Coulomb disorder in three-dimensional Dirac systems
Brian Skinner
Phys. Rev. B 90, 060202 — Published 20 August 2014
DOI: 10.1103/PhysRevB.90.060202
Coulomb disorder in three-dimensional Dirac systems

Brian Skinner

Materials Science Division, Argonne National Laboratory, Argonne, IL 60439, USA

(Dated: June 9, 2014)

In three-dimensional materials with a Dirac spectrum, weak short-ranged disorder is essentially irrelevant near the Dirac point. This is manifestly not the case for Coulomb disorder, where the long-ranged nature of the potential produced by charged impurities implies large fluctuations of the disorder potential even when impurities are sparse, and these fluctuations are screened by the formation of electron/hole puddles. In this paper I present a theory of such nonlinear screening of Coulomb disorder in three-dimensional Dirac systems, and I derive the typical magnitude of the disorder potential, the corresponding density of states, and the size and density of electron/hole puddles. The resulting conductivity is also discussed.

I. INTRODUCTION

Due to its long-ranged nature, Coulomb disorder often has dramatic consequences even in situations where short-ranged disorder does not. Thus, for example, Coulomb disorder has always deserved special consideration in the physics of semiconductors and such studies have revealed a great number of diverse and interesting scientific phenomena over the preceding half-century.

The large qualitative difference between short-ranged and Coulomb disorder is particularly pronounced for three-dimensional (3D) Dirac materials, in which the electron kinetic energy $\varepsilon$ is linearly proportional to the momentum $\hbar \vec{k}$ according to $\varepsilon = \hbar^2 |\vec{k}|$. To see this difference between short-ranged and Coulomb disorder qualitatively, one can compare the behavior of long-wavelength (small-$k$) electron states in a 3D Dirac system (3DDS) for the two cases. Suppose, for example, that the 3DDS has some concentration $N$ of positive and negative impurities per unit volume, each with random position and random sign, and that these are taken to be either short-ranged, with finite range $a$ and typical potential $\pm V_0$, or Coulomb, with charge $\pm e$. In the short-ranged case, an electron wavepacket with size $\lambda \gg N^{-1/3} \gg a$ experiences disorder from $\sim N\lambda^3$ impurities. The average value of the disorder potential created by these impurities is zero, since impurities with opposite signs are equally plentiful, but statistical fluctuations in the impurity concentration create a typical excess of $\sim \sqrt{N\lambda^3}$ impurities with one of the two signs. Thus, the volume-averaged disorder potential experienced by the electron is $\sim V_0 \sqrt{N\lambda^3}/(\lambda/a)^3 \sim 1/\lambda^{3/2}$. The electron kinetic energy, on the other hand, scales as $\varepsilon \propto k \propto 1/\lambda$. One can therefore conclude that short-ranged disorder has a perturbatively small effect on the electron energy for large-wavelength electron states (i.e., for states close to the Dirac point). This robustness of the 3D Dirac point against short-ranged disorder has long been understood theoretically and consequently “Dirac semimetal” phases with vanishing density of states (DOS) are generally predicted to survive short range disorder. (In fact, a very recent paper has suggested that rare resonances between short-ranged impurities can create an exponentially small DOS at the Dirac point.)

Now consider the case of disorder produced by long-ranged Coulomb impurities. As before, an electron wavepacket with large size $\lambda$ encloses many impurities of both signs, and the net charge of these is $\sim \pm e\sqrt{N\lambda^3}$. If one naively calculates the potential energy created by these impurities, one finds that the typical Coulomb potential energy experienced by the electron is $\sim e^2 \sqrt{N\lambda^3}/\kappa \lambda$ (in Gaussian units), where $\kappa$ is the dielectric constant. Thus, the disorder potential grows with increasing wavelength as $\lambda^{1/2}$, rather than falling off quickly and becoming irrelevant. Clearly, such Coulomb impurities must have a large and nonperturbative effect near the Dirac point at any finite concentration. As one might expect, the growth of the Coulomb potential at large length scales is in fact truncated by the formation of electron and hole puddles that screen the disorder potential, as is the case with narrow band gap semiconductors and two-dimensional Dirac systems like graphene and topological insulators. This disorder-induced puddling is shown schematically in Fig. 1.

It is the purpose of this paper to calculate the typical size and density of these puddles, as well as the corresponding disorder potential amplitude, DOS, and conductivity.

![Fig. 1. Schematic illustration of electron/hole puddles in a 3DDS at the Dirac point. The Coulomb potential energy $e\phi$ is shown as a function of some coordinate $r$. Coulomb impurities create potential fluctuations with typical size $r_s$ and amplitude $\Gamma$. Regions of positive $e\phi$ correspond to electron puddles, while negative $e\phi$ implies hole puddles.](image-url)
The question of disorder effects in 3DDSs has acquired a particular relevance in recent months, following the experimental discovery of two different 3D Dirac materials\[16,19\] not long after their theoretical prediction.\[20–24\] While the existence of a Dirac dispersion in these materials has been established, largely by photoemission experiments, it remains to be thoroughly understood how closely the Dirac point can be probed and to what extent its behavior is masked by disorder. As shown below, the presence of Coulomb impurities has the effect of “smearing” the Dirac point via the creation of electron and hole puddles, and this smearing typically occurs over tens of meV.

The structure and primary results of this paper are as follows. In Sec. II, a self-consistent theory is developed to describe the disorder potential based on the Thomas-Fermi (TF) approximation. Corresponding expressions are derived for the magnitude of the disorder potential [Eq. (7)], the DOS [Eq. (8)], the size of electron/hole puddles [Eq. (9)], and the concentration of electrons/holes in puddles [Eq. (10)]. While the primary focus of this section is on the behavior near the Dirac point, results are also presented for the case when the chemical potential is away from the Dirac point. In Sec. III, the zero-temperature conductivity is discussed, and a result is presented for the conductivity [Eq. (16)] and its minimum value [Eq. (17)]. Sec. IV concludes with a summary and some discussion of recent experiments.

II. SELF-CONSISTENT THEORY OF DISORDER AND SCREENING

This paper focuses on a model of disorder in which N monovalent Coulomb impurities per unit volume are randomly distributed throughout the bulk of a 3DDS. Such impurities create a random Coulomb potential, and this potential induces an electron/hole concentration \( n(\vec{r}) \), where \( n > 0 \) for electrons and \( n < 0 \) for holes. The value of \( n(\vec{r}) \) at a given spatial coordinate \( \vec{r} \) is related to the self-consistent magnitude of the Coulomb potential \( \phi(\vec{r}) \). In this paper, the primary tool for describing this relationship is the TF approximation:

\[
E_f[n(\vec{r})] - e\phi(\vec{r}) = \mu. \tag{1}
\]

Here, \( E_f(n) = \hbar \nu k_f(n) \text{sgn}(n) \) is the local Fermi energy, where \( k_f \) is the Fermi wave vector and \( \mu \) is the chemical potential of the 3DDS measured relative to the Dirac point. Assuming, generically, that the Dirac point has a degeneracy \( g \) (which for Weyl semimetals is equal to the number of degenerate Dirac points), the Fermi wave vector is \( k_f = (6\pi^2|n|/g)^{1/3} \), so that \( E_f(n) = (6\pi^2/g)^{1/3}\hbar \nu |n|^{1/3} \text{sgn}(n) \).

If the chemical potential \( \mu \) is large enough in absolute value that \( |e\phi| \ll |\mu| \), one can think that the electron density is relatively uniform spatially, and the corresponding electron DOS \( \nu = dn/dE_f = gE_f^2/(2\pi^2\hbar^3\nu^3) \) is also uniform. In this case, one can straightforwardly define a TF screening radius

\[
r_s = \sqrt{\frac{k}{4\pi\nu^2\nu}} = \sqrt{\frac{\pi}{2\alpha g}}k_f^{-1}. \tag{2}
\]

Here, \( \alpha = e^2/\hbar \nu \) is the effective fine structure constant. The TF approximation is valid in cases where the Fermi wavelength \( \sim k_f^{-1} \) is much shorter than the typical scale over which the potential varies, \( r_s \). As can be seen in Eq. (2), this corresponds to \( \alpha g \ll 1 \). For comparison, the 3DDS Cd$_3$As$_2$ has \( \hbar \nu \approx 5–10 \text{ eV} \), \( \kappa \approx 36 \), and \( g = 2 \), so that \( \alpha g \approx 0.08–0.16 \) and this approximation is justified. As shown below, the same criterion \( \alpha g \ll 1 \) justifies the use of the TF approximation for the case of \( \mu = 0 \).

If \( |\mu| \) is not large, so that the 3DDS is close to the Dirac point, then one cannot consider the electron density to be uniform and the typical screening radius \( r_s \) must be found self-consistently. In particular, one can assume that the disorder potential is screened with some unknown screening radius \( r_s \) and then calculate analytically the corresponding magnitude of the disorder potential and the resulting average density of states \( \langle \nu \rangle \). Inserting the result for \( \nu \) into Eq. (2), one arrives at a self-consistent relationship for \( r_s \), which can be solved to give a result for \( r_s, \langle \nu \rangle \), and the magnitude of the disorder potential.\[25\] This procedure is carried out explicitly in the remainder of the present section.

In a medium with screening radius \( r_s \), the screened potential produced by a single impurity with charge \( \pm e \) is the Yukawa-like potential:\[26\]

\[
\phi_1(r) = \pm \frac{e}{\kappa r} \exp[-r/r_s]. \tag{3}
\]

If one assumes that impurity positions are uncorrelated, then the mean squared value of the electron potential energy, \( \Gamma^2 \), can be found by integrating the square of the potential created by a single impurity, \( (e\phi_1)^2 \), over all possible impurity positions. This gives

\[
\Gamma^2 = \int (e\phi_1(r))^2 N d^3r = 2\pi e^2 N r_s/\kappa. \tag{4}
\]

For cases where \( r_s \gg N^{-1/3} \) (justified below), the potential at each point in space is the sum of the potentials produced by many independently-located impurities. By the central limit theorem, then, one can assume that the distribution of values of the potential across the system is Gaussian with variance \( \Gamma^2/e^2 \). Within the TF approximation, the value of the DOS at a point with potential \( \phi = \langle \nu \rangle = (g/2\pi^2\hbar^3\nu^3)(\mu + e\phi)^2 \), so that the spatially-averaged DOS is

\[
\langle \nu \rangle = \int_{-\infty}^{\infty} \nu(\phi) \frac{e^{-e^2\phi^2/2\Gamma^2}}{\sqrt{2\pi\Gamma^2/e^2}} d\phi = \frac{g}{2\pi^2\hbar^3\nu^3} (\Gamma^2 + \mu^2). \tag{5}
\]
Inserting this expression for \( \nu \) in Eq. (2) and plugging the resulting expression for \( r_s \) into Eq. (4) gives the following self-consistent expression for the amplitude \( \Gamma \) of the disorder potential:

\[
\Gamma^4 (\Gamma^2 + \mu^2) = \frac{2\pi^3}{g \alpha^3} \left( \frac{\epsilon^2 N^{1/3}}{\kappa} \right)^6.
\] (6)

Equation (6) can be solved for generic values of the chemical potential \( \mu \), but it is worth considering specifically the cases of \( \mu = 0 \) (when the 3DDS is at the Dirac point) and large \( |\mu| \). When \( \mu = 0 \), Eq. (6) gives

\[
\Gamma \equiv \Gamma_0 = \left( \frac{2\pi^3}{g \alpha^3} \right)^{1/6} \left( \frac{\epsilon^2 N^{1/3}}{\kappa} \right).
\] (7)

It is perhaps worth noting that this value for the disorder potential amplitude is smaller than the corresponding result for the surface of a disordered 3D topological insulator by a factor \( \sim \alpha^{1/6} \). This smaller disorder for 3DDSs is a consequence of stronger screening in three dimensions.

Inserting the value of \( \Gamma_0 \) from Eq. (7) into Eq. (5) gives the corresponding DOS at \( \mu = 0 \):

\[
\nu_0 = \left( \frac{\alpha^3 g^2}{4\pi^3} \right)^{1/3} \frac{N^{2/3}}{\hbar v}.
\] (8)

In this case, the screening radius \( r_s \), which is generically equal to the correlation length of the disorder potential, defines the typical size of electron and hole puddles (as illustrated in Fig. 1). By Eq. (2), its value is given by

\[
r_s = \left( \frac{1}{4g \alpha^3} \right)^{1/3} N^{-1/3}.
\] (9)

The typical concentration of electrons/holes in puddles \( n_p \) is found by equating \( \Gamma_0 \) with \( E_f(n_p) \), which gives

\[
n_p = \sqrt{\frac{g \alpha^3}{16\pi}} N,
\] (10)

so that the corresponding number of electrons/holes per puddle is

\[
M_p \approx \frac{4\pi}{3} r_s^3 n_p = \sqrt{\frac{\pi/162}{g \alpha^3}}.
\] (11)

When \( g \alpha \ll 1 \), there are many electrons per puddle: \( M_p \gg 1 \). Intriguingly, this value for \( M_p \) is independent of the impurity concentration, so that the number of electrons per puddle is independent of the details of the disorder. This universality is reminiscent of the problem of a single supercritical nucleus in a 3DDS, where the maximum observable “nuclear charge” also obtains a universal value \( \sim 1/\alpha^{3/2} \).

Notice also that at \( g \alpha \ll 1 \), the correlation length of the potential \( r_s \) is much longer than the typical Fermi wavelength \( k_f(n_p)^{-1} \sim N^{-1/3} g^{1/6}/\alpha^{1/2} \), so that the TF approximation is justified. This same condition also guarantees \( r_s \gg N^{-1/3} \), which validates the assumption of a Gaussian-distributed potential.

It is worth noting that Eqs. (7)–(11) can be derived qualitatively using the following very simple argument (which for simplicity uses \( g \sim 1 \)). Consider a volume of size \( r_s \) within the 3DDS; this volume is effectively a single electron/hole puddle. Those impurities within the volume can be said to contribute to the potential within it, while others are effectively screened out. The net charge of impurities in the volume is \( Q \sim e \sqrt{Nr_s^3} \) (with a random sign), and this impurity charge is compensated by the charge of electrons/holes, which have total number \( M_p \sim n_v r_s^3 \). Equating \( M_p \) with \( Q/e \) gives \( n_v^2 \sim N/r_s^3 \). Now one can note that the typical kinetic energy of electrons within the volume, \( \sim \hbar v r_s^{-1/3} \), must be similar in magnitude to the typical Coulomb energy \( \sim Qe/\kappa r_s \). This equality gives \( n_v^2 \sim \alpha^6 N^3 r_s^3 \). Combining the two equations for \( n_v \) gives \( r_s \sim N^{-1/3}/\alpha \), as in Eq. (9), and the other relevant quantities can be found by substitution.

As the chemical potential \( \mu \) is moved away from the Dirac point, the magnitude of the disorder potential decreases, as dictated by Eq. (6), and correspondingly the screening radius \( r_s \) shrinks. At \( |\mu| \gg \Gamma_0 \), puddles of electrons (for \( \mu < 0 \)) or holes (for \( \mu > 0 \)) dry up, and the system is well-described by linear screening with a spatially-uniform DOS. In this case the disorder potential magnitude becomes

\[
\Gamma \approx \left( \frac{2\pi^3}{g \alpha^3} \right)^{1/4} \left( \frac{\epsilon^2 N^{1/3}}{\kappa} \right)^{1/2} \frac{n_p}{|\mu|}.
\] (12)

The corresponding DOS approaches that of the non-disordered system,

\[
\nu \approx \frac{g}{2\pi^2} \frac{\mu^2}{(\hbar v)^3}
\] (13)

and the correlation length of the disorder potential is

\[
r_s \approx \sqrt{\frac{\pi \hbar v}{2\pi g |\mu|}}.
\] (14)

Equations (7)–(11) and (12)–(14) describe the system in the limits of \( \mu = 0 \) and \( |\mu| \gg \Gamma_0 \), respectively. The crossover between these two regimes can be described by evaluating Eqs. (5) and (6). The result of this process is shown in Fig. 2 where the variance of the disorder potential and the DOS are plotted as a function of the chemical potential \( \mu \). As one can see, these self-consistent equations predict a smooth, monotonic crossover from the puddle-dominated \( \mu = 0 \) result to the linear screening regime at large \( \mu \). It is worth noting, however, that \( \Gamma \) may in fact exhibit weakly nonmonotonic behavior as a function of \( \mu \), achieving a weak maximum at \( |\mu|/\Gamma_0 \sim 1 \). Such behavior is predicted theoretically for topological
insulators\textsuperscript{22} and arises because at $|\mu|/\Gamma_0 \sim 1$ the distribution of values of the Coulomb potential becomes skewed toward those values that bring the system locally closer to the Dirac point, where screening is poorer.

\[ \Gamma \]

\[ \text{FIG. 2. a) The variance in the disorder potential as a function of the chemical potential $\mu$. The dashed line corresponds to the linear screening regime of Eq. (12). b) The spatially-averaged density of states as a function of the chemical potential. The dashed line corresponds to Eq. (13), which is the DOS for a non-disordered system. $\Gamma_0$ and $\nu_0$ are given by Eqs. (7) and (8), respectively.} \]

III. CONDUCTIVITY

The previous section discussed the disorder potential produced by random Coulomb impurities using a self-consistent theory of screening. In this section I briefly discuss the implications of this screening for the low-temperature conductivity.

In situations where the mean free path $\ell$ for electron scattering is relatively large, $k_f\ell \gg 1$, the electron transport is well-described by the Boltzmann equation. In particular, the momentum relaxation time $\tau$ satisfies

\[ \frac{h}{\tau} = \pi \nu \int_0^{\infty} d\theta \sin \theta \left| \tilde{\phi}_1(q) \right|^2 (1 - \cos \theta) \frac{1 + \cos \theta}{2} \]

(see, e.g., Ref. \textsuperscript{29}). Here, $q = 2k_f \sin(\theta/2)$ is the momentum change resulting from scattering by an angle $\theta$ and $\tilde{\phi}_1(q) = 4\pi e^2/[k(q^2 + r^{*-2})]$ is the Fourier transform of $\phi_1(r)$ evaluated at wave vector $q$. The final factor of $(1 + \cos \theta)/2$ in Eq. (15) arises when backscattering is suppressed as a consequence of the spin texture at the Dirac point, as in Weyl semimetals; omitting this factor does not change the results to leading order in $k_f r_s \gg 1$.

Evaluating the integral in Eq. (15) gives

\[ \frac{h}{\tau} \approx \frac{8\pi^2 N e^4}{k^2 k_f^2} \ln(2k_f r_s) \]

assuming $k_f r_s \gg 1$, which is a condition of validity for the screened potential used here, and, again, corresponds to $\alpha g \ll 1$.

One can arrive at an expression for the conductivity $\sigma$ by combining the expression for $\tau$ with the Einstein relation for conductivity, $\sigma = e^2 \nu (v^2 \tau/3)$\textsuperscript{29}. If one assumes a relatively large and uniform electron density $n$ (i.e., a large chemical potential $|\mu| \gg \Gamma_0$), then $k_f \approx (6\pi^2 n/g)^{1/3}$ and

\[ \sigma = \left( \frac{3}{4\pi} \right)^{1/3} \frac{1}{\alpha^2 g^{1/3} \ln(2\pi/\alpha g)} \frac{n}{N} e^2 n^{1/3} \]

This expression was previously derived in Ref. \textsuperscript{29} for the case where $N = n$ (i.e., where electrons are provided by uncompensated donors).

As the electron density is reduced (the chemical potential is brought closer to the Dirac point), the number of carriers is reduced and the conductivity declines. At $\mu = 0$, the conductivity achieves a minimum whose value is determined by the concentration of electrons and holes in puddles. The value of this conductivity minimum, $\sigma_{\text{min}}$, can be estimated by inserting $n \sim n_p$ into Eq. (16), which gives

\[ \sigma_{\text{min}} \approx \frac{1}{6\pi (2g^2)^{1/3} \ln(2\pi/\alpha g)} \frac{e^2 N^{1/3}}{h} \]

Finally, one can check that the derived expressions indeed correspond to the Boltzmann semiclassical limit $k_f \ell \gg 1$. As expected, the value of $\tau$ corresponding to the minimum conductivity, where $n \sim n_p$, gives $k_f \ell = k_f v \tau \approx g^{2/3} / \alpha g \gg 1$. Thus, $k_f \ell \gg 1$, provided that $\alpha \ll 1$. At larger $n$ the mean free path only increases, so that the Boltzmann equation is a good description everywhere.

IV. CONCLUDING REMARKS

This paper has presented a simple picture of self-consistent screening of Coulomb impurities in 3DDSs through the formation of electron and hole puddles. Such effects are manifestly not perturbative near the Dirac point, regardless of the impurity concentration, and they have a prominent effect on both the observed DOS and the conductivity. The DOS, for example, vanishes only as the 2/3 power of the impurity concentration, which suggests that a true Dirac semimetal phase with vanishing DOS may remain frustratingly elusive experimentally.

While the experimental study of 3DDSs is still very young, one can get a sense of the typical scales of the disorder potential by using parameters for the recently-discovered bulk Dirac material Cd$_3$As$_2$, which has $\alpha \approx
0.04 and \( q = 2 \). Thus far, experimentally studied samples are \( n \)-type, apparently resulting from uncontrolled doping by As vacancies.\(^{117,118}\) For example, Ref. 19 reports relatively large \( \mu \approx 200 \text{meV} \), with a corresponding carrier concentration \( n \approx 2 \times 10^{18} \text{cm}^{-3} \); one can expect that the concentration of donor impurities \( N \) is similar in magnitude. Inserting these parameters into Eqs. (12) and (13) gives an estimated disorder potential of \( \Gamma \approx 20 \text{meV} \) and a screening radius \( r_s \approx 20 \text{nm} \). The latter seems consistent with the scale of disorder fluctuation sizes seen by scanning tunneling microscopy measurements.\(^{156}\) By Eq. (16) [see also Eq. (84) of Ref. 29], this level of disorder corresponds to a mobility \( \sigma/(en) \approx 30000 \text{cm}^2/\text{Vs} \), which also closely matches the value seen in experiment.\(^{175}\)

Future efforts to bring the bulk chemical potential of 3DDSs to the Dirac point will presumably require compensation of donors by acceptors. By Eqs. (10) and (11), the resulting disorder landscape can be expected to have a typical concentration \( n_p \approx 10^{-3} N \approx 10^{-15} \text{cm}^{-3} \) of electrons in puddles and \( \approx 10 \text{electrons/holes per puddle} \), with a disorder potential of magnitude \( \Gamma_0 \approx 45 \text{meV} \), assuming the impurity concentration \( N \) remains of order \( 10^{18} \text{cm}^{-3} \). Equation (17) suggests a corresponding minimum conductivity \( \sigma_{\text{min}} \approx 1 \text{S/cm} \).

Finally, one can note that existing 3DDS materials seem to have anisotropic Dirac cones, with a Dirac velocity in one particular direction, \( v_z \), that is as much as ten times smaller than the velocity in the transverse directions, \( v_{\perp} \).\(^{193}\) This anisotropy can be accounted for at the level of the present theory by substituting for \( v \) the geometric mean velocity \( (v_{\perp}^2 + v_z^2)^{1/2} \), so that the fine structure constant \( \alpha \approx 1/\text{v} \) is also modified.

**Acknowledgments**

I am grateful to R. Nandkishore, S. Gopalakrishnan, J. C. W. Song, B. I. Shklovskii, and E. B. Kolomeisky for helpful discussions and comments. Work at Argonne National Laboratory was supported by the U.S. Department of Energy, Office of Basic Energy Sciences under contract no. DE-AC02-06CH11357.