Emergence of quasimetallic state in a disordered two-dimensional electron gas due to strong interactions

B. Rosenstein*
National Center for Theoretical Sciences and Electrophysics Department, National Chiao Tung University, Hsinchu 30050, Taiwan and Physics Department, Bar Ilan University, Ramat Gan, Israel

Tran Minh-Tien
Electrophysics Department, National Chiao Tung University, Hsinchu 30050, Taiwan and Institute of Physics, National Center for Natural Sciences and Technology, Hanoi, Vietnam
(Received 28 April 2003; published 31 December 2003)

The interrelation between disorder and interactions in two-dimensional electron liquid is studied beyond weak-coupling perturbation theory. Strong repulsion significantly reduces the electronic density of states on the Fermi level. This makes the electron liquid more rigid and strongly suppresses elastic scattering off impurities. As a result the weak localization, although ultimately present at zero temperature and infinite sample size, is unobservable at experimentally accessible temperature at high enough densities. Therefore practically there exists a well-defined metallic state. We study diffusion of electrons in this state and find that the diffusion pole is significantly modified due to ''mixture'' with static photons similar to the Anderson-Higgs mechanism in superconductivity. As a result several effects stemming from the long-range nature of diffusion such as the Aronov-Altshuler logarithmic corrections to conductivity are less pronounced.

DOI: 10.1103/PhysRevB.68.245321 PACS number(s): 71.30.+h, 72.10.−d, 71.10.−w, 72.20.−i

I. INTRODUCTION

The question of mutual influence of long-range Coulomb interactions and disorder in two-dimensional electron gas (2DEG) attracted a great attention after an unexpected discovery of metallic state and clear metal-insulator transition by Kravchenko and co-workers.[1,2] The very existence of a metallic state with finite conductivity at zero temperature is in conflict with the weak-localization theory,[3] which predicts that in 2D even negligible amount of disorder localizes electrons at sufficiently low temperature. The theory however was firmly established at weak coupling or for short-range interactions only, while the metallic state exists and the transition was found for rather strong coupling $r_s = E_{ee}/E_F \sim 10$, where E_{ee} is the average interaction energy per electron and E_F is the Fermi energy. Therefore Coulomb interactions dominate the kinetic energy and cannot be considered ''small.'' In addition to an obvious difficulty to treat quantitatively or even qualitatively the strong coupling, it is not clear which one, disorder or Coulomb interactions, should be considered as a most important cause of the transition to an insulating state (the corresponding insulating state in these cases is of ''Anderson'' or ''Mott'' type[4]). Most probably it results from a nontrivial combination of these interactions.

The standard approach starts with a commonly accepted argument that a long-range Coulomb interaction after ''bubble resummation'' of the random-phase-approximation (RPA) type[5] becomes effectively short range. Therefore one can start the treatment of disorder after this resummation was performed. Disorder is treated within a similar approach in which ''rainbow'' diagrams[6] ''ladders and crossed ladders resummation''[7] (or, more systematically, the ''steepest descent'' approximations[8] in the path-integral language[9]) with interaction being already short ranged. In this way two kinds of massless modes determining the properties of the disordered electron gas are identified: diffusons (describing diffusive nature of the electron motion due to impurities) and Cooperons in the particle-particle channel. It is the last which lead to weak localization due to logarithmic infrared (IR) divergences in leading fluctuation contribution to conductivity[10] (diffusons can also lead to IR divergences at yet higher orders[11])

$$\Delta\sigma^{wl}(T) = \frac{e^2}{\pi h}\ln[T\tau/\hbar],$$

where τ is a free system relaxation time.

More sophisticated renormalization-group based methods using ''path integral''[12] and ''σ models''[7,13,14] with Coulomb interactions[15,16] were developed. Considering high-order vertex renormalization, it was found that there are additional logarithmic IR (Aronov-Altshuler[17]) divergencies:

$$\Delta\sigma^{ee}(T) = \frac{e^2}{\pi h}(1 - 3/4F^*)\ln[T\tau/\hbar],$$

where F^* is the Fermi-surface average of the screened Coulomb interaction, leading to a conclusion that long interactions increase tendency to weak localization.[18] This leads to a difficulty in understanding recent experiments in which apparently interactions do not necessarily lead to rapid increase of resistivity. Recent detailed experimental studies[19,20] clearly show that near the putative metal-insulator transition logarithmic terms either are suppressed or cancel each other (several arguments were put forward in Ref. 19 against such a fortuitous cancellation). The conductivity dependence on temperature follows the Gold-Dolgopolov[21] linear decrease, which at higher temperatures crosses over to the ballistic regime studied in detail recently in Ref. 22. Generally within this approach the Coulomb interaction is screened first and

the disorder effects are treated later. However recent electron-spin-resonance experiment[23] demonstrated that the screening length rapidly diverges when density is reduced towards the transition point. The density of states (DOS) at the Fermi level vanishes. It was noticed long time ago[24] that the diffusive motion of electrons slows down the process of screening. Therefore in the limit of small density and when disorder seems to play an important and possibly crucial role, it is reasonable to start from an approximation in which the interaction is not rendered short range.

With these experimental facts in mind, we reconsider the question of the interrelation of disorder and Coulomb interactions in 2DEG within a single consistent systematic approach without replacing it by a short-range potential from the beginning as is done in the clean limit or high density. The necessarily nonperturbative approach consists of two steps. First is a variational one (nonperturbative in coupling) and is similar in spirit to the Hartree-Fock for clean metals or the BCS approximation in superconducting metals. We find in Sec. II the ''best'' quadratic Hamiltonian representing the system. On this set of Hamiltonians a quasiparticle (and quasihole) Green's function is a variational parameter. There are possible contributions in the particle-particle (Cooper pairs) as well as the particle-hole channels due to Coulomb interactions (Hartree state in direct and Fock state in exchange channel), while interactions with disorder can be treated in a similar manner with the frequency dependent relaxation time being one of the variational parameters. Possible condensates in several channels do not realize that of course there is no condensate in the Cooper channel for a repulsive interaction and there is also no condensation in the direct channel due to the charge neutrality as is shown in Sec. II A. However the strong long-range exchange interaction creates (even in the clean case[25]) a dip in the DOS on the Fermi surface. At infinitely strong coupling the DOS on the Fermi surface approaches zero (to avoid confusion, this reduction is not related to the one found at higher orders for screened interaction in Ref. 18, see discussion of this topic in Sec. IV D). This makes the electron liquid very rigid and, as a result, the effects of disorder are greatly suppressed. This in turn leads to increase in conductivity at large coupling. The emergence of the above phenomenon can already be seen on perturbative level. The first-order quasiparticle energy shift due to exchange is (up to a constant)

$$\langle p|H_{int}|p\rangle = -\sum_{p'} v(p-p')\mathrm{sgn}\left(\mu - \frac{p'^2}{2m^*}\right).$$

It is easily shown (Sec. II B) that for purely repulsive $v(p)$ energy of states above the Fermi level is shifted up, while energy of states below the Fermi-level is shifted down. The logarithmic vanishing of the DOS is a direct consequence of the long range nature of the Coulomb interaction. It is important to note that the significant reduction of the DOS near the Fermi level does not mean that the effective mass is smaller than the band effective mass. On the contrary, it was shown in the clean case[26,25] that despite this the effective mass grows with coupling as was observed recently in Shubnikov–de Haas experiments.[27] We comment more on that in Sec. II B.

After the variational quadratic Hamiltonian (or variational quadratic action in the path-integral formalism) is found, we introduce in Sec. III A a systematic perturbation theory around it. In the path-integral language[9,8] it is a conventional steepest descent expansion with the variational action as a saddle point. First we introduce fields describing various possible kinds of fluctuations: diffusons, Cooperons, static photons (corresponding to the direct Coulomb interaction channel), the exchange and the Cooper channel interactions. The last two are evidently massive as well as half of diffusons and Cooperons.[8] However we find that there is a nontrivial mixing between photons and diffusons. The phenomenon is very reminiscent of the Anderson-Higgs mechanism in superconductivity[28] in which massless Goldstone boson of phase is mixing with (dynamical) photon. As a result both modes become ''massive.'' In the case of strongly coupled 2DEG the modes are not really massive, the density-density correlator describing diffuson becoming ''harder'':

$$\frac{1}{\omega\tau_r + e^2 D_r |p|/4\pi}, \tag{1}$$

compared to the noninteracting diffusion pole

$$\frac{1}{\omega\tau_r + D_r p^2}. \tag{2}$$

The (static) photon becomes RPA screened:

$$\frac{1}{2|p|/e^2 + 2\pi}.$$

The renormalized diffusion constant D_r increases (nonperturbatively) with coupling from its noninteracting value of $D = \mu\tau/m$. Similarly the renormalized relaxation time τ_r increases with r_s.

Equation (1) implies that electrons at large distances are no longer ''diffusive'': they obey diffusion equation with first space derivative only. The Cooperon on the other hand still retains its typical diffusive pole form, Eq. (2). The approximation scheme at higher orders therefore nontrivially combines the RPA and the disorder resummation on the same footing. It is important to emphasize that the scheme is manifestly ''gauge invariant.'' As was shown in Refs. 16 and 15, that it is very important to ensure gauge invariance at each stage in order not to miss important ''vertex corrections'' necessary to ensure charge conservation at each order of the expansion including the variational stage.

Next in Sec. IV we turn to the ''fluctuation corrections'' leading to weak localization. The leading fluctuation correction to conductivity is infrared divergent. However, as the coupling r_s grows, the correction grows slower than the main (Drude) contribution. This justifies the expansion even for small values of $D<1$ at which the standard $1/D$ expansion is invalid ($D_r \gg 1$ is however required). Therefore the crossover temperature at which conductivity starts approaching zero is

significantly lower than that for noninteracting electrons. This temperature is estimated as a temperature at which the perturbation theory in fluctuations breaks down, namely when the correction becomes a significant fraction of the leading-order contribution. The crossover temperature according to our analysis becomes unobservably small since it vanishes exponentially fast with coupling (due to logarithmic dependence of the fluctuation correction on temperature serving as an IR cutoff). Therefore one can practically (for samples of finite albeit large size) talk about stable metallic state in 2DEG. Our conclusions and discussion of the phase diagram as well as relations with other approaches are subject of Sec. V.

II. VARIATIONAL PRINCIPLE FOR COULOMB INTERACTIONS IN THE PRESENCE OF DISORDER

A. Model and the basic approximation

1. The model

We consider a system of electrons with effective band mass m^* confined to a plane interacting with each other and with random potential $U(x)$:

$$H=\int_x c^\dagger_{x\sigma}\left(-\frac{\nabla^2}{2m^*}-\mu+U(x)\right)c_{x\sigma} + \frac{1}{2}\int_{x,y} c^\dagger_{x\sigma_1}c_{x\sigma_1}V(x-y)c^\dagger_{x\sigma_2}c_{x\sigma_2}. \tag{3}$$

We set $\hbar=1$ throughout theoretical parts of the paper. σ is a spin and valley index and $v(x)$ is the 3D Coulomb interaction which has the following Fourier transform:

$$v(p)=\frac{e^2}{2\epsilon}\frac{1}{p}[1-\delta(p)]. \tag{4}$$

Here ϵ is the dielectric constant and the last term describes the background ensuring charge neutrality of the system, $v(p=0)=0$. Passing to the standard imaginary time path-integral formulation[9] and performing a well-known replica trick[8] one obtains the action

$$A[\psi,\overline{\psi}]=\int_{x,t}\overline{\psi}^{a\sigma}_{xt}\left(\partial_t-\frac{\nabla^2}{2m^*}-\mu\right)\psi^{a\sigma}_{xt} - \frac{1}{2\tau}\int_{x,t,s}\overline{\psi}^{a\sigma_1}_{xt}\psi^{a\sigma_1}_{xt}\overline{\psi}^{b\sigma_2}_{xs}\psi^{b\sigma_2}_{xs} + \frac{1}{2}\int_{x,y,t}\overline{\psi}^{a\sigma_1}_{xt}\psi^{a\sigma_1}_{xt}v(x-y)\overline{\psi}^{a\sigma_2}_{yt}\psi^{a\sigma_2}_{yt}. \tag{5}$$

Here $a,b=1,\ldots,N_r$ are replica indices, τ is the "bare" relaxation time describing strength of the random potential. Let us first consider, for the sake of simplicity, the spin-polarized case and only one "valley" (returning to the general case in Sec. IV A), which means that we drop the spin indices σ. The path-integral formulation of the variational principle, being completely equivalent to the standard methods such as summation of diagrams or Bogoliubov transformations, allows, in addition, a convenient treatment of quantum and thermal fluctuations. Transforming to the Matsubara frequency $\omega_n=(2n+1)\pi T$ and the momentum basis for the Grassmannian fields

$$\psi^a_{xt}=\sqrt{T}\sum_p\sum_n \exp[i(px-\omega_n t)]\psi^a_{pn},$$

and separating regions of phase space in which interaction connects electrons near the Fermi surface, one obtains[8]:

$$A=\sum_p\sum_n\overline{\psi}^a_{pn}\left(-i\omega_n+\frac{p^2}{2m^*}-\mu\right)\psi^a_{pn}+A_{dis}+A_C, \tag{6}$$

$$A_{dis}=-\frac{1}{2\tau}\sum_{p,q,r}\sum_{n,m}[\overline{\psi}^a_{p-q,n}\psi^a_{q,n}\overline{\psi}^b_{-p-r,m}\psi^b_{r,m} + \overline{\psi}^a_{p-q,n}\psi^a_{-p-r,n}\overline{\psi}^b_{r,m}\psi^b_{q,m} + \overline{\psi}^a_{p-q,n}\psi^a_{-p-r,n}\overline{\psi}^b_{q,m}\psi^b_{r,m}],$$

$$A_C=\frac{T}{2}\sum_{p,q,r}\sum_{n,m}[\overline{\psi}^a_{p-q,n}\psi^a_{q,n}v(p)\overline{\psi}^b_{-p-r,m}\psi^b_{r,m} + \overline{\psi}^a_{p-q,n}\psi^a_{-p-r,n}v(p)\overline{\psi}^b_{r,m}\psi^b_{q,m} + \overline{\psi}^a_{p-q,n}\psi^a_{-p-r,n}v(p)\overline{\psi}^b_{q,m}\psi^b_{r,m}].$$

All the fermion's momenta p, q, r are now considered to be around p_F. Both the disorder and the Coulomb interaction parts A_{dis} and A_C, respectively, have three terms corresponding to direct (Hartree), exchange (Fock) electron-hole channels and the electron-electron (Cooper) channel.

2. The most general quadratic Hamiltonian and the Hubbard-Stratonovich fields

A convenient way to look for the most general quadratic action is to perform a Hubbard-Stratonovich (HS) transformation introducing a field for each of the six channels. We should not consider the direct channel for disorder though since it is of higher order in number of replicas N_r which should approach zero (we assume that replica symmetry is not broken spontaneously). The effective action in terms of these fields is rather complicated:

$$A_{eff}=-\mathrm{Tr}\ln[G_N^{-1}]+\sum_{pnm}[Q^{ab}_{pnm}Q^{*ab}_{pnm}+\Delta^{*ab}_{pnm}\Delta^{ab}_{pnm}] + \frac{1}{2}\sum_{pp'n}[\Phi^*_{pn}v^{-1}(p-p')\Phi_{p'n}+\Theta^{*a}_{pp'n}v^{-1}(p-p') \times\Theta^a_{pp'n}+E^{*ab}_{pp'n}v^{-1}(p-p')E^{ab}_{pp'n}]. \tag{7}$$

Here G_N is the fermionic Green's function in Nambu space defined by

$$\eta^a_{pn}=\frac{1}{\sqrt{2}}\begin{pmatrix}\psi^a_{pn}\\ -\overline{\psi}^a_{-pn}\end{pmatrix},\quad \overline{\eta}^a_{pn}=\frac{1}{\sqrt{2}}(\overline{\psi}^a_{pn}-\psi^a_{-pn}). \tag{8}$$

The inverse fermionic propagator is a 2×2 matrix

$$G_N^{-1}=\begin{pmatrix} G^{-1} & F \\ F* & G^{-1}* \end{pmatrix}, \tag{9}$$

whose elements G^{-1} and F are themselves matrices in p,n, and a space. The diagonal element is

$$\langle {}^{a}_{p'n'}|G^{-1}|^{b}_{pn}\rangle=\delta_{nn'}\delta^{ab}\left[\left(i\omega_n-\frac{p^2}{2m^*}+\mu\right)\delta_{pp'}+i\Phi_{pn}\right]$$
$$+\frac{i}{\sqrt{\tau}}Q^{ab}_{p-p'nn'}-\delta_{nn'}E^{ab}_{pp'n}, \tag{10}$$

where the field Q describes the diffuson, Φ is a static photon field in the direct channel, while E is an "exchange field." The off-diagonal element

$$\langle {}^{a}_{p'n'}|F|^{b}_{pn}\rangle=\frac{i}{\sqrt{\tau}}\Delta^{ab}_{pp'nn'}+\delta^{ab}\Theta^{a}_{pp',n+n'} \tag{11}$$

contains the Δ and the Θ fields describing the Cooperon channel in the disorder part and the Cooper channel in the Coulomb interaction part (if the interaction were attractive this channel would have led to superconductivity), respectively. This effective action should be minimized as a function of all five HS fields determining the fermionic Green's function. The solution of this variational problem is discussed in the following section. Later in Sec. III we expand the path integral around the solution of the minimization equations to quadratic order (harmonic approximation) to determine elementary excitation modes and then in Sec. IV use Feynman rules to compute fluctuation corrections.

3. The saddle-point equations

Minimization of the effective action, Eq. (7), poses a nontrivial mathematical problem. Let us first remove obviously irrelevant fields and functional dependencies using symmetry arguments. Since the Coulomb interaction is purely repulsive, we assume that the electromagnetic U(1) gauge symmetry is unbroken (there is no condensation of the electron-electron pairs). Therefore $\Delta_{SP}=\Theta_{SP}=0$. This by no means indicates that there are no fluctuations in these channels. On the contrary fluctuations in the Cooperon channel play an important role in destroying the metallic state.

The translation invariance in space and time (we assume that the ground state is a liquid rather than Wigner crystal[29]) and the unbroken replica symmetry (assuming the ground state is a disordered, possibly overcooled liquid rather than electron glass[30]) implies

$$Q^{ab}_{pnmSP}=\delta^{ab}\delta_{nm}\delta_p q_n,$$
$$E^{ab}_{pp'nSP}=\delta^{ab}\delta_n\delta_{pp'}e_p, \tag{12}$$
$$\Phi_{pn}=\delta^{ab}\delta_n\delta_p\phi.$$

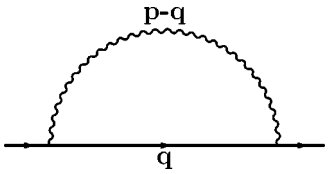

FIG. 1. The self-energy corresponding to the saddle-point equation in the clean limit.

We will comment on the last two (nontrivial) assumptions in Sec. V. Consequently the inverse Green's function of fermions simplifies to

$$F=0,$$
$$\langle {}^{a}_{p'n'}|G^{-1}|^{b}_{pn}\rangle=\delta^{ab}\delta^{nn'}\delta(p-p')(G^n_p)^{-1}, \tag{13}$$
$$(G^n_p)^{-1}=i\left(\omega_n+\frac{q_n}{\sqrt{\tau}}\right)+\mu-\frac{p^2}{2m^*}-\Sigma_p.$$

The minimization equation for the static photon condensate ϕ is

$$\phi=-i\sqrt{T}\delta(p)v(p)\sum_{q,n}G^n_q. \tag{14}$$

However due to neutralizing background $v(p=0)=0$, and the right-hand side of this equation vanishes. Therefore $\phi=0$. Minimization equations for q_n, e_p are

$$q_n=\frac{i}{\sqrt{\tau}}\sum_q G^n_q, \tag{15}$$

$$\Sigma_p=-T\sum_{p',n}v(p-p')G^n_{p'}. \tag{16}$$

We start with exact solution of these equations in the clean case.

B. The clean limit

1. A major simplification in the clean limit

In the absence of disorder $\tau\to\infty$, $q\to 0$ and we should consider the second equation (16) only. Substituting the Green's function, Eq. (13), it takes a form

$$\Sigma_p=-T\sum_{q,n}\frac{v(\mathbf{p}-\mathbf{q})}{i\omega_n+\mu-q^2/2m^*-\Sigma_q}, \tag{17}$$

corresponding to the diagram in Fig. 1. At zero temperature (to which we confine ourselves for most of the paper) the summation over Matsubara frequencies results in

$$\Sigma_p = -\frac{1}{2}\sum_q v(|\mathbf{p}-\mathbf{q}|)\mathrm{sgn}[\mu - q^2/2m^* - \Sigma_q]$$

$$= -\frac{1}{2(2\pi)^2}\int_{q=0}^{\infty}\int_{\varphi=0}^{\pi} qv[\sqrt{p^2+q^2-2pq\cos(\varphi)}]$$

$$\times \mathrm{sgn}[\mu - q^2/2m^* - \Sigma_q], \tag{18}$$

where φ is an angle between fermion's momenta vectors $\mathbf{p}$ and $\mathbf{q}$ and p, q are their lengths. Integral over the angle gives

$$\Sigma_p = -\frac{e^2}{(2\pi)^2}\int_{|q|\geq 0}\frac{qK\left[-\frac{4pq}{(p-q)^2}\right]}{|p-q|}\mathrm{sgn}[\mu - q^2/2m^* - \Sigma_q].$$

Here $K[x]$ is the full elliptic integral of the first kind.[31] The interaction with neutralizing background amounts to subtracting the term $q=p$. Now we switch to dimensionless variables describing deviations of particle's energy from Fermi surface,

$$p \cong \sqrt{2m^*\mu}(1+\varepsilon+\cdots). \quad q \cong \sqrt{2m^*\mu}(1+\varepsilon'+\cdots), \tag{19}$$

rescaling also the variational self-energy function $\Sigma_p = 2\mu e_\varepsilon$. In the resulting integral equation

$$e_\varepsilon = \frac{r_s}{\sqrt{2}\pi}\int_{\varepsilon'}\kappa[\varepsilon-\varepsilon']\mathrm{sgn}[\varepsilon'+e_{\varepsilon'}], \tag{20}$$

$$\kappa[\varepsilon] \equiv \frac{K[-4/\varepsilon^2]}{|\varepsilon|},$$

we extend the integration over ε' to $\{-\infty,\infty\}$ and observe that there exists a solution obeying physically reasonable property that the sign of e_ε is the same as that of ε (we checked that there are no other solutions in disordered case as well). This makes the right-hand side independent of e_ε and we obtain a solution

$$e_\varepsilon = \frac{\sqrt{2}r_s}{\pi}\int_{\varepsilon'=0}^{\varepsilon}\kappa[\varepsilon'] \equiv \frac{\sqrt{2}r_s}{\pi}\kappa_1[\varepsilon], \tag{21}$$

where function $\kappa_1[\varepsilon]$ can be expressed via Meijer function $G^{mn}_{pq}(z|^{a_1,\ldots,a_p}_{b_1,\ldots,b_q})$,[31]

$$\kappa_1[\varepsilon] = \tfrac{1}{4}\mathrm{sgn}[\varepsilon][M(\varepsilon)-M(0)], \tag{22}$$

$$M(\varepsilon) = G^{22}_{33}\left(\frac{4}{\varepsilon^2}\middle|^{1/2,1/2,1}_{0,0}\right).$$

It behaves as $\varepsilon(\ln[8/\varepsilon]+1)/2$ at small ε and as $\pi\ln[2\varepsilon]/2$ at large ε. The standard dimensionless coupling

$$r_s \equiv \frac{e^2}{4\pi\epsilon}\sqrt{\frac{m^*}{\mu}}$$

was introduced.

FIG. 2. The ''rainbow'' diagrams corresponding to the saddle-point equation. The solid lines denote the fermionic propagator, the wavy lines denote the Coulomb interactions.

Generally the solution of the saddle-point equation corresponds in terms of diagrams to summation of all the photonic rainbows, see Fig. 2. However, as it is well known,[5] the sum of the diagrams except that of Fig. 1 vanishes identically. This will no longer be the case in the disordered electron gas. The fermionic Green's function near the Fermi surface, namely at small ε, becomes nonanalytic:

$$G_\varepsilon^{-1} \sim i\hat{\omega}_n + \varepsilon\left(1+\frac{r_s}{\sqrt{2}\pi}\ln\frac{8}{\varepsilon}\right) \tag{23}$$

and it is dominated by the interaction ''corrections.'' Here $\hat{\omega}_n = \omega_n/2\mu$ is dimensionless frequency.

2. Depletion around the Fermi level

Although no energy gap (the Coulomb gap) has been opened within this approximation, it follows directly from Eq. (23) that the DOS right on the Fermi level vanishes logarithmically:

$$N(\varepsilon) \sim \frac{1}{\ln(1/\varepsilon)}. \tag{24}$$

Therefore one can term electron gas with such properties ''very marginal Fermi liquid.'' The logarithmic dip in the DOS is even weaker than a power normally associated with such a situation.[32] As we will see in the following section, this will naturally lead to effective reduction of scattering off impurities thereby increasing conductivity and making the transition to Anderson insulator at least more difficult. Of course this result is ''perturbative'' in a sense that for the inverse propagator only one diagram was taken, Fig. 1. Therefore the effect ''starts'' and can be understood at weak coupling. One can interpret the minimization equations [or the Hartree-Fock (HF) resummation] as a renormalization of energies of the one-particle states due to the collective effect of the many-body electron-electron repulsions. As we mentioned in the Introduction the reduction in DOS in perturbation theory can be seen from the eigenvalue shifts. Indeed the $\mathrm{sgn}[\varepsilon]$ factor in Eq. (21) makes it clear that energy of states above the Fermi level is shifted up, while energy of states below the Fermi level is shifted down. The logarithmic singularity is a direct consequence of the long-range nature of the Coulomb interaction.

Obviously, if there would not be a disorder to interfere with the screening, the RPA type of reasoning would imply that the singularity would be ''smoothed away'' or ''cured'' by the quantum fluctuation corrections. Higher orders in coupling are increasingly singular and the singularities should be ''resumed away.''[5] However even for the screened interactions there is a dip in the DOS.[33] It cannot approach zero, but might be reduced significantly. We calculated $N(0)$ for dif-

TABLE I. The density of states at Fermi level $N(0)$ (normalized to the ideal gas DOS $N_0=1/2\pi$) for various couplings r_s and diffusion constants λ.

	r_s					
λ	0.1	1	2	4	8	16
8	0.873	0.395	0.219	0.106	0.0476	0.0204
2	0.892	0.466	0.268	0.129	0.0570	0.0246
0.5	0.895	0.557	0.344	0.167	0.0714	0.0297
0.1	0.845	0.651	0.471	0.241	0.103	0.0395
RPA	0.961	0.817	0.709	0.623	0.506	0.393

ferent couplings assuming RPA potential instead of $v(p)$. For several couplings the reduction of DOS is given in the last line of Table I. It provides an indication at what degree of disorder we can continue to use the HF approach without encountering strong screening effects. As one can see, the RPA is apparently less important to the reduction of DOS already at r_s as low as 1.

This is consistent with results of more elaborate calculations of the renormalization constant Z defined in Eq. (26) below involving the Hubbard function in Ref. 25. The effect of screening is much smaller at large coupling due to small density of the polarizing electrons. As we will see in the following section, in the disordered case the ''marginality'' is replaced at large coupling by large nonperturbative renormalization of the parameters of the disordered Fermi liquid.

3. Effective mass increase versus the DOS drop

The reduction in the DOS due to the repulsive interaction is related and is sometimes confused with the issue of the vanishing in Hartree-Fock approximation of the ''renormalized mass'' and the Fermi-liquid parameter F_1^s defined by[6,5]

$$\frac{m_r^*}{m^*}=\frac{Z}{Z_F}=1+F_1^s, \tag{25}$$

where renormalization constants are defined by

$$Z^{-1}=1+\frac{m^*}{p_F}\frac{\partial}{\partial p}\mathrm{Re}\Sigma^{ret}(p,\omega=0)\big|_{p=p_F}\approx\frac{1}{2\pi N(\omega=0)}, \tag{26}$$

$$Z_F^{-1}=1-\frac{\partial}{\partial\omega}\mathrm{Re}\Sigma^{ret}(p=p_F,\omega)\big|_{\omega=0}.$$

Within the Hartree-Fock approximation (without the RPA resummation) the retarded self-energy $\Sigma^{ret}(p,\omega)$ does not depend on frequency. Consequently, $Z_F=1$ and since $Z<1$ for repulsive interactions[24] the renormalized mass m_r^* is smaller than the band mass m^* [m_r^* vanishes for the long-range interactions, see Eq. (24)]. Several groups tried to improve this beyond Hartree-Fock using the RPA and the Hubbard approximation.[26,25] Generally the wave-function renormalization $Z_F<1$ since it represents departure of the momentum distribution from the ideal Fermi gas one. Therefore the question whether the effective mass is larger than the band mass depends on which of the reduction factors spacial Z or temporary Z_F is smaller. Ting, Lee, and Quinn[26] obtained finite monotonically increasing $m_r^*>m^*$, while more recent calculation[25] indicates that at $r_s<1$, $m_r^*<m^*$ and become larger above $r_s=1$. Numerical simulations of the same system[34] indicate that, on the one hand, the renormalized mass is definitely smaller than m^* at least for $r_s<5$, but, on the other hand, it increases with r_s. In recent experiments the renormalized mass was measured using Shubnikov–de Haas oscillations in magnetic field.[27] Apparently the renormalized mass deduced that way monotonically increases with r_s. To conclude, the increase of the effective mass with coupling does not necessarily imply that the DOS at Fermi level cannot drop significantly. This is important for our approach since the drop in the DOS naively should result in suppression of the elastic scattering off impurities due to reduced phase space available. Now we return to the general case of disordered strongly coupled 2DEG.

FIG. 3. The ''rainbow'' diagrams involving both Coulomb interactions (wavy lines) and interactions with disorder (dashed lines) in the general case.

C. The saddle-point equations in the general case

1. Numerical solution

Following in the more complicated disordered case, Eqs. (15) and (16), the same steps as in the clean case, the equations for the scaled (dimensionless) quantities ($\Sigma_p=2\mu e_\varepsilon$, $q_n\equiv 2\mu\sqrt{\tau}\hat{q}_\omega$, $\omega\equiv 2\mu\hat{\omega}$) at zero temperature are

$$\hat{q}_\omega=\frac{1}{(2\pi)2\lambda}\int_\varepsilon\frac{\hat{\omega}+\hat{q}_\omega}{(\varepsilon+e_\varepsilon)^2+(\hat{\omega}+\hat{q}_\omega)^2}, \tag{27}$$

$$e_\varepsilon=\frac{r_s}{\sqrt{2}\pi^2}\int_{\omega,\varepsilon'}\kappa[\varepsilon-\varepsilon']\frac{\varepsilon'+e_{\varepsilon'}}{(\varepsilon'+e_{\varepsilon'})^2+(\hat{\omega}+\hat{q}_\omega)^2}, \tag{28}$$

where $\lambda=\mu\tau$ and function $\kappa[\varepsilon]$ was defined in Eq. (20). The approximation corresponds to summing up the whole set of rainbow diagrams involving both Coulomb interactions and interactions with disorder, Fig. 3.

The results for the real part of self-energy for $\lambda=0.5$, $r_s=1$ (where in the noninteracting case Anderson localization is very effective) and $\lambda=8$, $r_s=0.01$, 1, and 16 (still a quasimetal with small localization effects) are given in Fig. 4. The corresponding imaginary part of self-energy $\hat{q}_\omega$ is given in Fig. 5. We solved the equations for various values of coupling r_s between weak coupling up to $r_s=24$ [experimentally the metal-insulator transition is observed from $r_s=15$ (Ref. 1) to $r_s=25$ in GaAs/AlGaAs (Ref. 35) heterojunctions and around $r_s=10$ in Si metal-oxide-semiconductor field-effect transistor (Refs. 36 and 35)]. Solid lines are simulation results. The dashed lines are results

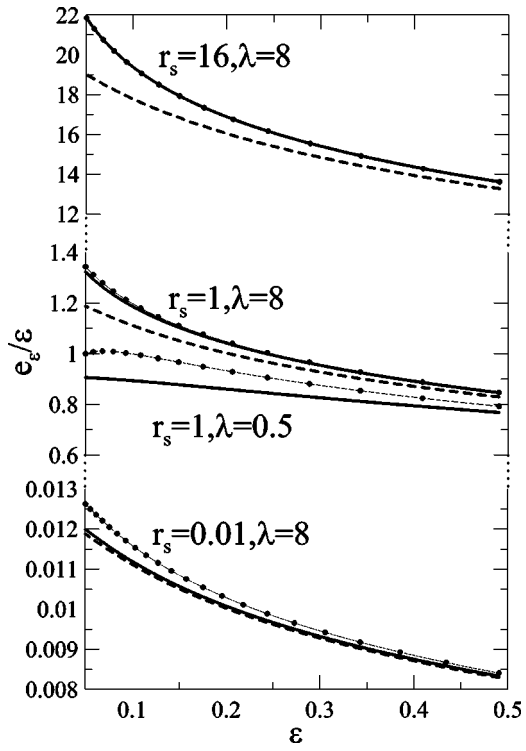

FIG. 4. The dependence of $e_\varepsilon/\varepsilon$ on ε for various values of r_s and λ. The solid lines are the numerical results, the dashed lines are the results of r_s expansion, and the dots are the results of $1/\lambda$ expansion.

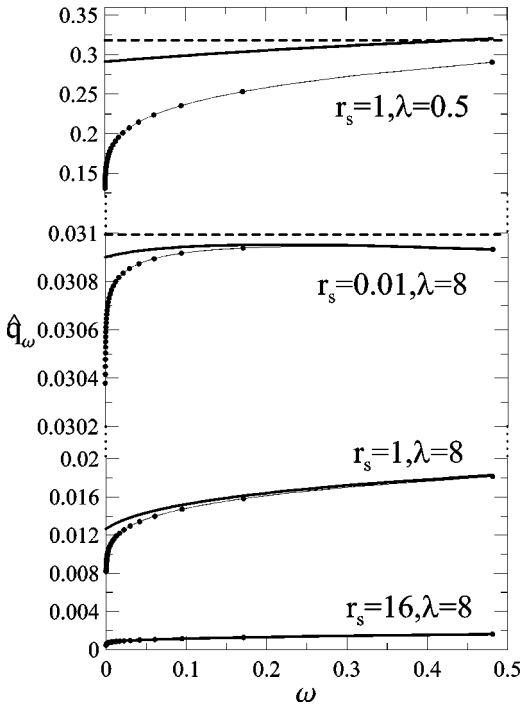

FIG. 5. The dependence of $\hat{q}_\omega$ on ω for various values of r_s and λ. The solid lines are the numerical results, the dashed lines are the results of r_s expansion, and the dots are the results of $1/\lambda$ expansion.

of expansion in r_s briefly described in the Appendix, while dots are the results of the next to leading order in $1/\lambda$ also summarized in the Appendix.

At zero frequency the large coupling value of $\hat{q}_{\omega=0}$ is much smaller than the noninteracting result $1/(4\lambda)$. We define the renormalized relaxation time via

$$\tau_r \equiv \frac{\sqrt{\tau}}{2q_{\omega=0}} = \frac{\tau}{4\lambda\hat{q}}. \tag{29}$$

One observes in Fig. 5 that for $r_s<10$ the disorder parameter $\hat{q}_\omega$ is almost independent of ω in the whole region where it is important for calculations of integrals over Green's functions (namely when $\hat{q}<\omega$). Therefore one can still consider the system as a disordered liquid without temporary dispersion. It is a good approximation to simplify the analysis by considering a variational principle with constant $\hat{q}_\omega=\hat{q}$.

The saddle-point equations in that case take a simpler form (after integration over ω)

$$1=\frac{1}{(2\pi)2\lambda}\int_\varepsilon \frac{1}{(\varepsilon+e_\varepsilon)^2+\hat{q}^2}, \tag{30}$$

$$e_\varepsilon=\frac{\sqrt{2}r_s}{\pi^2}\int_{\varepsilon'}\kappa[\varepsilon-\varepsilon']\mathrm{sgn}[\varepsilon']\left(\frac{\pi}{2}-\arctan\frac{\hat{q}}{|\varepsilon'+e_{\varepsilon'}|}\right). \tag{31}$$

From Fig. 4 we observe that at very large coupling it is similar to the very marginal Fermi liquid of the clean case. The interaction correction Σ_ε dominates and is not proportional to ε, therefore it does not reduce to a mere renormalization of the density of states on the Fermi level. At finite but large coupling the renormalization of the density of states is still large, see Table I. We calculated it as explained in Sec. II D, Eq. (33). At large coupling the quantity diverges very fast. The limiting values of $\hat{q}$ at zero frequency for various bare diffusion constants and couplings are given in Table II.

Before trying to further exploit the solutions of the saddle-point equations, we would like to explicitly show that even perturbatively the reduction in the DOS due to exchange can be clearly seen.

D. Reduction in the density of states

The fact that at small ε (deviation of fermion's energy from μ) and large coupling r_s, $\Sigma_\varepsilon \gg 2\mu\varepsilon$ leads to significant

TABLE II. Renormalization of the inverse relaxation time $\hat{q}$ for various couplings r_s and bare diffusion constants λ.

	r_s					
λ	0.1	1	2	4	8	16
8	0.028	0.013	0.0070	0.0034	0.0015	0.00067
4	0.057	0.028	0.015	0.0075	0.0033	0.0015
2	0.12	0.056	0.034	0.017	0.0073	0.0032
0.5	0.47	0.29	0.18	0.086	0.037	0.015
0.1	2.4	1.6	1.3	0.67	0.27	0.10

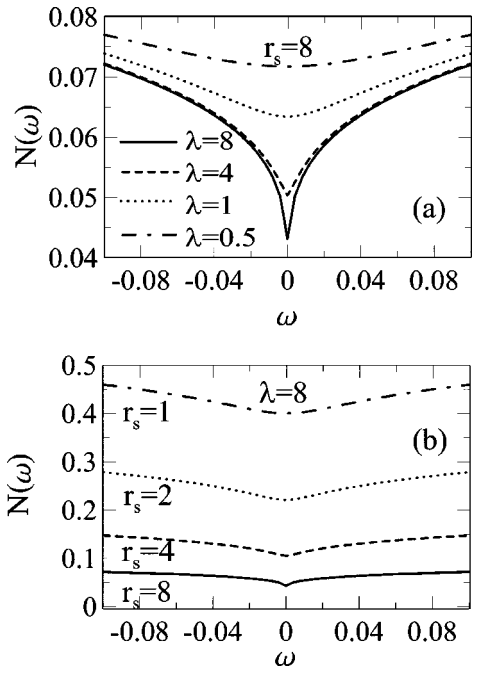

FIG. 6. The DOS $N(\omega)$ (normalized to the ideal gas DOS $N_0=1/2\pi$) for (a) fixed $r_s=8$ and various λ, (b) fixed $\lambda=8$ and various coupling r_s.

rearrangement of the DOS in the vicinity of the Fermi surface. The DOS $N(\omega)$ is related to the retarded Green's function by

$$N(\omega)=-\frac{1}{\pi}\mathrm{Im}\int_p G_p^{ret}(\omega). \tag{32}$$

For frequency independent q_ω it reduces to

$$N(\omega)=-\frac{1}{\pi}\mathrm{Im}\sum_p \frac{1}{\omega-2\mu\varepsilon_p-\Sigma_p+iq}$$
$$=\frac{1}{2\pi^2}\int_\varepsilon \frac{\hat{q}}{(\hat{\omega}-\varepsilon-e_\varepsilon)^2+\hat{q}^2}. \tag{33}$$

The DOS at $\omega=0$ for various couplings are given in Table I, while the DOS as a function of ω for fixed $r_s=8$ and various λ in Fig. 6(a) and for fixed $\lambda=8$ and various coupling r_s in Fig. 6(b). At large coupling the renormalization of limiting value of τ_r is very large and $\hat{q}$ approaches zero. The integrand in Eq. (33) becomes δ function and the DOS is determined by the derivative of e_ε with respect to ε at $\varepsilon=0$:

$$N(\omega)=\frac{1}{2\pi}\frac{1}{1+e'_\varepsilon}\bigg|_{\hat{\omega}=\varepsilon+\hat{e}_\varepsilon}. \tag{34}$$

Using the simplified saddle-point equation, Eq. (31), the derivative of e_ε is

$$e'\equiv\frac{d}{d\varepsilon}e\bigg|_{\varepsilon=0}=-\frac{2\sqrt{2}r_s}{\pi^2}\int_{\varepsilon\geqslant 0}\kappa'[\varepsilon]\left(\frac{\pi}{2}-\arctan\frac{\hat{q}}{e_\varepsilon}\right). \tag{35}$$

The asymptotic of the function $\kappa'[\varepsilon]$ is as follows. At small ε it is negative and divergent, $\kappa'[\varepsilon]\approx-1/2\varepsilon$. As we have seen in the clean case this leads to vanishing DOS. However in the presence of disorder the divergence is cut off since the second multiplier vanishes at small ε as $\hat{e}_\varepsilon/\hat{q}$. This in particular means that in the case of disorder the RPA improvement is not necessary as long as the cutoff due to disorder is larger than the cutoff due to screening (see discussion in Sec. II B).

At large ε the integral converges rapidly since $\kappa'[\varepsilon]\approx-\pi/2\varepsilon^2$. This can be seen perturbatively as well using results of the preceding section. To leading order in r_s,

$$e'=-\frac{2\sqrt{2}r_s^2}{\pi^2\hat{q}^{(0)}}\int_{\varepsilon\geqslant 0}\kappa'[\varepsilon]e_\varepsilon^{(1)},$$

where $\hat{q}^{(0)}$ and $e_\varepsilon^{(1)}$ are given in the Appendix by Eqs. (A1) and (A2). The frequency-dependent correction to the DOS (compared to the ideal gas $N_0=1/2\pi$) is perturbatively

$$\delta N(\omega)=N(\omega)-N_0=\frac{r_s}{8\pi^2\lambda^2}\int_\varepsilon\frac{(\hat{\omega}-\varepsilon)\hat{e}_\varepsilon^{(1)}}{[(\hat{\omega}-\varepsilon)^2+\hat{q}^{(0)2}]^2}.$$

In particular on the Fermi surface

$$\delta N(0)=\frac{32r_s\lambda^3}{\pi}\hat{q}^{(1)},$$

where $\hat{q}^{(1)}$ is given in the Appendix by Eq. (A3). The results for the DOS are given in Table I. The effect of reduction of the DOS at the Fermi level due to repulsion increases fast with increasing r_s and depends weaker on λ. As we mentioned, in the clean case the DOS approaches zero at any coupling no matter how small, which is clearly an artifact of neglecting screening at this stage. However the disorder effectively ''averages'' the distribution of states making it finite. At the limit of large disorder the DOS should approach the ideal-gas one. Repulsive interaction works to reduce the DOS near the Fermi surface and is expected to make scattering by impurities less effective. We will see in the following section that e' is related to the renormalized value of the diffusion constant: $D_r=\lambda/e'$.

III. A SYSTEMATIC EXPANSION IN FLUCTUATIONS AROUND THE SADDLE POINT

A. Classification of fluctuations around the ground state

All the Hubbard-Stratonovich fields correspond to elementary (harmonic) excitations of the system. The spectrum of these excitations is determined by a quadratic term in expansion of the effective action, Eq. (7), around the saddle point determined by a solution of the minimization equations (27) and (28). This quadratic form should be diagonalized to find the spectrum. We first sort out evidently massive modes

which do not play a role in subsequent discussion. These are exchange E and the Cooper exchange Θ fields. In the following two sections we address photons, diffusons, and Cooperons.

Since the ground state does not break spontaneously the electric charge U(1) symmetry the charged fields Θ and Cooperon Δ do not mix with neutral fields E, diffuson Q, and static photon Φ. In this sector the inverse propagator elements (second functional derivatives of the effective action with respect to relevant fields) are

$$\langle {}^{n'}_{P'p'}|D^{-1}_{\Theta\Theta}|^{n}_{Pp}\rangle = \delta(P-P')\,\delta^{nn'}[\,\delta(p-p')v(p)^{-1} - \delta(p-p')G^{n}_{P/2+p/2}G^{n}_{P/2-p/2} + \delta(p+p')G^{n}_{P/2+p/2}G^{n}_{P/2-p/2}],$$

$$\langle {}^{n'm'}_{P'p'}|D^{-1}_{\Delta\Theta}|^{n}_{Pp}\rangle = \delta(P-P')\,\delta^{n'+m',n}G^{n'}_{P/2+p/2}G^{m'}_{P/2-p/2},$$

$$\langle {}^{n'm'}_{p'}|D^{-1}_{\Delta\Delta^*}|^{nm}_{p}\rangle = \delta^{n+m,n'+m'}\,\delta(p-p') \times\left(\delta^{nn'}-\frac{1}{\tau}\sum_q\, G^{n}_{p+q}G^{m}_{q}\right), \quad (36)$$

where P, p are a total and relative momenta of the two-electron state, respectively. The relation to indices of the Θ fields used in the effective action, Eq. (7), is obvious:

$$P=p_1+p_2, \quad p=p_1-p_2.$$

The $\Theta\Theta$ element of the inverse fluctuation propagator indicated that in this channel there are no massless modes. This is evident at small coupling since the diagonal first term dominates, but since the Coulomb interaction is repulsive is true for any coupling. The mixing with Cooperon cannot turn it to a massless mode. We therefore discard the field Θ in what follows. Now we turn to the neutral fields.

Second derivative of the effective action with respect to E is

$$\langle {}^{n'}_{P'p'}|D^{-1}_{EE}|^{n}_{Pp}\rangle = \delta(P-P')\,\delta^{nn'}(\delta(p-p')v(p)^{-1} - \delta(p-p')G^{n}_{P/2+p/2}G^{n}_{P/2-p/2}). \quad (37)$$

It is again a massive mode and its mixing with other neutral field is unimportant. More interesting are the diffuson Q and the static photon Φ fields:

$$\langle {}^{n'm'}_{p'}|D^{-1}_{QQ}|^{nm}_{p}\rangle = \delta^{n-m,n'-m'}\,\delta(p-p') \times\left[\,\delta^{nn'}-\frac{1}{\tau}\sum_q\, G^{n}_{p+q}G^{m}_{q}\right], \quad (38)$$

$$\langle {}^{l}_{p'}|D^{-1}_{\Phi Q}|^{nm}_{p}\rangle = -\,\delta^{l,n-m}\delta(p-p')\,\frac{1}{\tau}\sum_q\, G^{n}_{p+q}G^{m}_{q}\,, \quad (39)$$

$$\langle {}^{l'}_{p'}|D^{-1}_{\Phi\Phi}|^{l}_{p}\rangle = \delta^{ll'}\,\delta(p-p')\left[v(p)^{-1}-\frac{T}{\tau}\sum_{qn}\, G^{n}_{p+q}G^{m}_{q}\right]. \quad (40)$$

To conclude there are three fields which might have potentially massless modes. While Cooperon cannot mix with the other two, diffusons and photons can. Even if both the diffuson and the photon are massless, after mixing the massless modes could turn massive. We study this phenomenon next.

B. Anderson-Higgs mechanism for diffusons

1. Matrix elements of the photon-diffuson inverse propagator matrix

At small frequency and momenta around the Fermi surface we will use the asymptotic expressions for the solution of the saddle-point equations:

$$q_\omega \approx q_{\omega=0}=2\mu\sqrt{\tau}\hat{q}=\frac{\sqrt{\tau}}{2\tau_r}, \quad (41)$$

$$\Sigma_\varepsilon \approx 2\mu e'\varepsilon \equiv 2\mu(Z^{-1}-1)\varepsilon,$$

where renormalizations of the ''disorder efficiency'' $1/\tau$ and of the inverse density of states at the Fermi level Z^{-1} were introduced in Eq. (26). As can be seen from Tables I and II at large coupling they are quite large.

Then one computes standard diagrams in Eq. (38) for matrix elements of the inverse propagators involving diffusons and static photons at certain fixed momentum p:

$$\langle n'm'|D^{-1}_{QQ}|nm\rangle = \delta^{n-m,n'-m'}\left[\,\delta^{nn'}-\frac{1}{\tau}\sum_q\, G^{n}_{p+q}G^{m}_{q}\right] \equiv \delta^{n-m,n'-m'}[\,\delta^{nn'}-\theta(-nm)B(p,n-m)], \quad (42)$$

where at small ω and p, it contains well-studied function ''bubble'' integral having the following asymptotic at small frequencies and momenta:

$$B(p,l)=\frac{1}{\tau}\sum_q\, G^{n}_{p+q}G^{n+l}_{q}\simeq 1-\omega_l\tau_r-D_r p^2, \quad (43)$$

for $n<0$, $n+l>0$. Within an approximation of ''renormalization,'' Eq. (41), one obtains

$$D_r=\frac{Z\tau_r^3}{\tau^3}\lambda.$$

Beyond this approximation the value of effective diffusion constant can be calculated from the values given in Tables I and II. As usual, we discard the massive same frequency sign $++$ and $--$ modes[8] and concentrate on different frequency sign excitations.

It is convenient to this end to rescale the photon field

$$\Phi_{pn}=\sqrt{v(p)T}\tilde{\Phi}_{pn}. \quad (44)$$

The mixing and the photon inverse propagators are

$$\langle {}^{l}_{p'}|D^{-1}_{\Phi Q}|^{nm}_{p}\rangle_l = -\,\delta^{l,n-m}\delta(p-p')\sqrt{v(p)T}\,\theta(-nm) \times B(p,n-m), \quad (45)$$

$$\langle {}^{l'}_{p'}|D^{-1}_{\Phi\Phi}|{}^{l}_{p}\rangle_{l'} = \delta^{ll'}\delta(p-p')[1-v(p)L(p,l)], \quad (46)$$

where the Lindhard function[24] has an asymptotic behavior

$$L(p,l) \equiv T\sum_n B(p,n,n-l) \approx (-1+\omega_l\tau_r+\cdots)/(2\pi). \quad (47)$$

One notices in Eq. (45) that, due to the factor $\theta(-nm)$, it is precisely the massless different frequency sign diffusons that mix with photon. The mixing strength is large since it is determined by the same bubble integral that appears in the diffuson's inverse propagator, Eq. (42). Now we will invert this matrix and find its eigenmodes.

2. Eigenvalues and eigenmodes: Physical photon and diffuson

The $Q\Phi$ inverse propagator matrix is "blocked" for different photon frequencies $l=n-m$ (we take $n>0$, $m<0$), namely we can consider a single value of l. For a fixed frequency $l>0$ the range of possible n is limited to $-l<n<0$ and the $(l+1)\times(l+1)$ matrix has the following form:

$$\begin{pmatrix} a & 0 & \dots & b \\ 0 & a & \dots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \dots & c \end{pmatrix}, \quad (48)$$

where

$$a = 1-B(p,l), \quad b = -\sqrt{v(p)T}B(p,l),$$
$$c = 1-v(p)TL(p,l). \quad (49)$$

Eigenvalues of this matrix are the $(l-1)$ times degenerate $a \sim \omega_l\tau_r + D_r p^2$ (original "massless" diffusons before mixing) and two nondegenerate eigenvalues

$$\lambda_\pm = \tfrac{1}{2}[a+c \pm \sqrt{(a-c)^2+4lb^2}],$$

corresponding to eigenvectors $\{1,1,1,\dots,1,a_\pm\}$ with $a_\pm = (\lambda_\pm - a)/b$. Their asymptotic at small momenta and frequency is

$$\lambda_+ \approx 1+v(p)/(2\pi),$$
$$\lambda_- \approx 2\pi v(p)^{-1}(\omega_l\tau_r + D_r p^2) + D_r p^2. \quad (50)$$

The form of the λ_+ eigenvalue means that physical photon mode $\overline{\Phi}^\omega_p = \Sigma^0_{n=-l} Q^n_p + a_+\Phi^\omega_p$ propagator [rescaling back by $\sqrt{v(p)T}$] is the RPA photon propagator exhibiting Debye screening:

$$\langle \overline{\Phi}^\omega_p \overline{\Phi}^{-\omega}_{-p}\rangle \sim \frac{1}{v^{-1}(p)+2\pi}. \quad (51)$$

The second is "symmetric" in n superposition of diffuson modes $\overline{Q}^\omega_p = \Sigma^0_{n=-l} Q^n_p + a_-\Phi^\omega_p$ and is no longer the usual diffusion pole:

$$\langle \overline{Q}^\omega_p \overline{Q}^{-\omega}_{-p}\rangle \sim \frac{1}{\omega\tau_r + D_r p^2 + D_r v(p)p^2/(2\pi)} \quad (52)$$

$$\simeq \frac{1}{\omega\tau_r + D_r e^2 p/(4\pi)}. \quad (53)$$

It becomes harder (less singular) than in the standard treatment in which the mixing between photon and diffuson is neglected. It is interesting to note that in the standard treatment at small couplings the mixing is not neglected as far as photon's propagator is concerned. One uses the RPA propagator, Eq. (51), despite the fact that the same off-diagonal matrix element, Eq. (39), is simultaneously responsible for the essential modification of the diffuson.

The inverse of a matrix of the type of Eq. (48) generally is

$$\frac{1}{ac-lb^2}\begin{pmatrix} c-\frac{lb^2}{a}+\frac{b^2}{a} & \frac{b^2}{a} & \cdots & -b \\ \frac{b^2}{a} & c-\frac{lb^2}{a}+\frac{b^2}{a} & \cdots & -b \\ \vdots & \vdots & \ddots & \vdots \\ -b & -b & \cdots & a \end{pmatrix}.$$

Matrix elements of the propagators therefore are

$$\langle n, n+l|D_{QQ}|n', n'+l\rangle = P_1\delta_{nn'} + P_2\frac{1}{2\pi l}, \quad (54)$$

$$\langle n, n+l|D_{Q\Phi}|l\rangle = P_3\frac{1}{\sqrt{2\pi l}}, \quad (55)$$

$$\langle l|D_{\Phi\Phi}|l\rangle = P_4, \quad (56)$$

where functions P and their asymptotic at small frequency and momentum are

$$P_1 = \frac{1}{a} \approx \frac{1}{\tau_r\omega + D_r p^2}, \quad (57)$$

$$P_2 = \frac{lb^2}{a(ac-lb^2)} \approx \frac{v(p)}{\tau_r\omega + D_r p^2 + D_r p^2 v(p)}, \quad (58)$$

$$P_3 = -\frac{\sqrt{2\pi l}b}{ac-lb^2} \approx \frac{\sqrt{\omega v(p)}}{\tau_r\omega + D_r p^2 + D_r p^2 v(p)}, \quad (59)$$

$$P_4 = \frac{a}{ac-lb^2} \approx \frac{\tau_r\omega + D_r p^2}{\tau_r\omega + D_r p^2 + D_r p^2 v(p)}. \quad (60)$$

We will see in Sec. III D that the diagonal part of the diffuson describes density fluctuations and therefore mixing with photon makes electrons nondiffusive at large densities.

Propagators for relevant modes supplemented by vertices, the fermion-fermion-diffuson

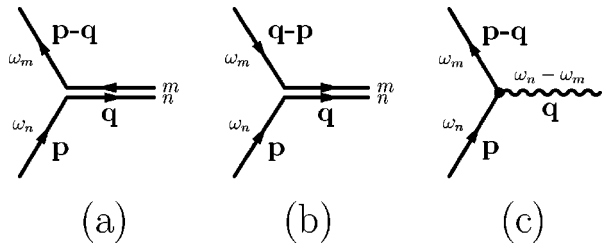

FIG. 7. The Feynman rules for (a) fermion-fermion-diffuson, (b) fermion-fermion-Cooperon, (c) fermion-fermion-photon vertex. The solid and wavy lines denote the fermion and photon propagator, respectively, the double solid lines with opposite (same) direction arrows denote the diffuson (Cooperon) propagator.

$$\Gamma^{-}_{\bar{\psi}\psi Q} = -\frac{i}{\tau}$$

and the fermion-fermion-photon Φ

$$\Gamma^{-}_{\bar{\psi}\psi\Phi} = -i$$

constitute Feynman rules shown in Figs. 7(a) and 7(c), respectively. However fields Φ and Q do not correspond to ''modes'' or ''bosonic excitations'' of the system due to mixing between them discussed in detail in the preceding section. One can still use these fields in calculation considering them as a vector and their propagator as a matrix.

C. The Cooperon propagator

As we mentioned in Sec. III A due to its charge the Cooperon does not mix with photon or any other neutral field. It is massless for different sign frequencies and massive for the same sign frequencies. The strong coupling however influences its propagator beyond the evident renormalization $\lambda \to D_r$, $\tau \to \tau_r$. Substituting the expression of the fermion propagators, Eq. (13), into Eq. (36) the inverse propagator of the excitation is

$$\langle nm|D^{-1}_{\Delta\Delta *}|nm\rangle = 1 - \theta(-nm)B(p, n-m). \qquad (61)$$

Numerical solution of the saddle-point equations substituted into Eq. (36) shows that the dependence is quadratic only till certain momentum at which it saturates, see Fig. 8 for $r_s = 1$, 2, and 4. This is of importance later when we estimate the quantum correction to conductivity in Sec. IV.

Note that the excitation remains massless even at strong coupling. This follows from very general considerations. Consider the bubble diagram

$$B(\omega_l, p) = \frac{1}{(2\pi)^2}\int_q G_{qn} G_{q+p,m}, \qquad (62)$$

where $l = m - n$. At zero momentum using the saddle-point equation (15) one obtains

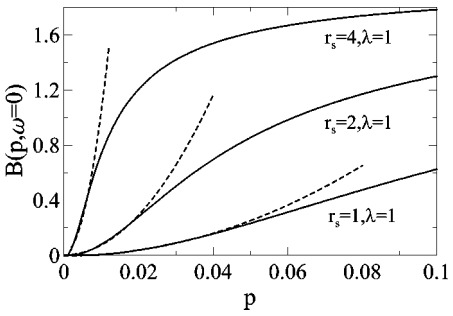

FIG. 8. The dependence of $B(p, \omega = 0)$ on p for various r_s and λ. The solid lines are the original values of $B(p, \omega = 0)$, while the dashed lines are their corresponding quadratic fitting values.

$$B(\omega_l, p=0) = \frac{1}{(2\pi)^2}\int_q (G_{qn} - G_{qm}) \frac{1}{i[\omega_l + q_m - q_n]}$$

$$= \frac{q_m - q_n}{\omega_l + q_m - q_n}. \qquad (63)$$

This approaches 1 in the limit of zero frequency as long as there is a jump in Q_ω from negative to positive frequencies. The propagator of the different frequency sign Cooperons at small frequencies and momenta is

$$D_{\Delta\Delta *} = \frac{1}{2q\omega + B' p^2}, \qquad (64)$$

where B' denotes a derivative of the bubble integral

$$B' = \left.\frac{\partial B(0,p)}{\partial p^2}\right|_{p=0}. \qquad (65)$$

Assuming quadratic momentum dependence of $B(p, \omega)$, direct calculation leads to

$$B' = \frac{1}{2\pi(2\mu)^2}\int_\varepsilon \frac{(\varepsilon + e'_\varepsilon)^2 - \hat{q}^2}{[\hat{q}^2 + (\varepsilon + e'_\varepsilon)^2]^3}(1 + e'_\varepsilon)^2. \qquad (66)$$

The values of B' for various couplings r_s and bare diffusion constants λ determining the Drude conductivity are given in Table III. The perturbative expression for this quantity is

$$B' = -\lambda\tau^2\left[1 + r_s\left(-12\lambda\hat{q}^{(1)} + \frac{2^{5/2}\lambda^3}{\pi^2}C\right)\right], \qquad (67)$$

where

$$C = \frac{1}{(4\lambda)^4} G^{22}_{33}((4\lambda)^2|^{3/2,3/2}_{1,1}) + G^{23}_{43}\left(\frac{1}{(4\lambda)^2}\Big|^{2,2,2}_{3/2,3/2,3}\right)$$

$$+ 2G^{23}_{43}\left(\frac{1}{(4\lambda)^2}\Big|^{2,2,2}_{3/2,3/2,4}\right), \qquad (68)$$

TABLE III. Derivative B'/τ^2 of the Lindhard function in the presence of disorder for various values of r_s and λ.

	r_s					
λ	0.1	1	2	4	8	16
8	12.2	3.13×10^2	3.33×10^3	6.06×10^4	1.40×10^6	3.12×10^7
4	5.78	1.14×10^2	1.15×10^3	2.11×10^4	5.14×10^5	1.29×10^7
2	2.74	40.4	3.77×10^2	7.02×10^3	1.81×10^5	4.93×10^6
0.5	0.620	4.75	34.8	6.35×10^2	1.90×10^4	6.18×10^5
0.1	0.117	0.605	1.79	25.3	9.15×10^2	4.10×10^4

with the standard notation $G^{mn}_{pq}(z|^{a_1,\ldots,a_p}_{b_1,\ldots,b_q})$ of the Meijer functions.[31] It rises fast with r_s, while being weakly dependent on λ. Finally the fermion-fermion-Cooperon vertex is

$$\Gamma^{-}_{\bar\psi\psi\Delta}=-\frac{i}{\tau}$$

and it completes the Feynman rules shown in Fig. 7(b).

D. Density-density correlator and conductivity to leading order

1. Modification of diffusive motion due to strong interaction

One of the most important characteristics of 2DEG is the density-density correlator describing the diffusive nature of the charge carrier's motion in a disordered medium. It is closely related to dielectric function and polarizability of 2DEG. The correlator is given in Matsubara formalism by

$$\chi(\omega,p)=\sum_{q_1,q_2}\int_0^{1/T}d\tau\, e^{i\omega\tau}\times\langle T_\tau[\overline{\psi}_{p+q_1}(\tau)\psi_{q_1}(\tau)\overline{\psi}_{p+q_2}(0)\psi_{q_2}(0)]\rangle.$$

First we use the Feynman rules stated above to calculate the density-density correlator at the leading order. In the limit of small frequencies the contributions come from diagrams (a)–(d) in Fig. 9. They are $-L(p,l)$, $-\omega_l B(p,l)^2 P_1(1+P_4)/2\pi$, $2\sqrt{\omega_l v(p)/2\pi}L(p,l)B(p,l)P_3$, and $-v(p)L(p,l)^2P_4$, respectively, and can be combined into

$$\chi(\omega_l,p)=\frac{\chi_0(\omega_l,p)}{1+v(p)\chi_0(\omega_l,p)}.$$

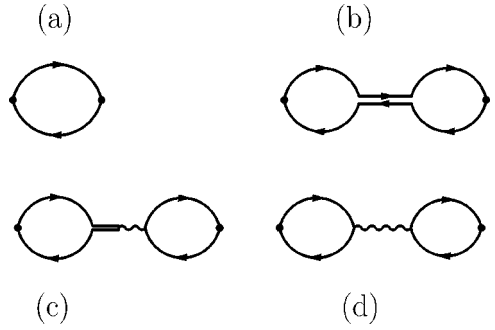

FIG. 9. Contributions to the density-density correlator at the leading order.

Here expressions for L, B, and P_4 are given in Eqs. (47), (62), (57)–(60) and the "noninteracting" correlator is defined by

$$\chi_0(\omega_l,p)=-\left(L(p,l)+\frac{\omega_l}{2\pi}\frac{B(p,l)^2}{1-B(p,l)}\right).$$

Its asymptotic at small ω and p is

$$\chi(\omega_l,p)=\frac{D_r p^2}{\omega_l\tau_r+D_r p^2+D_r p^2 v(p)/(2\pi)}.$$

Therefore not surprisingly it is proportional to propagator of the "diagonal" diffuson defined in Eq. (52). The diffusive behavior dominates short-range fluctuations only on scale smaller than $s=2\epsilon/e^2$. On a larger scale the last term in the denominator is linear in p and is therefore larger than the standard diffusion term. This makes diffusion less long range although in 2D it does not become a short-range one. The scale was introduced by Si and Varma[38] and we will comment on connection to their work in Sec. V.

2. The leading (Drude) contribution conductivity

The dc conductivity can be read off the density-density correlator using the relation

$$\sigma=\lim_{\omega\to0}\lim_{p\to0}\frac{e^2\omega}{p^2}\chi(\omega,p)=-\lim_{\omega\to0}\frac{e^2\omega^2}{2\pi}\frac{B(\omega,0)^2}{[1-B(\omega,0)]^2}B'=\frac{e^2}{2\pi}4q^2B'/\tau, \quad (69)$$

which follows from the Kubo formula.[10] Here we used the asymptotic of the bubble integral B, Eq. (47), and the imaginary part of self-energy q and derivative of the bubble integral B' defined in Eqs. (15) and (65), respectively. Results are given in Table IV for various r_s and λ. One observes that at large coupling the "Drude" conductivity increases considerably compared with the noninteracting one. We explain this by reduction of interaction with disorder (q is much smaller than its noninteracting value of $1/2\tau$) despite the reduction in density of states (B' larger than its noninteracting value of $-\mu$). At small r_s using Eqs. (A2) and (A3) of the Appendix and Eq. (67) one obtains

TABLE IV. The Drude conductivity $\sigma/2\pi e^2$ for various values of r_s and λ.

	r_s					
λ	0.1	1	2	4	8	16
8	9.85	51.4	1.68×10^2	7.15×10^2	3.39×10^3	1.53×10^4
4	4.79	22.0	70.3	3.01×10^2	1.47×10^3	7.18×10^3
2	2.33	9.31	28.7	1.24×10^2	6.22×10^2	3.20×10^3
0.5	0.553	1.61	4.44	19.0	1.03×10^2	5.77×10^2
0.1	0.107	0.246	0.470	1.79	10.5	68.3

$$\sigma=\frac{e^2}{2\pi}\lambda\left(1-4r_s\lambda\hat{q}^{(1)}+r_s\frac{2^{5/2}\lambda^3}{\pi^2}C\right),$$

where C is given in Eq. (68). The leading-order contribution dominates at small couplings and disorder. However, the theory has zero modes—Cooperons. Therefore possible IR divergencies might render the leading-order results invalid at large coupling or disorder. In principle for zero temperature and infinite samples the results are invalid for all couplings. Our next task is to find a range of parameters and temperature (or sample sizes) in which the IR divergencies at the next order are still small compared to the main contribution.

IV. SUPPRESSION OF WEAK LOCALIZATION BY THE LONG-RANGE INTERACTION EFFECTS

A. The saddle-point expansion and the spin-singlet approximation

In this section we describe some of the corrections around the variational ground state found in Sec. II and used in Sec. III to calculate several physical quantities. The steepest descent expansion in terms of Feynman diagrams is quite standard.[7,8] We briefly describe it introducing N_s identical ''spin'' components to show that the expansion might be interpreted as ''$1/N_s$'' expansion (spin might include other degeneracies such as multiple valleys in Si). The action including the spin indices σ is given in Eq. (5). It is a peculiar feature of the disordered Coulomb problem that the leading order in $1/N_s$ vanishes for two entirely unrelated reasons. The direct contribution in the disorder part vanishes due to the fact that it is of higher (second) order in replicas N_r as well:

$$\sum_{a,b}\sum_{\sigma_1,\sigma_2}G^{aa}_{\sigma_1\sigma_1}G^{bb}_{\sigma_2\sigma_2}\sim N_r^2N_s^2.$$

The direct N_s^2 contribution to the Coulomb part

$$\int_{x,y}\sum_{a}\sum_{\sigma_1,\sigma_2}G^{aa}_{\sigma_1\sigma_1}(x,x)v(x-y)G^{aa}_{\sigma_2\sigma_2}(y,y)=0$$

vanishes due to neutralizing background $\int_y v(x-y)=0$ (and under assumption of homogeneity). Therefore leading terms are of order N_s. The free theory action is also of order N_s. Therefore all the terms in action are of the order N_s and it plays a role of the ''loop expansion parameter'' and comes always in combination with $1/\hbar$. This Hartree-Fock logic is not rigorous however. There is an assumption involved: it is assumed that all the Hubbard-Stratonovich fields which in general are tensorial are dominated by their singlet part:

$$Q^{\sigma_1\sigma_2}\sim\delta^{\sigma_1\sigma_2}Q.$$

In this paper we will make such an assumption. Therefore we neglect, for example, triplet channels in the physical case of $N_s=2$.

The partition function (suppressing for simplicity the replica indices and writing explicitly just one of the HS fields) or any observable is expanded around the saddle point

$$Z=\int_Q e^{N_sA_{eff}[Q]}\approx\int_Q\exp(N_sA_{eff}[Q_{SP}]+\tfrac{1}{2}QD^{-1}[Q_{SP}]Q+\Delta A[Q]),$$

where $\Delta A[Q]$ contains all the cubic, quartic, and higher-order terms in Q. From this Feynman rules are read and they scale compared to Fig. 7 in the following way: HS fields' propagators are proportional to N_s, fermion loop also has N_s, while fermion-fermion-boson vertex is $1/\sqrt{N_s}$. The leading-order contribution of conductivity considered so far is of order N_s and we will consider order $N_s^0=1$ in the following section.

B. Fluctuation correction to the density-density correlator and conductivity

1. Density-density correlator

The correction to density-density correlator at two-loop order which contributes to the small frequency limit (the only ones needed for subsequent calculation of the conductivity) is given in Fig. 10,

$$\begin{aligned}\delta\chi(\omega_l,p)=T\sum_{q,r}\sum_{nmn'm'}&\theta(-n(n+l))\theta(-n'(n'+l))\\&\times B(p,l)^2\langle n,n+l|D_{QQ}|m,m+l\rangle\\&\times G^m_qG^{m+l}_{p+q}G^{m'}_{q+r}G^{m'+l}_{p+q+r}\\&\times\langle n,m'|D_{\Delta\Delta r}|m+l,m'+l\rangle\\&\times\langle m',m'+l|D_{QQ}|n',n'+l\rangle.\end{aligned}$$

All the other diagrams are regular as $\omega\to0$, hence they do not give contributions to the dc conductivity. Near the Fermi surface one ''disentangles'' the momenta flowing in the central loop, see Fig. 10,

$$\delta\chi=T\sum_{n=0}^{l}\frac{B(p,l)^2}{[1-B(p,l)]^2}B_4(p,l)\sum_r\frac{1}{1-B(r,l)},\quad(70)$$

where

$$B_4(p,l)\equiv\sum_q G^n_qG^{n+l}_{-q}G^n_{p-q}G^{n+l}_{q-p}.\quad(71)$$

The integral over r is logarithmically infrared divergent in 2D and, as usual,[10] signals breakdown of naive perturbation

TABLE V. The values of B_4'/τ^4 for various r_s and λ.

	r_s					
λ	0.1	1	2	4	8	16
8	3.77	4.64×10^2	1.60×10^4	1.25×10^6	1.47×10^8	2.11×10^{10}
4	3.46	2.86×10^2	9.05×10^3	7.21×10^5	8.92×10^7	1.28×10^{10}
2	3.18	1.70×10^2	4.78×10^3	3.85×10^5	5.14×10^7	7.78×10^9
0.5	2.73	53.6	1.04×10^3	8.10×10^4	1.35×10^7	2.58×10^9
0.1	2.44	27.4	1.23×10^2	6.65×10^3	1.51×10^6	4.72×10^8

theory and appearance of weak-localization effects. We assume an IR cutoff (to be defined more explicitly below) and will use this expression to calculate conductivity.

2. Weak localization

The fluctuation correction to conductivity using Kubo formula is

$$\delta\sigma=\lim_{\omega\to0}\frac{e^2\omega^2}{2\pi}\frac{B(0,l)^2}{[1-B(0,l)]^2}B_4'\sum_r\frac{1}{1-B(r,l)},$$

where the derivative of B_4, Eq. (71), is defined by

$$B_4'\equiv\left.\frac{\partial B_4(p,l=0)}{\partial(p^2)}\right|_{p=0}.$$

After some algebra it takes a form

$$B_4'=\frac{1}{2\pi(2\mu)^4}\int_\varepsilon\frac{(\varepsilon+e_\varepsilon)^2}{[\hat{q}^2+(\varepsilon+e_\varepsilon)^2]^4}(1+e_\varepsilon')^2.$$

Values of the coefficient B_4' for various couplings and disorder strength are given in Table V.

Before discussing physical implications of the correction we provide perturbative result for B_4':

$$B_4'=2\lambda\tau^4\left[1+r_s\lambda\left(-2^2 5\hat{q}^{(1)}+\frac{2^{5/2}\lambda^2}{3\pi^2}C_4\right)\right],$$

where

$$C_4=\frac{3}{2^7\lambda^4}G_{33}^{22}((4\lambda)^2|_{1,1}^{3/2,3/2})-9G_{43}^{23}\left(\frac{1}{(4\lambda)^2}\Big|_{3/2,3/2,3}^{2,2,2}\right)+12G_{34}^{23}\left(\frac{1}{(4\lambda)^2}\Big|_{3/2,3/2,4}^{2,2,2}\right)+4G_{43}^{23}\left(\frac{1}{(4\lambda)^2}\Big|_{3/2,3/2,5}^{2,2,2}\right).$$

The quantity is increasing very fast with coupling r_s and decreases slowly with disorder strength.

Returning to conductivity one obtains

$$\delta\sigma=\frac{e^2}{2\pi}\frac{2q^2B_4'}{\pi\tau}\int_0^\infty\frac{r\,dr}{1-B(r,l)}.$$

In 2D the integral is dominated by small momenta. Therefore at very low temperature we can use approximation of Eq. (64),

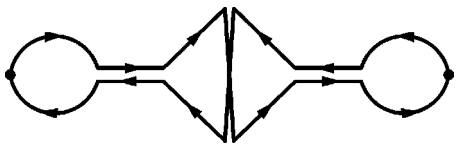

FIG. 10. The correction to the density-density correlator at two-loop order.

$$\delta\sigma=\frac{e^2}{2\pi}\frac{2q^2B_4'}{\pi\tau B'}\int_{p_{IR}}^{p_{UV}}\frac{r\,dr}{r^2}=\frac{e^2}{2\pi}\frac{q^2B_4'}{\pi\tau B'}\ln\frac{p_{UV}^2}{p_{IR}^2}. \quad (72)$$

As usual[10] it is cut off in both infrared and ultraviolet. The infrared cutoff for the weak-localization logarithmic divergence can be set by finite temperature

$$p_{IR}^2=\frac{2\pi Tm^*/\tau}{2(q/\sqrt{\tau})(B'/\tau^2)}$$

or finite size L of the sample

$$p_{IR}=\frac{2\pi}{L}.$$

The ultraviolet cutoff is[39]

$$\frac{p_{UV}^2}{2m^*}=\min\{\mu,1/\tau\}. \quad (73)$$

C. Crossover temperature

Let us find a temperature at which the perturbation theory in ''loops'' or $1/N_s$ breaks down. At this temperature the Drude conductivity is significantly reduced by fluctuations and one conservatively estimates it as settling of the weak-localization (the Anderson insulator) regime. It is estimated by equating leading [Eq. (69)] and the fluctuation correction [Eq. (72)] to conductivity times a factor R of order 1 at finite temperature (or sample size):

$$\sigma_0=\frac{e^2}{2\pi}4N_sq^2B'/\tau=R\,\delta\sigma$$
$$=\frac{e^2}{2\pi}\frac{q^2B_4'}{\pi\tau B'}\ln\left[\left(\frac{2m^*}{\tau}\right)\Big/\left(\frac{2\pi T_{wl}m^*\tau^{3/2}}{2qB'}\right)\right], \quad (74)$$

where we used the large disorder value in Eq. (73). Therefore

$$T_{wl}=\frac{2qB'}{\pi\tau^{5/2}}e^{-\mathrm{s}},$$

with a dominant argument of the exponential being

$$\mathrm{s}=\frac{4\pi N_s(B')^2}{RB_4'}.$$

Considering first the bare diffusion constant as fixed one observes that as r_s increases the temperature first rises due to preexponential factor, but then after reaching a maximum exponentially drops at large r_s. On the other hand fixing r_s

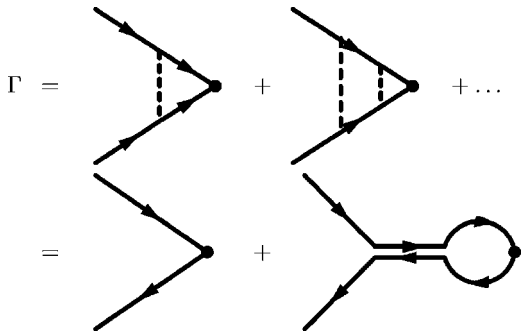

FIG. 11. Vertex correction due to interactions with disorder (dashed lines) in the conventional notations (Ref. 18) (first line) and in the present paper notations (second line).

the temperature quickly drops as λ increases. In experiment what is usually varied is density of electrons. In this case as density gets lower both the diffusion constant becomes smaller and r_s becomes larger. The overall effect is that for clean samples and relatively large r_s the quasimetallic state is stable to very long temperatures due to reduction of the DOS at the Fermi level. Note that the trajectory in the λ, r_s space of the experimental setup is itself dependent on the DOS.[40] This complicates the actual comparison since effects of screening cannot be neglected in experiments to date, as was discussed in Sec. II. Qualitatively however the picture is that at large coupling the metallic state survives effect of scattering off impurities due to the reduction in the DOS.

D. Higher-order effects: Aronov-Altshuler effect revisited

1. Higher-order corrections to the vertex function and conductivity

It was shown in Ref. 37 that in perturbation theory when one sums up all the corrections to the vertex part (the $\psi\psi\rho$ condensate where ρ is the density field that couples to static photon) shown in Fig. 11, it becomes proportional to the diffusive pole $1/(\omega\tau + Dp^2)$. The same expressions are also shown in Fig. 11 in our notations as a sum of leading and the next to leading terms in the steepest descent expansion. In this picture however the diffuson propagator is considered in the noninteracting theory. The vertex part enters high-order diagrams creating logarithmically divergent corrections which strengthen (in the singlet sector) weak localization. The major diagrams involving the singular vertex part contributing to conductivity are given in Fig. 12 (some other contributions cancel, see Ref. 18). The physical interpretation of this phenomenon is that electrons scatter coherently on Friedel oscillations due to density fluctuations.[22]

Without the crucial pole factor the Aronov-Altshuler corrections do not diverge in the infrared. In the strong-coupling regime considered in this paper, due to mixing of diffusons with static photons, the vertex given in Fig. 13 has much softened small momentum asymptotic: $1/(\omega\tau + Dp^2)$. It can be easily seen that this softening is quite enough to render the contributions such as those in Fig. 12 (which is of the order $1/N_s$, namely higher than the weak-localization one) finite. One therefore would ask how this can be understood diagrammatically in terms of conventional disorder coupling.[6] The point is that the mixing effectively sums up diagrams to all orders in Coulomb coupling r_s, Fig. 13. Each one of these is divergent, while their sum is not. This is quite analogous to the disappearance of IR divergencies due to long-range photon ''chains'' after the RPA diagrams are summed.

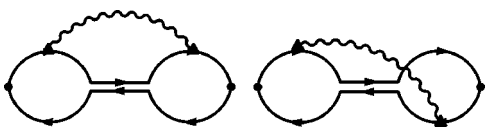

FIG. 12. Major contributions to the conductivity in perturbation theory.

2. Density of states near the Fermi-level at large coupling

The Aronov-Altshuler corrections to conductivity are directly related to the downturn cusp in DOS due to Coulomb interaction. In 2D the cusp is given by[18]

$$\delta N(\varepsilon) \propto \ln[\varepsilon]$$

again due to the renormalization of the vertex, see Fig. 11. It was claimed that this is precisely what was observed in tunneling junction experiments in disordered metal films.[41,42] In our approach, due to Anderson-Higgs mechanism, this renormalization is greatly reduced. An alternative explanation at very strong coupling and significant disorder might be the leading-order reduction of DOS discussed in Sec. II D, see Fig. 6. This does not contradict the experiments in metals since in these experiments disorder is large (even very large), while the coupling r_s is quite small. Of course if the density is sufficiently high the screening can no longer be neglected and the Aronov-Altshuler effect becomes dominant. However, as we mentioned before, in very clean 2DEG samples the disorder can reduce the screening and our approach of neglecting the screening at the leading order becomes more appropriate. In this case higher orders will not be large enough to undermine this assumption. One will get again a reduction in the DOS, but for entirely different reason. A related issue is emergence of the Coulomb gap commented on in the following section.

V. SUMMARY AND DISCUSSION

To summarize we present a consistent gauge-invariant approach to disordered strongly interacting electron gas in 2D. Physically the basic phenomenon is the reduction in the DOS at Fermi level due to strong Coulomb repulsion. This in turn suppresses both screening and scattering of impurities stabilizing the metallic state against weak-localization effects. Formally the approach consists of two steps. The first is variational (or ''self-consistent''): the most general quadratic states were considered and one with minimal energy identified. At this stage the ''RPA'' screening is neglected assuming it is sufficiently weakened by disorder so that higher orders are small. The second step is steepest descent perturbative expansion (which also can be identified as an expansion in parameter $1/N_s$ with N_s number of spin components

or valleys). Although the general philosophy of the steepest descent expansion is not drastically different from the one adopted in other works (see for example Ref. 8), two rather independent observations were made. The first is that the exchange part of Coulomb interaction leads at strong coupling to a significant reduction of the DOS near the Fermi surface and to via the reduction suppresses the disorder effects. The second is that when the steepest descent expansion procedure is followed consistently, mixing between static photons and diffusons not only causes Debye screening of the photon, but leads in addition to a softening of the diffusion pole. This in turn leads to a number of observable consequences such as a significant modification of the vertex part and consequently of the Aronov-Altshuler contribution to conductivity. The contribution becomes regular in the strong coupling, not logarithmic [the absence of the effect of the DOS reduction due to interaction in Ref. 17 can be traced to an approximate calculation of the exchange diagram in their Eq. (16)].

In this section we discuss several general questions and assumptions and relation of our work to other attempts to incorporate the long-range Coulomb interactions into the theory of disordered electron gas.

The theory of disordered electron gas relies to a large extent on the existence of massless collective modes, diffusons and Cooperons. It is tempting to interpret these excitations as Goldstone bosons of some symmetry breaking (with several complications arising from the quenched disorder, see Ref. 43). The σ-model approach initiated by Wegner[7] and others[14,13] long ago and developed and applied to the Coulomb interaction case recently by Baranov *et al.*[16] starts from an assumption that the $G=\mathrm{Sp}(2N)\otimes\mathrm{Sp}(2N)$ symmetry of free disordered electron gas is spontaneously broken down to diagonal subgroup $H=\mathrm{Sp}(2N)$, where N enumerates replica, spin, and Matsubara indices. We have shown in Sec. III however that diffusons mix with photon and become ''harder'' than standard Goldstone bosons. This might signal that the σ-model approach should be modified to incorporate the Anderson-Higgs mechanism. Actually the need for such a modification can be found in recent remarkable work of Ref. 16 and we comment on this now.

Unfortunately the presence of strong Coulomb interactions explicitly breaks a subgroup of G. An example of the explicitly broken-symmetry transformation is $\delta\psi_{pn}=\overline{\psi}_{-p,n}$, $\delta\overline{\psi}_{p,n}=-\psi_{-p,-n}$ for positive Matsubara frequencies $n>0$ and $\delta\psi_{pn}=\overline{\psi}_{-p,-n}$, $\delta\overline{\psi}_{p,n}=-\psi_{-p,-n}$ for negative Matsubara frequencies $n<0$. The symmetry is broken by both the frequency term $\Sigma_{p,n}\overline{\psi}^a_{pn}(-i\omega_n)\psi^a_{pn}$ and by the Coulomb interaction term. However while the breaking by the frequency term is ''soft'' and insignificant, as far as static quantities such as dc conductivity are concerned, it was shown[16] that the Coulomb interaction effectively represented on the σ model level by the ''square of trace'' operator [Eq. (2.1) in Ref. 16] is relevant and cannot be reduced to a soft breaking. Baranov *et al.* notice that at large distances the diffusion is suppressed which coincides with our Eq. (52). At short distance scales the electrons are diffusive. We believe that the Anderson-Higgs mechanism can be treated approximately within the σ-model approach as long as the mixing is small. This is a well-known problem in quantum-field theory[44] under the name of ''gauged σ models.'' These issues might become clearer when the present approach is extended beyond 2D (say to $2+\varepsilon$). Work on this is in progress. A related issue is understanding the difference between diffusons and Cooperons. Within the σ model approach the $\mathrm{Sp}(2N)\otimes\mathrm{Sp}(2N)$ symmetry forces the Cooperon and the diffuson propagators to be the same. It is precisely an explicit (not spontaneous) symmetry breaking due to Coulomb interactions that leads to the hardening of the diffuson (mixing with photon), while leaving Cooperon intact possible.

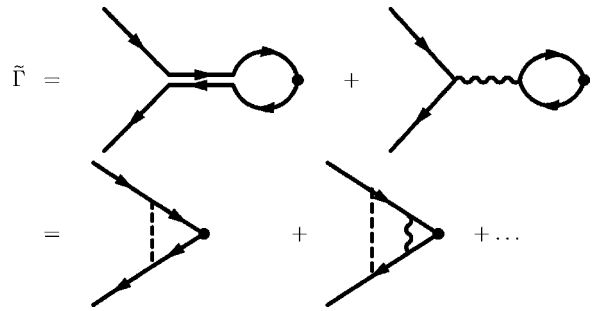

FIG. 13. Vertex corrections due to mixing of diffusons with static photons in the present paper notations (first line) and in the conventional notations (Ref. 6) (second line).

The fact that Coulomb interaction modifies diffusion at large distances was also discussed in Ref. 38 and this might lead also to suppression of weak localization. The work however was criticized,[45,16,22] that the vertex parts were not taken into account or alternatively the treatment is not gauge invariant. Our work explicitly shows that despite the fact that diffuson is harder (although still massless) at large distances, the Cooperon is not. Therefore although weak localization is suppressed, the suppression is much weaker and completely different. The logarithmically divergent contribution to conductivity includes Cooperon.

It was shown by Efros and Shklovskii[46] on the basis of a heuristic argument with plausible assumptions about the nature of the localized electronic states (neglecting their overlaps) that there should be a Coulomb gap in the strongly interacting electron gas. As in the case of σ models with Coulomb interactions,[16] it is not clear from the first orders in our scheme whether the reduction of the density of states is along the line of their argument. We believe this is unlikely due to the fact that they neglect the effect of exchange on the ''states'' ψ_i defined there. It is also not clear at this point whether the opening of Coulomb gap that Baranov *et al.*[16] deduce on the basis of the $2+\varepsilon$ expansion is related to the reduction in density of states due to exchange.

One can extend the approach presented here to the ''self-consistent'' scheme initiated by Vollhardt and developed to include Coulomb interactions by Sadovsky.[39] This will allow quantitative study of the insulating state and of the Coulomb gap. Note however that the ''gap'' equations of Ref. 39 follow the perturbative Aronov-Altshuler contributions, while the self-consistent form of our conductivity contributions,

Eqs. (69) and (72), will contain different diagrams. The work on this is in progress.

At last we briefly comment on two general assumptions made. The first is the spatial homogeneity. It is clear[47] that clean and even disordered 2DEG (Ref. 48) at sufficiently strong coupling become inhomogeneous Wigner crystals or ''glass.'' It was even speculated[29] that the Wigner crystallization (which occurs around $r_s=40$ in clean systems) might be related to the observed metal-insulator transition. Following general argument can be advanced against such a scenario. It has been observed recently that in several clean systems of thermally fluctuating repelling objects the homogeneous state (liquid or gas) exists down to zero temperature. One such system is the one-component classical plasma.[49] Another is a system of vortex lines in type-II superconductors.[50] The latter is quite analogous to 2DEG. The difference is that thermal fluctuations should be replaced by quantum and bosonic field by fermionic (statistics is quite unimportant in the low-density limit though). To be sure the energy of the solid is lower, so below the melting point the liquid state is metastable (in conventional liquids for which in addition to repulsive interaction there is a long-range attractive force, the metastable state ceases to exist at spinodal point). It is reasonable to assume (and it was demonstrated recently[51]) that disorder favors homogeneous state over a structured crystal). Therefore transition to a Wigner crystal or glass state would occur at much higher couplings than metal-insulator transition and the relevant state is homogeneous as was assumed in the present paper.

Another assumption commonly made is that the replica symmetry used to derive our starting point, Eq. (5), was assumed to be unbroken. This means that we neglected a possibility of ''electron glass.''[30] This is a distant possibility in the quasimetallic state since, as we argued in the paper, reduction in the DOS due to long-range interactions makes disorder less favored. Eventually in the insulating state glassy behaviors will eventually prevail.

ACKNOWLEDGMENTS

We are grateful to our colleagues in Hsinchu, Professor T. K Lee, Professor A. Voskoboynikov, Professor H. H. Lin, Professor S. Y. Hsu, Professor C. Y. Mou, and Professor V. Gudmundsson and especially Professor J. J. Lin for numerous conversations and encouragement. One of us (B.R.) thanks Professor S. V. Kravchenko, Professor M. P. Sarachik, Professor A. Gold, and Professor Q. Si for sharing their insight during their visits to NCTS. He also thanks especially E. Kogan for important clarification of details of Altshuler-Aronov work and other members of the Theoretical Physics Group of Bar Ilan University, where part of this work has been done, for encouragement. The work was supported by the NSC of Taiwan Grant No. 91-2811-M-009-006 and NCTU-Bar Ilan cooperation grant. One of us (T.M.-T.) also acknowledges the support of the Vietnam National Program of Basic Research on Natural Science, Project No 4.1.2.

APPENDIX: APPROXIMATE SOLUTION OF THE MINIMIZATION EQUATIONS

1. Expansion in small r_s

From Eq. (31) we observe that at small coupling $r_s \ll 1$, $\hat{e}_\varepsilon$ starts from the first order: $e_\varepsilon = \mathrm{sgn}[\varepsilon](r_s e_\varepsilon^{(1)} + \cdots)$. Substituting this into Eq. (30) immediately gives the leading term for q_ω:

$$\hat{q}_\omega^{(0)} = \mathrm{sgn}[\omega]\frac{1}{4\lambda} = \mathrm{sgn}[\omega]\hat{q}^{(0)}. \tag{A1}$$

Therefore we expand $\hat{q}_\omega = \mathrm{sgn}[\omega](\hat{q}^{(0)} + r_s \hat{q}_\omega^{(1)} + \cdots)$. Then the leading-order contribution to e_ε can be computed:

$$e_\varepsilon^{(1)} = \frac{\sqrt{2}}{\pi^2}\int_{\omega>0,\varepsilon'} \kappa[\varepsilon-\varepsilon']\frac{\varepsilon' + e_{\varepsilon'}}{(\varepsilon' + e_{\varepsilon'})^2 + (\hat{\omega}+\hat{q}^{(0)})^2}$$
$$= \frac{\sqrt{2}}{\pi^2}\int_{\varepsilon'} \kappa[\varepsilon-\varepsilon']\mathrm{sgn}(\varepsilon')\left(\frac{\pi}{2} - \arctan\frac{\hat{q}^{(0)}}{|\varepsilon'|}\right). \tag{A2}$$

We use here the constant $\hat{q}_\omega$ approximation. Substituting this into Eq. (30) one obtains

$$\hat{q}^{(1)} = -\frac{3\sqrt{2}(\hat{q}^{(0)})^2}{2\pi^2}\int_\varepsilon \frac{\kappa[\varepsilon]}{(\varepsilon^2 + 4\hat{q}^{(0)2})}$$
$$= -\frac{3}{2^{15/2}\pi^3\lambda^2} G_{33}^{22}((4\lambda)^2|_{1,1}^{3/2,3/2}), \tag{A3}$$

where the Meijer function $G_{pq}^{mn}(z|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p})$ is defined in Ref. 31. We observe that it is negative, namely the long-range interaction reduced the effect of disorder. Diagrammatically the minimization equations sum all the rainbows including both disorder and interaction, Fig. 2. In perturbation theory one evaluated a diagram with one photonic rainbow and arbitrary number of disorder lines, namely diagram in Fig. 1 with disordered Green's function. Its imaginary part is precisely $\hat{q}^{(1)}$. The actual expansion parameter is $\sqrt{2}r_s/\pi^2$ rather than r_s as can be seen from comparison of the perturbative and exact solutions. Therefore perturbation theory breaks down completely at $r_s \sim 10$.

2. Expansion in small $1/\lambda$

It is important also to obtain analytical solutions of the minimization equations at large r_s. This turns out to be possible for relatively clean case in which we can expand in $1/\lambda$. The leading order is the clean solution already discussed in Sec. II B: $e_\varepsilon^{[0]} = e_\varepsilon^{[0]} + e_\varepsilon^{[1]}/\lambda + \cdots$ with $e_\varepsilon^{[0]} = \sqrt{2}r_s\kappa_1[\varepsilon]/\pi$. The leading term for $\hat{q}_\omega = \hat{q}_\omega^{[1]}/\lambda + \hat{q}_\omega^{[2]}/\lambda^2 + \cdots$ is

$$\hat{q}_\omega^{[1]} = \frac{1}{4\pi}\int_\varepsilon \frac{\hat{\omega}}{(\varepsilon + e_\varepsilon^{[0]})^2 + \hat{\omega}^2}. \tag{A4}$$

The correction to dispersion relation can also be computed in a quite compact form:

$$e_{\varepsilon}^{[1]} = -\frac{r_s}{2\sqrt{2}\pi^2}\int_{\varepsilon',\varepsilon''>0}\frac{\kappa[\varepsilon-\varepsilon']-\kappa[\varepsilon+\varepsilon']}{(\varepsilon'+e_{\varepsilon'}^{[0]}+\varepsilon''+e_{\varepsilon''}^{[0]})^2}. \quad \text{(A5)}$$

The actual expansion parameter is $1/4\pi\lambda$ rather than λ as can be seen from comparison of the perturbative and exact solutions. Therefore perturbation theory breaks at $\lambda \sim 0.1$. The results are marked by dotted lines in Figs. 4 and 5. Generally numerical results agree with both perturbative expansions.

*Corresponding author. Email address: baruchro@hotmail.com

[1] S.V. Kravchenko *et al.*, Phys. Rev. B **50**, 8039 (1994); **51**, 7038 (1995).

[2] E. Abrahams, S.V. Kravchenko, and M.P. Sarachik, Rev. Mod. Phys. **73**, 251 (2001).

[3] E. Abrahams, P.W. Anderson, D.C. Licciardello, and T.V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979); P.A. Lee and T.V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).

[4] E. Abrahams and G. Kotliar, Science **274**, 1893 (1996).

[5] G.D. Mahan, *Many-Particle Physics*, 3rd ed. (Kluwer Academic, Dordrecht, 2002).

[6] A.A. Abrikosov, L.P. Gor'kov, and I. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963).

[7] F. Wegner, Z. Phys. B **25**, 327 (1976); Nucl. Phys. B **280**, 193 (1987).

[8] D. Belitz and T.R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).

[9] J.W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, Reading, MA, 1988).

[10] D. Vollhardt and P. Wolfle, Phys. Rev. Lett. **48**, 699 (1982); Phys. Rev. B **22**, 4666 (1980).

[11] T.R. Kirkpatrick and D. Belitz, Phys. Rev. B **50**, 8272 (1994).

[12] A.M. Finkelstein, Zh. Eksp. Teor. Phys. **84**, 168 (1983) [Sov. Phys. JETP **57**, 97 (1983)]; Z. Phys. **56**, 189 (1984).

[13] A.M.M. Pruisken, Nucl. Phys. B **235**, 277 (1984).

[14] S. Hikami, Phys. Rev. B **24**, 2671 (1981).

[15] A.M.M. Pruisken, M.A. Baranov, and B. Skoric, Phys. Rev. B **60**, 16 807 (1999); M.A. Baranov, A.M.M. Pruisken, and B. Skoric, *ibid.* **60**, 16 821 (1999); A.M.M. Pruisken, B. Skoric, and M.A. Baranov, *ibid.* **60**, 16 838 (1999).

[16] M.A. Baranov, I.S. Burmistrov, and A.M.M. Pruisken, Phys. Rev. B **66**, 075317 (2002).

[17] B.L. Altshuler and A.G. Aronov, Sov. Phys. JETP **50**, 968 (1979).

[18] B.L. Altshuler, A.G. Aronov, and P.A. Lee, Phys. Rev. Lett. **44**, 1288 (1980); B.L. Altshuler and A.G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, edited by A.L. Efros and M. Pollak (Elsevier Science, New York, 1985); B.L. Altshuler, A.G. Aronov, D.E. Khmelnitskii, and A.I. Larkin, in *Quantum Theory of Solids*, edited by I.M. Lifshitz (Mir, Moscow, 1982).

[19] P.T. Coleridge, A.S. Sachrajda, and P. Zawadzki, Phys. Rev. B **65**, 125328 (2002).

[20] S.S. Safonov *et al.*, Phys. Rev. Lett. **86**, 272 (2001).

[21] A. Gold and V.T. Dolgopolov, J. Phys. C **18**, L463 (1985); Phys. Rev. B **33**, 1076 (1986).

[22] G. Zala, B.N. Narozhny, and I.L. Aleiner, Phys. Rev. B **64**, 201201(R) (2001); *ibid.* **64**, 214204 (2001).

[23] Z. Wilamowski *et al.*, Phys. Rev. Lett. **87**, 026401 (2001).

[24] G.F. Giuliani and J.J. Quinn, Phys. Rev. B **29**, 2321 (1984).

[25] Y.-R. Jang and B.I. Min, Phys. Rev. B **48**, 1914 (1993).

[26] C.S. Ting, T.K. Lee, and J.J. Quinn, Phys. Rev. Lett. **34**, 870 (1975); T.K. Lee, C.S. Ting, and J.J. Quinn, *ibid.* **35**, 1048 (1975).

[27] S.A. Vitkalov, M.P. Sarachik, and T.M. Klapwijk, Phys. Rev. B **65**, 201106(R) (2002); V.M. Pudalov *et al.*, Phys. Rev. Lett. **88**, 196404 (2002); A.A. Shashkin, S.V. Kravchenko, V.T. Dolgopolov, and T.M. Klapwijk, Phys. Rev. B **66**, 073303 (2002).

[28] P.W. Anderson, *Basic Notions of Condensed Matter Physics* (Benjamin/Cummings, Menlo Park, CA, 1984).

[29] S. Chakravarty, L. Yin, and E. Abrahams, Phys. Rev. B **58**, R559 (1998).

[30] A.A. Pastor and V. Dobrosavljevic, Phys. Rev. Lett. **83**, 4642 (1999).

[31] I.S. Gradstein and I.M. Ryzhik, *Table of Integrals, Series and Products* (Academic Press, San Diego, 1994).

[32] C.M. Varma, Z. Nussinov, and W. van Saarloos, Phys. Rep. **361**, 267 (2002).

[33] S. Yarlagadda and G.F. Giuliani, Phys. Rev. B **40**, 5432 (1989).

[34] Y. Kwon, D.M. Ceperley, and R.M. Martin, Phys. Rev. B **50**, 1684 (1994).

[35] A.R. Hamilton *et al.*, Phys. Rev. Lett. **82**, 1542 (1999); M.Y. Simmons *et al.*, *ibid.* **84**, 2489 (2000).

[36] V.M. Pudalov *et al.*, Phys. Rev. B **60**, R2154 (1999).

[37] C. Castellani, C. Di Castro, P.A. Lee, and M. Ma, Phys. Rev. B **30**, 527 (1984).

[38] Q. Si and C.M. Varma, Phys. Rev. Lett. **81**, 4951 (1998).

[39] M.V. Sadovskii, Phys. Rep. **282**, 226 (1997).

[40] J.H. Davis, *The Physics of Low-Dimensional Semiconductors* (Cambridge University Press, Cambridge, 1997).

[41] W.L. McMillan, and J. Mochel, Phys. Rev. Lett. **46**, 556 (1981); Y. Imry and Z. Ovadyahu, Phys. Rev. Lett. **49**, 841 (1982).

[42] V.Y. But'ko, J.F. DiTusa, and P.W. Adams, Phys. Rev. Lett. **84**, 1543 (2000).

[43] A.J. McKane and M. Stone, Ann. Phys. (San Diego) **131**, 36 (1981).

[44] M. Bando, T. Kugo, and K. Yamawaki, Phys. Rep. **164**, 217 (1998).

[45] P. Schwab and C. Castellani, Phys. Rev. Lett. **84**, 4779 (2000).

[46] A.L. Efros and B.I. Shklovskii, J. Phys. C **8**, L49 (1975).

[47] B. Tanatar and D.M. Ceperley, Phys. Rev. B **39**, 5005 (1989).

[48] R. Asgari and B. Tanatar, Phys. Rev. B **65**, 085311 (2002).

[49] R.J.F. Leote de Carvalho, R. Evans, and Y. Rosenfeld, Phys. Rev. E **59**, 1435 (1999).

[50] D. Li and B. Rosenstein, Phys. Rev. B **65**, 220504 (2002).

[51] D. Li and B. Rosenstein, Phys. Rev. Lett. **90**, 167004 (2003).