The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/91824

Please be advised that this information was generated on 2020-03-12 and may be subject to change.
First-principles Theory of Nonlocal Screening in Graphene

M. van Schilfgaarde\textsuperscript{1} and M. I. Katsnelson\textsuperscript{2}

\textsuperscript{1}\textit{School of Materials, Arizona State University, Tempe AZ}

\textsuperscript{2}\textit{Radboud University Nijmegen, Institute for Molecules and Materials, 6525AJ Nijmegen, The Netherlands}

Abstract

Using the quasiparticle self-consistent \textit{GW} (QSGW) and local-density (LD) approximations, we calculate the $q$-dependent static dielectric function, and derive an effective 2D dielectric function corresponding to screening of point charges. In the $q \to 0$ limit, the 2D function is found to scale approximately as the square root of the macroscopic dielectric function. Its value is $\approx 4$, a factor approximately 1.5 larger than predictions of Dirac model. Both kinds of dielectric functions depend strongly on $q$, in contrast with the Dirac model. The QSGW approximation is shown to describe QP levels very well, with small systematic errors analogous to bulk $sp$ semiconductors. Local-field effects are rather more important in graphene than in bulk semiconductors.

PACS numbers: 73.22.Pr, 71.27.+a, 73.22.-f
Graphene is a first truly two-dimensional (2D) crystal, with unique electronic and structural properties (for review, see Refs. 1–5). Screening of electron-electron and electron-impurity interactions in graphene is an important theoretical issue crucial for both many-body effects in electronic structure and for transport properties, especially for electron scattering by charge impurities. Here we develop a definition for an effective 2D dielectric function in an ab initio context, and calculate it within the Quasiparticle self-consistent GW (QS\textit{GW}) and local-density (LD) approximations. The former takes into account many-body effects beyond the density functional GGA or LDA schemes essential for correct description of excited states and thus screening effects. There are several \textit{GW} calculations for graphene, where \textit{G} and the screened Coulomb interaction \textit{W} are computed from the LDA. They all predict a notable (20-40\%) increase of the Fermi velocity \textit{v}_{\text{F}} at the Dirac point \textit{K} relative to the LDA(GGA) value, with \textit{v}_{\text{F}} between 1.1 and 1.2·10\textsuperscript{6} m/s, in very good agreement with experiment. The dielectric function and optical conductivity as a function of frequency \textit{\omega} for zero wave vector \textit{q}=0 was also calculated in Refs. 20,21. Here we focus on the static dielectric function (\textit{\omega}=0) as a function of \textit{q}. This quantity is relevant for calculations of resistivity via charge impurities, as well as for the problem of supercritical Coulomb centers and possible exciton instabilities.

The inverse dielectric function \textit{\epsilon}\textsuperscript{-1}(\textit{r}, \textit{r}', \textit{\omega}) relates the change in total potential \textit{\delta}\textit{V} to an external perturbing potential \textit{\delta}\textit{V}_{\text{ext}} as

\begin{equation}
\delta\textit{V}(\textit{r}, \textit{\omega}) = \int d\textit{r}' \epsilon\textsuperscript{-1}(\textit{r}, \textit{r}', \textit{\omega}) \delta\textit{V}_{\text{ext}}(\textit{r}', \textit{\omega}).
\end{equation}

\textit{\epsilon}\textsuperscript{-1} is obtained from a convolution of the polarization operator \textit{\Pi} and the bare Coulomb interaction \textit{v} as

\begin{equation}
\epsilon\textsuperscript{-1} = (1 - \textit{v}\textit{\Pi})\textsuperscript{-1}.
\end{equation}

In a system with translation symmetry, \textit{\epsilon}\textsuperscript{-1}, \textit{\Pi}, and \textit{v} can be expanded in Bloch functions \{\textit{B}_{\textit{I}}(\textit{r})\}, e.g.

\begin{equation}
\epsilon\textsuperscript{-1}(\textit{r}, \textit{r}', \textit{\omega}) = \sum_{\textit{q} \textit{I} \textit{J}} \textit{B}_{\textit{I}}(\textit{r}) \epsilon\textsuperscript{-1}_{\textit{I} \textit{J}}(\textit{\textit{q}}, \omega) \textit{B}_{\textit{J}}^\ast(\textit{r}').
\end{equation}
The most common choice of $\{B^q_I(r)\}$ are plane waves,

$$B^q_I(r) \rightarrow B^q_G(r) = \exp(i(q + G) \cdot r),$$

(3)

$G$ being reciprocal lattice vectors.

Quantities of interest are coarse-grained averages of $\epsilon^{-1}_{GG'}(q, \omega)$. The “macroscopic” response to a plane wave perturbation is

$$\epsilon_M(q, \omega) = \left[\epsilon^{-1}_{G=0,G'=0}(q, \omega)\right]^{-1}$$

(4)

The matrix structure of $\epsilon^{-1}$ with $G \neq G'$ reflects local field effects in terms of classical electrodynamics. The quantity $\epsilon_M(q)$ is commonly approximated by just $\epsilon(q)$; that is, the Umklapp processes, or local field effects are neglected. This is not such a bad approximation in $sp$ semiconductors but as we show here, it is a rather poor approximation in graphene. $\epsilon_M(q)$ corresponds to screening potential $\delta V^\text{ext}$ with a single Fourier component $q$. Selecting $G=G'=0$ averages $\epsilon^{-1}$ over the unit cell, restricting the spatial variation to the envelope $\exp(iq \cdot r)$. While $\epsilon_M$ is a quantity of relevance to some experiments, perhaps the most relevant is screening of a point charge in the graphene sheet, which governs e.g., scattering from impurities.

As graphene is a 2D system, we need to consider how the impurity potential $v(q) = 4\pi/q^2$ is screened in the sheet. The (statically) screened potential from a point charge at the origin may be written in cylindrical coordinates $r=(\rho, z, \theta)$ and $q=(\bar{q}, q_z, \theta_q)$ as

$$W(\rho, z) = \frac{1}{2\pi} \int_0^\infty d\bar{q} \bar{q} J_0(\bar{q}\rho) W^{2D}(\bar{q}, z)$$

(5)

$$W^{2D}(\bar{q}, z) = 4 \int_0^\infty dq_z e^{iq_z z} \frac{\epsilon^{-1}(\bar{q}, q_z)}{q^2 + q_z^2}$$

(6)

Thanks to graphene’s hexagonal symmetry, $\epsilon^{-1}$ does not depend on $\theta_q$ for small $\bar{q}$. $W^{2D}(\bar{q}, z)$ is the 2D (Hankel) transform of $W(r)$, the analog of the 3D transform $W(q)=\epsilon^{-1}(q)v(q)$. In the absence of screening $\epsilon^{-1} = 1$ and $W^{2D}(\bar{q}, z)$ reduces to the bare coulomb interaction $v^{2D}(\bar{q}, z)$:

$$v^{2D}(\bar{q}, z) = 4 \int_0^\infty dq_z e^{iq_z z} \frac{1}{q^2 + q_z^2} = \frac{2\pi}{\bar{q}} e^{-\bar{q}z}$$

(7)

An appropriate definition of an effective 2D dielectric function is then

$$\epsilon^{2D}(\bar{q}, z) = v^{2D}(\bar{q}, z)/W^{2D}(\bar{q}, z)$$

(8)

Graphene wave functions have some extent in $z$ which must be integrated over to obtain a scattering matrix element. But the largest contribution originates from $z=0$, so $W^{2D}(\bar{q}, 0)$ is a reasonable estimate for the scattering potential. This is particularly so for small $\bar{q}$ of primary interest here.
In practice we carry calculations in a periodic array of graphene sheets in the \( xy \) plane, spaced by a distance large enough that the sheets interact negligibly. To calculate \( \epsilon_{G=G'=0}^{-1}(\mathbf{q}, \omega) \) we adopt the all-electron, augmented wave implementation that was developed for the quasiparticle self-consistent \( GW \) (QSGW) approximation, described in Ref.\textsuperscript{26}. It makes no pseudo- or shape-approximation to the potential, and does not use PWs (Eq. \textsuperscript{3}) for the product basis \( \{B\} \), but a mixed basis consisting of products of augmented functions in augmentation spheres, and plane waves in the interstitial region. The all-electron implementation enables us to properly treat core states. We calculate \( \epsilon_{G=G'=0}^{-1}(\mathbf{q}, \omega) \) in the random phase approximation, using Bloch functions for eigenstates\textsuperscript{18}. These are obtained from single-particle eigenfunctions \( \Psi_{kn} \) and eigenvalues \( \epsilon_{kn} \) in both the LDA and QSGW approximations. In both cases the generalized LMTO method is used\textsuperscript{27,28}.

QSGW has been shown to be an excellent predictor of materials properties for many classes of compounds composed of elements throughout the Periodic Table, with unprecedented ability to consistently and reliably predict materials properties over a wide range of materials\textsuperscript{26,29-32}. Nevertheless there are small, systematic errors: in particular bandgaps in insulators such as GaAs, \( \text{SrTiO}_3 \) and NiO, are systematically overestimated. Its origin can be traced to a large extent to the RPA approximation to the polarizability, \( \Pi^{RPA}=iGG \). The RPA bubble diagrams omit electron-hole interactions in their intermediate states. Short-range attractive (electron-hole) interactions induce redshifts in \( \text{Im} \epsilon_{\mathbf{q}, \omega} \) at energies well above the fundamental bandgap; see e.g. Fig. \textsuperscript{6} in Ref.\textsuperscript{26}. That ladder diagrams are sufficient to remedy most of the important errors in \( \Pi^{RPA} \) was demonstrated rather convincingly in \( \text{Cu}_2\text{O} \), by Bruneval et al.\textsuperscript{33}. Moreover Shishkin et al.\textsuperscript{34} incorporated these ladder diagrams in an approximate way for several \( sp \) semiconductors, and established that they do in fact largely ameliorate the gap errors. Yang et al. investigated the effect of ladder diagrams in graphene and graphite, and showed that in a manner very analogous to ordinary semiconductors, these diagrams induce a redshift in the peak of \( \text{Im} \epsilon_{RPA}(\mathbf{q}, \omega) \) near 5 eV\textsuperscript{21} of \( \sim 0.6 \) eV. They found a strong similarity with conventional semiconductors, namely that the redshift from ladder diagrams approximately cancels the error in the LDA joint density of states.

A redshift in the peak of \( \text{Im} \epsilon(\omega) \) increases the static dielectric constant \( \epsilon_{\infty} \), as can be readily seen through the Kramers-Kronig relations. Remarkably, \( \epsilon_{\infty} \) as calculated by the RPA in QSGW, is underestimated by a nearly \textit{universal factor} of 0.8, for many kinds of insulators and semiconductors\textsuperscript{30}, including transition metal oxides such as \( \text{NiO} \textsuperscript{26}, \text{CeO}_2 \), and \( sp \) semiconductors\textsuperscript{34}. (This error
TABLE I: Energy gap $E_G$ and valence bandwidth $\Gamma_{1v}$ in diamond (eV); Fermi velocities $v_F$ in graphite and graphene ($10^6$ m/sec). There is a significant renormalization of the bandgap from the electron-phonon interaction in diamond, estimated to be 370 meV. Thus $QSGW$ overestimates $E_G$ by a slightly smaller amount than in other semiconductors, and the scaling of $\Sigma$ as described in the text results in a slightly underestimated gap. The electron-phonon interaction also reduces the Fermi velocity in graphene, estimated to be 4 to 8% in an LDA-linear response calculation. The calculated Fermi velocities should be reduced by this much when comparing to experiment. $v_F$ calculated by $QSGW$ is slightly overestimated, for much the same reason semiconductor gaps are overestimated. $v_F$ calculated from the scaled-$\Sigma$ potential, is slightly larger than $v_F$ calculated LDA-based $GW$, i.e. $G^{LDA}W^{LDA}$, just as semiconductor bandgaps are slightly larger. When renormalized by the electron-phonon interaction, $v_F$ agrees very well with the measured value.

|                  | LDA   | QSGW  | scaled $\Sigma$ | Expt  |
|------------------|-------|-------|-----------------|-------|
| $E_G$, diamond   | 4.09  | 5.93  | 5.56            | 5.50  |
| $\Gamma_{1v}$, diamond | 21.3  | 23.1  | 22.7            | 23.0 ± 0.2 |
| $\Gamma_{1v}$, graphene | 19.4  | 22.9  | 22.2            |       |
| $v_F$(H), graphite | 0.77  | 0.99  | 0.94            | 0.91 ± 0.15 |
| $v_F$(K), graphene | 0.82  | 1.29  | 1.20            | 1.1   |

*Ref. 39*

is often approximately canceled in the LDA, fortuitously. As Yang et al. noted, the cancellation seems to apply to graphene in a manner similar to ordinary semiconductors.) Because $\epsilon$ is systematically underestimated, $W = \epsilon^{-1}v$ and $\Sigma = -iGW$ are systematically overestimated; therefore QP excitation energies are also systematically overestimated. We have found that simply scaling by 0.8 (the nearly universal ratio $\epsilon_{QSGW}^{\infty}/\epsilon_{expt}^{\infty}$) largely eliminates discrepancies between $QSGW$ and measured QP levels in a wide range of $spd$ systems, including all zincblende semiconductors, and many other kinds of insulators. For graphene, we find that the $QSGW$ macroscopic ($q \rightarrow 0$) dielectric constant was found to be 80% of the LDA one, consistent with the universal pattern in bulk insulators noted above. The many points of consistency with 3D behavior, both in the $QSGW$ QP levels and the dielectric response suggest that $QSGW$ will exhibit the same reliable description of the 2D graphene system, with similar systematic errors. To confirm this, some band
parameters for three pure (undoped) carbon compounds calculated by QSGW and QSGW with Σ scaled by 0.8 are shown in Table I. Scaling QSGW has a minor effect on the quasiparticle levels: e.g. it reduces $v_F$ by 7%. As Table I shows, $v_F$ falls in very close agreement with experiment when Σ is scaled and the electron-phonon interaction is taken into account, consistent with agreement in gaps in the bulk insulators. Even though the QSGW and LDA work functions are similar (Fig. 1), the valence band is significantly widened relative to LDA, more so in graphene than in diamond.

Careful checks for convergence were made in various parameters. To check supercell artifacts, a “small” 3D unit cell with the graphene planes repeated at a spacing equivalent to 4 atomic layers of graphite (25 a.u.) was compared against a “large” cell, with graphene planes spaced at 8 layers. The bands from $-\infty$ to $E_F+5$ eV were found to be a very similar, with a slight increase in $v_F$ ($1.23\rightarrow1.29 \cdot 10^6$ m/s). $k$ convergence in the construction of Σ was monitored by comparing QP levels generated on a $6\times6\times2$ $k$ mesh to a $9\times9\times2$ mesh. QP levels were nearly identical: $v_F$ differed by <1% in the both the small and large 3D cells.

$\epsilon_{00}^{-1}(q, \omega)$ must be integrated with a fine $k$ mesh. To deal with the delicate $q\rightarrow0$ limit, we calculated $\epsilon^{-1}$ integrating on a standard $k$ mesh including Γ, and an offset mesh (Eqns. 47 and 52 in Ref. 26), and averaged them. We present data for averaged $18\times18\times4$ meshes. Calculations
without local fields were also performed for a pair of 24×24×4 meshes. $\epsilon(q, z=0, \omega=0)$ calculated by 18- and 24- (averaged) mesh integrations were essentially indistinguishable for $q>0.1 \times 2\pi/a$, and differed by a few percent for $q>0.02 \times 2\pi/a$.

$$
    \epsilon^0(q, q_z = 0, \omega = 0) \quad \text{and} \quad \omega \rightarrow 0 \quad \text{in Fig. 11 of Ref. 40.}
$$

$\epsilon^0(q, q_z = 0)$ was calculated on a grid of points $\{q, q_z\}$; the $q_z=0$ case is shown in the first panel of Fig. 2. It was found that $\epsilon^0$ is well parametrized (max error <0.1) by

$$
    \epsilon^{-1}_0(q, q_z) = \frac{a^2(q) + q_z^2}{\epsilon_0(q, 0) a^2(q) + q_z^2}
$$

$$
    a^2(q) = \frac{a_0 a_1 q^2}{a_1 + q^2} \approx a_0 q^2
$$

where $a_0=1.3$ and 1.2 for QSGW and LDA, respectively, and $a_1=1.6(2\pi/a)^2$. The approximate form for $a$ in Eq. 10 is sufficient for any $q$ where $\epsilon^0$ differs significantly from unity. With Eq. 9 $W^{2D}$ can be integrated analytically. Taking the approximate expression for $a^2(q)$ we obtain

$$
    \epsilon^{2D}(q, z) = \frac{\gamma(\gamma^2 - 1)}{\gamma(a_0 - 1) + (\gamma^2 - a_0)e^{(1-\gamma)q_z}}
$$

where $\gamma = \sqrt{a_0 \epsilon_0(q, 0)}$.

Fig. 2 shows both kinds of dielectric functions, $\epsilon_M$ corresponding to the macroscopic polarizability, and the effective 2D static dielectric function $\epsilon^{2D}(q, z=0)$ calculated from Eq. 9. Local
fields reduce the strength of the screening. The difference between LDA and QSGW results are modest; and as noted earlier, the LDA results are likely to be slightly better because they benefit from a fortuitous cancellation of errors. As $\bar{q}\rightarrow 0$, $\gamma$ is significantly larger than $a_0$ and unity. Keeping only the leading order in $\gamma$, we obtain the surprising result that $\epsilon^{\text{2D}}(0,z=0) \approx \sqrt{a_0\epsilon_{00}(\bar{q},q_z=0)}$. $\epsilon^{\text{2D}}(0,z=0)$ is roughly a factor 1.5 times larger than the Dirac Hamiltonian result at zero doping. Such a model predicts $\epsilon(q) \approx 2.4$ independent of $q$, as shown by Ando$^7$. We find $\epsilon^{\text{2D}}(\bar{q}, z=0) \approx 3.5$ for $\bar{q}\rightarrow 0$, but $\epsilon^{\text{2D}}$ is a very strong function of $\bar{q}$.

Although virtual transitions involving Van Hove peaks of the density of states might strongly enhance$^{16,17} \epsilon^{\text{2D}}$ were they sufficiently close to the Fermi level, apparently lie too far away in graphene. The case of small $\bar{q}$ ($\bar{q}\sim k_F \leq 10^{-2}$ Å$^{-1}$) is relevant for transport properties. In this region our first-principles calculations do not dramatically contradict predictions of the Dirac model. At the same time, for the problem of supercritical Coulomb centers and relativistic collapse (fall on the center)$^{10,23,24}$ distances of order of several inverse lattice constants are essential (this is the radius of screening cloud, according to renormalization group analysis$^{10}$), which corresponds to larger $q$. For this region our results show that the Dirac model overestimates the screening.

**Acknowledgments**

MIK acknowledges support from Stichting voor Fundamenteel Onderzoek der Materie (FOM), the Netherlands. MvS was supported by ONR contract N00014-7-1-0479 and NSF QMHP-0802216.
A. K. Geim and K. S. Novoselov, Nature Mater. 6, 183 (2007).

M. I. Katsnelson, Mater. Today 10, 20 (2007).

A. H. C. Neto, F. Guinea, N. M. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. 81, 109 (2009).

A. K. Geim, Science 324, 1530 (2009).

S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, preprint arXiv:1003.4731.

J. González, F. Guinea, and M. A. H. Vozmediano, Nucl. Phys. B 424, 596 (1994).

T. Ando, J. Phys. Soc. Japan 75, 074716 (2004).

B. Wunsch, T. Stauber, F. Sols, and F. Guinea, New J. Phys. 8, 318 (2006).

M. I. Katsnelson, Phys. Rev. B 74, 201401(R) (2006).

A. V. Shytov, M. I. Katsnelson, and L. S. Levitov, Phys. Rev. Lett. 99, 236801 (2007).

M. M. Fogler, D. S. Novikov, and B. I. Shklovskii, Phys. Rev. B 76, 233402 (2007).

M. Polini, A. Tomadin, R. Asgari, and A. H. MacDonald, Phys. Rev. B 78, 115426 (2008).

E. Rossi and S. Das Sarma, Phys. Rev. Lett. 101, 166803 (2008).

L. Brey and H. A. Fertig, Phys. Rev. B 80, 035406 (2009).

M. M. Fogler, Phys. Rev. Lett. 103, 236801 (2009).

M. I. Katsnelson and A. V. Trefilov, Phys. Lett. A 109, 109 (1985).

M. I. Katsnelson and A. V. Trefilov, Phys. Rev. B 61, 1643 (2000).

F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys 61, 237 (1998).

G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. 74, 601 (2002).

P. E. Trevisanutto, C. Giorgetti, L. Reining, M. Ladisa, and V. Olevano, Phys. Rev. Lett. 101, 226405 (2008).

L. Yang, J. Deslippe, C.-H. Park, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. 103, 186802 (2009).

C. Attacalite and A. Rubio, Phys. Stat. Sol. (b) 246, 2523 (2009).

A. V. Shytov, M. I. Katsnelson, and L. S. Levitov, Phys. Rev. Lett. 99, 246802 (2007).

V. M. Pereira, J. Nilsson, and A. H. Castro-Neto, Phys. Rev. Lett. 99, 166802 (2007).

J. Sabio, F. Sols, and F. Guinea, Phys. Rev. B 81, 045428 (2010).

T. Kotani, M. van Schilfgaarde, and S. V. Faleev, Phys. Rev. B 76, 165106 (2007).

M. van Schilfgaarde, T. Kotani, and S. V. Faleev, Phys. Rev. B 74, 245125 (2006).
To ensure completeness in the interstitial floating orbitals were placed in the “empty sites” where nuclei would fall if graphene were turned into graphite. An $spdfspd$ basis was used for C atoms.

28 S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. 93, 126406 (2004).

29 M. van Schilfgaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. 96, 226402 (2006).

30 A. N. Chantis, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. 96, 086405 (2006).

31 A. N. Chantis, M. van Schilfgaarde, and T. Kotani, Phys. Rev. B 76, 165126 (2007).

32 F. Bruneval, N. Vast, L. Reining, M. Izquierdo, F. Sirotti, and N. Barrett, Phys. Rev. Lett. 97, 267601 (2006).

33 M. Shishkin, M. Marsman, and G. Kresse, Phys. Rev. Lett. 99, 246403 (2007).

34 M. Cardona and M. L. W. Thewalt, Rev. Mod. Phys. 77, 1173 (2005).

35 C.-H. Park, F. Giustino, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. 99, 086804 (2007).

36 P. E. Trevisanutto, C. Giorgetti, L. Reining, M. Ladisa, and V. Olevano, Phys. Rev. Lett. 101, 226405 (2008).

37 Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, Nature 438, 201 (2005).

38 I. Jiménez, L. J. Terminello, D. G. J. Sutherland, J. A. Carlisle, E. L. Shirley, and F. J. Himpsel, Phys. Rev. B 56, 7215 (1997).

39 A. G. Marinopoulos, L. Reining, A. Rubio, and V. Olevano, Phys. Rev. B 69, 245419 (2004).