Dynamical gap generation in graphene nanoribbons: An effective relativistic field theoretical model

A. J. Chaves,[1] G. D. Lima,[2] W. de Paula,[1] C. E. Cordeiro,[3] A. Delfino,[3] T. Frederico,[1] and O. Oliveira[1,4]
[1]*Departamento de Física, Instituto Tecnológico de Aeronáutica, 12228-900 São José dos Campos, Brazil*
[2]*Departamento de Física, Universidade Federal do Piauí, 64049-550 Teresina, Piauí, Brazil*
[3]*Instituto de Física, Universidade Federal Fluminense, Avenida Litorânea s/n, 24210-340 Niterói, Rio de Janeiro, Brazil*
[4]*Departamento de Física, Universidade de Coimbra, 3004 516 Coimbra, Portugal*
(Received 20 December 2010; published 7 April 2011)

We show that the assumption of a nontrivial zero band gap for a graphene sheet within an effective relativistic field theoretical model description of interacting Dirac electrons on the surface of graphene describes the experimental band gap of graphene nanoribbons for a wide range of widths. The graphene band gap is dynamically generated, corresponding to a nontrivial gapless solution, found in the limit of an infinitely wide graphene ribbon. The nanoribbon band gap is determined by the experimental graphene work function.

DOI: 10.1103/PhysRevB.83.153405 PACS number(s): 72.80.Vp, 11.10.Kk

Introduction. Graphene sheets, carbon nanotubes (CNTs) and graphene nanoribbons (GNRs) are very closely related subjects governed by quantum mechanical effects in a constrained two-dimensional (2D) surface. In this compactified dimensional world, symmetry and scale drive the astonishing properties that these one-atom-thick materials exhibit. GNRs are recognized today as promising nanosheets to be used in nanoelectronic and spintronic devices.[1] Such material has attracted a lot of attention since free-standing graphene sheets were reported by using different experimental techniques. Experiments confirmed that the charge carriers can be described as massless Dirac fermions,[2] opening a solid basis for its relativistic treatment. It is expected that massless fermions moving through graphene have an approximate ballistic transport behavior with very small resistance due backscattering suppression.[3] Graphene is also a good thermal conductor, when compared with other semiconductor junctions, due to the high mobility of carriers at room temperature.[4,5]

Energy band gaps for GNRs were measured in Refs. 6 and 7 using different methods. First-principles *ab initio* methods,[8] density functional theory (DFT),[9] and tight binding (TB) models[10] were applied to calculate GNR band gaps for a reasonable range of not too small widths. In opposition to TB models, first-principles calculations do not support metallic nanoribbons.[8] On the other hand DFT calculations predict energy gap oscillations as a function of the GNR for very small widths.[9]

The two-dimensional graphene sheet is essentially a zero-gap semiconductor (with the Fermi level E_F precisely at $E = 0$), with a linear chiral carrier kinetic energy dispersion relation given by $E = \hbar v_F k^2$, where k is the 2D carrier wave vector and v_F is the Fermi velocity (independent of carrier density).

The fermion behavior on the graphene sheet has been discussed within the free Hamiltonian, $\hat{H}_0$, picture using four different approaches:[11] (i) as Schrödinger fermions $\hat{H}_0 = \hat{p}^2/2m^*$, where m^* is an effective mass; (ii) as ultrarelativistc Dirac fermions, where $\hat{H}_0 = c\vec{\sigma}\,\hat{p}$; (iii) as massless Dirac fermions with $\hat{H}_0 = v_F\vec{\sigma}\,\hat{p}$, where $\vec{\sigma}$ are the Pauli matrices; and (iv) as massive chiral fermions with $\hat{H}_0 = \vec{\sigma}\,\hat{p}^2/2m^*$, where $\vec{\sigma}$ is a pseudospin matrix describing the two sublattices of the honeycomb lattices.[12]

This work is based on a relativistic effective field theory of electrons containing a local four-fermion interaction. The effective interaction Lagrangian parametrizes the actual forces acting on the electrons. It assumes the relevant dynamics from the atom-electron and electron-electron interactions within the many-body system (including, e.g., electron-phonon interactions). The rationale to use effective contact interactions goes back to the claim[13,14] that the Hartree-Fock approximation of the single-particle energy of the many-fermion system with the point-coupling interaction can be compared to the first-order contribution of the surface density in a density functional theory.[15]

The relativistic picture of interacting fermions has to be understood together with the concept of an effective mass. In the case of graphene the effective mass vanishes, corresponding to a gapless material. A nonvanishing effective mass for GNRs means that a gap is open. Our model provides a mechanism for dynamical mass generation for the nanoribbons, even for a gapless graphene.

We assume a homogeneous 2D system of electrons (or holes),[2] where the free part of the electron Hamiltonian corresponds to case (iii), i.e., $\hat{H}_0 = v_F\vec{\sigma}\,\hat{p}$. The two sublattices of graphene, associated with the Dirac points K and K', are taken into account by introducing a factor 2 in the electron multiplicity.

The model is based on a relativistic Lagrangian field theoretical approach in which a four-fermion scalar and a vector pointlike interaction are considered and the many-fermion problem solved in the Hartree-Fock approximation on a 2D surface. The model has only two free parameters, one related to the strengths of the four-fermion interaction and the Fermi momentum cutoff Λ. Exploring the connection of the model with graphene, i.e., by assuming that graphene is a nontrivial gapless solution of the dynamical equations, we are able to relate the strengths of the interactions and cutoff to the graphene work function. This procedure allows us to fix the parameters and calculate the gaps of the nanoribbons without any further assumptions.

Our work is organized as follows. In Sec. II, we present the model and derive the gap equations. In Sec. III, we present the solution of the gap equations for the nanoribbon

and compare our results with the experimental data.[6] Our concluding remarks are given in the last section.

The model. The electron-hole dynamics is described by a relativistic quantum field effective Lagrangian in a $2+1$ space-time, where the interaction has a pairwise and contact form, which generalizes a nonrelativistic 2D model[16]

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi + G_s(\bar{\Psi}\Psi)^2 + G_v(\bar{\Psi}\gamma^\mu\Psi)^2, \quad (1)$$

where s and v refer to scalar and vector couplings, respectively, and Ψ refers to fermion fields. In the following, we use a system of units such that $\hbar = v_F = 1$.

Our method of solving the model for the nanoribbon starts from a nanotube geometry where the fermions are constrained to move on the cylindrical surface. The appropriate coordinates are the position along the nanotube symmetry axis (z) and the angle in the transverse plane (θ)—see Ref. 16 for details. The GNRs are obtained by *unfolding* an infinite-length carbon nanotube. The resulting sheet is a GNR with a width W, after eliminating dangling bonds. The procedure to obtain GNRs from CNTs is described in Ref. 9.

The Dirac modes are quantized along the transversal direction after demanding that the spinors vanish at the edges of the nanoribbon. From this condition it follows that the allowed fermion transverse momenta are given by $k_n = n\pi/W$, where $n = \pm 1, \pm 2, \ldots, n_{\max}$ with $n_{\max}$ being the largest integer smaller than $\Lambda W/\pi$. The maximum momentum of the negative energy electrons filling the valence band, as illustrated in Fig. 1, is Λ.

The fermion mass, i.e., the electron self-energy, is generated dynamically and is computed from the gap equations for the GNRs, using a self-consistent Schwinger-Dyson equation at one-loop level (Hartree-Fock approximation). Recall that the Schwinger-Dyson equation has a cutoff in the momentum of the negative electron energy in the valence band given by Λ.

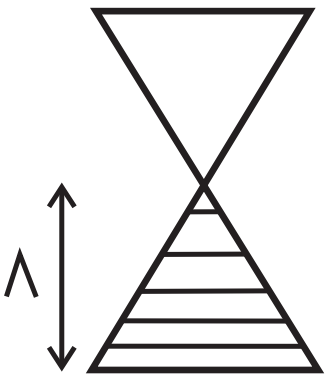

FIG. 1. Schematic graphene bands for conducting (upper) and valence (lower) fermions. For the discussion of Λ see the text.

The gap equations for the self-energies, after integration on the longitudinal momentum, are given by

$$\Sigma_s = \frac{2G}{\pi W}(\Sigma_s + m)\sum_{n=-n_{\max}}^{+n_{\max}} S(k_n), \quad (2)$$

$$\Sigma_v^0 = \frac{2G\Lambda^2}{\pi^2} f(x), \quad (3)$$

$$\Sigma_v^\theta = -\frac{2G}{\pi W}\sum_{n=-n_{\max}}^{+n_{\max}} S(k_n)\left(k_n - \Sigma_v^\theta\right), \quad (4)$$

where

$$S(k) = \ln\left[\sqrt{\frac{\Lambda^2 - k^2}{a^2}} + \sqrt{\frac{\Lambda^2 - k^2}{a^2} + 1}\right], \quad (5)$$

$$f(x) = \frac{1}{x}\sum_{n=-n_{\max}}^{+n_{\max}} \sqrt{1 - \frac{n^2}{x^2}}, \quad (6)$$

and $a^2 = (k_n - \Sigma_v^\theta)^2 + (\Sigma_s + m)^2$.

In Eqs. (2)–(4) we define $x = \Lambda W/\pi$ and $G = G_s - 3G_v$ and a factor of 2 was included to take into account electrons at the K and K' points. Note that the scalar G_s and the vector G_v coupling constants combine into a single coupling constant G. Furthermore, the self-energies given by Eqs. (2)–(4) include a rest fermion mass m. However, hereafter we consider only the $m = 0$ case.

The scalar self-energy Σ_s is identified with the half-band-gap nannoribbon. The self-energy vector part is actually a quadrivector containing space-time components. However, due to symmetry only the time component given by Σ_v^0 survives. This component is directly related to the number of particles N through

$$N = \frac{2LW}{G}\Sigma_v^0, \quad (7)$$

where L is the longitudinal length of the GNR.

In a relativistic picture, the infinite graphene sheet ($W \to \infty$) has no electrons in the conducting band. The valence band is fully occupied and there is no gap between the two bands, i.e., $\Sigma_s = 0$. Then, it remains to parametrize only Eq. (3), since $\Sigma_s = 0$ is a trivial solution for Eq. (2). Recall that we are considering massless fermions $m = 0$. Anyway, when considering very large but finite values for W, (2) admits other solutions besides the trivial $\Sigma_s = 0$ and one has an opening gap generated dynamically.

Let us look for the nontrivial solution of (2). This defines a unique coupling constant, a function of Λ, $G = G_{\text{crit}}$, which will be called the critical graphene limit. If one takes the continuum limit of Eq. (2) with $m = 0$, after setting $y = k_n/\Lambda$ it follows that

$$\frac{2\Lambda G_{\text{crit}}}{\pi^2}\int_0^1 dy\, \ln\left(\sqrt{\frac{1}{y^2} - 1} + \frac{1}{y}\right) = 1. \quad (8)$$

Given that the above integral is equal to $\pi/2$, then the critical graphene limit requires a

$$G_{\text{crit}} = \frac{\pi}{\Lambda}. \quad (9)$$

The Λ dependence can be eliminated in favor of the graphene experimental work function value $\mathcal{W} = 4.8$ eV.[17] The work function is related to the chemical potential μ, which needs to be computed in a consistent thermodynamic way.[16,18] The pressure P and the density energy $\mathcal{E}$ of the system determine the chemical potential by $\mu = (P + \mathcal{E})/\sigma$, with $\sigma = N/A$ and where A is the area of the GNR. P and $\mathcal{E}$ are given by the diagonal terms of the energy-momentum tensor constructed from the model Lagrangian (1) in the usual way. In a mean field approach,[19]

$$\begin{aligned}\mathcal{E} &= \langle T^{00}\rangle \\ &= \frac{1}{\pi W}\sum_n \Omega\left(k_n^2\right)\left[2\Sigma_v^0 - \Omega\left(k_n^2 - a^2\right) - a^2 S(k_n)\right] \\ &\quad + \frac{1}{2G}\left(\Sigma_v^0\right)^2, \end{aligned} \tag{10}$$

where $\Omega(x) = \sqrt{\Lambda^2 - x}$, and

$$\begin{aligned} P &= \langle T^{zz}\rangle \\ &= \frac{-1}{\pi W}\sum_n \left[\Omega\left(k_n^2\right)\Omega\left(k_n^2 - a^2\right) - a^2 S(k_n)\right] - \frac{1}{2G}\left(\Sigma_v^0\right)^2\Sigma_s^2, \end{aligned} \tag{11}$$

where z is the longitudinal direction of the GNR. Then, the expression for the chemical potential reads

$$\mu = \frac{G\,\Lambda^2}{\pi^2} f(x) - \sqrt{\Lambda^2 + \Sigma_s^2} = -\mathcal{W}, \tag{12}$$

where $\mathcal{W}$ is the work function. As a side remark, according to Ref. 18, the chemical potential computed in the way described above contains already both the bulk and surface contributions from the energy density and pressure.

Graphene is recovered in the limit $W \to \infty$ and, in this case, $x = \Lambda W/\pi$ also diverges, while $f(x \to \infty) = \pi/2$. Therefore, for graphene if one sets $G = G_{\text{crit}}$ and $\Sigma_s = 0$, one defines the parameters of the model. Then, using G_{crit}, it follows from Eq. (12) that $\Lambda = 2\mathcal{W}$. Given that the experimental value for the graphene work function is 4.8 eV,[17,20] this determines that $\Lambda = 9.6$ eV. From now on, we will always use these results for computing the nanoribbon gaps within the effective relativistic field theoretical model given by the Lagrangian density (1).

The above definitions fully parametrize Eqs. (2)–(4), allowing the computation of a self-consistent solution for Σ_s as a function of the nanoribbon width W. Furthermore, given a value for Λ, one can compute the surface density $\sigma = \Lambda^2/\pi$.

Results. In order to solve the model numerically, besides G_{crit} and $\Lambda = 9.6$ eV, we use $v_F = c/300$.[12] Our numerical procedure is as follows: (i) For a given W, $x = \Lambda W/\pi$ and $k_n = n\pi/W$ are defined, constraining the integer $n_{\max} \leqslant x$; (ii) given that $\Sigma_v^\theta = 0$, $a^2 = k_n^2 + \Sigma_s^2$ and Eq. (2) can be solved iteratively to compute $E_g = 2\Sigma_s$. Σ_v^0 is calculated directly from Eq. (3).

In Fig. 2 the results of the model for the band gap E_g as a function of the nanoribbon width are compared with the experimental data of Ref. 6. As observed in the figure, the model reproduces the GNR's experimental gap for a large range of nanoribbon widths. We call to the reader's attention, that the curve in Fig. 2 is the direct outcome from the model without any *ad hoc* normalization to reproduce the right E_g scale. This striking agreement between experimental data and theoretical predictions seems to indicate that the model, as defined above, captures the essential physics required to understand the gap formation in nanoribbons. Moreover, the results summarized in Fig. 2 give support to the idea that the concept of the effective interacting valence Dirac electrons which generate dynamically a band gap is useful to describe the electronic properties of GNRs.

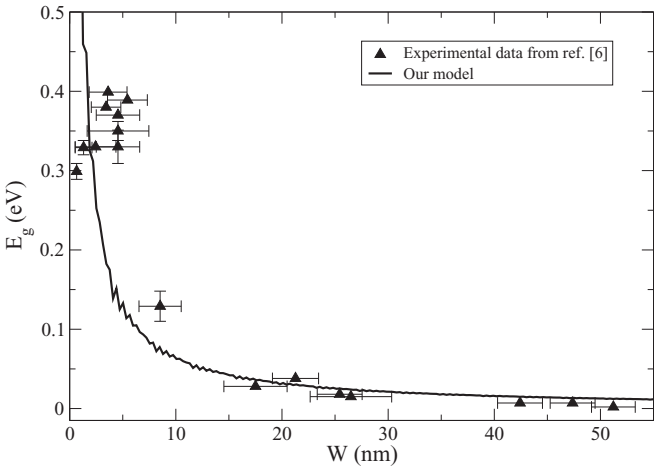

FIG. 2. GNR band gap as a function of ribbon width W.

The experiments reported in Refs. 17 and 20 show that for CNTs with radius between 0.5 and 1.5 nm, the work function oscillates with an amplitude of ~0.5 eV or less around the graphene work function $\mathcal{W}$. As the CNT radius increases, the oscillations rapidly become very smooth. This consideration may help in understanding why our model still works for GNRs with small widths. Moreover, it also shows that the graphene scale, used in the definition of our model, is robust regarding the GNR finite-size details. Possibly, the massless fermions picture reinforces this point, since the localization of carriers is strongly suppressed favoring ballistic behavior for them.

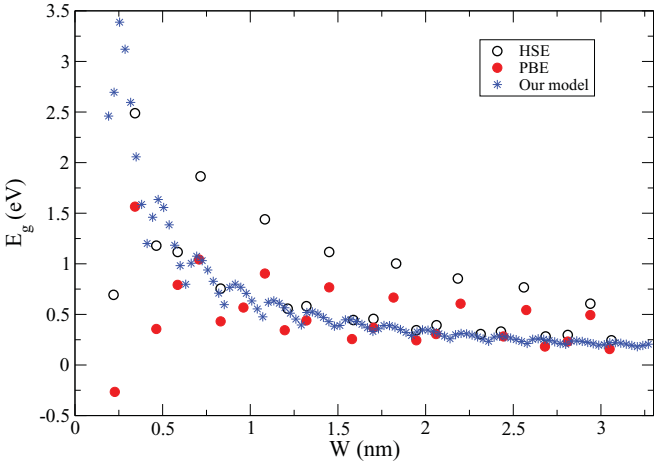

FIG. 3. (Color online) GNR band gap as a function of ribbon width. Theoretical calculations from PBE and HSE density functionals (Ref. 9).

Let us now compare the model with other theoretical predictions. In Ref. 8, first-principles calculations for three types of armchair-edged GNRs have been reported. Their results are similar to ours. For smaller widths, previous theoretical calculations[9] have been done using DFT functionals from the works of Perdew *et al.* (PBE)[21] and Heyd *et al.* (HSE).[22] According to the authors, the band gap oscillates as a function of W. A similar behavior is observed in our model—see Fig. 3. Typically, our results are in between those of the two DFT (PBE, HSE) calculations.

From the point of view of the relativistic model, the oscillations in the physical quantities are related to $n_{\max}$. As the width W decreases, $x = \Lambda W/\pi$ decreases and higher values for the integer n become possible. Whenever $n_{\max}$ increases by unity, due to the behavior of the derivative of N relative to the single-particle energy, a Van Hove singularity shows up. This explains the oscillatory behavior of the GNR band gaps observed in our model. As seen in Fig. 3, whenever W gets larger, the amplitude of oscillations diminishes and they become sizable only for a width of the order of a few nanometers or smaller.

Concluding remarks. Our results for the effective relativistic field theoretical model show that it accounts reasonably well for some of the physics require to understand the GNR from small to large widths. Moreover, the graphene scale is used to defined the model parameters and it seems to be a robust scale regarding the GNR finite-size details. In this sense, our results are very auspicious for further investigations of GNRs with the present model which may, as claimed before, represent a first-order approximation to a DFT calculation. A nice feature of the relativistic field approach is its simplicity. Indeed, it allows for a detailed analytical description that, eventually, can guide our intuition in the investigation of GNRs and in looking for general driving physics trends.

To conclude, we would like to make general comments on how to improve the present effective model. For small widths, due to the increase in the kinetic energy, some of the fermions can jump into the conduction band. Therefore, for small W's the model, instead of having a single cutoff Λ, will require two, one for the valence band Λ_1 and another for the conduction band Λ_2. The presence of the new scale could lead to larger amplitude oscillations as the *ab initio* calculations suggest—see Fig. 3. Moreover, when W becomes smaller, the edges distinguish armchair and zigzag GNRs, as discussed in Ref. 23. This requires refined boundary conditions for the spinors at the GNR at the boundaries. In general Σ_v^θ vanishes due to the symmetry between k_n and $-k_n$. However, if this symmetry is broken due to the boundary conditions, then one should solve Eqs. (2)–(4) self-consistently. In any case, the relativistic model still keeps its simplicity and the way we have solved the model is, essentially, unchanged.

Acknowledgments. The authors express their thanks to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Nível Superior (CAPES) and to Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for partial financial support.

[1] A. K. Geim and K. S. Novoselov, Nature Mater. **6**, 183 (2007); K. S. Novoselov *et al.*, Science **306**, 666 (2004).
[2] K. S. Novoselov *et al.*, Nature (London) **438**, 197 (2005).
[3] D. A. Abanin and L. S. Levitov, Science **317**, 641 (2007).
[4] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College Press, London, 1998).
[5] D. A. Areshkin, D. Gunlycke, and C. T. White, Nano Lett. **7**, 204 (2007).
[6] Xiaolin Li, Xinran Wang, Li Zhang, Sangwon Lee, and Hongjie Dai, Science **319**, 1229 (2008).
[7] Melinda Y. Han, Barbaros Ozyilmaz, Yuanbo Zhang, and Philip Kim, Phys. Rev. Lett. **98**, 206805 (2007).
[8] Y. -W. Son, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **97**, 216803 (2006).
[9] V. Barone, O. Hod, and G. E. Scuseria, Nano Lett. **6**, 2478 (2006).
[10] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejon, Phys. Rev. B **66**, 035412 (2002).
[11] A. K. Geim, Science **324**, 1530 (2009).
[12] A. H. Castro Neto , F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
[13] M. Brack, Helv. Phys. Acta **58**, 715 (1985).
[14] R. J. Furnstahl, J. Phys. G **31**, S1357 (2005).
[15] W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133A (1965).
[16] C. E. Cordeiro, A. Delfino, and T. Frederico, Carbon **47**, 690 (2009); Phys. Rev. B **79**, 035417 (2009).
[17] S. Suzuki, Y. Watanabe, Y. Homma, S. Fukuba, and A. Locatelli, Appl. Phys. Lett. **85**, 127 (2004).
[18] K. Morawetz, N. H. March, and R. H. Squire, Phys. Lett. A **372**, 1707 (2008).
[19] B. D. Serot and J. D. Walecka, Adv. Nucl. Phys. **16**, 1 (1986).
[20] S. Suzuki, S. Bower, Y. Watanabe, and O. Zhou, Appl. Phys. Lett. **76**, 4007 (2000).
[21] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[22] J. Heyd, G. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).
[23] L. Brey and H. A. Fertig, Phys. Rev. B **73**, 235411 (2006).