Screening of a hypercritical charge in graphene

M. M. Fogler,[1] D. S. Novikov,[2] and B. I. Shklovskii[2]

[1]*Department of Physics, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA*
[2]*W. I. Fine Theoretical Physics Institute, University of Minnesota, Minneapolis, Minnesota 55455, USA*
(Received 19 October 2007; published 5 December 2007)

Screening of a large external charge in graphene is studied. The charge is assumed to be displaced away or smeared over a finite region of the graphene plane. The initial decay of the screened potential with distance is shown to follow the 3/2 power. It gradually changes to the Coulomb law outside of a hypercritical core whose radius is proportional to the external charge.

Recent discovery of graphene—a two-dimensional (2D) form of carbon[1]—brought an exciting link between solid-state physics and quantum electrodynamics (QED). The half-filled π-band of graphene has a relativistic massless Dirac spectrum $\epsilon = \pm \hbar v |\mathbf{k}|$ where $\epsilon > 0$ for the electrons and $\epsilon < 0$ for holes, $\mathbf{k}$ is the deviation of the quasimomentum from the Brillouin zone corner, and $v \approx 10^6$ m/s. The role of the fine-structure constant is played by the dimensionless parameter

$$\alpha = e^2/\kappa \hbar v, \quad e^2/\hbar v \approx 2.2, \qquad (1)$$

where κ is the dielectric constant at the interface of substrate and vacuum. For conventional SiO_2 substrates $\kappa \approx 2.4$; hence Coulomb interaction is strong, $\alpha \sim 1$.

In this work, we consider the problem of screening of a Coulomb potential $V_0 = eZ/\kappa r$ that can be induced in graphene by a group of charged impurities in the substrate, by a nearby gate, or by a cluster of dopants. This problem is important for a number of properties of graphene nanostructures, including transport,[2–5] local gating,[6–9] controlled doping,[10,11] and chemical sensing.[12] Not surprisingly, it has attracted much attention.[2,4,5,13–19] In particular, it has been noted[16–18] that at half-filling the problem of a Coulomb charge in graphene has an interesting parallel with that of a hypothetical supercritical atom with $Z > \hbar c/e^2 \approx 137$. For such an atom, the standard solution[20] of the Dirac equation breaks down and a physically acceptable atomic structure is obtained only after accounting for a finite radius of the nucleus.[21] This structure is characterized by a vacuum reconstruction: a certain number of electrons is spontaneously created (liberating positrons), they bind to the nucleus, and render it subcritical. In graphene the critical charge[16–18,22] $Z_c \simeq 1/2\alpha$ is much smaller than in QED; hence solid-state analogs of supercritical atoms may be realizable even at $Z \sim 1$.

According to all prior investigations, screening properties of an undoped graphene resemble those of a dielectric: the screened potential V of a supercritical charge has been argued not to deviate much from the Coulomb law,

$$V(r) = \frac{e}{\kappa r} \frac{F(r)}{2\alpha^2}. \qquad (2)$$

Here $F(r)$ is a slow logarithmic function. Such a conclusion follows from the standard linear response theory—random phase approximation (RPA)[2,14]—and was supposedly confirmed by calculations within the Thomas-Fermi (TF)

method[13,15,16] that is able to go beyond the linear response. Below we reexamine these conclusions for the case of a hypercritical charge $Z \gg 1$, which lets itself to a controlled treatment and adds new physics. Without loss of generality we assume that the external charge attracts electrons, $Z > 0$.

Since it was not always made clear previously, we emphasize that the problem is ill-defined unless one explicitly regularizes the strong Coulomb singularity at the origin. This is as crucial as introducing a finite size of a nucleus in QED. Therefore, the charge Z must be either displaced away from graphene plane by some distance d or spread over the area of some radius r_0 in this plane. In order to deal exclusively with Dirac fermions the smearing parameter $\max\{d, r_0\}$ must exceed $a\alpha\sqrt{Z}$ where $a = 2.5$ Å is the graphene lattice constant; otherwise, the quasiparticle energy shift due to the potential V would exceed the modest energy separation 4 eV $\sim e^2/a$ of the Dirac point and the nearest σ bands.[23] These other bands would then also need to be included, leading one to a three-dimensional (3D) problem that has little to do with special properties of graphene.

Our main result is that the induced 2D electron density and the screened potential have the form

$$n(r) \simeq \frac{1}{4\pi\alpha^2} \frac{r_1}{r^3}, \quad V(r) \simeq \frac{e}{2\alpha^2 \kappa} \sqrt{\frac{r_1}{r^3}}, \quad r_1 \equiv 2\alpha^2 Zd, \qquad (3)$$

in the range of distances $\max\{d, r_0\} \ll r \ll r_1$. Based on electrostatics, the law (3) is robust and universal. How then does one reconcile it with Eq. (2)? As we clarify below, the situation is as follows. In the strongly interacting case, $\alpha \sim 1$, Eq. (3) controls the entire supercritical core, i.e., the circle around the origin where the net charge exceeds Z_c. This fact has eluded previous studies. However, if α is small, the domain of validity of Eq. (3) narrows down, opening up a window where Eq. (2) is realized. Although current experiments are not in this regime, small α can be achieved using large κ substrates, e.g., HfO_2 (Ref. 6) or simply liquid water, $\kappa \sim 80$.

The three-line derivation of Eq. (3) can be given if, as discussed above, the charge Z is pointlike but removed from the graphene plane[13] by an appropriate distance d. The key idea is that if we treat the graphene sheet as a perfect metal, then classical electrostatics dictates that the induced charge density is given by $n = n_{cl}$, where

$$n_{cl}(r) = \frac{1}{2\pi} \frac{Zd}{(r^2+d^2)^{3/2}} = \frac{1}{4\pi\alpha^2} \frac{r_1}{(r^2+d^2)^{3/2}}. \quad (4)$$

At $r \gg d$ we get the first formula in Eq. (3). To derive $V(r)$ we employ the TF approximation,

$$\mu[n(r)] - eV(r) = 0. \quad (5)$$

Combined with the formula for the chemical potential,

$$\mu(n) = \text{sgn}(n) \sqrt{\pi} \hbar v |n|^{1/2}, \quad (6)$$

specific for the 2D Dirac spectrum, it yields the second formula in Eq. (3), concluding the derivation. The rest of our paper is needed mainly to explain why the above reasoning is correct, why Eq. (3) is completely general rather than restricted to the case of a remote charge, and finally, where the room may still exist for the differing predictions for n and V advocated in Refs. 13, 15, and 16.

First, let us clarify why it was legitimate to approximate the density response of graphene—a complicated quantum system—simply by that of an ideal metal. The reason is this. At $r \ll r_1$ the local screening length $r_s = (\kappa/2\pi e^2)(d\mu/dn) \sim \alpha^{-1}|n|^{-1/2}$ is much smaller than the characteristic scale $\max\{r,d\}$ over which the potential $V(r)$, or equivalently, the effective background 2D charge density $n_{cl}(r)$ vary. Therefore, the unscreened charge density, $\sigma(r) \equiv n_{cl}(r) - n(r)$, is smaller than the background one, $n_{cl}(r)$, by some large factor related to the ratio of r_s and $\max\{r,d\}$. [The precise relation is expressed by Eqs. (11) and (12) below.]

The next step is to explain why or rather *where* the TF approximation can be trusted. This is determined by the conditions that $\max\{r,d\}$ exceeds the local Fermi wavelength $\lambda_F(r) \sim n^{-1/2}(r)$. For $\alpha \sim 1$ we can use $n(r)$ from Eq. (3) to write this condition as $r \lesssim r_2 = Zd$. Thus, for $\alpha \sim 1$ the domains of validity of the TF and the perfect screening approximations coincide, $r_1 \sim r_2$. At $r \ll r_2$ all corrections to Eq. (3), both smoothly varying with r and Friedel oscillations[16,24] are subleading.

Let us briefly discuss the nature of screening at $r > r_2$ where the TF approximation breaks down. Define $Q(r)$ to be the net effective charge inside the circle of radius r,

$$Q(r) \equiv \int_0^r 2\pi\sigma(r')r'dr'. \quad (7)$$

At $r = r_2$, Q drops to a number of the order of the critical one $Z_c \sim 1/2\alpha$. Consideration of screening now requires a detailed analysis of the eigenstates of the Dirac equation[16–18] in the potential created by the charge $Q(r_2)$. According to Ref. 16, some amount of charge, in fact, exactly the critical one remains unscreened: $Q(\infty) = Z_c$. The saturation of Q at this value occurs near a certain $r = r_*$. However for $\alpha \sim 1$, r_* and r_2 must coincide up to a factor of the order of unity. Thus Eq. (3) governs the entire supercritical core except perhaps a nonparametrically wide outer region $r \sim r_2$ where a more complicated dependence[16] may apply. At even larger distances the potential $V(r)$ follows the RPA prediction

$$V(r) \simeq eZ_c/\varepsilon r, \quad r \gg r_2, \quad (8)$$

where $\varepsilon = \kappa[1 + (\pi/2)\alpha]$ is the RPA dielectric constant.[2,14] A more careful examination of the behavior of $V(r)$ at such r requires accounting for the infrared renormalization of α (which enters ε),[14,19,27] that is not directly related to the problem at hand.

Let us return to the analysis of the supercritical region and show how to refine our results by computing corrections to Eq. (3). For this we complete the set of the TF Eqs. (5) and (6) by adding another one for $V(r)$:

$$\frac{\kappa}{e}V = \int \frac{d^2\mathbf{r}' \sigma(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = \int_0^\infty dq\, J_0(qr)\tilde{\sigma}(q) = \int_0^r \frac{g(s)ds}{\sqrt{r^2-s^2}}, \quad (9)$$

where $J_0(z)$ is the Bessel function[25] and $\tilde{\sigma}$, aptly parametrized by $\tilde{\sigma}(q) = \int_0^\infty ds\, g(s)\cos qs$,[26] is the 2D Fourier transform of σ. Inverting the last equation of Eq. (9), we get $g(u) = (2\kappa/\pi e)(d/du)\int_0^u V(s)s\, ds/\sqrt{u^2-s^2}$ and

$$Q(r) = Q(\infty) - \frac{2}{\pi} \frac{\kappa}{e^2} \int_r^\infty \frac{u\, du}{\sqrt{u^2-r^2}} \frac{d}{du} \int_0^u \frac{eV(s)s\, ds}{\sqrt{u^2-s^2}}. \quad (10)$$

The leading correction to the perfect screening can be obtained by substituting $\mu[n_{cl}(s)]$ in lieu of $eV(s)$; cf. Eq. (5). The resultant expression is cumbersome, and so we quote only the limiting forms:

$$\frac{\sigma(r)}{n_{cl}(r)} \simeq \begin{cases} -\Gamma^2(5/4)\sqrt{\dfrac{8d}{\pi r_1}}, & r \ll d, \quad (11) \\[2ex] \dfrac{16}{\pi^2}\Gamma^4(5/4)\sqrt{\dfrac{r}{r_1}}, & d \ll r \ll r_1, \quad (12) \end{cases}$$

where the Gamma-function[25] $\Gamma(5/4) \approx 0.906$. In agreement with the above physical argument, the deviation from the perfect screening at all $r \ll r_1$ is small.

These analytical predictions were verified by numerical simulations. To this end we solved the TF equations (5)–(9) inside a finite square of the 2D plane. The integrals were replaced by discrete sums over a uniform 256×256 grid defined therein and the periodic boundary conditions were imposed. The solution for $n(r)$ and $V(r)$ was found by a standard iterative method, using underrelaxation to ensure convergence. As shown in the inset of Fig. 1, the analytical and the numerical results agree extremely well for a suitably large hypercritical charge $\alpha^2 Z = 20$.

Small-α regimes. Let us now show that Eq. (2) can be reconciled with our theory under the condition $\alpha \ll 1$, i.e., $\kappa \gg 1$. In this case there is a gap between the above defined characteristic length scales r_1 and r_2. This gap is filled by an additional regime where the TF approximation regime is still valid but the screening is ineffective.

To see that consider first moderately small α, such that $1/\sqrt{Z} \ll \alpha \ll 1$. Since the screening is weak, Eq. (10) is no longer convenient. Instead, the derivation of V and n can be done along the lines of Ref. 15 but with several important refinements. First, we trade the two last equations of Eq. (9) for
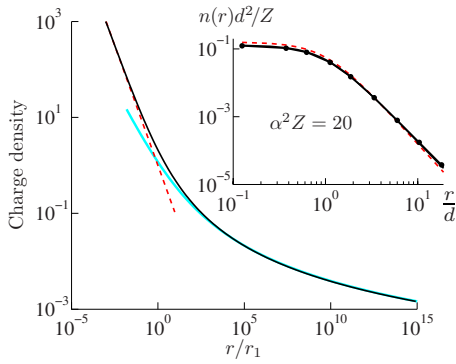
FIG. 1. (Color online) Main panel: Density profile in the limit of $d \to 0$ at fixed r_1. The quantity plotted on the vertical axis is $4\pi\alpha^2 r_1^2 n(r) = F^2(r) r_1^2 / r^2$. The thin line is the numerical solution of Eq. (15); the dashed line is the perfect screening, $F = r_1/r$; the thick line is for F from Eq. (16) with $c = 0.6$. Inset: expanded view of $n(r)$ inside the hypercritical core. The thin line and the dashed line have the same meaning as before; the dots correspond to an analytical formula whose limits are given by Eqs. (11) and (12).

$$\frac{\kappa}{e} V(r) = \int_0^\infty \frac{4r' dr'}{r' + r} K\left(\frac{2\sqrt{rr'}}{r + r'} \right) [n_{cl}(r') - n(r')], \quad (13)$$

where $K(z)$ is the complete elliptic integral of the first kind.[25] Next, we treat Eq. (2) as the definition of yet unknown function F and use Eqs. (5) and (6) to obtain

$$n(r) = F^2(r)/4\pi\alpha^2 r^2. \quad (14)$$

Taking the limit $d \to 0$ at fixed r_1, we get the equation

$$F(t) = \int_{-\infty}^\infty du [\theta(t - u) + \phi(u - t)][e^{-u} - F^2(u)], \quad (15)$$

where $t = \ln(r/r_1)$ and $\theta(t)$ is the unit step function. Function $\phi(t)$, defined by Eq. (12) of Ref. 15, has the following properties: It is a logarithmically divergent at $t = 0$, is exponentially small at $|t| \gg 1$, and satisfies $\int \phi(t) dt = \ln 4$. It is easy to see that at large negative t, we must have perfect screening, $F^2(t) \simeq e^{-t}$. The asymptotic behavior of F at large $t > 0$ can be deduced by replacing $\phi(t)$ with $(\ln 4)\delta(t)$.[28] After this, we can differentiate the integral equation (15) to get

$$F^{-1} - (2\ln 4)\ln F = \ln(r/r_1) + c, \quad c = \text{const}, \quad (16)$$

where we returned to the original linear coordinate r. The direct numerical solution of Eq. (15) shows in excellent agreement with Eq. (2) and Eq. (16) if the constant c is set to 0.6; see Fig. 1. At $r \sim r_1$, this solution crosses over to the strong screening regime, Eq. (12).

The range of $r \ll r_2$ where Eq. (16) is valid is again determined by the condition $r \gg n^{-1/2}(r)$, which yields

$$r_2 \sim r_1 \exp(1/2\sqrt{\pi}\alpha). \quad (17)$$

At $r \gg r_2$ the screened potential is given by Eq. (8).

Consider even smaller α, such that $1/Z \ll \alpha \ll 1/\sqrt{Z}$. Here the Coulomb interactions are so weak that the smearing of the external charge is no longer necessary: The "dangerous"

region $r < a\alpha\sqrt{Z}$ is smaller than the lattice constant. In addition, the domain of the perfect screening, $r < r_1$, which is the region of validity of Eq. (3) disappears. (Weak interactions entail poor screening.) In this case $c \to (2\alpha^2 Z)^{-1} + \ln(r_1/a)$, so that the solution

$$V(r) \simeq \frac{eZ}{\kappa r} \frac{1}{1 + 2\alpha^2 Z \ln(r/a)}, \quad (18)$$

advocated in Ref. 15 actually applies, at $\ln(r/a) < 1/\alpha$.

In-plane charge. In the concluding part of the paper we wish to return to the structure of the hypercritical core and show that Eq. (3) remains valid if the charge Z resides within the 2D plane. To gain some intuition consider first an artificial scenario where the external charge is highly localized yet the σ bands of graphene can be neglected. In this case the maximum possible electron density (relative to that of the half-filled π band) is $n_{\max} = 2/\sqrt{3}a^2$. This density is indeed reached at r smaller than some radius b as a result of attraction of electrons to the hypercritical charge Z. At $r > b$, electron density is gradually decreases, which can be thought of as appearance of "holes" at the top of the conduction band.

Incidentally, the charge profile of these holes within the perfect screening approximation is known exactly. It can be read off the results of Ref. 29 where the structure of a depletion region around a disk of a negative charge in a semiconductor was studied. For a high density of the external charge these authors found that $b = (Z/2\pi n_\infty)^{1/2}$, where n_∞ is the uniform electron density far away from the depletion. They also found[29] that the density profile at large r is given by $n(r) = n_\infty - (Zb/2\pi r^3)$ at $r \gg b$. Adopting these results to our problem, we get

$$n(r) = Zb/2\pi r^3, \quad r \gg b = (Z/2\pi n_{\max})^{1/2}, \quad (19)$$

leading to Eq. (3) with $r_1 = 2\alpha^2 Zb \sim aZ^{3/2}$ for $\alpha \sim 1$.

Consider now a more realistic setup where the external charge $n_{ext}(r)$ is distributed over a disk of radius $r_0 \gg a\sqrt{Z}$. Then $n \lesssim n(0) \sim Z/\pi r_0^2 \ll n_{\max}$, so that σ bands can indeed be disregarded. Let us show that

$$n(r) \sim \frac{Zr_s}{2\pi r^3}, \quad r_0 \ll r \ll r_1 = 2\alpha^2 Zr_s, \quad (20)$$

where the screening length $r_s \sim 1/\alpha\sqrt{n(0)}$.

Based on the near-perfect screening framework used in the first part of the paper [and justified *a posteriori* by Eq. (20)] we can claim that $V(r)$ is substantial only in the region $r < r_0$ and is greatly reduced at $r > r_0$. This implies that the Fourier transform of V is nearly wave-vector-independent over a range of q,

$$\tilde{V}(q) = c_1 Zer_s/\kappa + \mathcal{O}(Z^{1/2}), \quad r_1^{-1} \ll q \ll r_0^{-1}.$$

The first term, with $c_1 \sim 1$, follows from Eq. (5). In turn, the Fourier transform of the charge density, $\tilde{\sigma}(q) = \tilde{n}_{ext}(q) - \tilde{n}(q) = \tilde{V}(q)/(2\pi e/\kappa q)$, that produces this potential behaves as $\tilde{\sigma}(q) = c_2 Zr_s q + \mathcal{O}(Z^{1/2})$, where $c_2 \sim 1$. After the inverse Fourier transform, the net charge density $\sigma(r)$ is seen to be

dominated by the term $-c_2 Z r_s / 2\pi r^3$ at $r_0 \ll r \ll r_1$. Since $n_{\text{ext}}(r)=0$ for such r, this term is entirely due to n, proving our statement.

In summary, we considered the problem of nonlinear screening of a large charge by the massless electrons in graphene. The consistent formulation of the problem requires the charge to be either displaced from the graphene plane or to be spread over a disk of finite radius r_0. In both cases the screening is nonlinear within a region of a parametrically large radius r_1. In the interval between r_0 and r_1 the screened potential decays as $1/r^{3/2}$. Our results are relevant for current and future experiments that involve local charging or doping of graphene. Thus, if small α can be achieved experimentally, it may be possible to verify the predicted crossover from Eq. (3) to Eq. (2) and finally to Eq. (18) by using scanned probe techniques.[9]

[1] For a review, see A. K. Geim and K. S. Novoselov, Nat. Mater. **6**, 183 (2007).

[2] T. Ando, J. Phys. Soc. Jpn. **75**, 074716 (2006).

[3] K. Nomura and A. H. MacDonald, Phys. Rev. Lett. **98**, 076602 (2007).

[4] E. H. Hwang, S. Adam, and S. Das Sarma, Phys. Rev. Lett. **98**, 186806 (2007).

[5] S. Adam, E. H. Hwang, V. M. Galitski, and S. Das Sarma, arXiv:0705.1540, Proc. Natl. Acad. Sci. USA (to be published).

[6] B. Özyilmaz, P. Jarillo-Herrero, D. Efetov, D. A. Abanin, L. S. Levitov, and P. Kim, Phys. Rev. Lett. **99**, 166804 (2007).

[7] B. Huard, J. A. Sulpizio, N. Stander, K. Todd, B. Yang, and D. Goldhaber-Gordon, Phys. Rev. Lett. **98**, 236803 (2007).

[8] J. R. Williams, L. DiCarlo, and C. M. Marcus, Science **317**, 638 (2007).

[9] J. Martin, N. Akerman, G. Ulbricht, T. Lohmann, J. H. Smet, K. von Klitzing, and A. Yacoby, arXiv:0705.2180 (unpublished).

[10] T. Ohta, A. Bostwick, T. Seyller, K. Horn, and E. Rotenberg, Science **315**, 951 (2006).

[11] B. Uchoa, C.-Y. Lin, and A. H. Castro Neto, arXiv:0706.1237 (unpublished).

[12] F. Schedin, A. K. Geim, S. V. Morozov, D. Jiang, E. H. Hill, P. Blake, and K. S. Novoselov, Nat. Mater. **6**, 652 (2007); E. H. Hwang, S. Adam, and S. Das Sarma, arXiv:cond-mat/0610834 Phys. Rev. B (to be published); T. O. Wehling, K. S. Novoselov, S. V. Morozov, E. E. Vdovin, M. I. Katsnelson, A. K. Geim, and A. I. Lichtenstein, arXiv:cond-mat/0703390 (unpublished).

[13] D. P. DiVincenzo and E. J. Mele, Phys. Rev. B **29**, 1685 (1984).

[14] J. González, F. Guinea, and M. A. H. Vozmediano, Nucl. Phys. B **424**, 595 (1994); Phys. Rev. B **59**, R2474 (1999).

[15] M. I. Katsnelson, Phys. Rev. B **74**, 201401(R) (2006).

[16] A. V. Shytov, M. I. Katsnelson, and L. S. Levitov, arXiv:0705.4663 (unpublished).

[17] D. S. Novikov, arXiv:0706.1391 (unpublished).

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

[19] R. R. Biswas, S. Sachdev, and D. T. Son, arXiv:0706.3907 (unpublished).

[20] C. G. Darwin, Proc. R. Soc. London, Ser. A **118**, 654 (1928); W. Gordon, Z. Phys. **48**, 11 (1928).

[21] I. Pomeranchuk and Y. Smorodinsky, J. Phys. (USSR) **9**, 97 (1945); Y. B. Zeldovich and V. S. Popov, Usp. Fiz. Nauk **105**, 403 (1971) [Sov. Phys. Usp. **14**, 673 (1972)].

[22] V. R. Khalilov and C. L. Ho, Mod. Phys. Lett. A **13**, 615 (1998).

[23] R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **46**, 1804 (1992).

[24] V. V. Cheianov and V. I. Falko, Phys. Rev. Lett. **97**, 226801 (2006); D.-H. Lin, Phys. Rev. A **73**, 044701 (2006).

[25] I. S. Gradshteyn and I. M. Ryzhik, in *Table of Integrals, Series, and Products*, 6th ed., edited by A. Jeffrey and D. Zwillinger (Academic, San Diego, 2000).

[26] I. N. Sneddon, *Mixed Boundary Value Problems in Potential Theory* (Wiley, New York, 1966), Sec. 3.2.

[27] D. T. Son, Phys. Rev. B **75**, 235423 (2007).

[28] Dropping $\phi(t)$ altogether (Ref. 15) gives poor numerical accuracy even at the largest t used in the graph of Fig. 1.

[29] T. Deruelle, K. Ensslin, P. M. Petroff, A. L. Efros, and F. G. Pikus, Phys. Rev. B **45**, 9082 (1992).