PHYSICAL REVIEW B VOLUME 57, NUMBER 16 15 APRIL 1998-II

Metallic phase and metal-insulator transition in two-dimensional electronic systems

C. Castellani and C. Di Castro
Istituto di Fisica della Materia e Dipartimento di Fisica, Università di Roma ''La Sapienza,'' Piazzale Aldo Moro 2, 00185 Roma, Italy

P. A. Lee
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 5 January 1998)

The recent experimental observation of a metal-insulator transition in two dimensions prompts a reexamination of the theory of disordered interacting systems. We argue that the existing theory permits the existence of a metallic phase and propose a number of experiments such as magnetoconductance and tunneling in the presence of a parallel field, which should provide diagnostic tests as to whether a given experimental system is in fact in this regime. We also comment on a generic flow diagram which predicts a maximum metallic resistivity. [S0163-1829(98)52616-1]

The discovery by Kravchenko and co-workers[1,2] of a metal-insulator transition (MIT) in a two-dimensional (2D) system (Si-MOSFET), where MOSFET represents metal-oxide-semiconductor field-effect transistor, and its confirmation by other workers using different device designs[3] and materials[4,5] have generated much excitement because the conventional wisdom has been that all states are localized in two dimensions. Up to now the discussion of this phenomenon has been mainly based on the one parameter scaling theory of localization of noninteracting particles,[6] even though the relevance of the interaction has been stressed within a phenomenological approach to scaling.[7] The possibility of unusual superconductivity[8] or spin-orbit scattering[9] has also been raised. On the other hand, within the scaling theory which includes the combined effect of interaction and disorder[10,11] a 2D disordered system may remain metallic even in the limit of zero temperature.[12] In two dimensions the expansion parameter is the dimensionless resistance per square $R_\Box$ defined as $g=(e^2/\pi h)R_\Box$. For weak disorder ($g \ll 1$) the scaling is towards a metallic state ($dR_\Box/dT>0$).[11,12] Furthermore, the theory predicts that a magnetic field, via the Zeeman splitting, will drive the system towards an insulating state.[11,13] This is in agreement with experiment.[2] It is, therefore, useful to revisit this theory in light of the recent experimental development. One reason why the theory has not received general acceptance is that the scaling equations have the peculiar feature that the scaling variables diverge at some finite value of the length scale and the theory becomes uncontrolled. While this is certainly true in the vicinity of the MIT where $g \approx 1$, in this paper we reconsider the problem of 2D metallic behavior and argue that for weak disorder the theory remains under control over a large temperature range, provided the renormalization of the energy scale (relative to the length scale) is taken into account. In fact, this renormalization allows the possibility of a metallic state with finite resistance in two dimensions, in contrast to the scaling theory of localization, which permits only an insulator or a perfect metal ground state. We then study the magnetoresistance and tunneling density of states in the presence of a magnetic field, and point out that these are excellent diagnostic tools to extract key parameters and to test the applicability of the theory. At the end we shall discuss the MIT within the context of our theory of the metallic phase and comment on the effects of various symmetry breaking perturbations on the scenario we are proposing. Our main goal is to stimulate experimentalists to further study the metallic state both in the systems which have been studied up to now and possibly in other promising materials which we will discuss.

We begin by summarizing the results of the scaling theory of interacting disordered systems.[10–15] In addition to the dimensionless resistance g, the theory is characterized by the coupling constants γ_2, γ_c, and Z which obey the following scaling equations:

$$\frac{dg}{dy} = g^2\left[1+1+3\left(1-\frac{1+\gamma_2}{\gamma_2}\ln(1+\gamma_2)\right)-\gamma_c\right], \quad (1)$$

$$\frac{d\gamma_2}{dy} = g\left[\frac{1}{2}(1+\gamma_2)^2+\gamma_c(1+3\gamma_2+2\gamma_2^2)\right], \quad (2)$$

$$\frac{dZ}{dy} = gZ\left(-\frac{1}{2}+\frac{3}{2}\gamma_2+\gamma_c\right), \quad (3)$$

$$\frac{d\gamma_c}{dy} = g\left(\frac{1}{2}+\frac{3}{2}\gamma_2+\frac{\gamma_c}{2}-\frac{3}{2}\gamma_2\gamma_c\right)-\gamma_c^2, \quad (4)$$

where $y=-\ln\lambda$ describes a rescaling of the length scale so that momenta in the range $\lambda k_0^2 < k^2 < k_0^2$ are integrated out, where $k_0 \approx (v_F\tau)^{-1}$ is the short distance cutoff with τ being the elastic scattering time. The parameter Z describes a rescaling of the energy scale, $Z\gamma_2$ is related to the scattering amplitude in the triplet particle-hole channel, while $Z\gamma_c$ is related to the singlet particle-particle (Cooper channel) amplitude. These parameters can be interpreted in the context of the Fermi liquid theory.[16,17] For example, the specific heat linear T coefficient is modified by Z, so that Z plays the role of m^*/m. The uniform magnetic susceptibility is given by $\chi_s/\chi_s^0 = Z(1+\gamma_2)$ so that γ_2 plays the role of the Landau

 © 1998 The American Physical Society

parameter $-A_a^0$. The key quantity in this theory is the diffusion propagation, which has a pole of the form $(Dq^2 - iZ\omega)^{-1}$, where D is related to the conductivity σ (which equals $R_\square^{-1}$ in 2D) by $\sigma = \nu_0 D$; ν_0 is the bare density of states. In the context of the Fermi liquid theory, the diffusion pole can be written in the form $(D_Q q^2 - i\omega)^{-1}$, where $D_Q = D/Z$ has the interpretation of the quasiparticle diffusion constant. Equations (1)–(3) are derived to linear order in g and in the Cooper amplitude γ_c but include all orders in the interaction amplitude γ_2. The exception is Eq. (4) for γ_c where the last term is quadratic in γ_c and independent of g. This term renormalizes γ_c downwards, so that for $\gamma_c > 0$, γ_c becomes less important with scaling and can be neglegted for much of our subsequent discussions. The term $1+1$ in Eq. (1) is written in a way to remind us that weak localization and the singlet particle-hole channel in the case of Coulomb interaction give equal contributions to the enhancement of resistivity upon scaling. The next term is the contribution from the triplet particle-hole amplitude which has the opposite effect of reducing resistivity. According to Eqs. (2) and (3) both γ_2 and Z grow upon scaling. In fact, the growth is so rapid that they diverge at a finite scale y_0, so that near y_0 they behave as $\gamma_2 \sim (y_0 - y)^{-1}$ and $Z = (y_0 - y)^{-3}$. This divergence signals the breakdown of the perturbative scaling equations. Here we want to make two important points: (1) the divergence of Z is, in fact, a necessary condition for the existence of a metallic state in two dimensions; and (2) due to the rapid growth of Z there is a wide range of temperature where the scaling equations are valid and the system behaves like a metal. The key point is that the growth of Z forces us to perform scaling in an anisotropic manner in k space and energy space, a familiar situation in dynamical scaling. As we mentioned earlier, the key quantity is the diffusion pole $(Dq^2 - iZ\omega)^{-1}$. The scaling procedure then consists of integrating out the following regions in momentum space and energy space:[13]

$$\lambda k_0^2 < k^2 < k_0^2; \quad \lambda k_0^2 < \frac{Z}{D}\omega < k_0^2.$$

For Z growing with scaling, the energy or temperature scale decreases rapidly with scaling, and is given by

$$T = \lambda D k_0^2 / Z(\lambda). \tag{5}$$

Strictly speaking, this formula needs further correction when $Z_2 = Z(1+\gamma_2)$ becomes much greater than Z, because the energy denominator $(Dq^2 - iZ_2\omega)$ also appears in some intermediate steps. However, the qualitative point that the temperature scale can go all the way to zero remains valid. This is important because in one parameter scaling, the point has been made that the theory scales to either an insulator or a perfect metal ($R_\square \to 0$) in two dimensions, because the β function is always nonzero.[7] The diverging Z at $y = y_0$ allows us to escape from this conclusion because, in principle, one can reach the point $y = y_0$ with g finite, so that according to Eq. (5) the system maintains a finite $R_\square$ as $T \to 0$.

The next question is whether a metallic state can be realized in a region of parametric space and temperature where Eqs. (1)–(3) are valid. From Eqs. (2) and (3) it is apparent that the effective expansion parameter in the theory is $g\gamma_2$. Then by starting with a sufficiently small g, it is possible to integrate Eqs. (1)–(3) until $g\gamma_2$ becomes of order unity. Since Z diverges as $(y_0 - y)^{-3}$, much faster than $\gamma_2 \sim (y_0 - y)^{-1}$, the scaling can proceed to a rather low temperature before $g\gamma_2 \approx 1$ and the perturbative equations break down. By making the assumptions that g approaches a constant linearly in $(y - y_0)$ we conclude, using Eq. (5), that the low-temperature behavior of the resistivity is given by $R_\square(T) = R_0 + cT^{1/3}$ with $c > 0$. [Note that at very low temperature, when $g\gamma_2 \approx 1$, the assumption that γ_c is negligible is no longer valid and indeed γ_c approaches a fixed point value $\gamma_c^* = 1$ for $\gamma_2 \to \infty$. This would change the behavior of Z, leading to $Z \sim (y_0 - y)^{-3/5}$. This, in turn, modifies the temperature dependence of $R_\square = R_0 + c'T^{5/3}$ when the regime $\gamma_c \simeq 1$ is reached before getting out of the range of validity of Eqs. (1)–(4).] To summarize, for sufficiently small g, we expect that initially g will exhibit $\ln T$ correction over a broad temperature range. If γ_2 is sufficiently large to begin with, the $\ln T$ correction is metalliclike. If γ_2 starts out small, the ln correction resembles weak localization, but will change sign below a certain temperature scale when γ_2 has grown sufficiently to overwhelm the localization term and the singlet contribution in Eq. (1). At a still lower temperature the resistivity drops rapidly, perhaps as $T^{1/3}$ (and possibly crossing over to $T^{5/3}$) before the one loop scaling equations break down.[18] This qualitative behavior has been confirmed[20] by numerical integration of Eqs. (1)–(4). The point we wish to emphasize is that these equations predict a metallic behavior down to very low temperature in a region of parameter space where the one loop scaling equations remain reliable. Thus the existence of a metallic state over an experimentally accessible temperture range should not in itself be a great surprise.

We have seen that the key ingredient in arriving at a metallic state is the existence of a large γ_2. The question is whether γ_2 can be directly measured experimentally. We have mentioned that the uniform magnetic susceptibility provides a measurement of $Z(1+\gamma_2)$. However, this is a difficult, though not impossible, experiment in a two-dimensional electron gas.[21] Instead, we find that magnetoresistance and tunneling in the presence of a parallel field provide direct measurements of γ_2. A parallel field provides a Zeeman splitting of the spin states which cut off the $S_z = \pm 1$ parts of the triplet particle-hole channel as well as the $S_z = 0$ part of the triplet and singlet particle-particle channel. This gives rise to positive magnetoresistance. The contribution coming from the particle-hole channel was calculated in the weak-coupling limit in Ref. 22. This calculation was later extended to strong scattering amplitudes.[23] Here we further extend this calculation to include the effect of the energy renormalization Z. In analogy with the Fermi liquid theory, we expect the spin splitting of the quasiparticle to be given by $\tilde{\Omega}_s = (1+\gamma_2)\Omega_s$, where $\Omega_s = g_L \mu_B H$. Therefore the diffusion pole should be modified to $(D_Q q^2 - i\omega - i\tilde{\Omega}_s S_z)^{-1}$ for the $S_z = \pm 1$ components of the triplet particle-hole channel. Inserting this modification into the expression for the $S_z = \pm 1$ contribution to the conductivity, we find

$$\delta\sigma(T,H)=\frac{ie}{h}\int_{-\infty}^{\infty}d\omega\frac{d}{d\omega}\left(\omega\coth\frac{\omega}{2kT}\right)\int\frac{d^2k}{(2\pi)^2}D_Q^2k^2 \times\sum_{S_z=\pm1}\frac{1}{(D_Qk^2-i\omega-i\tilde{\Omega}_sS_z)^2} \times\frac{2\gamma_2}{D_Qk^2-i(1+\gamma_2)\omega-i\tilde{\Omega}_sS_z}. \quad (6)$$

The parameters D, Z, and γ_2 in this equation are scale dependent. Noting that the contributions for small H are dominated by small k and ω, we evaluate these parameters at the scale λ given by Eq. (5). The integrals are then performed following Ref. 22. In particular, we find that for small H,

$$\sigma(H,T)-\sigma(0,T)=-0.084\frac{e^2}{\pi h}\gamma_2(\gamma_2+1)\left(\frac{g_L\mu_BH}{kT}\right)^2. \quad (7)$$

We recover the weak-coupling limit by setting $\gamma_2\to F/2$ where $F\ll1$ is the interaction parameter in Ref. 22. If we include the Cooper channel contribution, we will find an additional contribution of $-0.084(e^2/\pi h)\,\gamma_c(\gamma_2+1)^2(g_L\mu_BH/kT)^2$. The above treats the effect of spin splitting only and is appropriate for H parallel to the plane. For perpendicular field we have, in addition to Eq. (7), the usual weak localization negative magnetoresistance. In this case there is an additional contribution proportional to γ_c but now the orbital field scale given by $\Omega_H=4DeH/c$ also enters as a cutoff and the magnetic-field dependence from this term is more complicated. Since in the weak-coupling regime we expect γ_c to scale to weak coupling, we shall concentrate on Eq. (7). The main point is that the quadratic in the H term in parallel field magnetoresistance provides a measurement of the parameter γ_2. It will be very interesting to see if this parameter is indeed large in the metallic MOSFET samples and whether it increases with decreasing temperature. The available data are not systematic enough to answer these questions in the metallic regime. Most of the experiments on magnetoresistance are close to the MIT and for fields with $\Omega_s\gtrsim kT$. Qualitatively, the (positive) magnetoresistance increases as one moves away from the MIT.[3] This is in agreement with our expectation that γ_2 should consistently increase in order to establish a metallic phase.

Another way to measure γ_2 is by tunneling experiment. It was pointed out that the tunneling density of states exhibits additional structure between the energy scales of the bare spin splitting $g_L\mu_BH$ and the enhanced spin splitting due to interaction effects.[23] Following the Fermi liquid analogy, this second energy scale should be given by $\tilde{\Omega}_s$. In particular, in two dimensions the derivative of the tunneling density of states has logarithmic singularities at $\omega=g_L\mu_BH$ and $\omega=(1+\gamma_2)g_L\mu_BH$. Thus tunneling gives a direct measurement of γ_2. Recently a new technique has been developed to tunnel into a 2D electron gas.[24] It will be very interesting to apply it to the new metallic samples.

As the field is increased, we expect a crossover to the strong Zeeman splitting universality class. The detailed crossover is complicated, but the high field limit is one of the few fixed points that is controlled. The system always scales to an insulator, and in the weak disorder limit, a universal logarithmic temperature dependence was predicted:[13] $\sigma(T)=\sigma_0+(e^2/\pi h)(2-2\ln2)\ln(T\tau)$. As far as we know, this prediction has never been tested. The new MOSFET samples offer an ideal testing ground for this prediction.

Up to now, we have limited our discussion to the weak disorder case, when Eqs. (1)–(4) remain valid. We now comment on the possibility of the existence of a nontrivial fixed point if somehow the scaling equations can be extended to strong coupling. In Ref. 19 the two loop contribution to the scaling equations was evaluated under the assumption of $\gamma_2\gg1$ but for small $g\gamma_2$. The two loop scaling equations of Ref. 19 indeed exhibit a nontrivial fixed point. From this fixed point two separatrices originate ending at $\gamma_2=0$ and $\gamma_2=\infty$. Since the interesting part of the flow diagram is not in the weak-coupling regime, the scaling equations and the details of the flow cannot be trusted. Nevertheless, the structure of the flow may be generic. Here we wish to make some general comments. If the initial γ_2 is not too large, the system exhibits a metal-to-insulator transition. An interesting feature of this flow is that on the metallic side of the separatrix the system reaches infinite γ_2 and Z at a finite scale λ as in one loop order. Thus the discussion we gave earlier in this paper still holds and a metallic state with finite $R_\square$ is possible at $T=0$. In fact, the metallic state in the low T limit exhibits a maximum metallic resistivity given by $\rho_M=(\pi h/e^2)g_M$, where g_M is the value of g on the separatrix at $\gamma_2=\infty$. This g_M is, in general, smaller than the value g^* at the fixed point. Experimentally $\rho^*=(\pi h/e^2)g^*$ is determined as the resistance which separates the metallic and insulating states at higher temperature. This feature seems to be consistent with currently available data. For example, the data of Ref. 1 yields $\rho_M\approx0.1(h/e^2)$ and $\rho^*\approx2(h/e^2)$.

The scaling behavior near the MIT will be controlled both by the existence of a fixed point at finite g^* and γ_2^* and by the runaway towards $g\simeq g_M$ and $\gamma_2=\infty$. Then one can show that $R_\square=\tilde{\rho}[T/(\delta n)^{\nu z}]$, where δn is the deviation from the critical density and the critical indices ν and z are determined by the fixed point. $\tilde{\rho}$ is a scaling function and according to the previous discussion $\tilde{\rho}(\infty)=(\pi h/e^2)g^*$ and $\tilde{\rho}(0)=(\pi h/e^2)g_M$.

Besides the magnetic field, other symmetry breaking perturbations have relevant effects on our picture of the 2D metallic phase. Spin flip scattering by magnetic impurities will cause a crossover to a low T insulating phase. The effect of spin-orbit (SO) scattering is more intriguing. In $d=2$, intrinsic SO coupling or SO scattering by impurities only affects the out of plane component of the spin.[25] In this case the one loop equations[26] still lead to a diverging behavior of the $(S_z=0)$ triplet amplitude and a metallic phase at low T. We suggest that the above discussion on the MIT applies in this case even though the 2D SO could result into a different universality class. A much more dramatic effect on our theory of the metallic phase is the SO scattering deriving from possible asymmetry of the confining potential since it is equivalent to a 3D SO coupling and cutoff all triplets.[27] If this coupling is sizable, the theory predicts an insulating behavior at zero temperature,[28] at least in the limit in which the

SO band splitting is less than the inverse elastic scattering time. In our opinion, evidences of 2D or 3D SO are still lacking.

The scenario we outlined in this paper has the advantage of permitting a metallic state in two dimensions and therefore a metal-insulator transition. However, given the uncertainties of the strong-coupling theory, a good strategy is to approach the MIT from the metallic side and try to gain a thorough understanding of the metallic state. This motivates us to propose magnetic susceptibility, magnetoresistance, and tunneling experiments as ways to directly measure the key parameters of the theory γ_2 and Z. We also worked out the qualitative behavior of the temperature dependence of the resistivity, in a regime where the theory is valid. Here our results do not compare favorably with experiments. The data of Refs. 1 and 4 have been fitted to the form $\rho(T)=\rho_0+\rho_1\exp(-T_0/T)$. This is very different from the $\ln T$ dependence followed by a low-temperature power law that we predict. Furthermore, the parameter T_0 appears to scale with the Fermi energy which is relatively small in these low density systems. Thus the possibility remains that some physics on the scale of the Fermi energy is playing the dominant role and the data are far from the low-energy scaling regime we considered here. We believe these questions can be addressed by more detailed studies of the metallic state along the lines suggested in this paper. Yet another possible research direction to confirm the theory here presented is to study 2D systems where γ_2 is expected to be large to begin with, such as almost ferromagnetic metallic thin films. Examples are weak ferromagnets such as MnSi or $TiBe_2$, if the ferromagnetism can be suppressed by alloying.[29,30]

We thank Michael Ma and Olav Syljuåsen for helpful discussions. P.A.L. acknowledges support of the NSF under Grant No. DMR−9523361.

[1] S. V. Kravchenko *et al.*, Phys. Rev. B **51**, 7038 (1995); Phys. Rev. Lett. **77**, 4938 (1996).
[2] D. Simonian, S. V. Kravchenko, and M. P. Sarachik, Phys. Rev. Lett. **79**, 2304 (1997).
[3] D. Popovic, A. B. Fowler, and S. Washburn, Phys. Rev. Lett. **79**, 1543 (1997).
[4] Y. Hanien *et al.*, cond-mat/9709814 (unpublished).
[5] M. Y. Simmons *et al.*, Phys. Rev. Lett. **80**, 1292 (1998).
[6] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[7] V. Dobrosavljevic, E. Abrahams, E. Miranda, and S. Chakravarty, Phys. Rev. Lett. **79**, 455 (1997).
[8] D. Belitz and T.R. Kirkpatrick, cond-mat/9795923 (unpublished); F.C. Zhang and T.M. Rice, cond-mat/9708050 (unpublished); P. Phillips *et al.*, cond-mat/9709168 (unpublished).
[9] V.M. Pudalov, cond-mat/9707076 (unpublished).
[10] A. M. Finkelstein, Zh. Eksp. Teor. Fiz. **84**, 168 (1983) [Sov. Phys. JETP **57**, 97 (1983)].
[11] A. M. Finkelstein, Z. Phys. A **56**, 189 (1984).
[12] C. Castellani *et al.*, Phys. Rev. B **30**, 1596 (1984).
[13] C. Castellani, C. Di Castro, P. A. Lee, and M. Ma, Phys. Rev. B **30**, 527 (1984).
[14] A. M. Finkelstein, Sov. Sci. Rev., Sect. A **14**, 1 (1990).
[15] For review, see P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 787 (1985).
[16] C. Castellani and C. Di Castro, Phys. Rev. B **34**, 5935 (1986).
[17] C. Castellani, G. Kotliar, and P. A. Lee, Phys. Rev. Lett. **59**, 323 (1987).
[18] It is worth noting that the two loop equations of Ref. 19 would lead to a low-temperature behavior $R_\square(T)-R_\square(0)\propto T^{1/2}\exp(-C/T^{1/2})$. However, since at the same time $g\gamma_2\gg 1$, there is no control on this result.
[19] T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **41**, 11 082 (1990); D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994), Sec. VI.
[20] O. Syljuåsen (unpublished).
[21] See S. E. Barrett *et al.*, Phys. Rev. Lett. **74**, 5112 (1995).
[22] P. A. Lee and T. V. Ramakrishnan, Phys. Rev. B **26**, 4009 (1982).
[23] R. Raimondi, C. Castellani, and C. Di Castro, Phys. Rev. B **42**, 4724 (1990).
[24] H. B. Chan, P. I. Glicofridis, R. C. Ashoori, and M. R. Melloch, Phys. Rev. Lett. **79**, 2867 (1997).
[25] P. A. Lee, J. Non-Cryst. Solids **35**, 21 (1980).
[26] A. M. Finkelstein, Solid State Commun. **54**, 611 (1984).
[27] See the discussion in Ref. 14, p. 69.
[28] C. Castellani *et al.*, Solid State Commun. **52**, 261 (1984).
[29] J. Smith, Physica B&C **107**, 251 (1981).
[30] D. M. Paul *et al.*, J. Magn. Magn. Mater. **42**, 201 (1984), and references therein.