(k_x + ik_y) \\ v_F(k_x - ik_y) & -\xi_1\sigma_z \end{pmatrix},$$

$$H_{eff}^{[K^*]} \approx \xi_1 + \begin{pmatrix} -\xi_1\sigma_z & v_F(k_x - ik_y) \\ v_F(k_x + ik_y) & \xi_1\sigma_z \end{pmatrix}. \quad (14)$$

The off-diagonal terms in the above equations come from the well-known form of $\mathbf{H}_\pi$, and v_F is just the Fermi velocity of π electrons at the Dirac points. The effective SOC ξ_1 in our TB scheme has the explicit form

$$\xi_1 \approx |\xi_0|^2(2\Delta_\varepsilon)/(9V_{sp\sigma}^2). \quad (15)$$

Equation (15) is the key result from our tight-binding calculation. Equation (14) leads to a spectrum $E(\vec{k}) = \pm\sqrt{(v_F k)^2 + \xi_1^2}$. Taking the values of the corresponding parameters from Ref. 11, one can estimate ξ_1 to be of order 10^{-3} meV, so the energy gap is $2\xi_1$ at the Dirac points.

Equations (14) are similar to those in Ref. 8, except that the SOC constant ξ_1 is three orders of magnitude smaller than their estimate. We can also consider the SOC of π orbitals between next nearest neighbors (NNNs) which is not forbidden by the symmetry. In this case the electron moving between NNNs will be accelerated by atoms other than the two NNN ones, which provides the corresponding SOC. This will involve three-center integrals, i.e., two orbital centers and a potential center, which are different with each other. Generally speaking, such integrals are very small, which leads to SOC of order at most 10^{-3} meV by our estimate, and it may actually be smaller.

The argument above is supported by accurate first-principles calculations based on density-functional theory. The relativistic electronic structure of graphene was calculated self-consistently by the plane-wave method[13] using a relativistic fully separable pseudopotential in the framework of noncollinear magnetism.[14] The exchange-correlation potential is treated by the local density approximation (LDA) whose validity for the system considered here has been demonstrated by many other studies. The experimental lattice parameter $a=2.456$ Å is used in the calculation. The convergence of calculated results with respect to the number of **k** points and the cutoff energy has been carefully checked.

Figure 1 shows the band structure of graphene. We can see that the gap induced by SOC for the σ orbit is 9.0 meV at the Γ point. The figure also indicates that there is a gap induced by SOC for the π orbit at the K point, and the magnitude of the splitting gap is 0.8×10^{-3} meV, which is in good agreement with the estimate obtained from the tight-binding model discussed above. Since the number discussed here is so small, a few notes are necessary: (1) the calculations are valid within the LDA; (2) the numeric accuracy of the present calculations reaches 10^{-6} meV per atom; (3) the convergence of gap size with respect to the number of **k** points and cutoff energy is better than 1×10^{-4} meV; (4) the Kramer doublet degeneracy can be reproduced down to 10^{-5} meV. Nevertheless, it is clear that the gap induced by SOC at the K point is of order 10^{-3} meV. Considering that graphene may be typically deposited on substrates, the graphene layers are generally strained due to small lattice mismatches; thus the lattice mismatch strain can tune the splitting gap at the K point. We have calculated the band structure for different lattice constants of graphene, and have found that the splitting gap increases a little with compression while the gap decreases with tensile strain.

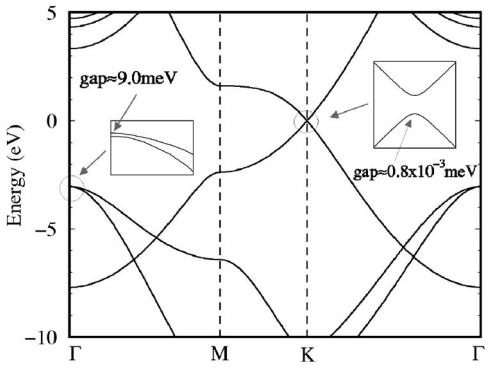

FIG. 1. Calculated relativistic band structure of graphene.

In conclusion, we provided a careful calculation on the spin-orbit gap of graphene, which leads to the same mass term for the relativistic Dirac fermions in the continuum limit,[8] but with a much smaller magnitude of the gap 10^{-3} meV. The physical reason for the smallness of the spin-orbit gap can also be understood from the tight-binding model as coming from the lattice C_3 symmetry, which leads to the vanishing of the leading-order contributions. Such a small gap is consistent with the experimental observation of semimetallic behavior of graphene. It shows that the proposed quantum spin Hall effect in graphene cannot be observed until temperatures as low as $T \ll 10^{-2}$ K. In addition, impurity scattering in a disordered sample may also destabilize the effect.

This work is supported by the Knowledge Innovation Project of the Chinese Academy of Sciences, the NSFC under Grants No. 10404035, No. 10674163, and No. 10534030 (for Y.G.Y.), No. 90303022, No. 60576058, and No. 10425418 (for Z.F.), and No. 10374058 and No. 90403016 (for X.L.Q.). S.C.Z. is supported by the NSF through Grant No. DMR-0342832 and by the U.S. Department of Energy, Office of Basic Energy Sciences under Contract No. DE-AC03-76SF00515. We would like to acknowledge helpful discussions with C. Kane, A. MacDonald, Z. B. Su, and B. S. Wang.

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

[2] 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).

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

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

[5] V. P. Gusynin and S. G. Sharapov, Phys. Rev. B **73**, 245411 (2006).
[6] N. M. R. Peres, F. Guinea, and A. H. Castro Neto, Phys. Rev. B **73**, 125411 (2006).
[7] S. V. Morozov, K. S. Novoselov, M. I. Katsnelson, F. Schedin, L. A. Ponomarenko, D. Jiang, and A. K. Geim, Phys. Rev. Lett. **97**, 016801 (2006).
[8] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[9] S. Murakami, N. Nagaosa, and S.-C. Zhang, Phys. Rev. Lett. **93**, 156804 (2004).
[10] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
[11] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College Press, London, 1998).
[12] T. Ando, J. Phys. Soc. Jpn. **69**, 1757 (2000).
[13] Z. Fang and K. Terakura, J. Phys.: Condens. Matter **14**, 3001 (2002).
[14] G. Theurich and N. A. Hill, Phys. Rev. B **64**, 073106 (2001).