Electron-hole duality during band-to-band tunneling process in graphene-nanoribbon tunnel-field-effect-transistors

Deblina Sarkar,[a] Michael Krall, and Kaustav Banerjee[a]
Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106, USA

(Received 3 October 2010; accepted 26 November 2010; published online 29 December 2010)

This letter provides clear insight into the interplay between electron and hole characteristics of carriers within the forbidden gap during the band-to-band tunneling process, taking graphene-nanoribbons as an example. Accurate numerical models are presented and analytical formulas for tunneling probabilities are derived for both source/drain to channel and direct source-drain tunneling based on the Wentzel–Kramers–Brillouin (WKB) method. It is shown that not considering the electron-hole duality can lead to significant errors in numerical calculations, and more importantly, lack of proper understanding of the phenomenon gives rise to seriously misleading conclusions. Furthermore, the regime of validity of the WKB approximation for graphene-nanoribbon tunnel-field-effect-transistors is discussed in light of the electron-hole duality concept. © *2010 American Institute of Physics*. [doi:10.1063/1.3528338]

Recently, the tunnel-field-effect-transistors (TFETs), employing band-to-band tunneling (BTBT), have attracted a lot of attention[1–6] due to their steep subthreshold slope characteristics. Though rigorous quantum mechanical treatment of tunneling can be achieved through the *nonequilibrium Green's function* (NEGF) formalism, computationally efficient models are required for fast calculations. WKB approximation provides such a tool, through which the transmission coefficient can be calculated from the wave-vector, which is dependent on the type of barrier. Often the barrier for tunneling is considered to be E_C-E,[1,2,7–11] which is the barrier for electrons, or E-E_V,[12] which is the barrier for holes. Such consideration is not only conceptually incorrect but can also lead to substantial errors in numerical calculations, as shown in this paper.

An analytical formula for the wave-vector within the forbidden gap has been developed by Kane[13] with the assumption of identical effective masses for electrons (m_e^*) and holes (m_h^*). However, a common incorrect use of Kane's formula when m_e^* and m_h^* are significantly different[3–5] arises from the lack of appreciation of the electron-hole participation during the tunneling process. The reduced mass[14,15] gives equal weight to electron and hole contributions to tunneling, which is also not correct since the proportion should obviously depend on the respective effective masses. Flietner's formula[16] is applicable for $m_e^* \neq m_h^*$ and reduces to that of Kane for $m_e^*=m_h^*$, but it has been developed specifically for materials with parabolic band-structures. For graphene based materials with nonconventional band-structures, tunneling probability for constant force has been reported[17] using the simplified E-k relation. However, lack of perception of electron and hole interplays has led to incorrect conclusions.[17]

In this letter, we illustrate the electron-hole duality (EHD) within the forbidden gap during the BTBT process taking a graphene-nanoribbon (GNR) TFET as an example. The dispersion relation of GNRs can be derived from the mode space Hamiltonian as[18] $E=\sqrt{t_0^2+t_n^2-2t_0t_n\cos(3k_xa/2)}$ where $t_0=-2.77$ eV, a is the C–C bond length, and t_n is related to the quantized y directed wave-vector k_{yn},[19] as $t_n=t_0[1+\exp(i\sqrt{3}k_{yn}a/2)]$ giving rise to a bandgap of $E_G=2\sqrt{t_0^2+t_n^2-2t_0t_n}$. Under the influence of two different forces, as shown in Fig. 1(a), the barrier for electrons (E_C-E) and holes (E-E_V) is different and thus would lead to different tunneling probabilities. Let us first consider only electron tunneling, in which case we can write

$$\sqrt{t_0^2+t_n^2-2t_0t_n\cos\left(\frac{3k_ea}{2}\right)}-\frac{E_G}{2}=-(E_C-E) \quad \text{for} \quad 0<x<x_0. \tag{1}$$

Here, the subscript x for the wave-vector (k_e) has been dropped for convenience and henceforth any k used will refer to the x directed momentum. $x=0$ and $x=x_0$ are defined as the turning points at which $E_C=E$ and $E_V=E$, respectively, as shown in Fig. 1(a). Thus E_C-E increases from 0 to E_G as x increases from 0 to x_0. However, an interesting situation occurs for E_C-$E>E_G/2$ for which we observe that the square root term on the left hand side of Eq. (1) becomes negative,

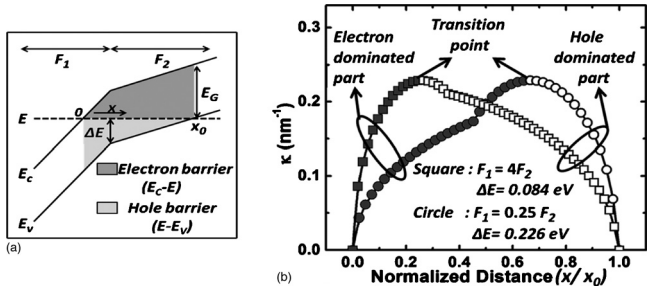

FIG. 1. (a) Band diagram of a graphene-nanoribbon under the influence of two different forces (F_1 and F_2). (b) Imaginary part of the wave-vector (κ) as a function of normalized distance between the two turning points ($x=0$ and $x=x_0$). A GNR with 38 carbon atoms along the width, giving a bandgap of 0.27 eV is considered. The squares represent κ for $F_1=20$ MV/m, $F_2=5$ MV/m, and $\Delta E=0.084$ eV, while the circles represent that for $F_1=5$ MV/m, $F_2=20$ MV/m, and $\Delta E=0.226$ eV. When $F_1>F_2$, the barrier for electrons increases steeply and hence reduces the electron dominated part of the wave-vector, and the transition point shifts to the left. Reverse situation occurs for $F_1<F_2$.

[a]Authors to whom correspondence should be addressed. Electronic addresses: deblina@ece.ucsb.edu and kaustav@ece.ucsb.edu.

which is unphysical. The reason is the following. Let us define the point at which E_C-E=$E_G/2$ as x=x_t. From 0 to x_t, the barrier for electrons is less than that for holes, and hence the electron properties should dominate. However, for $x>x_t$, the situation reverses and the use of the equation for electron tunneling in this region gives rise to the unphysical situation. The correct equations considering electron-hole duality are

$$\sqrt{t_0^2+t_n^2-2t_0t_n\cos\left(\frac{3k_ea}{2}\right)}-\frac{E_G}{2}=-(E_C-E)$$

$$\text{for}\quad 0<x\le x_t, \tag{2a}$$

$$-\sqrt{t_0^2+t_n^2-2t_0t_n\cos\left(\frac{3k_ha}{2}\right)}+\frac{E_G}{2}=E-E_V$$

$$\text{for}\quad x_t\le x<x_0. \tag{2b}$$

Using the WKB approximation, the effective tunneling probability (T_{EHD}) can then be written as

$$T_{\text{EHD}}=\exp\left\{-2\left(\int_0^{x_t}\kappa_e dx+\int_{x_t}^{x_0}\kappa_h dx\right)\right\};\quad k_e=i\kappa_e,\quad k_h=i\kappa_h. \tag{3}$$

In Fig. 1(b), the EHD within the forbidden gap is illustrated.

An analytical formula for the tunneling probability in case of a constant force can be derived by assuming a simplified conical dispersion relation near the Dirac points.[21] Employing the EHD concept we can write

$$\sqrt{(\hbar v_f k_e)^2+\left(\frac{E_G}{2}\right)^2}-\frac{E_G}{2}=-qFx\quad\text{for}\quad 0<x\le x_t, \tag{4a}$$

$$-\sqrt{(\hbar v_f k_h)^2+\left(\frac{E_G}{2}\right)^2}+\frac{E_G}{2}=E_G-qFx$$

$$\text{for}\quad x_t\le x<x_0. \tag{4b}$$

For a constant force, the transition point occurs in the middle, i.e., x_t=$x_0/2$. Using Eqs. (3) and (4) we obtain T_{EHD}=$\exp(-\pi E_G^2/4v_fq\hbar F)$. It is instructive to note that due to some miscalculations, Jena *et al.*[17] arrived at the same expression for the tunneling probability as in T_{EHD} above, even after using a triangular barrier (i.e., barrier only for electrons), which led to completely erroneous conclusions in the comparison between GNR and III-V based TFETs. However, the error does not arise because of the reason offered by Vandenberghe *et al.*,[22] who attributed the assumption of parabolic nature of band-structure in the entire Brillouin zone to be the cause. The root of the error is due to the fact that Jena *et al.*[17] compared the tunneling probability improperly derived with the consideration of electron barrier of GNRs with the T_{electron} of III-V materials.

Figure 2 shows the tunneling probability for the case of a constant force [e.g., F_1=F_2=F in Fig. 1(a)] applied to the GNR. The EHD model shows close match with the rigorous quantum mechanical simulation using NEGF. Use of the barrier of the form given by Flietner in Eq. (1) or consideration of either electron or hole tunneling can give rise to significant errors.

To take into account the tunneling between source (or drain) and channel in a TFET, the band bending can be approximately taken to be that due to a constant force. However, this approximation is not valid for the case of direct tunneling between source and drain [Fig. 3(a)]. Figure 3(b) shows the accuracy of the EHD model in calculating the direct source-to-drain tunneling probability. Here, we also derive an analytical expression for it. For that we divide the area between the drain and source into three distinct regions [R_1, R_2, and R_3 in Fig. 3(a)]. We assume constant force F_1 and F_3 in the drain-channel (R_1) and source-channel (R_3) junction, respectively. We assume the bands in the region (R_2) between R_1 and R_3 to be flat. Now using the EHD model, we can find the wave-vector in the three regions, and the tunneling probability (T_{SD}) can be written as

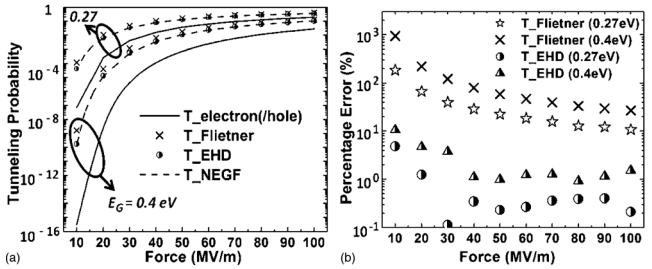

FIG. 2. (a) Tunneling probability as a function of the applied force. $T_{\text{electron(/hole)}}$ is calculated using the effective mass approximation (Ref. 20) [as the use of E-k relation for either electron (hole) tunneling gives rise to the unphysical situation], taking E_C-E (E-E_V) as the barrier. It is clear that using only the electron or the hole barrier leads to considerable deviations from NEGF calculations. (b) The percentage errors for the EHD model and Flietner's model compared to NEGF simulations are plotted separately for clarity for two different values of E_G (0.27 and 0.4 eV). It is seen that the EHD model, expressed as Eqs. (2) and (3), shows very good agreement with simulations based on the NEGF; while the use of Flietner's form of barrier can lead to significant errors. All models have been numerically solved to provide meaningful comparison.

$$T_{SD}=\exp\left\{-2\left(\int_0^{x_1(E)}\kappa_1 dx+\int_0^{L_2}\kappa_2 dx+\int_0^{x_3(E)}\kappa_3 dx\right)\right\}. \tag{5}$$

It is to be noted that the lengths of integration in R_1 (x_1) and R_3 (x_3) are energy dependent. To find x_1, we note from Fig. 3(a) that traversing from A-B-C brings us to the same energy. To find x_3, we move along M-N-O-P. Hence we can write

$$qF_1x_1(E)-E_G+E=0, \tag{6a}$$

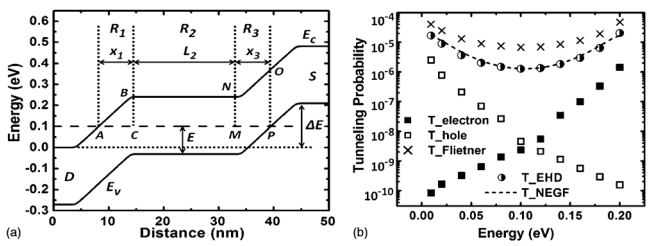

FIG. 3. (a) Band-structure during direct tunneling between source and drain in a GNR TFET. E_G is taken to be 0.27 eV and F_1=F_2=24 MV/m. (b) Tunneling probability as a function of the energy level E shown in (a). The EHD model, expressed as Eq. (5), shows close match with the NEGF simulations. Consideration of either electron or hole tunneling underestimates the tunneling probability by several orders of magnitude, while the use of Flietner's formula leads to considerable overestimation of the same. All models have been numerically solved to provide meaningful comparison.

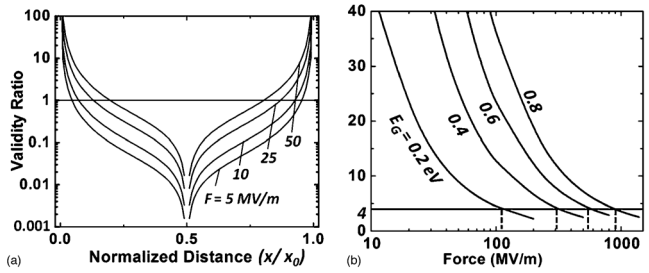

FIG. 4. (a) The factor $k^{-2}dk/dx$, which we define as WKB validity ratio for different forces applied to a GNR (E_G=0.4 eV) as a function of the normalized distance between the two turning points. As force increases, the region where the validity condition is not satisfied increases. (b) Regime in which WKB can be used, as a function of E_G. The projection of the points at which the curve for a particular bandgap intersects the line y=4; on the x-axis define the maximum force up to which WKB would be applicable for that bandgap.

$$E_G - E + qF_3x_3(E) - E_G = 0. \tag{6b}$$

Let I_1, I_2, and I_3 represent the first, second, and third integrations in Eq. (5), respectively. Using Eqs. (2), (5), and (6) we derive the expressions for I_1, I_2, and I_3 and finally for the tunneling probability as

$$I_1 = \frac{1}{\hbar v_f}\frac{1}{qF_1}\left\{\frac{\pi E_G^2}{16} + \frac{\sqrt{(E_G-E)E}(E_G-2E)}{4} + \frac{E_G^2}{8}\sin^{-1}\left(\frac{E_G-2E}{E_G}\right)\right\},$$

$$I_2 = \frac{L_2}{\hbar v_f}\sqrt{(E_G-E)E},$$

$$I_3 = \frac{1}{\hbar v_f}\frac{1}{qF_3}\left\{\frac{\pi E_G^2}{16} - \frac{\sqrt{(E_G-E)E}(E_G-2E)}{4} - \frac{E_G^2}{8}\sin^{-1}\left(\frac{E_G-2E}{E_G}\right)\right\},$$

$$T_{SD} = \exp\{-2(I_1+I_2+I_3)\}. \tag{7}$$

While the above analytical formulas can provide better physical insight, more accurate results can be obtained by using direct numerical solutions to Eq. (3) or Eq. (5), as employed in Figs. 2 and 3(b).

Finally, it is important to note that the large error in the tunneling probability attributed to the use of the WKB method[6] is mainly due to its improper use. It is shown here that with proper understanding of the EHD concept and use of accurate E-k relation, WKB can actually yield results in very good agreement with rigorous NEGF calculations. However, since the effective barrier for tunneling is reduced due to the interplay of both electrons and holes compared to that due to consideration of only one of them, it is important to know the regime of validity of the WKB method in light of the EHD concept. The WKB equation is derived under the condition that $k^{-2}dk/dx \ll 1$. It can be observed from Fig. 4(a) that the required condition for WKB is not satisfied in the regions near the turning points and that this region expands with increasing force. However, it has been argued that the errors introduced by the use of WKB will not be significant for any arbitrary barrier given that the condition $x_0\bar{\phi}^{1/2} \geq 4$ holds[23] where x_0 is the barrier width in Å and $\bar{\phi}$ is the mean barrier height in eV. Since consideration of either electron or hole tunneling overestimates the tunneling barrier, it will also overestimate the regime of applicability of WKB while the actual case considering EHD is shown in Fig. 4(b), as a function of bandgap and applied force.

In summary, we have highlighted an underlying physical concept behind the BTBT process that has been mostly overlooked in literature. It has been shown that ignoring the dual nature of electrons and holes during the BTBT phenomenon can not only lead to substantially erroneous results but also to misleading conclusions. During the tunneling phenomenon, particles transit through a forbidden gap with imaginary wave-vectors. If only electron (hole) tunneling and thereby only the barrier for electrons (holes) are considered, it implies that the electron (hole) faces a barrier at the valence (conduction) band edge, and hence its wave-vector becomes imaginary there, which is unphysical because the wave-vector can be imaginary only within the bandgap and not on the valence/conduction band edge. Understanding the EHD concept is also vital for accurately determining the regime of validity of the WKB method and can thereby dispel criticisms that seem to have smeared its reputation. While all the results in this paper are presented for GNRs, the EHD concept is inherent to the physics of the BTBT process and hence is applicable to any other material.

This work was supported by the National Science Foundation, Grant No. CCF-0811880.

[1] J. Knoch and J. Appenzeller, Device Research Conference, 2005, p. 153.
[2] M. T. Björk, J. Knoch, H. Schmid, H. Riel, and W. Riess, Appl. Phys. Lett. **92**, 193504 (2008).
[3] K. K. Bhuwalka, S. Sedlmaier, A. K. Ludsteck, C. Tolksdorf, J. Schulze, and I. Eisele, IEEE Trans. Electron Devices **51**, 279 (2004).
[4] S. H. Kim, H. Kam, C. Hu, and T.-J. K. Liu, Dig. Tech. Pap. - Symp. VLSI Technol. **2009**, 178.
[5] N. N. Mojumder and K. Roy, IEEE Trans. Electron Devices **56**, 2193 (2009).
[6] Y. Gao, T. Low, and M. Lundstrom, Dig. Tech. Pap. - Symp. VLSI Technol. **2009**, 180.
[7] S. M. Sze, *Physics of Semiconductor Devices*, 3rd ed. (Wiley, New York, 2007).
[8] S. Banerjee, W. Richardson, J. Coleman, and A. Chatterjee, IEEE Electron Device Lett. **8**, 347 (1987).
[9] R. B. Fair and H. W. Wivell, IEEE Trans. Electron Devices **23**, 512 (1976).
[10] Q. Zhang, W. Zhao, and A. Seabaugh, IEEE Electron Device Lett. **27**, 297 (2006).
[11] K. Boucart and A. Ionescu, IEEE Trans. Electron Devices **54**, 1725 (2007).
[12] D. Jimenez, X. Cartoix, E. Miranda, J. Sune, F. A. Chaves, and S. Roche, J. Comput. Electron. **5**, 361 (2006).
[13] E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).
[14] R. A. Logan and A. G. Chynoweth, Phys. Rev. **131**, 89 (1963).
[15] L. Wang and P. Asbeck, IEEE Nanotechnology Materials Device Conference, 2009, p. 196.
[16] H. Flietner, Phys. Status Solidi **54**, 201 (1972).
[17] D. Jena, T. Fang, Q. Zhang, and H. Xing, Appl. Phys. Lett. **93**, 112106 (2008).
[18] J. Guo, S. Datta, M. Lundstrom, and M. P. Anantram, Int. J. Multiscale Comp. Eng. **2**, 257 (2004).
[19] L. Brey and H. A. Fertig, Phys. Rev. B **73**, 235411 (2006).
[20] D. Gunlycke and C. T. White, Phys. Rev. B **77**, 115116 (2008).
[21] D. P. DiVincenzo and E. J. Mele, Phys. Rev. B **29**, 1685 (1984).
[22] W. Vandenberghe, B. Soree, W. Magnus, and G. Groeseneken, J. Phys.: Conf. Ser. **193**, 012111 (2009).
[23] K. H. Gundlach and J. G. Simmons, Thin Solid Films **4**, 61 (1969).