Tunneling of Dirac electrons through spatial regions of finite mass

J. Viana Gomes and N. M. R. Peres
Center of Physics and Department of Physics, University of Minho, P-4710-057, Braga, Portugal

Abstract. We study the tunneling of chiral electrons in graphene through a region where the electronic spectrum changes from the usual linear dispersion to a hyperbolic dispersion, due to the presence of a gap. It is shown that contrary to the tunneling through a potential barrier, the transmission of electrons is, in this case, smaller than one for normal incidence. This mechanism may be useful for designing electronic devices made of graphene.

PACS numbers: 72.10.-d,73.23.-b,73.40.-c,73.21.Cd

Submitted to: J. Phys.: Condens. Matter

1. Introduction

A new and exciting field in condensed matter physics started when graphene - a two-dimensional, one carbon-atom thick material - was isolated for the first time.[1, 2] It was experimentally shown that the charge carriers in graphene could be controlled by a bottom gate setup-up; the charge carrier were shown to be either holes or electrons depending on the sign of the bottom gate voltage. In the transition from hole-based to electron-based transport the conductivity shows a minimum (not zero) value, $\sigma_{\text{min}}$. Its experimental value is of the order of $\sigma_{\text{min}} \simeq 4e^2/h$,\cite{[1, 2, 3, 4]} but the actual value seems to be somewhat sample dependent.\cite{[5]}

This value for $\sigma_{\text{min}}$ imposes therefore a limitation on the minimum value of the current a field effect transistor made of graphene can transport. The existence of a conductivity minimum in graphene is a consequence of the fact that the elementary excitations of graphene are Dirac fermions, with a linear dispersion relation, instead of usual electrons with parabolic-like dispersion, characteristic of ordinary semi-conductors. Interestingly enough, the calculated value of the conductivity of graphene at the neutrality point is off the experimental value by the factor $1/\pi$.\cite{[6, 7, 8]} Although this value is the more common result for the theoretical calculation of $\sigma_{\text{theo}}$, there are, however, several different values available in the literature.\cite{[9]}

It is also interesting that a clean graphene sample with metallic leads and smooth edges has a value of $\sigma = \sigma_{\text{theo}}$ as long as it width ($w$) is much larger than its length ($L$), being smaller than $\sigma_{\text{theo}}$ in the opposite limit.\cite{[10]} Considering the case of metallic armchair edges, it found
that $\sigma > \sigma_{\text{theo}}^{\text{min}}$ for $w/L \ll 1$ and that $\sigma \rightarrow \sigma_{\text{theo}}^{\text{min}}$ for $w/L \gg 1$.\[10\] This shows that disorder is not needed for having $\sigma \simeq \sigma_{\text{theo}}^{\text{min}}$.

Another characteristic of Dirac electrons in graphene is their ability to tunnel through a potential barrier with probability one.\[11, 12\] This so-called Klein tunneling of chiral particles has long ago been proposed in the framework of quantum electrodynamics,\[13, 14, 15\] but was never observed experimentally. Graphene opens up a route to observe this effect in a tabletop experiment, where the potential is created by some electrostatic gate potential. The manifestation of Klein tunneling is also present when electrons in graphene are forced to transverse a $n - p$ junction, leading to a selective transmission of those electrons approaching the junction perpendicularly.\[16\] Other unusual effects, such as the focusing of an electric current by a single $p - n$ junction are also characteristic of Dirac electrons in graphene.\[17\]

As appealing as the Klein tunneling may sound from the point of view of fundamental research, its presence in graphene is unwanted when it comes to applications of graphene to nanoelectronics. This comes about because the pinch-off the field effect transistor may be very ineffective. The same may occur because of the minimum conductivity of graphene at the neutrality point (as discussed above). One way to overcome these difficulties is by generating a gap in the spectrum. From the point of view of Dirac fermions this is equivalent to the generation of a mass term. There are two known forms of generating gaps in the spectrum of graphene. The first one is by patterning graphene nanoribbons.\[18, 19\] The mechanism of producing these gaps depends on the nature of the termination of these nanoribbons. For armchair nanoribbons the gap comes from quantum confinement of Dirac fermions induced by the finite nature of the ribbons in the transverse direction. For zigzag nanoribbons the gap stems from the formation of polarized spin edge-states characteristic of these type of ribbons. The formation of these polarized states is also possible in bilayer graphene.\[20\] It is interesting to notice that Klein tunneling can also be circumvented by using a graphene bilayer.\[11\] The value of the induced gaps depend on the width of the ribbons, but for large widths it is of the order of 0.1 eV.

Another possibility of generating gaps in the graphene spectrum is to deposit graphene on top of hexagonal boron nitride (BN).\[21\] This material is a band gap insulator with a boron to nitrogen distance of the order of 1.45 Å,\[23\] (in graphene the carbon-carbon distance is 1.42 Å) and a gap of the order of 4 eV. It was shown that in the most stable configuration, where a carbon is on top of a boron and the other carbon in the unit cell is centered above a BN ring, the value of the induced gap is of the order of 53 meV. Depositing graphene on a metal surface with a BN buffer layer leads to $n-$doped graphene with an energy gap of 0.5 eV.\[22\]

The two mechanisms described above can be used to produce arrangements of graphene where in some spatial zones of the material the Dirac electrons will have gaps in the spectrum. The first possibility is to pattern graphene planes such that in several areas of the graphene flake narrow nanoribbons may exist. Another possibility it to combine wafers of silicon oxide and hexagonal boron nitride, such that in the region where the BN is located the local spectrum of graphene will present a finite gap. We shall explore in this paper this latter possibility and the way it can prevent Klein tunneling from occurring. The
two mechanisms just mentioned can then be at the heart of future nanoelectronics built of graphene. The second method is also related to junctions of graphene with other kind of systems, being them superconducting, normal-conductor/graphene/normal-conductor, or multiterminal junctions. Also the study of electron transport in disordered graphene samples is of interest, specially because the tunneling may be assisted by impurities, which is a manifestation of Klein tunneling.

For a review on the experimental aspects of graphene physics see the work of Geim and Novoselov. Some of the theoretical aspects of graphene physics are reviewed qualitatively by Castro Neto et al. by Katsnelson, and by Geim and MacDonald; a more comprehensive review is given by Castro Neto et al. For a review on Klein tunneling see the work by Beenakker.

2. Basic definitions.

As described in the previous section, we assume that it is possible to manufacture slabs with SiO$_2$-BN interfaces, on top of which a graphene flake is deposit. This will induce spatial regions where graphene has a vanishing gap intercalated with regions where the BN will cause a finite gap.

In the following we will consider the graphene physics in two different regions: the $k$–region, where the graphene sheet is standing on top of SiO$_2$, and a $q$–region, where a mass-like term is present, caused by BN, inducing an energy gap of value $2t'$ (for all numerical purposes we use $t' = 0.1$ eV). The wavefunctions in these two regions will be referred by $\psi_k$ and $\psi_q$, respectively. The geometry of the scattering process is represented in Fig. 1.

The Hamiltonian for massless Dirac electrons in graphene, around the $K$-point in the Brillouin zone is given by

$$ H_g = v_F \sigma \cdot p, $$

where $\sigma = (\sigma_x, \sigma_y)$, $p = -i\hbar \nabla$, with $i = x, y, z$, is the $i$ Pauli matrix, and $v_F = 3ta/(2\hbar)$, with $t$ the nearest neighbor hopping matrix in graphene and $a$ the carbon-carbon distance. Therefore, in the massless wave function, in the $k$–region ($t' = 0$), is given by

$$ \psi_{k,s} = \frac{1}{\sqrt{2}} \left( \begin{array}{c} 1 \\ u(k, s) \end{array} \right) e^{i k \cdot r}, $$

with

$$ u(k, s) = s e^{i \theta}, $$

where $s = \text{sign}(E)$ and $\theta = \arctan(k_y/k_x)$. The corresponding energy eigenvalue is

$$ E = \pm v_F \hbar \sqrt{k_x^2 + k_y^2} = \pm \hbar v_F k, $$

with $k$ the absolute value of the wavevector.

In a region of finite mass the Hamiltonian for Dirac electrons is

$$ H_g = v_F \sigma \cdot p + t' \sigma_z, $$

and the wavefunction is

$$ \psi_q = \frac{1}{\sqrt{2}} \left( \begin{array}{c} u_q \\ -i\hbar \nabla \cdot \psi_q \end{array} \right) e^{i q \cdot r}, $$

with $u_q$ given in Eq. (2).
Tunneling of Dirac electrons through spatial regions of finite mass

with \( mv_F^2 = t' \) the mass term (\( m \) is the effective mass); as a consequence the electronic spectrum will present a finite energy gap of value \( 2t' \). In the \( q \)-region (the gaped region, \( t' \neq 0 \)), the wave function is

\[
\psi_{q,s} = \frac{1}{\sqrt{2}} \left( \frac{1}{v(q,s)} \right) e^{i q \cdot r},
\]

where

\[
v(q,s) = \frac{E - t'}{\hbar v_F(q_x - i q_y)}.
\]

Due to momentum conservation, electrons propagating through a \( k-q \) interface will conserve their wavevector component parallel to the interface. Thus, taking this interface to be located along the \( \hat{y} \) axis, we will always \( k_y = q_y \). The \( q \)-region eigenenergy, associated with the eigenstate (6) and the Hamiltonian (5), is

\[
E = \pm \sqrt{(q_x^2 + k_y^2)(\hbar v_F)^2 + t'^2}.
\]

It is amusing to notice that the spectrum (8) has the same form as for the electrons in a graphene bilayer, when the two graphene planes are at different electrostatic potentials. Using Eqs. (8) and (4), we write

\[
v_F \hbar q_x = \sqrt{E^2 \cos^2(\theta) - t'^2}
\]

and, depending on \( E^2 \cos^2(\theta) \) being larger or smaller than \( t'^2 \), \( q_x \) may take a real or a pure imaginary value. Wave propagation follows for the former case, evanescent waves in the latter.

For a real \( q_x \), and since \( q_y = k_y \), we have

\[
\sqrt{E^2 - t'^2} \sin(\phi) = |E| \sin(\theta)
\]

where \( \phi \) is the angle of propagation of the electron in the \( q \)-region (see Fig.1). Equation (10) is just the usual Snell’s law, for electrons being refracted at the interface separating the \( k \)- and \( q \)-regions. We see that \( \phi \geq \theta \) whenever \( |E| > t' \).

2.1. Forward and backward propagation.

We consider now the simple reflection in the interface, with the incident and the reflected waves both on the \( k \)- or on the \( q \)-region.

In the \( k \)-region, the \( \hat{x} \)-component of the wavevector of the reflected wave is symmetrical with respect to the incident wave. Thus, for this case, we have the following transformations under a reflection (see also Fig.1)

\[
k_x \rightarrow -k_x \quad \text{and} \quad e^{i \theta} \rightarrow e^{i(\pi - \theta)} = -e^{-i\theta}.
\]

This leads to the generalization of Eqs. (2) and (3)

\[
u_{\pm} = \pm s e^{\pm i \theta},
\]

\[
\psi_k^\pm = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ u_{\pm} \end{pmatrix} e^{\pm ik_x x + ik_y y}.
\]
Tunneling of Dirac electrons through spatial regions of finite mass

Figure 1. Top figure: graphene band structure for the massless and massive cases. In this latter, the quasi-parabolic bands have a gap energy $2t'$. Bottom figure: geometry in the reflection in the $k-q$ interface. An incident wavefunction $\psi_k^+$ with wavevector $k_+$ is reflected and refracted into the $\psi_k^-$ and the $\psi_q^+$ wavefunctions with wavevectors $k_-$ and $q_+$, respectively. Since the momentum is conserved at the interface, one has that $q_y = k_y$. The refracted wave propagates with an angle $\phi$, which is slightly larger than the incident and reflected angles $\theta$ with $|q_x| < |k_x|$, a consequence of energy conservation.

where the plus and minus signs refers to waves propagating, respectively, in the positive and negative directions of the $\hat{x}$-axis.

A similar reasoning leads to the generalization of (7) to $v(q, s)$,

$$v_{\pm} = \frac{E - t'}{\hbar v_F (\pm q_x - ik_y)}.$$  \hspace{1cm} (14)

and, also, of the $q-$region wavefunction to

$$\psi_{q}^{\pm} = \frac{1}{\sqrt{2}} \left( \begin{array}{c} 1 \\ v_{\pm} \end{array} \right) e^{\pm iq_x x + ik_y y}.$$  \hspace{1cm} (15)

The differences we have just highlighted on the wavefunctions and coefficients for forward and backward propagating particles can be also seen in the differences in positive and negative angles of incidence in the interface. This changes are useful, when a guiding-wave kind of device is made. Let us therefore analyze the case when $k_y \to -k_y$. If in Eq. (11) we keep $k_x$ unaltered and “reflect” instead $k_y$ we would obtain

$$k_y \to -k_y \quad \text{and} \quad e^{i\theta} \to e^{-i\theta},$$  \hspace{1cm} (16)

with similar relations for $\phi$, the angle in the q-region. For this cases, we get

$$v_{\pm}(-\phi) = -v_{\mp}(\phi) \quad \text{and} \quad u_{\pm}(-\theta) = -u_{\mp}(\theta).$$  \hspace{1cm} (17)

Apart from a minus sign, these relations shows that the the operation of changing the sign of $k_y$ (i.e., the angle of incidence) is equivalent to the one of changing the sign of $k_x$ ($q_x$ in the q-region). Of curse that the extra minus sign in the right hand side of both expressions in Eq. (17) are of no consequence within the calculation of reflection and transmission coefficients that follow.
2.2. Real and evanescent waves in the $q$-region

Since in the $q$-region there is a gap in the energy spectrum then $q_x$ can take both real and pure imaginary values. In the first case, we have wave propagation in this region, in the latter just evanescent waves. No simple expression as the one given by Eq. (12) can be written in this case. Instead, we need to consider separately the cases where $q_x$ is a real or a pure imaginary number.

2.2.1. For $q_x$ real. For real $q_x$, we can write a similar expression to the one in Eq. (12),

$$v_{\pm} = \pm v e^{\pm i\phi},$$

(18)

with $\phi$ given by Eq. (10) and

$$v = \frac{E - t'}{v_F \hbar |q|} = \frac{E - t'}{\sqrt{E^2 - t'^2}},$$

(19)

where Eq. (8) was used.

2.2.2. For $q_x$ pure imaginary. Since $k_y$ is always a real number, Eq. (14) implies that, if $q_x$ is a pure-imaginary, $v_{\pm}$ also is, and then

$$v_{\pm} = \mp i\nu_{\pm},$$

(20)

where

$$\nu_{\pm} = \pm \frac{E - t'}{\hbar v_F(\pm k_y - \alpha)},$$

(21)

with the real absorption coefficient $\alpha$ defined as $q_x = i\alpha$, and $\alpha$ given by

$$\alpha = (v_F \hbar)^{-1} \sqrt{t'^2 - E^2 \cos(\theta)^2}.$$

(22)

2.2.3. Complex conjugate of the $u_{\pm}$ and $v_{\pm}$ coefficients. For the calculation of the intensity reflection and transmission coefficient we will need to deal with the complex conjugate of the $u_{\pm}$ and $v_{\pm}$ coefficients. The definition (12) for $u_{\pm}$ implies that

$$u_{\pm}^* = -u_{\mp}.$$  

(23)

In the case of $v_{\pm}$, its complex conjugate depends on the the fact of having a real or imaginary $q_x$,

$$v_{\pm}^* = \begin{cases} -v_{\mp} & \text{if } q_x \text{ is real} \\ -v_{\pm} & \text{if } q_x \text{ is imaginary} \end{cases}$$

(24)

3. Transmission and reflection at the interface: the step case.

3.1. Reflection and transmission amplitude coefficients.

We compute now the reflection and transmission amplitude coefficients for electrons crossing an interface between a $k-$ and a $q-$ region. Unlike what happens in optics and due to the differences on back and forward propagation, we will need to consider not only two but four
Tunneling of Dirac electrons through spatial regions of finite mass

3.1.1. Propagation from a \( k \)- into a \( q \)-region. We start by deriving the amplitude reflection and transmission coefficients, which will be denoted as \( r_{kq}^\pm \) and \( t_{kq}^\pm \) respectively, for the case of the propagation from a \( k \)- into a \( q \)- region. This situation is described in Fig. 1 and also in Fig. 2a).

Since there is a partially reflected wave, the total wave function in the \( k \)-region must be written as an superposition of one associated with the incident electrons and other with those that are reflected,

\[
\Psi_k(x) = A \psi_k^+ + B \psi_k^-.
\]  

(25)

\( A \) and \( B \) are the normalized amplitudes for the incident and reflected wave functions. In the \( q \)-region, with \( C \) the amplitude of the transmitted wave function, we have

\[
\Psi_q(x) = C \psi_q^+.
\]  

(26)

Using Eqs. (25) and (26), and imposing the continuity condition of the particle’s wave function at the interface, i.e. \( \Psi_k(x, y = 0) = \Psi_q(x, y = 0) \), we find

\[
r_{kq}^+ = \frac{B}{A} = \frac{v_+ - u_+}{u_- - v_-} \quad \text{and} \quad t_{kq}^+ = \frac{C}{A} = 1 + r_{kq}^+,
\]  

(27)

where the superscript \( + \) recall that the incident wave function is, in this case, traveling in the positive direction of the \( x \)-axis.

Had we considered the case were the particles travel in the backward direction, represented by Fig. 2c), we would have obtained

\[
r_{kq}^- = \frac{v_- - u_-}{u_+ - v_+} \quad \text{and} \quad t_{kq}^- = 1 + r_{kq}^-.
\]  

(28)

This result can be obtain simply by exchanging the plus by the minus signs in Eq. (27).
3.1.2. Propagation from the \( q^- \) into the \( k^- \) region. For computing the reflection and transmission coefficients for the cases where the electrons come from the \( q^- \) region into the \( k^- \) region, \( r_{qk}^\pm \) and \( t_{qk}^\pm \), we need only to exchange \( u \leftrightarrow v \) in the corresponding backward and forward expressions (27) and (28). The result is

\[
r_{qk}^\pm = \frac{u_\pm - v_\mp}{v_\pm - u_\mp} \quad \text{and} \quad t_{qk}^\pm = 1 + r_{qk}^\pm. \tag{29}
\]

3.2. Amplitude coefficients: general algebraic relations.

It will be very useful in the following, for expression simplification purposes, the derivation of simple relations between the reflection and transmission amplitude coefficients. Similar relations to those we present here exist also for the photons’ optics case. For instance, we may write \( \pm r_{12} + t_{12} = 1 \) when a light beam is reflected and refracted in a diopter between regions 1 and 2, with the plus or minus sign corresponding respectively to the cases where \( n_1 \), the index of refraction of medium 1, is smaller or larger that the one of region 2.

Here, however, we have that in general \( r_{mn}^\pm \neq r_{mn}^- \) (and similarly for the transmission coefficients) and these relations are less trivial (we have used the notation \( m \neq n = \{k, q\} \)).

Using the definitions in Eqs. (27), (28) and (29), we can write

\[
R + T = 1, \tag{30}
\]

where

\[
R = r_{kq}^+ t_{kq}^- = r_{qk}^+ v_{qk}^- \quad \text{and} \quad T = t_{qk}^+ t_{qk}^- = t_{qk}^+ t_{qk}^- . \tag{31}
\]

These relations are general and don’t depend on the \( q_x \) being real or imaginary. Another general relations, useful to simplify expressions of the transmission of multi-layered structures, are

\[
r_{mn}^\pm r_{mn}^\mp = -R \times \frac{t_{nm}^\mp}{t_{nm}^\pm}. \tag{32}
\]

3.3. Intensity reflection and transmission coefficients.

The general definitions for the intensity reflection and transmission coefficients are

\[
R_{mn}^\pm = r_{mn}^\pm (r_{mn}^\pm)^* \quad \text{and} \quad T_{mn}^\pm = t_{mn}^\pm (t_{mn}^\pm)^* = 1 - R_{mn}^\pm, \tag{33}
\]

where we keep the same notation as before. We will consider now, separately, the cases were \( q_x \) is a real number or a pure imaginary.

3.3.1. For \( q_x \) real. For \( q_x \) a real number, we note first that for any \( m \neq n = \{k, q\} \),

\[
(r_{mn}^\pm)^* = r_{mn}^\pm \quad \text{and} \quad \mathcal{R} = \mathcal{R}^*. \tag{34}
\]
These relations are just a consequence of the definitions of \( r_m^{\pm} \) in Eqs. (27), (28), and (29), and in Eqs. (23) and (24). Using Eq. (34) in Eq. (33) results in

\[
R = R \quad \text{and} \quad T = T \quad \text{for } q_x \text{ real},
\]

which is valid, for both interfaces and both directions of propagation. Furthermore, using Eq. (30) we get

\[
R + T = 1,
\]

an expected result.

Explicitly, the \( R \) coefficient defined in Eq. (31) is given by

\[
R = \frac{(v_+ v_- - 1) - (u_- v_+ + u_+ v_-)}{(v_+ v_- - 1) - (u_- v_+ + u_+ v_-)}.
\]

Making use of Eqs. (14) and (12), we may write

\[
v_+ v_- = -v^2, \quad u_\mp v_\mp = -v e^{\pm i(\theta - \phi)}
\]

and

\[
u_\pm v_\pm = s v e^{\pm i(\theta + \phi)},
\]

where \( v \) is given by Eq. (7) and \( \phi \) by Eq. (10). Using these expressions we obtain

\[
R = \frac{1 + v^2 - 2 sv \cos(\theta - \phi)}{1 + v^2 + 2 sv \cos(\theta + \phi)},
\]

where

\[
v \cos(\theta \pm \phi) = \frac{v_F}{E + v'} (q_x \cos(\theta) \mp k_y \sin(\theta)).
\]

Finally, after algebraic simplification, we obtain

\[
R = \mathcal{R} = \frac{k_x - q_x}{k_x + q_x}.
\]

### 3.3.2. For \( q_x \) pure imaginary

For \( q_x \) a pure imaginary, we see that

\[
\left\{ \begin{array}{l}
(v_{mn}^{\pm})^* r_{mn}^{\pm} = 1 \\
\mathcal{R} \mathcal{R}^* = 1
\end{array} \right.
\]

Using these relations along with Eq. (33), we straightforwardly obtain

\[
R_{kq}^{\pm} = 1 \quad \text{and} \quad T_{kq}^{\pm} = 0 \quad \text{with } q_x \text{ a pure imaginary}.
\]

This is an expected result since the transmission \( T = 1 - R \) must be zero in the case where the wave in the \( k \)-region enters in the gap of the \( q \)-region. If the incident wave propagates in the gap region, i.e. it is an evanescent wave, the coefficients \( R_{kq}^{\pm} \) and \( T_{kq}^{\pm} \) are physically meaningless.

We see from the second expression in (36) that \( \mathcal{R} \) is a modulo 1 complex quantity. It may be written as

\[
\mathcal{R} = e^{i2\varphi},
\]
with $2\varphi$ a convenient definition of its argument. To compute this angle, in the spirit of Eq. (22), we replace $q_x$ by $i\alpha$ in Eq. (35) to obtain

$$\mathcal{R} = \frac{k_x - i\alpha}{k_x + i\alpha}. $$

Computing the real part of this quantity, we get

$$\cos (2\varphi) = \frac{2E^2 \cos(\theta)^2}{t^2} - 1$$

and, after straightforward manipulation,

$$\cos (\varphi') = \cos (\theta) \frac{|E|}{v} \quad \text{or else} \quad \tan (\varphi) = \frac{\alpha}{k_x}, \quad (39)$$

a Snell type expression for the $q_x$ pure imaginary case.

Since $R = \mathcal{R}\mathcal{R}^* = 1$ for $q_x$ a pure imaginary, a general expression for the intensity reflection and transmissions coefficients (valid for $q_x$ both real and pure imaginary) is given by

$$R = 1 - T = \left| \frac{k_x - q_x}{k_x + q_x} \right| = \frac{1 + v^2 - 2sv \cos (\theta - \phi)}{1 + v^2 + 2sv \cos (\theta + \phi)}. \quad (40)$$

Naturally, Eq. (40) depends on $t'$, since both $v$, $\theta$, and $\phi$ depends on this quantity. When one considers the case $\theta = \phi = t' = 0$ one obtains $R = 0$. This expression is plotted in a density/contour plot in Fig. [3].

4. The barrier.

With the above definitions, the computation of transmission and reflections coefficients for any type of multi-interface device follows similar expressions as those found in normal optics. To illustrate this, we consider in the following a heterostructure made of a $q-$region of width $w$ placed between two semi-infinite slabs of $k$-regions, as shown in the Fig. [4]. Our goal will be the derivation of the intensity transmission coefficient for this case, which we will denoted by $T_b$. We notice that results for barriers of the same height when the spectrum of the electrons is linear in every spatial regions was considered in Ref. [38].

In Figure[3] the wave function $\Psi_1$ describes an electron, traveling in the positive direction of $\hat{x}$, just before crossing the diopter $q - k$. This wave function can be seen as resulting from the coherent superposition of two wave functions, one being itself after a round trip in the $q-$region, given by $\Psi_1 t_{qk}^* t_{kq} e^{i2q_k w}$, and another one which is the incident wave function $\Psi_0$ after crossing the first interface $k - q$, equal to $\Psi_0 t_{qk}^* e^{iq_k w}$. Adding these two contribution and solving in order to $\Psi_1$ we obtain,

$$\Psi_1 = \Psi_0 \frac{t_{qk}^* e^{iq_k w}}{1 - r_{qk}^* r_{qk} e^{i2q_k w}}.$$

† There is no analog, however, for the gap-region with normal incidence.
Tunneling of Dirac electrons through spatial regions of finite mass

Figure 3. Intensity transmission for particles crossing the interface from a $k$-region into a $q$-region. The black region corresponds to a zero transmission, a case that corresponds to the total internal reflection in the usual photonic optics.

Figure 4. Barrier: scheme for the computation of the transmission.

If we denote the amplitude transmission coefficient for this barrier as $t_B = \Psi_2 / \Psi_0$ and using the fact that $\Psi_2 = t^+_q \Psi_1$, we finally obtain

$$t_B = \frac{T e^{i q_x w}}{1 - R e^{i 2 q_x w}}$$

(41)

where the definitions (30) were used.

4.1. $q_x$ real: free propagation.

If there is wave propagation in the $q$-region, $q_x$ is real, $R = R$ and $T = T$, and

$$T_B = t_B t_B^* = 1 - R_B = \left[ 1 + \left( \frac{2}{\pi} F \right)^2 \sin^2 (q_x w) \right]^{-1}.$$  

(42)

where we used the finesse definition

$$F = \frac{\pi \sqrt{R}}{T} = \frac{\pi t'}{2 q_x},$$

(43)
Tunneling of Dirac electrons through spatial regions of finite mass

Figure 5. Transmission of a barrier for a \( q \)-region width of \( w = 50a \) (top) and \( w = 300a \) (bottom). For a sufficient narrow width, the wave tunnels across the \( q \)-region resulting in a non null transmission. In optics, this behavior is known as frustrated total internal reflection. The dashed lines marks the region where in the step, the transmission is zero.

to highlight the similarity with an Fabry-Pérot solid etalon (made of glass, for example) in usual optics. However, this similarity is elusive. In the solid etalon case, in general, the finesse is almost a constant coefficient since the interfaces’ reflectivities (e.g., in a diopter glass-air) has a small dependence on the energy (optical resonances are typically far from the visible part of the spectrum and there’s no gap as in the case of the graphene with a mass term). In the case treated here, \( F \) has a strong dependence on the energy \( E \) of the particles and, furthermore, there is also a gap present. We will revisit a Fabry-Pérot type of device further in this work.

4.2. Inside the gap: frustrated total internal reflection.

Inside the gap, \( q_x \) is a pure imaginary and there’s no wave propagation. This is similar to the total internal reflection in optics, where only an evanescent wave exists that carries no energy
Tunneling of Dirac electrons through spatial regions of finite mass

Figure 6. Transmission coefficient for a simple step and several barriers for normal incidence. In the barriers case, there is a frustrated total internal reflection and, in the gap, the transmission is non zero and increases with decreasing values of the width of the $q$-region.

Figure 7. Left: Transmission coefficient for a barrier with a width $50a$ and for different angles of incidence $\theta$. At $E = 0$, all curves have the same value and for $\theta = \pi/2$, the transmission is a delta function at $E = 0$. Right: Zero energy transmission coefficient of a barrier as function of its width $w$.

(since in here, the coefficient $R_{kq} = 1$) and decays exponentially in the $x$ direction (although keeping the phase term $e^{i k_y y}$ in the $y$ direction).

However, by placing a $k$-region nearby the evanescent wave, some of the energy of the totally reflected wave tunnels throughout the gap region, a phenomena known in optics as frustrated total internal reflection. This phenomena is also described by Eq. (42), whenever $q_x$ becomes a pure imaginary. In this case, replacing $q_x = i \alpha$ and using the definitions (22)
and (38) we may simplify Eq. (42) to
\[ T_B = t_B t_B^* = 1 - R_B = \frac{1}{1 + \zeta}, \] (44)
where we have used the definition
\[ \zeta = \frac{\sinh^2(\alpha w)}{\sin^2(\varphi)}, \] (45)
which will be used in the following. If \( E = 0 \), Eq. (39) implies that \( \varphi = -\pi/2 \) and \( \alpha = t'/v_F\hbar \). \( T_B \) is in this case independent of \( \theta \) and is equal to
\[ T_B(E = 0) = \cosh^{-2} \left( w \frac{t'}{v_F\hbar} \right). \] (46)
This behavior is clearly shown in left panel of Fig. 7. In the right panel of Fig 7 it is shown how this tunneling transmittance at zero energy varies with the barrier width. A 50% reduction is accomplished for a barrier with a width of approximately 36 \( a \).

5. Transfer matrices

The method used in the last Section for computing \( t_B \), although being simple becomes very difficult to handle for more complex hetero-structures, with more than two interfaces. These type of cases are usually treated with the use of transfer matrices. These will be computed in the following. Figure 8 shows the scheme used to compute the transfer matrix in an interface \( k - q \) and \( q - k \). In both cases, our goal is to derive \( \Psi_1 \) and \( \Psi'_1 \) from the knowledge of \( \Psi_0 \) and \( \Psi'_0 \). Defining
\[ \left( \begin{array}{c} \Psi_1 \\ \Psi'_1 \end{array} \right) = M_{mn} \left( \begin{array}{c} \Psi_0 \\ \Psi'_0 \end{array} \right), \]
Tunneling of Dirac electrons through spatial regions of finite mass

where $M_{mn}$ is the transfer matrix for the generic $m - n$ interface, and using Eq. (31), we obtain the result

$$M_{mn} = \begin{bmatrix} \frac{1}{t_{nm}^+} & \frac{r_{nm}^-}{t_{nm}^-} \\ \frac{r_{nm}^+}{t_{nm}^+} & \frac{1}{t_{nm}^-} \end{bmatrix}$$

(47)

The determinant of this matrix is given by

$$Det(M_{kq}) = [Det(M_{qk})]^{-1} = \frac{t_{kq}^+}{t_{qk}^-}.$$  

(48)

As expected, $Det(M_{kq} \times M_{qk}) = 1$. The free propagation of a particle in a $k-$ and in a $q-$region of width $\xi$ is given, respectively, by

$$L_{k}(\xi) = \begin{bmatrix} e^{ik_\xi \xi} & 0 \\ 0 & e^{-ik_\xi \xi} \end{bmatrix}; \quad L_{q}(\xi) = \begin{bmatrix} e^{iq_\xi \xi} & 0 \\ 0 & e^{-iq_\xi \xi} \end{bmatrix}.$$  

(49)

6. The diode.

We consider now a more complex system composed by a sandwich of two $q-$regions of width $w$ separated by a $k-$region with width $d$, inside two semi-infinite $k-$regions. To derive the amplitude transmission coefficient of such a structure we need to compute the expression,

$$t_d = M_{qk} L_q(w) M_{kq} L_k(d) M_{qk} L_q(w) M_{kq}. $$

The result of this expression can be simplified using Eq. (32), resulting in

$$t_D = \frac{\mathcal{T}^2 e^{2iq_\xi w}}{\left(\mathcal{R} e^{2iq_\xi w} - 1\right)^2 - \mathcal{R} (e^{2iq_\xi w} - 1)^2 e^{2ik_\xi d}}.$$  

For the most important case where the $q-$regions are barriers with energy higher than the energy of the particles, we have a resonant diode. In this case and using the definitions (22) and (38) we get

$$t_D = t_B^2 \times \left[1 - \frac{\sinh^2(\alpha w)}{\sinh^2(\alpha w + i\varphi)} e^{2k_\xi d}\right]^{-1},$$  

(50)
where \( t_B \) is the amplitude transmission for a simple barrier, given by Eq. (41).

We may simplify Eq. (50) by expanding the term \( \sinh (\alpha w + i\varphi) \) and expressing the result in a complex polar representation. Doing this and using the definition for \( \zeta \) in Eq. (45) we obtain

\[
\frac{\sinh^2(\alpha w)}{\sinh^2(\alpha w + i\varphi)} = \frac{\zeta}{1 + \zeta} \exp(i2\tilde{\varphi})
\]

with the phase term argument given by

\[
\tilde{\varphi} = -\arctan \left[ \coth (\alpha w) \tan (\varphi) \right].
\]

The intensity transmission coefficient can now be easily computed, being equal to

\[
T_D = \frac{1}{1 + (\frac{2}{\pi F_D})^2 \sin^2 (\tilde{\varphi} + k_x d)},
\]

where now, the diode finesse \( F_D \) is given by

\[
F_D = \pi \sqrt{\zeta(1 + \zeta)}.
\]
6.1. Revisiting the Fabry-Pérot: etalon made with "mirrors".

The expression (51) results in the simple case of a Fabry-Pérot etalon if:

(i) $\alpha w \gg 1$, which implies $\coth (\alpha w) \approx 1$ and $\tilde{\varphi} \approx \varphi$;

(ii) $E \ll t'$, which implies that $\varphi \approx \pi/2$.

With these approximations we get

$$F_D = \frac{\pi}{2} \sinh (2\alpha w)$$

and then

$$T_D = \frac{1}{1 + \sinh^2 (2\alpha w) \cos^2 (k_x d)},$$

an expression that can be derived from a delta function potentials treatment, treated in Section 7.

6.2. The diode tunneling conductance.

Let us now compute the tunneling conductance of the device as function of the potential bias $V$, the chemical potential of the leads, and the length $w$ of the barrier. We shall assume that the device is operating in the region where the chemical potential of the leads lies inside the gap of the barrier. The total tunneling current density (i.e. the current per unit of cross-sectional length) through the device is given by

$$J(V,w) = -\frac{2e}{4\pi^2 v_F \hbar^2} \int d\theta E dE \cos \theta T(E,\theta)$$

$$\times \left[ f(E - \mu_L) - f(E - \mu_R) \right]$$

where $f(x) = (1 + e^{x/(k_B T)})^{-1}$, $\mu_L = \mu + eV/2$, $\mu_R = \mu - eV/2$, and $\mu$ the chemical potential of leads in the equilibrium state. The linear response conductance per unit of cross-sectional length is given, at zero temperature, by

$$G(\mu, w) = \frac{e^2}{\hbar} \frac{|\mu|}{3\pi^2 a t} \int_{-\pi/2}^{\pi/2} d\theta \cos \theta T(\mu, \theta).$$

In Fig. (11) we plot $G(\mu, w)$ as function of $\mu$ for several widths $w$. It is clear that the value of $G(\mu, w)$ may change by several orders of magnitude, close to $\mu = 0$ by a small change of $\mu$. Naturally for wider barriers one obtains a smaller conductance.

6.3. A limiting case.

The limiting case of barrier can be represented by a delta-function potential, $V(x, y) = g \sigma_z \delta(x)$. (See next section for complete discussion of delta function potentials in the Dirac equation.) It is interesting to compute the reflected flux for both Schrödinger and Dirac electrons for this potential. In the first case one obtains

$$R = \frac{(2mg/\hbar^2)^2}{4k_F^2 \cos^2 \theta + (2mg/\hbar^2)^2},$$

where $k_F = \sqrt{2m_e e^2/(\hbar^2)}$. For Dirac electrons, the expression is different and involves the mass $m_d$, the Fermi energy $E_F$, and the angle $\theta$. The details of the computation are given in the next section.

The next section discusses the complete treatment of delta function potentials in the Dirac equation, including the reflection and transmission coefficients for both Schrödinger and Dirac electrons. It involves solving the Dirac equation with a delta function potential and deriving expressions for the transmission and reflection coefficients.
whereas for Dirac electrons the result is

$$ R = \tanh^2 \left[ g / (\hbar v_F) \right]. $$

(57)

It is clear that for electrons in graphene $R$ is angular and energy independent. For the case $g \gg \hbar v_F$ the reflexion tends to unity.

### 7. The diode: a limiting case.

Finally we want to discuss a limiting case of the resonant tunneling diode made of graphene. The device is represented in Fig. 9. The corresponding study for Schrödinger electrons was done by Tsu and Esaki.\[43\]

A limiting situation of the device described in Fig. 9 is one where the barriers are are describe by a scalar Lorentz potential of the form

$$ V(x, y) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \left[ 1 - \theta(|x| - \epsilon) \right] \sigma_z $$

$$ + \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \left[ 1 - \theta(|x - d| - \epsilon) \right] \sigma_z $$

$$ = g \sigma_z [\delta(x) + \delta(x - d)]. $$

(58)

The connection with the true barrier is made by identifying $g$ with $\alpha t' w a$, with $\alpha$ a numerical constant of dimensions inverse of length. This form of the potential is equivalent to a mass term and therefore to a gap in the spectrum. However, given the short range nature of the potential, its effect comes only in the boundary conditions imposed on the wave function at the potential position.

The problem of Dirac electrons in delta function potentials has been studied in the past\[44, 45\] and is not without subtleties.\[46, 47, 48\] The subtleties can be traced back to the problem of evaluating the integral

$$ \int_{-\epsilon}^{\epsilon} f(x) \delta(x) dx, $$

(59)
where \( f(x) \) is a discontinuous function at \( x = 0 \). If we try to solve the problem of Dirac electrons with a delta function potential using the same trick one uses for Schrödinger electrons we face the problem defined by the integral. This is so because the wave function of Dirac electrons in a delta function potential is discontinuous at the point where the delta function is located. There are several strategies to overcome this difficulty. The most straightforward was devised by McKellar and Stephenson and generalized by Dominguez-Adame and Maciá. In short, the Dirac equation along the \( x \) direction can be written as

\[
\frac{d\phi(x)}{dx} = \hat{G}(x)\phi(x),
\]

where \( \phi(x) \) is a spinor wave function. This problem can be formally solved as

\[
\phi(\epsilon) = T_x e^\int_{x_0}^{\epsilon} \hat{G}(x)dx \phi(x_0),
\]

where the operator \( T_x \) is the position order operator such that

\[
T_x [\hat{G}(x)\hat{G}(y)] = \hat{G}(x)\hat{G}(y)\theta(x-y) + \hat{G}(y)\hat{G}(x)\theta(y-x).
\]

Since we are interested in determining the boundary conditions obeyed by the wave function \( \phi(x) \) at the delta function position we consider the infinitesimal interval \( x \in [-\epsilon, \epsilon] \) obtaining

\[
\phi(\epsilon) = e^{-i\frac{g}{v_F}\hbar\sigma_x \sigma_z} \phi(-\epsilon).
\]

To evaluate how the exponential acts on \( \phi(-\epsilon) \) we use the Lagrange-Silvester formula for a function \( f(M) \) of a matrix \( M \)

\[
f(M) = f(\lambda_1)\frac{1\lambda_2 - M}{\lambda_2 - \lambda_1} + f(\lambda_2)\frac{1\lambda_1 - M}{\lambda_1 - \lambda_2},
\]

where \( \lambda_{1,2} \) are the eigenvalues of \( M \). For the problem at hand, Eq. (65) leads to the following boundary condition around \( x = 0 \)

\[
\begin{pmatrix}
\phi_a(0^+) \\
\phi_b(0^+)
\end{pmatrix} = \cosh \tilde{g} \begin{pmatrix}
\phi_a(0^-) \\
\phi_b(0^-)
\end{pmatrix} + i \sinh \tilde{g} \begin{pmatrix}
\phi_b(0^-) \\
-\phi_a(0^-)
\end{pmatrix},
\]

where \( \tilde{g} = \frac{g}{v_F \hbar} \) is an adimensional interaction constant and \( 0^\pm \) represent positive and negative infinitesimals. A similar boundary condition holds for \( x = d \). For the potential (58) we can now define three different regions, I, II, and III, defined as \( x < 0 \), \( 0 < x < d \), and \( x > d \), respectively. In each of these regions the wave function is a sum of two plane waves of opposite momentum along the \( x \)-direction, with each plane wave multiplied by the coefficients \( A_{\Gamma} \) and \( B_{\Gamma} \), where \( \Gamma = I, II, \) and III labels the three regions defined above.

Once the matrix \( T \) has been computed (see appendix Appendix A) the reflection coefficient is obtained from

\[
r = -\frac{T_{12}^*}{T_{11}^*},
\]
and the transmitted flux $T_f$ from

$$T_f = 1 - rr^*.$$  \hspace{1cm} (68)

For the case of zero electrostatic potentials, $U_i = 0$, we obtain

$$T_f = \frac{1}{1 + \sinh^2(2\tilde{g}) \cos^2(2dk)},$$  \hspace{1cm} (69)

which is a similar expression to the one in (53) if $\tilde{g} = \alpha w$. It is simple to identify the limit $T \rightarrow 1$, which occurs when $2kd = (2n + 1)\pi$, with $n = 0, 1, 2, \ldots$.

In Figure 12 we show the transmitted flux $T_f(E, \theta)$ as function of the energy and of the angle $\theta$. The barrier between the leads and the center of the device is represented by a delta function potential, therefore wider barriers are represented by larger values of $\tilde{g}$. From Fig. 12 we can see that for larger values of $\tilde{g}$ the transmission in the forward direction is essentially zero except at some resonant energies, where the transmission goes to one. As function of the angle we see that there are some angles for which the transmission is also one. When the length of central part of the device is increase (larger $d$) the resonances become closer to each other and more resonances appear.

In Figure 13 we present an intensity plot of $T_f(E, \theta)$ for a device with $\tilde{g} = 0.5$ and $d = 200a$. In this figure we can follow the evolution of the resonances in the $E$ versus $\theta$ plane. The six lines of larger transmission are associated with the resonances we see in Fig. 12 for $\tilde{g} = 2$ and $d = 100a$.

As before, the linear response conductance per unit of cross-sectional length is given, at zero temperature, by Eq. (55) and represented in Fig. 14. For small values of $\tilde{g}$ the conductance shows smooth oscillations whereas for larger $\tilde{g}$ values strong resonances are observed. The number of observed resonances depends on the length $d$. 

**Figure 12.** (color on line) Top panels: transmitted flux $T_f(E = 0.1, \theta)$ for fixed $\tilde{g} = 2$ and two $d = 100a, 200a$ and for fixed $d = 100a$ and three $\tilde{g} = 0.5, 1, 2$. Bottom panels: transmitted flux $T_f(E, \theta = 0)$ for the same previous cases.
Tunneling of Dirac electrons through spatial regions of finite mass

8. Final remarks.

In this paper we discussed the tunneling properties of Dirac electrons in two-dimensions when they transverse regions of space where the spectrum presents a finite energy gap. In the case we considered here the gap is induced by depositing graphene on top of Boron Nitride, rendering, in this way, the sub-lattices A and B non-equivalent. The consequence is the opening of a gap in the energy spectrum, that we have parametrized by the parameter $t' = 0.1t$. We have shown that the existence of an energy gap prevents the Klein paradox from taking place, a necessary condition for building nanoelectronic devices made of graphene. We have also shown that basic devices like a resonant tunneling diode can be made of graphene.
Tunneling of Dirac electrons through spatial regions of finite mass

by intercalating two regions where the spectrum of graphene presents a gap. We have also shown that simple analytical expressions can be derived for the tunneling through these types of heterostructures. In addition we have showed that a limiting case of the resonant tunneling diode can be understood by using Dirac delta function potentials.

Clearly that one is lead to think that a full description of the ballistic (no impurities) transport process in the system should also include the effect of temperature and phonons. We note, however, that the electron-phonon interaction has been shown to have a small effect in the optical conductivity of graphene. [49] Which means that phonons should not be very important in the description of the transport process. Also the polaronic effect leads to a renormalization of the velocity $v_F$ in the $k$–region and to the renormalization of the effective mass $m v_F^2 = t'$ in the $q$–region. But since these two parameters can be considered as effective ones there is no point in including the polaronic effect explicitly. In concerns the temperature, clearly it will be of no importance when the chemical potential is above the gap. When the energy is in the gap there will certainly be temperature activated transport adding on top of the tunneling current. For small temperatures this will be a small effect.

We believe our results of relevant for future nanoelectronics applications of graphene.

Acknowledgments

NMRPs acknowledge the financial support from POCI 2010 via project PTDC/FIS/64404/2006.

Appendix A. Details of section 7 results.

We now give the details that allows one to derive the result 69 Applying boundary conditions derived in Sec. 7 and a $T$–matrix description[42] for the scattering problem we obtain

(\begin{array}{c}
A_{II} \\
B_{II}
\end{array}) = T(\begin{array}{c}
A_I \\
B_I
\end{array}) = (T_{11} T_{12}) (T_{11}^* T_{12}^*) (A.1)

with

\[ T = V_3^{-1} M_2 V_2^{-1} M_1. \]  \hspace{1cm} (A.2)

The several matrices involved in Eq. (A.2) are defined as

\[ M_1 = \cosh \tilde{g} \begin{pmatrix}
1 & 1 \\
s_1 e^{i\theta_1} & -s_1 e^{-i\theta_1}
\end{pmatrix} + i \sinh \tilde{g} \begin{pmatrix}
s_1 e^{i\theta_1} & -s_1 e^{-i\theta_1} \\
-1 & -1
\end{pmatrix}, \]  \hspace{1cm} (A.3)

\[ V_2^{-1} = \frac{1}{2 \cos \theta_2} \begin{pmatrix}
e^{-i\theta_2} & s_2 \\
e^{i\theta_2} & -s_2
\end{pmatrix}, \]  \hspace{1cm} (A.4)

\[ M_2 = \cosh \tilde{g} \begin{pmatrix}
e^{ik_2 d} & e^{-ik_2 d} \\
s_2 e^{i(\theta_2 + k_2 d)} & -s_2 e^{-i(\theta_2 + k_2 d)}
\end{pmatrix} \]
Tunneling of Dirac electrons through spatial regions of finite mass

\[ + i \sinh \tilde{g} \begin{pmatrix} s_2 e^{i(\theta_2 + k_2 d)} & -s_2 e^{-i(\theta_2 + k_2 d)} \\ -e^{ik_2 d} & -e^{-ik_2 d} \end{pmatrix}, \]  

(A.5)

and

\[ V_3^{-1} = \frac{1}{2 \cos \theta_3} \begin{pmatrix} e^{-i(\theta_3 + k_3 d)} & s_3 e^{-ik_3 d} \\ e^{i(\theta_3 + k_3 d)} & -s_3 e^{ik_3 d} \end{pmatrix}. \]  

(A.6)

The momenta \( k_2 \) and \( k_3 \) are given by

\[ k_i = \frac{1}{v_F \hbar} \sqrt{(E - U_i)^2 - (E - U_1)^2 \sin^2 \theta_1}, \]  

(A.7)

with \( i = 2, 3 \). The angles \( \theta_i \) are defined as

\[ \theta_i = \arctan \frac{k_y}{k_i}, \]  

(A.8)

with \( k_y = |E - U_1| \sin \theta_1/(v_F \hbar) \). The \( s_i \) functions are given by \( s_i = \text{sign}(E - U_i) \), with \( i = 1, 2, 3 \). The potential energies \( U_i \) represent some electrostatic potential created in the corresponding region. Although an analytical expression for \( T \) can be produced by carrying out the four matrix multiplications, the resulting expression is too cumbersome to be given here. In the special case that all \( U_i = 0 \), the matrix elements of \( T \) have a simple form given by

\[ T_{11} = \cosh^2 \tilde{g} + e^{-2ikd} \sinh^2 \tilde{g}, \]  

(A.9)

and

\[ T_{12} = -s_i e^{-i\theta} \left( 1 + e^{-2ikd} \right) \cosh \tilde{g} \sinh \tilde{g}, \]  

(A.10)

with \( s = \text{sign} E, k = |E| \cos \theta/(v_F \hbar) \), and \( \theta \) the incident angle in the barrier.

References

Tunneling of Dirac electrons through spatial regions of finite mass


