



Electrical observation of subband formation in graphene nanoribbons

Yu-Ming Lin,* Vasili Perebeinos, Zhihong Chen, and Phaedon Avouris
 IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA

(Received 30 April 2008; revised manuscript received 3 October 2008; published 30 October 2008)

We report the experimental observation of subband formation in graphene nanoribbons, where conductance becomes quantized due to the lateral quantum confinement. We show that this quantization in graphene nanoribbons can be observed at temperatures as high as 80 K and channel lengths as long as 1.7 μm . The observed quantization is in agreement with that predicted by theoretical calculations.

DOI: 10.1103/PhysRevB.78.161409

PACS number(s): 73.23.-b, 73.63.-b

The isolation of graphene,¹ a single sheet of graphite, has inspired numerous studies because of its fundamental physics interest^{2,3} and the promising applications of carbon-based electronics.⁴ As a truly two-dimensional (2D) system and a zero-gap semiconductor where the electrons and holes behave as massless fermions, graphene possesses distinctly different transport properties from that of conventional 2D and three-dimensional (3D) electronic materials. However, in order to utilize their remarkable electrical characteristics in certain applications, it would be highly desirable to produce a band gap in graphene, and therefore, intense efforts are being made to explore the properties of low-dimensional [one dimensional (1D) and zero dimensional (0D)] graphene nanostructures.⁵⁻⁹

By patterning graphene into a narrow ribbon structure, the carriers are laterally confined to form a quasi-one-dimensional system, similar to the case of carbon nanotubes. Due to the linear dispersion relation $E=v_f\hbar k$ and high Fermi velocity ($v_f\sim 10^6$ m/s) in graphene, the quantization energy of graphene nanoribbons (GNRs) can be substantially larger than that of conventional semiconducting materials of the same dimension and parabolic dispersion. The formation of 1D subbands in GNRs is expected to yield an energy gap for certain ribbon widths and crystallographic directions.¹⁰ Recently, sub-10 nm GNRs with smooth edges were obtained and demonstrated to possess a band gap at room temperature.⁹ In terms of transport, this quantum confinement can also lead to quantized conductance associated with the formation of 1D subbands, one of the most important transport characteristics of mesoscopic physics.

Recently, experimental studies on GNRs obtained by plasma etching also revealed a thermally-activated conductivity^{5,6} and the presence of a size-dependent energy gap. However, there remains some controversy in understanding the transport behavior in realistic GNR devices, where the imperfect and unknown edge configurations may lead to Coulombic-blockade-type transport at low temperatures.⁸ Furthermore, an energy gap could also be induced by substrate interactions.¹¹ Despite numerous theoretical predictions on GNRs,¹²⁻¹⁴ conductance quantization has yet to be reported experimentally.

Here we present electrical transport measurements for GNR devices with lateral widths of 30 nm, and report conductance plateau features in GNRs at temperatures as high as 80 K. By modulating the Fermi energy by a back gate, we observe conductance quantization plateaus for electron and holes in the same device, both possessing comparable trans-

mission probability for each 1D conduction mode. Our results provide direct experimental evidence of quantum size confinement effects and the formation of subbands for 1D graphene nanostructures.

The GNR devices in our study, schematically shown in Fig. 1(a), were fabricated from mechanically exfoliated graphene sheets on a *p*-doped Si substrate covered with 300-nm-thick SiO₂. Raman spectroscopy and atomic force microscope (AFM) measurements were employed to identify single-layer graphene. For the details of the device fabrication, we refer to Refs. 6 and 15. The source and drain contacts are made of Pd, and the Si substrate acts as the back gate. Narrow ribbons of graphene are formed by oxygen plasma reactive ion etching using a patterned hydrogen silsesquioxane (HSQ) layer as the protective mask. Figure 1(b) shows the AFM image of an array of 30-nm-wide GNRs after this HSQ layer is removed in hydrofluoric (HF) acid solution. We found that electrical properties of the GNRs are

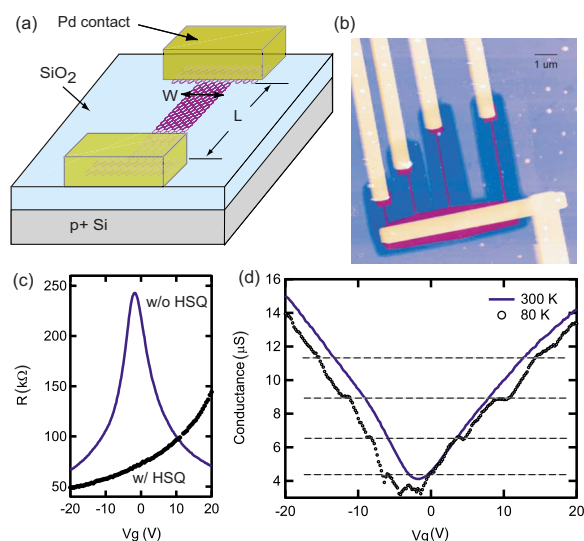


FIG. 1. (Color online) (a) Schematics of the GNR device where the graphene is contacted by two Pd metal leads and the *p*-doped Si substrate acts as the gate electrode. (b) AFM image of GNR devices with different channel lengths. (c) Resistance of a GNR device measured as a function of gate voltage before and after the HSQ layer is removed, showing the impact of HSQ on electrical behaviors. Both measurements were performed in vacuum after annealing at 135 °C. (d) Conductance of a GNR device ($W=30$ nm and $L=850$ nm) measured as a function of gate voltage at 300 and 80 K. The bias voltage is 10 mV.

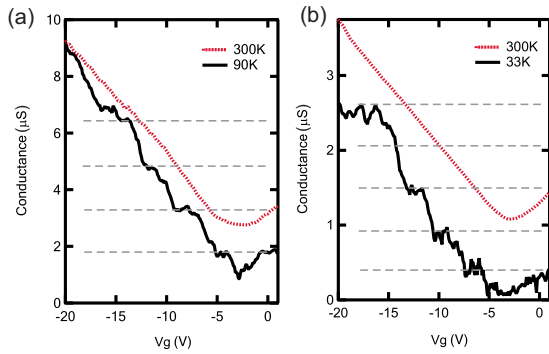


FIG. 2. (Color online) Conductance measured as a function of gate voltage for 30-nm-wide GNR devices with different channel lengths. The GNR channel lengths are 900 nm and $1.7 \mu\text{m}$ for (a) and (b), respectively. At low temperatures, both devices show equally spaced conductance plateaus with $\Delta G \approx 1.7$ and $0.6 \mu\text{S}$ for (a) and (b), respectively. The bias voltage is kept at 10 mV.

strongly affected by this HSQ layer, which needs to be removed in order to reveal their intrinsic transport properties and, more importantly, the conductance quantization behaviors at low temperatures, as discussed below. Figure 1(c) shows the resistance as a function of gate voltage V_g of a GNR device before and after the HSQ layer is removed. We note that the presence of HSQ leads to a significant positive shift of the Dirac point voltage and a reduction in carrier mobility in graphene. In contrast, after removing the HSQ layer, the GNR exhibits ambipolar behavior with the resistance maximum associated with the Dirac point occurring at $V_g \approx 0$ V, indicating a negligible unintentional doping in the final device.

Transport measurements of GNRs were performed in vacuum ($\sim 10^{-7}$ torr), and the devices studied all possess a Dirac voltage near $V_g \sim 0$ V after annealing at 135°C . Figure 1(d) shows the conductance G of a GNR device ($L=850$ nm and $W=30$ nm) measured as a function of gate voltage V_g at a dc bias of 10 mV. At room temperature, the conductance curve $G(V_g)$ in Fig. 1(d) resembles that of bulk graphene,² showing the characteristic “V” shape that reflects symmetric hole and electron transports at negative and positive gate voltages, respectively. While the overall conductance curve $G(V_g)$ of the GNR displays little variation as temperature decreases, several plateau features start to appear in the measured $G(V_g)$ curve and become apparent for $T < 100$ K [see Fig. 1(d)]. The slight asymmetry in the slope of the n and p branches is likely associated with the gate oxide hysteresis. Nevertheless, we note that these conductance plateaus are observed in both electron and hole branches with nearly the same conductance values and an equal spacing [see dashed line in Fig. 1(d)].

Conductance plateau features similar to those found in Fig. 1(d) are also observed in other 30-nm-wide GNR devices for channel lengths L up to $1.7 \mu\text{m}$ (see Fig. 2). These conductance plateau structures are highly reproducible under different thermal cycles, and they generally become more well defined as T decreases, in particular for longer channel devices. However, as the temperature further decreases below 10 K, the $G(V_g)$ traces of these GNRs are usually overwhelmed by a pronounced fluctuation background as a func-

tion of gate voltage. Unlike the plateau structures, the low- T background fluctuations are not retraceable in different thermal cycles, indicating that the low- T background fluctuations may be due to the universal conductance fluctuation (UCF) phenomena.⁶

In order to understand the conductance plateaus in GNRs, here we consider two possible transport scenarios in nanostructures: the charge hopping through a series of quantum dots and the conduction associated with multiple 1D subbands. While it has been suggested that for GNRs with highly disordered edges, the transport behavior may become dominated by the former, our experimental results here cannot be accounted for by this Coulomb-blockade-(CB)-dominated transport mechanism, as explained below. First, in a disordered GNR, the renormalized charging energy E_C of a quantum dot is estimated ~ 5 meV for $W \approx 30$ nm.⁸ This small E_C would smear out any charging effects even at 50 K. In addition, we note that while the spacing ΔG between conductance plateaus decreases with increasing channel length L , the number of plateaus is roughly constant in the same gate voltage span (~ 20 V) for these GNRs of the same width. This is also incompatible with the CB-dominated transport, where both the number and the size of the quantum dots are expected to vary with the length. Furthermore, in the unlikely situation that the entire GNR channel is dominated by one single dot, the transport would be distinguished by a series of periodic conductance peaks as a function of gate voltage, whereas in a multiple-dot configuration, these peaks are expected to occur randomly due to the nature of the uncontrolled dot sizes. Nevertheless, neither could lead to such well-defined and reproducible conductance plateaus observed in multiple devices here. Therefore, we attribute the conductance plateaus in GNRs to the formation of 1D conduction modes in nanoribbons.

There have been numerous theoretical studies on the GNR electronic structure and its dependence on the ribbon width and the cutting angle using π -orbital tight-binding models^{12,13} and the first principle calculations.^{14,16} To a first-order approximation, the band structure of GNRs can be described by zone folding the graphene band structure. In a graphene nanoribbon of width W , the wave vector perpendicular to the transport direction has a quantization requirement $k_\perp W = \pi m$, giving rise to various 1D subbands, each with the dispersion relation given by

$$E_m(k_\parallel) = \pm \hbar v_f \sqrt{k_\parallel^2 + (m + \alpha)^2 \pi^2 / W^2}, \quad (1)$$

where $m=0, \pm 1, \pm 2, \dots$ is an integer for the subband index and k_\perp and k_\parallel are the wave vectors perpendicular and parallel to the transport direction, respectively. Here $0 \leq |\alpha| < 0.5$ depends on the crystallographic orientation of the GNR, and yields a band gap $E_g = 2|\alpha| \Delta E$, where $\Delta E = \hbar v_f \pi / W$. We note that in GNRs, the wave function has to vanish at the ribbon edges, in contrast to the periodic boundary condition in the case of carbon nanotubes (CNTs). This leads to a different quantization requirement in CNTs, namely, $k_\perp W_{\text{CNT}} = 2\pi m$, and as a result twice as large energy separation in CNTs for the same circumference length W . In addition, the 1D subbands of GNRs are singly degenerate as opposed to those of CNTs,¹⁷ where the orbital degeneracy associated

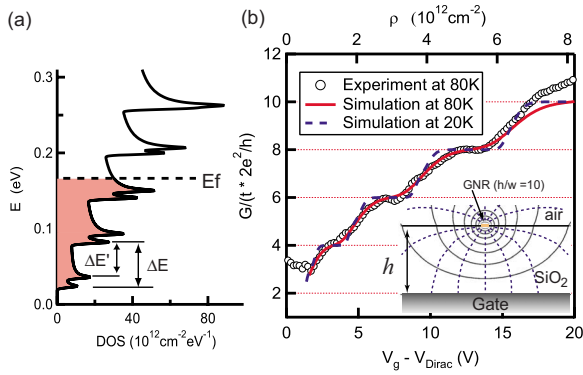


FIG. 3. (Color online) (a) The density of states of a graphene nanoribbon. (b) Calculated normalized conductance $G/(2e^2/h)$ as a function of carrier density for graphene nanoribbon devices of $W=31$ nm, using armchair GNR with $N=253$. t is the transmission coefficient for each subband. The inset illustrates schematically the field lines and the equipotential surfaces of the GNR, demonstrating a significant discrepancy from the case of a parallel-plate capacitor.

with the K-K' bands is lifted with the splitting determined by the GNR chirality.

For a more realistic model of the GNR band structure, we adopt the tight-binding method with the next-neighbor hopping $t_C=2.7$ eV. The Fermi velocity is related to the hopping integral t according to $v_f=3at_C/2\hbar \approx 10^6$ m/s, where $a \approx 0.246$ nm is a lattice constant of graphene. At the ribbon edge, we assume enhanced hopping between the neighboring C atoms by 12% to account for the C-C bond-length contraction.¹⁸ We find that in armchair GNRs (unrolled zigzag CNTs), $\alpha \approx 0.27, 0.4,$ and 0.066 in the families $N=3p, 3p+1,$ and $3p+2,$ respectively, where p is an integer and N is the number of dimer lines across the ribbon width. From Eq. (2), the K-K' orbital splitting $\delta E_{kk'}$ in families $N=3p+1$ and $N=3p+2$ is about $0.2\Delta E$ and $0.13\Delta E,$ respectively. This results in a series of closely spaced energy level pairs with an interpair spacing of $\Delta E' = \Delta E - \delta E_{kk'}$ [see Fig. 3(a)]. On the other hand, for the $N=3p$ family, the K-K' orbital splitting is $0.46\Delta E$ (or $0.54\Delta E$), such that the resulting subbands are roughly equally spaced with half the quantization energy $\Delta E.$

Using the Landauer approach, the device conductance at a finite temperature can be expressed by

$$G = (2e^2/h) \sum_i \int T_i(E) [- (\partial f_{\nu} / \partial E)] dE, \quad (2)$$

where $T_i(E)$ is the transmission probability of carriers in each subband at energy $E,$ and the factor 2 comes from the spin degeneracy. At zero temperature, Eq. (2) can be reduced to $G=2e^2/h \sum_i t_i$ where $t_i=T_i(E_f)$ and the summation includes all the 1D modes below the Fermi energy. While the transmission coefficient is a function of energy and the subband index,¹⁹ the quasiequidistant steps in measured conductance indicate a nearly constant transmission probability at low energies.

Now we compare simulations of the GNRs with the experimental results. For a 30-nm-wide GNR, we estimate the quantization energy ΔE to be 60 meV. From simulations

based on Eq. (2), we find that an energy spacing of at least $6kT$ is required in order to observe conductance quantization features at a given temperature $T.$ Since the conductance steps are clearly visible in Fig. 1 at 80 K, corresponding to a minimum-energy separation of 40 meV, the energy splitting at the K-K' bands must be smaller than 20 meV. These facts can be accounted for, within the tight-binding model, by adopting a $(3p+1)$ family armchair ribbon to model the electronic structure of our 30-nm-wide GNRs. Figure 3(a) shows the calculated density of states of an armchair GNR with $N=253,$ corresponding to $W=31$ nm. For this ribbon width, $\Delta E \approx 58$ meV and the separation between the nearly degenerate pairs of bands is $\Delta E' = 0.8\Delta E \approx 46$ meV. As shown in Fig. 3(b), we find an excellent agreement between the measurement and the simulation at 80 K. In order to experimentally estimate the band gap, we utilize the Arrhenius plot of the minimum conductance at the Dirac point as a function of temperature and extracted an energy gap of ~ 42 meV for one of our GNR devices. We note that this value is not only within the energy range $E_g \leq \Delta E$ but also in close agreement to the calculated $E_g = 2|\alpha| \cdot \Delta E = 46$ meV with $\alpha = 0.4.$

In Fig. 3(b), the step height $4e^2/h t$ corresponds to four conduction channels because the steps associated with individual occupation of nearly degenerate K-K' subbands, which are separated by only 11 meV, cannot be resolved at this temperature. From Fig. 3(b), the overall transmission probability, including the contributions from the channel and the contacts, is $t \approx 0.016$ for this GNR. The discrepancy between the simulated and measured conductances of the fourth plateau around $V_g \geq 17$ V may be due to the enhanced transmission probability at high energies. The transmission probability of the GNR, in principle, depends on the subband energy, which can be calculated by considering a potential barrier between the contact and the channel.¹⁹ However, for a reasonably small contact barrier height ≤ 0.2 eV, the transmission coefficient is found to be essentially constant for each subband, consistent with our experimental results.

The transmission coefficient in our graphene nanoribbons is much smaller than unity as in the case of perfect quantum point contacts, and t decreases with increasing channel length. This indicates that in addition to the contacts, the transmission coefficients are also affected by the channel quality. Although the origin of the small transmission coefficient in our graphene nanoribbons is currently unclear to us, the fact that the quantization phenomena are preserved in a nonballistic channel indicates the reduction in transmission coefficient is mainly caused by intraband scattering processes. We note that, since the device conductance at room temperature is comparable to its low-temperature value, the scattering mechanism is not temperature sensitive in these GNRs, and therefore, the dominant scattering process is likely due to the impurity scattering or edge-induced scattering in the channel.²⁰ Further theoretical studies are required to understand the impact of the realistic edges and the charged impurities on the transmission coefficient of graphene 1D channels.

Here we consider the gating efficiency of a back-gated graphene nanoribbon device, which is found to be very different from the case of 2D graphene. The carrier density in the GNR is related to the gate voltage as $\rho = C_g(V_g - V_{\text{Dirac}}),$

where C_g is a gate capacitance and V_{Dirac} is a gate voltage corresponding to the Dirac point. As shown by Fig. 3(b), we observe four conductance plateaus over a V_g span of 20 V, corresponding to a shift of the Fermi level by $E_f \approx 260$ meV and a change in carrier density of 8.4×10^{12} cm $^{-2}$. This is comparable to the carrier density in 2D graphene at the same Fermi energy E_f given by $\rho_{2D} = E_f^2 / \pi v_f^2$. The dependence of ρ on V_g leads to an effective gate capacitance of $C_{\text{GNR}}/e = 4.2 \times 10^{11}$ cm $^{-2}$ V $^{-1}$, which is significantly larger than the value expected for 2D graphene, given by $C_{2D}/e \approx 7.2 \times 10^{10}$ cm $^{-2}$ V $^{-1}$ for the gate dielectric of SiO $_2$ with thickness $h = 300$ nm. This difference is due to the fact that when the aspect ratio of the GNR width to the oxide thickness (w/h) becomes much smaller than one, the field lines between the GNR and the gate deviate from that of a parallel-plate capacitor [see inset of Fig. 3(b)], resulting in an increasing C_{GNR}/C_{2D} ratio as (w/h) decreases. We have numerically calculated the capacitance of the GNR of infinite length L in the back gate configuration where the SiO $_2$ dielectrics and the vacuum below and above the ribbon are explicitly considered. For $h/w = 300/30$, we find $C_{\text{GNR}}/C_0 \approx 10$,²¹ which is within a factor of 2 of the estimated values based on experiments.²²

In addition to the Fermi-level modulation by the gate control, the number of 1D conduction modes that contribute to the transport in GNRs also depends on the bias voltage between the source and drain contacts, yielding a V_d -dependent device conductance in GNRs.¹² Figure 4(a) shows measured conductance of a GNR device ($W = 30$ nm and $L = 1.7$ μm) as a function of V_g for two bias voltages V_d of 50 and 100 mV at 15 K, both exhibiting conductance plateaus with the same quantization spacing ΔG as V_g varies. However, in Fig. 4(a), it is noted that the conductance curve $G(V_g)$ at $V_d = 100$ mV is always higher than that of $V_d = 50$ mV by roughly one unit of quantized conductance ΔG for all gate voltages. This dependence of conductance on the drain bias can be understood through the schematic shown in Fig. 4(b).

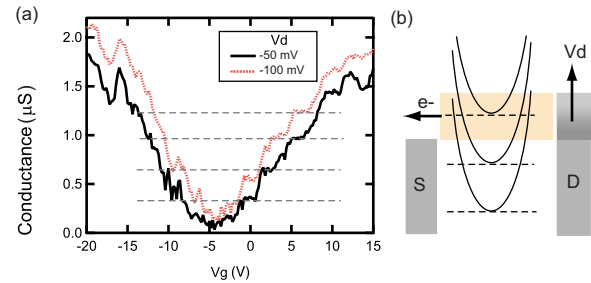


FIG. 4. (Color online) (a) Conductance of a GNR device measured as a function of gate voltage at $T = 15$ K for two drain biases, 50 and 100 mV. The GNR device has a channel length and width of 1.7 μm and 30 nm, respectively. (b) Schematic energy-band diagram of a GNR with a drain bias voltage applied. The conductance is determined by the number of 1D subbands accessible in the bias window. For a given gate voltage, the number of 1D modes contributing to the transport can be increased at a sufficiently large drain bias.

As the source-drain bias windows become larger than the energy spacing of the 1D subbands, the number of 1D conduction modes is increased by one regardless of the gate voltage, resulting in a higher conductance level by ΔG .

In conclusion, we have fabricated graphene nanoribbon devices and reported a experimental observation of subband formation in GNRs and conductance quantization phenomena at temperatures as high as 80 K and channel lengths up to 1.7 μm . We have performed temperature and bias dependence studies, and shown that the energy spacing between 1D subbands of 30-nm-wide GNRs is around 50 meV. The experimental results are in excellent agreement with theoretical calculations within the tight-binding approximation. The experimental findings here provide an important step toward developing graphene-based quantum devices.

The authors thank Bruce Ek for expert technical assistance.

*yming@us.ibm.com

¹K. S. Novoselov *et al.*, Nature (London) **438**, 197 (2005).

²A. K. Geim and K. S. Novoselov, Nature Mater. **6**, 183 (2007).

³A. H. Castro Neto *et al.*, arXiv:0709.1163, Rev. Mod. Phys. (to be published).

⁴P. Avouris *et al.*, Nat. Nanotechnol. **2**, 605 (2007).

⁵M. Y. Han *et al.*, Phys. Rev. Lett. **98**, 206805 (2007).

⁶Z. Chen *et al.*, Physica E **40**, 228 (2007).

⁷C. Stampfer *et al.*, Appl. Phys. Lett. **92**, 012102 (2008).

⁸F. Sols *et al.*, Phys. Rev. Lett. **99**, 166803 (2007).

⁹X. Wang *et al.*, Phys. Rev. Lett. **100**, 206803 (2008).

¹⁰K. Nakada *et al.*, Phys. Rev. B **54**, 17954 (1996).

¹¹S. Zhou *et al.*, Nature Mater. **6**, 770 (2007).

¹²N. M. R. Peres *et al.*, Phys. Rev. B **73**, 195411 (2006).

¹³D. Gunlycke *et al.*, Appl. Phys. Lett. **90**, 142104 (2007).

¹⁴J. Fernandez-Rossier *et al.*, Phys. Rev. B **75**, 205441 (2007).

¹⁵Y.-M. Lin and P. Avouris, Nano Lett. **8**, 2119 (2008).

¹⁶L. Yang *et al.*, Phys. Rev. Lett. **99**, 186801 (2007).

¹⁷C. T. White *et al.*, Nano Lett. **7**, 825 (2007).

¹⁸Y.-W. Son *et al.*, Phys. Rev. Lett. **97**, 216803 (2006).

¹⁹J. Tworzydło *et al.*, Phys. Rev. Lett. **96**, 246802 (2006).

²⁰J. H. Chen *et al.*, Nat. Nanotechnol. **4**, 377 (2008).

²¹The gate capacitance of GNRs can also be approximated by the following semi analytical approach. Assuming a homogeneous dielectric medium ϵ below and above the GNR, the capacitance due to the back gate can be expressed by $\frac{C_0}{C} = \frac{2}{\pi} \arctan \frac{W}{4h} + \frac{W}{4h\pi} \ln(1 + \frac{16h^2}{W^2})$, where $C_0 = \epsilon WL/h$. The solution assumes a constant charge density on the GNR. The gate capacitance of the device is then approximated by $C_{\text{GNR}} \approx (C_0^{\text{air}} + C_0^{\text{SiO}_2})/2$, yielding an enhancement factor of 9.

²²The actual gate capacitance of the GNR is expected to be smaller than the numerical simulations because of (i) the nonideal dielectrics used and (ii) the screening from source and drain electrodes.