Zener tunneling in semiconducting nanotube and graphene nanoribbon *p−n* junctions

Debdeep Jena,[a)] Tian Fang, Qin Zhang, and Huili Xing
Department of Electrical Engineering, University of Notre Dame, Indiana 46556, USA

A theory is developed for interband tunneling in semiconducting carbon nanotube and graphene nanoribbon *p−n* junction diodes. Characteristic length and energy scales that dictate the tunneling probabilities and currents are evaluated. By comparing the Zener tunneling processes in these structures to traditional Group IV and III–V semiconductors, it is proved that for identical bandgaps, carbon-based one-dimensional (1D) structures have higher tunneling currents. The high tunneling current magnitudes for 1D carbon structures suggest the distinct feasibility of high-performance tunneling-based field-effect transistors. © *2008 American Institute of Physics.*
[DOI: 10.1063/1.2983744]

Carbon-based one-dimensional (1D) materials such as nanotubes and graphene nanoribbons (GNRs) are currently under extensive investigation for the fundamental physics they exhibit as well as for possible applications they might have in the future.[1,2] A large class of traditional semiconductor devices rely on the quantum mechanical tunneling of carriers through classically forbidden barriers. Among these, the Esaki diode and the resonant tunneling diode are the prime examples.[3] The phenomena of tunneling have been studied extensively for traditional parabolic-bandgap semiconductors in three-dimensional (3D) bulk as well as for quasi-two-dimensional (2D) and quasi-1D heterostructures. Semiconducting carbon nanotubes (CNTs) and GNRs do not have parabolic bandstructures, and the carrier transport in them approaches the ideal 1D case. In that light, it is timely to examine the phenomena of tunneling in these materials.

Tunneling currents in semiconducting CNT *p−n* junctions have been measured and analyzed recently (see Refs. 4–6). For tunneling probabilities, an energy-dependent carrier effective mass has been used[7] to take advantage of previously existing results of parabolic bandstructure semiconductors. In this work, we evaluate the Zener tunneling probabilities of CNT and GNR based *p−n* diodes starting from their intrinsic bandstructures, which removes the need to define an effective mass. In addition to interband tunneling probabilities, a number of fundamental associated parameters characterizing the tunneling process are found.

The bandstructure of the *n*th subband of a semiconducting CNT or GNR is given by[8]

$$\mathcal{E} = s\hbar v_F \sqrt{k_x^2 + k_n^2}, \qquad (1)$$

where $2\pi\hbar$ is Planck's constant and $v_F \sim 10^8$ cm/s is the Fermi velocity characterizing the bandstructure of graphene. $s=+1$ denotes the conduction band and $s=-1$ denotes the valence band. The electron momentum along the CNT or GNR axis is $\hbar k_x$.

For GNRs, the transverse momentum is quantized by the ribbon width,[9] $k_n = n\pi/3W$, where $n = \pm 1, \pm 2, \pm 4, \pm 5, \pm 7, \pm 8, \ldots$ for a GNR of dimensions $(x, y) = (L, W)$, where $W \ll L$. The corresponding bandgap is

$\mathcal{E}_g = 2\hbar v_F k_1 = 2\pi\hbar v_F/3W \sim 1.3/W$ eV (where W is in nanometers). For comparison, a semiconducting CNT of diameter D has the same bandstructure with $k_1 = 2/3D$ and a bandgap of $\mathcal{E}_g = 4\hbar v_F/3D \sim 0.8/D$ eV, where D is in nanometers. If $W = \pi D/2$, the properties (bandgap and bandstructure) of semiconducting CNTs and GNRs are similar. The results derived below are applicable to GNRs and CNTs on equal footing. It is assumed that the length of the GNR (CNT) is much larger than the width (diameter) such that the longitudinal momentum of carriers in the ribbon is quasicontinuous.

We now evaluate the interband tunneling probability in a $n^+ - p^+$ GNR or CNT diode of bandgap $\mathcal{E}_g$. We consider that the doping of the *n*- and *p*-sides is such that the equilibrium Fermi level is at the conduction band edge ($\mathcal{E}_c$) in the *n*-side and at the valence band edge ($\mathcal{E}_v$) on the *p*-side. Such doping could be either chemical or electrostatic.[1] Under this situation, a forward bias would not lead to current flow (this is similar to the "backward diode"[10]). When a reverse-bias eV is applied, the band diagram looks as shown in Fig. 1. Let the electric field in the junction region be $\mathcal{F}$. We assume that the depletion region thickness and the net electric field do not change appreciably from the equilibrium values (true under small bias voltages). Then, the potential energy barrier seen
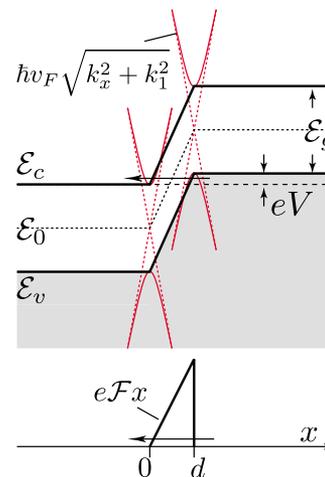

FIG. 1. (Color online) Interband tunneling in a reverse-biased GNR/CNT *p−n* junction and the potential barrier seen by tunneling electrons.
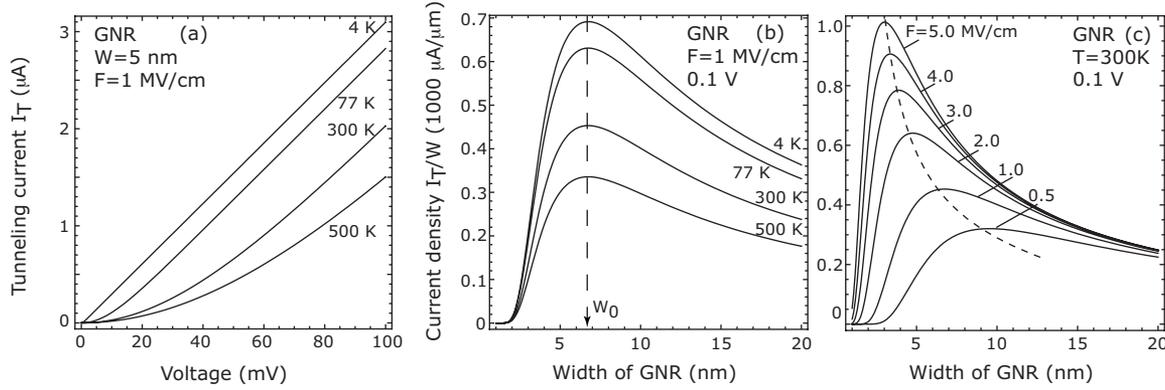
FIG. 2. Tunneling currents in GNR $p-n$ junctions: (a) $W=5$ nm device, voltage dependence at various temperatures. Tunneling currents per unit width for different GNR widths at (b) various temperatures at $\mathcal{F}=1$ MV/cm and (c) for various $\mathcal{F}$ at 300 K. To maximize the tunneling current density, an optimum $(\mathcal{F}, W)$ combination exists as illustrated by the dashed lines.

by electrons in the valence band of the p^+ side is

$$V_0(x) = e\mathcal{F}x \tag{2}$$

in the range $0 < x < d$ such that $d = \mathcal{E}_g/e\mathcal{F}$. This is indicated in Fig. 2. $\mathcal{E}_0$ is the Dirac point of the underlying graphene bandstructure for both GNRs and CNTs and serves as a convenient reference of energy in the problem.

Since the energy of carriers near the band edge ($k_x \approx 0$) is conserved during the tunneling of electrons from the valence to the conduction band, the condition

$$-\sqrt{(\hbar v_F k_x)^2 + \left(\frac{\mathcal{E}_g}{2}\right)^2} + \mathcal{E}_g - e\mathcal{F}x$$

$$= +\sqrt{(\hbar v_F k_x)^2 + \left(\frac{\mathcal{E}_g}{2}\right)^2} \tag{3}$$

holds for the wavevector k_x at all x. Within the tunneling barrier, the wavevector is imaginary. Denoting this by $k_x = i\kappa_x$, where κ_x is real, we obtain

$$\kappa_x(x) = k_0\sqrt{1 - \left(1 - \frac{x}{d}\right)^2}, \tag{4}$$

where $k_0 = \mathcal{E}_g/2\hbar v_F$ is a characteristic wavevector for tunneling. Since $\int_0^d \sqrt{1 - (1 - \frac{x}{d})^2} = \pi d/4$, the Wentzel-Kramer-Brillouin (WKB) band-to-band tunneling (BTBT) probability for the $p^+ - n^+$ junction given by $T \sim \exp[-2|\int_0^{+d} \kappa_x(x)dx|]$ leads to

$$T_{\text{WKB}}(k_x \approx 0) \sim \exp\left[-\frac{\pi}{4} \cdot \frac{\mathcal{E}_g^2}{\hbar v_F e\mathcal{F}}\right], \tag{5}$$

which can be expressed as $T_{\text{WKB}} \sim \exp(-\mathcal{F}_0/\mathcal{F})$, where $\mathcal{F}_0 = \pi\mathcal{E}_g^2/4e\hbar v_F$ denotes the characteristic electric field at the junction for the onset of strong tunneling. The corresponding characteristic barrier thickness for the onset of strong tunneling is $d_0 \sim \mathcal{E}_g/e\mathcal{F}_0 = 4\hbar v_F/\pi\mathcal{E}_g$.

Using the value of the Fermi velocity, the characteristic field evaluates to $\mathcal{F}_0 \sim 12.6 \times (\mathcal{E}_g)^2$ MV/cm and the characteristic tunneling distance evaluates to $d_0 \sim 0.8/\mathcal{E}_g$ nm, where $\mathcal{E}_g$ is the bandgap of the GNR or CNT expressed in electron volt in both these expressions. As an example, for a GNR with $W=5$ nm, $\mathcal{E}_g \sim 0.275$ eV, the characteristic tunneling field is ~ 0.9 MV/cm and the characteristic tunneling barrier thickness is ~ 11 nm.

The tunneling probabilities evaluated above are for the first subband ($n=1$). For the nth conduction and valence subbands of CNTs and GNRs, the effective subbandgap scales as $\sim n\mathcal{E}_g$. The tunneling probabilities of carriers near the respective subband edges are then given by $T_{\text{WKB},n} \sim \exp(-\pi\hbar v_F k_n^2/\mathcal{F})$, which decay as $\exp(-n^2)$, indicating a rather strong damping of the tunneling probabilities of higher subbands. This result turns out to be identical to one for the nth transverse mode of a zero-gap 2D graphene $p-n$ junction, as first evaluated by Cheianov and Fal'ko (see Ref. 11). Except for the narrowest bandgap CNTs and GNRs, the tunneling is primarily from the first valence subband in the p-side to the first conduction subband on the n-side.

Note that the above derivation uses a triangular barrier approximation. It has been shown by Kane[12] that a parabolic barrier more accurately represents the physics of the tunneling process and leads to an exponential factor with a different coefficient than from the triangular barrier approximation. The difference is small. For the rest of this work, we will use the result derived above.

The tunneling current may now be written in the form $I_T = e(g_s g_v/L)\Sigma_{k_x}[f_v - f_c]T_{\text{WKB}} \times v_g(k_x)$, where $g_s = 2$ is the spin degeneracy, g_v is the valley degeneracy [=2 for CNTs and =1 for GNRs (Ref. 1)], L is the length of the CNT or GNR, f_v, f_c are the occupation functions of the valence band states in the p-type side and the conduction band in the n-side, respectively, and $v_g(k_x)$ is the group velocity. The summation over the k_x states $\Sigma_{k_x}(\ldots)v_g(k_x)$ can be converted into an integral over energy by using

$$\sum_{k_x}(\ldots)v_g(k_x) \rightarrow L/(2\pi)\int dk_x(\ldots)v_g(k_x)$$

$$\rightarrow L/(2\pi)\int d\mathcal{E}(\ldots)v_g(k_x)(\partial k_x/\partial \mathcal{E}).$$

Since $v_g(k_x) = \hbar^{-1}\partial\mathcal{E}/\partial k_x$ for any bandgap of the CNT or GNR, the tunneling current is then given by the equivalent integral

$$I_T = \frac{2g_v e}{h}\int_0^{eV}[f_v(\mathcal{E}) - f_c(\mathcal{E})]T_{\text{WKB}}d\mathcal{E}, \tag{6}$$

where $f_v(\mathcal{E}) = 1/(1 + \exp[(\mathcal{E} - eV)/k_BT])$ and $f_c(\mathcal{E}) = 1/(1 + \exp[\mathcal{E}/k_BT])$. Here, k_B is the Boltzmann constant. For the

band diagram shown in Fig. 2, the net tunneling current evaluates to

$$I_T = \frac{2g_v e^2}{h} T_{\text{WKB}} \times V_T \ln\left[\frac{1}{2}\left(1 + \cosh\frac{V}{V_T}\right)\right], \quad (7)$$

where $V_T = k_B T/e$. This expression captures the temperature and bias voltage dependence of the tunneling current. If the applied bias is much greater than the thermal energy ($V \gg V_T$), the current reduces to a form similar to the Landauer expression $I_T \approx (2g_v e^2/h)T_{\text{WKB}}(V - V_T \ln 4)$. Thus, the tunneling current has a negative temperature coefficient at high bias conditions. The dependence of tunneling currents in GNRs on voltage and temperature is illustrated in Fig. 2(a).[13]

The tunneling current *per unit width* of GNRs is maximized for $W_0 = \sqrt{2\pi^3 \hbar v_F / 9e\mathcal{F}}$. For example, for $\mathcal{F} \sim 1$ MV/cm, the GNR width for maximum current drive is $W_0 \sim 6.5$ nm and the current density at that width is ~ 450 μA/μm, comparable to traditional field-effect transistors (FETs). For thinner ribbons at higher fields, the current densities can approach 1000 μA/μm, which is a typical benchmark for Si metal-oxide semiconductor FETs. The dependence of the current densities on temperature, electric field, and GNR widths is shown in Figs. 2(b) and 2(c).

The BTBT probability in traditional parabolic band-structure semiconductors in the triangular barrier approximation depends on the bandgap as[3] $T_{\text{parabolic}} \sim \exp(-4\sqrt{2m^\star}\mathcal{E}_g^{3/2}/3e\hbar\mathcal{F})$, where $m^\star$ is a reduced carrier effective mass. The probability for GNRs and CNTs retains the same dependence on electric field. The effective mass does not appear in the tunneling probability of the GNR or CNT diodes since their bandstructure is not parabolic at any energy. If one compares the Zener tunneling probabilities in diodes made of CNTs or GNRs with other direct-bandgap semiconductors of the *same* bandgap, then the ratio

$$\frac{T_{\text{carbon}}}{T_{\text{parabolic}}} \sim \exp\left[-\frac{\mathcal{E}_g^{3/2}}{e\hbar\mathcal{F}}\left(\frac{\pi\sqrt{\mathcal{E}_g}}{4v_F} - \frac{4\sqrt{2m^\star}}{3}\right)\right] \quad (8)$$

indicates that the GNR or CNT $p-n$ diode will have a higher interband tunneling probability if the relation $\mathcal{E}_g < (16v_F/3\pi)^2 \times 2m^\star$ is satisfied. From the **k·p** theory for traditional direct-bandgap semiconductors, the electron effective mass (in the conduction band) is related to the bandgap by the approximation[14] $m_c^\star \approx (\mathcal{E}_g/20)m_0$, where the bandgap is in electron volts and m_0 is the free electron mass. This leads to the requirement

$$m_0 v_F^2 > \left(\frac{3\pi}{16}\right)^2 \times 10 \text{ eV}, \quad (9)$$

which is satisfied since the left-hand side is $m_0 v_F^2 \approx 5.7$ eV and the right-hand side is 3.5 eV. By this comparison, the CNT or GNR $p-n$ diode will have a higher reverse-bias Zener tunneling probability than a traditional semiconductor of the same bandgap. However, the effective mass used for BTBT probability in traditional semiconductors is typically a *reduced* effective mass, which for narrow-gap semiconductors is equal to $m_r^\star \approx m_c^\star/2$. In that case, the WKB probability

of the GNR is slightly smaller. The above analysis shows that the WKB tunneling probabilities of GNRs and traditional narrow-gap semiconductors are similar in magnitude when they are bandgap matched.

However, two facts tilt the tunneling *currents* decisively in favor of CNTs and GNRs. First, for bulk 3D $p-n$ junctions, the transverse kinetic energy of carriers can be large and leads to a further exponential decrease in carrier tunneling probability,[3] which is avoided in 1D structures. Second, if normal parabolic-bandgap semiconductors are shrunk to length scales comparable to those of CNTs and GNRs, their bandgaps and effective masses increase further due to quantum confinement.

Although Zener tunneling currents are detrimental in traditional devices such as rectifiers, field-effect, and bipolar transistors, it is important to note that the fundamental switching action in these devices is controlled by thermionic emission over barriers, which requires a minimum of $(k_B T/e)\ln(10) \sim 60$ mV per decade change in current at 300 K (the "classical" limit). However, a crop of tunneling FETs has been recently proposed and demonstrated,[4,15,16] which rely on the very mechanism studied in this work. These devices are capable of reaching far below the classical limit for switching by exploiting quantum mechanical tunneling. It is for such devices that high interband tunneling current drives in carbon-based 1D nanostructures hold a distinct advantage and much promise in the future.

[1]P. Avouris, Z. Chen, and V. Perebeinos, Nat. Nanotechnol. **2**, 605 (2007).
[2]A. K. Geim and K. S. Novoselov, Nat. Mater. **6**, 183 (2007).
[3]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), pp. 513–565.
[4]J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris, Phys. Rev. Lett. **93**, 196805 (2004).
[5]J. Appenzeller, M. Radosavljevic, J. Knoch, and P. Avouris, Phys. Rev. Lett. **92**, 048301 (2004).
[6]K. Bosnick, A. Gabor, and P. McEuen, Appl. Phys. Lett. **89**, 163121 (2006).
[7]B. Bourlon, D. C. Glattli, P. Placais, J. M. Berroir, C. Miko, L. Forro, and A. Bachtold, Phys. Rev. Lett. **92**, 026804 (2004).
[8]T. Fang, A. Konar, H. Xing, and D. Jena, Appl. Phys. Lett. **91**, 092109 (2007).
[9]L. Brey and H. A. Fertig, Phys. Rev. B **73**, 235411 (2006).
[10]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), pp. 537–539.
[11]V. V. Cheianov and V. Fal'ko, Phys. Rev. B **74**, 041403(R) (2006).
[12]E. O. Kane, J. Appl. Phys. **32**, 83 (1961).
[13]The temperature dependence of the bandgap of GNRs has been neglected [for CNTs, it varies $\leq \pm 10$ meV from 0–300 K as shown in R. Capaz, C. D. Spataru, P. Tangney, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **94**, 036801 (2005)].
[14]P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, 2nd ed. (Springer, New York, 1999), pp. 64–66.
[15]Q. Zhang, W. Zhao, and A. Seabaugh, IEEE Electron Device Lett. **27**, 297 (2006).
[16]W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. King, IEEE Electron Device Lett. **28**, 743 (2007).