The transparency of graphene and other direct-gap two-dimensional materials

Daniel J. Merthe and Vitaly V. Kresin

Department of Physics and Astronomy,
University of Southern California, Los Angeles, California 90089, USA

(Dated: September 27, 2016)

Abstract

Graphene and other two-dimensional materials display remarkable optical properties, including a simple light transparency of $T \approx 1 - \pi \alpha$ for light in the visible region. The physical origin of these properties requires accurate treatment in order to make analysis and applications reliable. Most theoretical rationalizations of the "universal" graphene opacity employ a model coupling light to the electron’s crystal momentum and put emphasis on the linear dispersion of the graphene bands. However, such a formulation of interband absorption is not allowable within band structure theory. We demonstrate that the observed optical behavior of graphene, as well as of semiconductor membranes, can be straightforwardly explained within the correct use of canonical momentum coupling. The essence of the low-frequency behavior lies in the two-dimensional character of the density of states rather than in the precise dispersion relation. At higher excitation energies the calculation predicts a specific asymmetry in the absorption spectrum of graphene.
I. INTRODUCTION

Continuously improving fabrication and characterization techniques have revealed interesting and elegant properties of graphene and other two-dimensional (2D) materials. The optical properties of graphene [1–4] as well as membranes of InAs [5] have been measured in recent studies. Remarkably, the transparency of both materials in the visible range is simply

\[ T \approx 1 - \pi \alpha, \]

where \( \alpha = e^2/\hbar c \) is the fine structure constant. This suggests that the optical conductivity of 2D semiconductors and semimetals possesses certain universal features.

The opacity of graphene due to interband transitions has been calculated analytically by a number of authors [6–9]. The majority of these treatments proceed by writing down the tight-binding Hamiltonian in the pseudo-relativistic form, \( \hat{H} = v_F \sigma \cdot \hbar \mathbf{k} \), where \( \sigma \) is the Pauli vector and \( \mathbf{k} \) is the in-plane crystal momentum of the basis states, and expressing the electron’s coupling to the electromagnetic field via the standard vector potential substitution (sometimes called “minimal coupling”) \( \hbar \mathbf{k} \to \hbar \mathbf{k} + e \mathbf{A}/c \). In this presentation, the linear Dirac-cone dispersion of graphene is viewed as a key element underlying its optical properties.

However, the above procedure has a fundamental conceptual flaw [10]. Specifically, \( \hbar \mathbf{k} \) is not the momentum of the electron but the crystal momentum, i.e. a state label (quantum number) rather than an operator. The actual full non-relativistic graphene Hamiltonian is quadratic in the momentum operator \( \hat{\mathbf{p}} \), and it is the latter which must be augmented with the vector potential.

It is indeed possible to describe the evolution of the Bloch states of carriers within a single band \( n \) under the influence of an external DC field \( \hat{H}' \) by an effective Hamiltonian

\[ \hat{H}_{\text{eff}} = E_n(\hat{\mathbf{p}}/\hbar) + \hat{H}' \]

where \( E_n(\mathbf{k}) \) is the band energy [11]. This formalism underlies the well-known semiclassical equations of carrier transport. However, it is strictly limited to intraband dynamics and there is absolutely no justification for applying such a procedure to interband transitions induced by AC fields. Indeed, one finds that every textbook discussion of interband absorption necessarily reverts to matrix elements of \( \hat{\mathbf{p}} \) [12–14].

The aforementioned finding [5] that a 2D nanomembrane of a direct-gap semiconductor also displays the same \( \pi \alpha \) quantum of absorption attests that it is the dimensionality of these systems (specifically, the electronic density of states in 2D) that underlies their beautiful optical characteristics. In this paper we show that by applying the "minimal coupling" substitution to the canonical momentum \( \hat{\mathbf{p}} \) and using Fermi’s Golden Rule, one can straight-
forwardly reproduce the experimentally observed constant opacity for low frequencies and also the qualitative features of graphene’s absorption spectrum for higher frequencies, which are noticeably different from what one obtain by using a “k-substitution”.

II. LOW FREQUENCY LIGHT ABSORPTION BY 2D ELECTRONS

Only general considerations are needed to characterize the low frequency absorption rate of electrons in 2D systems with a conical or higher power-law band structure. Introducing an electromagnetic field invokes the substitution $\hat{p} \rightarrow \hat{p} + e \mathbf{A}/c$, rendering a perturbation Hamiltonian, $\hat{H}' = (e/mc)\mathbf{A} \cdot \hat{p}$, where $-e$ and $m$ are the charge and mass of the electron. We shall consider a normally incident circularly polarized plane wave, which in the $z = 0$ plane takes the form,

$$A(z = 0, t) = (cE/2\omega_0)(\hat{x} + i\hat{y})e^{-i\omega_0 t}$$

where $c$ is the speed of light and $E$ is the electric field amplitude. At sufficiently low temperature, only vertical transitions $\Delta \mathbf{k} = 0$ are allowed, and these interband transitions dominate the absorption process. Then, according to Fermi’s Golden Rule, the rate of vertical transitions from the valence band (“−”) to the conduction band (“+”) at a point $\mathbf{k}$ in the Brillouin Zone (BZ) under such a perturbation is given by

$$R(\mathbf{k}) = \frac{\pi e^2 \mathcal{E}^2}{2\hbar m^2 \omega_0^2} ||\langle+, \mathbf{k}|\hat{p}|-, \mathbf{k}\rangle||^2 \delta(E_+(\mathbf{k}) - E_-(\mathbf{k}) - \hbar\omega_0)$$

To calculate the momentum matrix element for low energy excitations of a general material with a dispersion relation of the form

$$E_{\pm}(\mathbf{k}) = \pm \left[\Delta^2 + C_\pm q^\lambda\right]$$

where $q \equiv |\mathbf{k} - \mathbf{K}|$ is the distance in $k$-space from the band gap minimum, $E_+(\mathbf{K}) - E_-(\mathbf{K}) = \Delta$, and $C_\pm$ and $\lambda$ are material dependent parameters, we use a well known result of $\mathbf{k} \cdot \mathbf{p}$ perturbation theory. This calculation is similar to that of Fang et. al [5]. The effective mass in the $n^{th}$ band is given by [12]

$$\frac{1}{m_n(\mathbf{k})} = \frac{1}{m} + \frac{2}{m^2} \sum_{n' \neq n} \frac{||\langle n, \mathbf{k}|\hat{p}|n', \mathbf{k}\rangle||^2}{E_n(\mathbf{k}) - E_{n'}(\mathbf{k})}$$

In the limit of small $q$ and $\Delta$, the small energy difference between the valence and conduction bands causes the corresponding term to dominate the sum. In this approximation, we obtain

$$\frac{1}{m_+(\mathbf{k})} + \frac{1}{m_-(\mathbf{k})} \approx \frac{4}{m^2} \frac{||\langle+, \mathbf{k}|\hat{p}|-, \mathbf{k}\rangle||^2}{E_+(\mathbf{k}) - E_-(\mathbf{k})}$$
where \( m_+ (m_-) \) is the effective mass of electrons (holes) in the conduction (valence) band. Using the dispersion relation in Eq. (3) to obtain the effective masses, the modulus square of the momentum matrix element is

\[
||\langle +, k | \hat{p} | -, k \rangle ||^2 = \frac{\lambda^2 m^2}{4\hbar^2} \left[ \Delta + (C_+ + C_-)q^\lambda \right] (C_+ + C_-)q^{\lambda-2} \quad (6)
\]

The total transition rate is given by integrating Eq. (2) over the Brillouin zone. In order to transform this integration over \( k \) into an integration over energy difference, \( \epsilon(k) \equiv E_+(k) - E_-(k) = \Delta + (C_+ + C_-)q^\lambda \) (7)

the joint density of states \( \rho(\epsilon) \) for a direct-gap spectrum is defined by the relation,

\[
\rho(\epsilon) = \frac{2gS}{4\pi^2} \int d^2k \delta(E_+(k) - E_-(k) - \epsilon)
\]

where \( S \) is the area of the material surface and \( g \) is the valley degeneracy factor (the total angle around the band extrema divided by \( 2\pi \)). From this and Eq. (3) we obtain,

\[
\rho(\epsilon) = \frac{gS}{\pi\lambda(C_+ + C_-)} \left( \frac{C_+ + C_-}{\epsilon - \Delta} \right)^{1-\frac{\lambda}{2}} \Theta(\epsilon - \Delta) \quad (9)
\]

where \( \Theta \) is the step function. Combining this with Eqs. (2) and (6), the total rate of transitions is

\[
R = \int d\epsilon \frac{\pi e^2 \rho(\epsilon)}{2m^2 \omega_0} ||\langle +, k | \hat{p} | -, k \rangle ||^2 \rho(\epsilon) \delta(\epsilon - \hbar \omega_0)
= \frac{g\pi}{2} \frac{e^2}{\hbar c} \frac{cS\epsilon^2}{\pi \hbar \omega_0} \Theta(\hbar \omega_0 - \Delta)
= \frac{g\lambda}{2} \pi \alpha R_0 \Theta(\hbar \omega_0 - \Delta) \quad (10)
\]

where \( \alpha = e^2/\hbar c \) is the fine structure constant and \( R_0 = cS\epsilon^2/4\pi \hbar \omega_0 \) is the incidence rate of photons. In going from the first to the second equality, Eq. (7) is used to express the modulus square of the momentum matrix element in terms of \( \epsilon \). Remarkably, the material dependent parameters \( C_\pm \) cancel out between the density of states and the momentum matrix element. However, the power parameter \( \lambda \) still affects the total transition rate.

For a semiconductor membrane with a band gap at the center of the BZ, \( \lambda = 2, g = 1 \), we find an opacity of

\[
R/R_0 = \pi \alpha \quad (11)
\]

for frequencies above the band gap. For graphene with its band gaps at the corners, we have \( \lambda = 1, g = 2 \) (six corners, each contributing an angle of \( 2\pi/3 \); see Fig. 3), we obtain from
Eq. (10) exactly the same value. This value, instead of being "universal" as suggested in Ref. [5], should more appropriately be viewed as a numerical confluence.

For multilayers of ABC-stacked graphene, the electronic energy eigenvalues near the Fermi level can be written in the form,

$$E_{\pm}(k) \propto \pm q^N$$

where $N$ is the number of layers [15]. Thus, with $g = 2$ and $\lambda = N$, Eq. (10) yields a low frequency opacity of

$$R/R_0 = N\pi\alpha$$

which is in good agreement with the observations in Ref. [3] for $N$ up to 4.

Thus based on the 2D character of the electron system and using only a general dispersion relation, we are able to explain in a correct quantum band-structure framework the opacity of monolayer graphene, multilayer graphene, and direct gap semiconductor membranes [1, 2, 4, 5]. However, this calculation is only valid in the limit $E_+(k) - E_-(k) \to 0$. In the case of graphene, the absorption rate can be readily evaluated for higher electromagnetic frequencies within the tight-binding approximation. This more complete calculation also avoids questions which may arise regarding the $\Delta, q \to 0$ limit in view of Eq. (5).

III. ABSORPTION IN THE GRAPHENE STRUCTURE

In order calculate the transition rate for a broader frequency range, the momentum matrix elements $(\pm, k | p | -, k)$ need to be evaluated explicitly.

We proceed in the tight-binding approximation as usual, using linear combinations of $2p_z$ orbitals at each carbon atom in the lattice as a basis set for the Bloch wave functions. For reference, the lattice geometry is shown in Fig 1a. Through the standard analysis [16], one finds that the energy eigenvalues of the valence and conduction bands are given in the nearest-neighbor approximation by

$$E_{\pm}(k) = \pm \gamma_0 |f(k)|$$

$$f(k) = 1 + e^{-ik_{1}a_1} + e^{-ik_{2}a_2}$$

with "hopping" energy $\gamma_0 \approx 2.8$ eV and the upper (lower) sign indicating the conductance (valence) band, as above. The corresponding eigenfunctions are
FIG. 1. (a) Lattice geometry of graphene and (b) band structure in the tight binding approximation, with conical dispersion near the Dirac points (circled). The orange surface is the valence band and the green surface is the conduction band.

\[
\psi_{\pm,k}(x) = \frac{1}{\sqrt{2N}} \sum_R e^{i\mathbf{k} \cdot \mathbf{R}} \left[ \chi(x - \mathbf{R}) \pm e^{i\phi} \chi(x - \mathbf{R} - \mathbf{h}) \right] \tag{15}
\]

where \( \chi(x) \) is the carbon \( 2p_z \) orbital, the phase factor is \( e^{i\phi(k)} \equiv f(k)/|f(k)| \), and the sum is over all Bravais lattice points \( \mathbf{R} \) (red points in Fig. 1a).

For the diagonal (intraband) momentum matrix elements we already have the relation,

\[
\frac{1}{m} \langle -, k | \hat{p} | -, k \rangle = \frac{1}{\hbar} \frac{\partial E_-(k)}{\partial k} \equiv \bar{v}_-(k) \tag{16}
\]

for the group velocity \( \bar{v}_-(k) \) of electrons or holes in the valence band [12], but as will be seen below it is useful to calculate this matrix element in parallel. Generally, the matrix elements are given by
\[ \langle \pm, k|\hat{p}|-k \rangle = \frac{1}{2N} \sum_{\mathbf{R},\mathbf{R}'} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} \times \left[ \int d^3x \chi(\mathbf{x} - \mathbf{R}')\hat{p}\chi(\mathbf{x} - \mathbf{R}) - e^{i\phi} \int d^3x \chi(\mathbf{x} - \mathbf{R}')\hat{p}\chi(\mathbf{x} - \mathbf{R} - \mathbf{h}) \right. \\
\left. \pm e^{-i\phi} \int d^3x \chi(\mathbf{x} - \mathbf{R}' - \mathbf{h})\hat{p}\chi(\mathbf{x} - \mathbf{R}) \right] \] (17)

By the parity symmetry of \( \chi(\mathbf{x}) \) and assuming only wavefunction overlap between nearest neighbors, the first and last integrals are zero. Also assuming only nearest neighbor overlap, the remaining two integrals reduce to

\[ \int d^3x \chi(\mathbf{x} - \mathbf{R}')\hat{p}\chi(\mathbf{x} - \mathbf{R} - \mathbf{h}) = \delta_{\mathbf{R},\mathbf{R}'} \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} - \mathbf{h}) + \delta_{\mathbf{R},\mathbf{R}'} + a_1 \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} - \mathbf{h} - a_1) + \delta_{\mathbf{R},\mathbf{R}'} + a_2 \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} - \mathbf{h} - a_2) \] (18)

\[ \int d^3x \chi(\mathbf{x} - \mathbf{R}' - \mathbf{h})\hat{p}\chi(\mathbf{x} + \mathbf{R}) = \delta_{\mathbf{R},\mathbf{R}'} \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} + \mathbf{h}) + \delta_{\mathbf{R},\mathbf{R}'} - a_1 \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} + \mathbf{h} + a_1) + \delta_{\mathbf{R},\mathbf{R}'} - a_2 \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} + \mathbf{h} + a_2) \] (19)

where \( \delta_{\mathbf{w},\mathbf{w}'} \) is the Kronecker symbol. These integrals can be further simplified with the following consideration. Let \( \mathbf{s} = s\hat{\mathbf{u}} \) be a displacement in the \( x - y \) plane, where \( \hat{\mathbf{u}} \) is a unit vector. Then,

\[ \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} - \mathbf{s}) = -i\hbar \int d^3x \chi(\mathbf{x}) \frac{\partial}{\partial \mathbf{s}} \chi(\mathbf{x} - \mathbf{s}) = i\hbar \frac{\partial}{\partial \mathbf{s}} \int d^3x \chi(\mathbf{x})\chi(\mathbf{x} - \mathbf{s}) \] (20)

By the azimuthal symmetry of \( \chi(\mathbf{x}) \), this latter overlap integral is independent of the direction of \( \mathbf{s} \), and is a function \( F(\mathbf{s}) \) of only its magnitude \( s \),

\[ \int d^3x \chi(\mathbf{x})\chi(\mathbf{x} - \mathbf{s}) = \int d^3x \chi(\mathbf{x})\chi(\mathbf{x} - s\hat{\mathbf{u}}) \equiv F(s) \] (21)

Therefore, defining \( D(s) \equiv (\partial/\partial s)F(s) \), the integral in Eq. (20) is

\[ \int d^3x \chi(\mathbf{x})\hat{p}\chi(\mathbf{x} - \mathbf{s}) = i\hbar \hat{\mathbf{u}}D(s) \] (22)

With this applied to Eqs. (18) and (19) and the nearest-neighbor separation \( a \), Eq. (15)
can be rewritten as

\[
\langle -, k|\hat{p}|-, k \rangle = \hbar D(a) \left[ \frac{\hbar}{a} \sin(\phi(k)) + \frac{\hbar + a_1}{a} \sin(\phi(k) + k \cdot a_1) + \frac{\hbar + a_2}{a} \sin(\phi(k) + k \cdot a_2) \right]
\] (23)

and

\[
\langle +, k|\hat{p}|-, k \rangle = -i\hbar D(a) \left[ \frac{\hbar}{a} \cos(\phi(k)) + \frac{\hbar + a_1}{a} \cos(\phi(k) + k \cdot a_1) + \frac{\hbar + a_2}{a} \cos(\phi(k) + k \cdot a_2) \right]
\] (24)

The quantity \( D(a) \) can in principle be evaluated by explicitly computing the integral in Eq. (21) and taking the derivative. However, with knowledge of \( \langle -, k|\hat{p}|-, k \rangle = m\bar{v}_-(k) \) from the band structure, \( D(a) \) can be eliminated using Eq. (23). Specifically, near each of the Dirac points \( K_D \) where the dispersion becomes conical, we have

\[
\bar{v}_-(k \rightarrow K_D) = -v_F \frac{k - K_D}{|k - K_D|} = \frac{3\hbar D(a)}{2m} \frac{k - K_D}{|k - K_D|}
\] (25)

where \( v_F \) is the Fermi velocity and the expression on the right comes from expanding Eq. (23) about the point \( k = K_D \). Therefore, \( D(a) = -2m v_F/3\hbar \) and the off-diagonal momentum matrix element is

\[
\langle +, k|\hat{p}|-, k \rangle = \frac{2im v_F}{3} \left[ \frac{\hbar}{a} \cos(\phi(k)) + \frac{\hbar + a_1}{a} \cos(\phi(k) + k \cdot a_1) + \frac{\hbar + a_2}{a} \cos(\phi(k) + k \cdot a_2) \right]
\] (26)

The norm of this matrix element as a function of \( k \) is plotted in Fig. 2. The Dirac points are saddle points, where the norm has the value \( m v_F \), with the global maxima occurring at the midpoints between Dirac points, where the norm is \( 2m v_F \). Using the expression for the momentum matrix element in Eq. (26), the opacity can be found by integrating Eq. (2) over the entire BZ for a given light frequency \( \omega_0 \). This is straightforward near the Dirac points where the constant-energy contours are circular, but becomes less manageable farther away. Therefore, we performed the integration by the Monte Carlo method, using a narrow rectangular gate in place of the delta function. The red curve in Fig. 3 shows the calculated opacity.

The blue dashed curve in Fig. 3 shows the opacity calculated via a \( k \)-substitution to the tight-binding Hamiltonian from Ref. [17]. The two curves are qualitatively similar, with a
FIG. 2. Norm of the momentum matrix element $\langle +, k|\hat{p}|-, k \rangle$ as function of $k$ in the BZ. The black dots are the Dirac points $K_D$. Red indicates higher values and purple lower values, going to zero at $k = 0$.

peak near 5.6 eV due to the high joint density of states at this energy (as shown in Fig. 3). At low frequencies both curves show that the opacity rises gradually from $\pi \alpha$, matching the results of Refs. [1, 3, 4]. However, as the frequency increases the two curves diverge noticeably. With the $k$-substitution shortcut the opacity levels off near $\pi \alpha$ after the peak in a symmetric manner. But the present calculation based on the full evaluation of the matrix elements of $p$ shows the opacity decreasing for high frequencies, and is more asymmetric as a result. In this, it has a closer similarity to the experimental profile in Refs. [3, 4], which was also observed to have a high degree of asymmetry about the maximum, see Fig. 3. (The experimental peak was found to be red-shifted to 4.62 eV; this behavior was assigned to the influence of excitonic effects).

Thus a correct calculation of the optical transparency of graphene, performed by applying the vector potential coupling to the electrons’ canonical momentum, is straightforward and produces results which are in good agreement with the available experimental data. The two-dimensional character of the system, rather than the particular dispersion relation, is the dominant factor responsible for the low-frequency “universality” of a constant opacity. It is straightforward to extend the calculation to doped graphene [4], and to apply it to other 2D direct-gap materials. It would be very interesting if the measurement range of
FIG. 3. Opacity of graphene calculated with ’p-substitution’ (red solid curve) and ’k-substitution’ (blue dashed curve) as a function of photon energy, and the calculated joint density of states (JDOS) for valence-conductance band transitions as a function of energy difference. For comparison, the measured opacity of graphene from Ref. [4] is also shown with brown squares.

Graphene’s optical transmission could be extended to higher frequencies, so as to explore its behavior and shape as it dips back below the low-frequency value.

IV. ACKNOWLEDGMENTS

We would like to thank Prof. Gerd Bergmann and Prof. Dmitry Budker for valuable discussions. This research was supported by the U.S. National Science Foundation under grant CHE-1213410.

[1] R. R. Nair, P. Blake, A. N. Grigorenko, K. S. Novoselov, T. J. Booth, T. Stauber, N. M. R. Peres, and A. K. Geim, Science 320, 1308 (2008).
[2] K. F. Mak, M. Y. Sfeir, Y. Wu, C. H. Lui, J. A. Misewich, and T. F. Heinz, Phys. Rev. Lett. 101, 196405 (2008).
An elementary argument illustrates that the $k$-substitution procedure is fundamentally inconsistent. In the mentioned pseudo-relativistic tight-binding Hamiltonian, $\hat{H} = v_F \hat{\sigma} \cdot \hbar \mathbf{k}$, only $\hat{\sigma}$ is an operator while $\mathbf{k}$ is simply a vector. Therefore in the eigenfunction (i.e., band structure) basis, the Hamiltonian becomes diagonal ($\hat{H} = v_F \sigma_z \hbar k$). Now a $k$-substitution yields a completely diagonal perturbation Hamiltonian, meaning that no absorption can take place. However, in a proper quantum formulation, a similarity transformation cannot affect the physical description.