Screening Breakdown on the Route toward the Metal-Insulator Transition in Modulation Doped Si/SiGe Quantum Wells

Z. Wilamowski,[1,2] N. Sandersfeld,[1] W. Jantsch,[1] D. Többen,[3] and F. Schäffler[1]

[1]*Institut für Halbleiter- und Festkörperphysik, Johannes Kepler Universität, A-4040 Linz, Austria*
[2]*Institute of Physics, Polish Academy of Sciences, Al Lotnikow 32/46, PL 0668 Warsaw, Poland*
[3]*Walter-Schottky-Institut, Technische Universität München, D-85748 Garching, Germany*

(Received 26 April 2000; published 20 June 2001)

Exploiting the spin resonance of a two-dimensional (2D) electron in SiGe/Si quantum wells, we determine the carrier density dependence of the magnetic susceptibility. Assuming weak interaction, we evaluate the density of states at the Fermi level, $D(E_F)$, and the screening wave vector, q_{TF}. Instead of the constant values of an ideal 2D system, we observe a gradual decrease towards the band edge. Calculating the mobility from q_{TF} yields good agreement with experimental values justifying the approach. The decrease in $D(E_F)$ is explained by potential fluctuations which lead to tail states that make screening less efficient and, in a positive feedback, cause an increase of the potential fluctuations.

DOI: 10.1103/PhysRevLett.87.026401 PACS numbers: 71.30.+h, 73.21.Fg, 73.40.–c

Experimental evidence for a low-temperature metallic state in a variety of 2D carrier systems [1–3] (2DCS) has newly highlighted the long-lasting discussions about the nature of the metal-to-insulator transition (MIT). Since scaling theory for noninteracting carrier systems predicts localization in 2D at $T = 0$ [4], recent experimental and theoretical efforts concentrate on interacting 2DCSs. The most important parameter, the carrier density n_s, determines the average carrier distance in units of the effective Bohr radius a_B (= 3.2 nm in Si) as $r_s = 1/(a_B\sqrt{\pi n_s})$, which is proportional [5] to the ratio between the e^--e^- Coulomb energy $E_{e\text{-}e}$ and the kinetic energy E_{kin}. The other important quantity is disorder which we characterize here by the amplitude of potential fluctuations, δV, and the dimensionless ratio, $r_d = e\delta V/E_{\mathrm{kin}}$. (Here e is the electron charge.) If disorder is negligible ($r_d \ll r_s$) collective localization (Wigner crystallization) is predicted [6] at around $r_s = 37$, whereas increasing r_d leads to carrier localization at lower r_s via the defect potential. Most of the experiments reported [6] employed disordered systems with r_s values around 10 at the MIT. Moreover, the experiments almost exclusively concentrated on transport, which allows the identification of an MIT, but does not reveal the mechanisms that cause it. It is therefore not surprising that the origin of the apparent metallic phase is under debate [6–8].

To gain additional experimental information, we propose a method that gives access to r_d and the screening properties of a 2DCS. It is based on conduction electron spin resonance (CESR) in strained Si quantum wells, and on the magnetic susceptibility χ_m of the 2DCS determined from CESR. Far enough on the metallic side a weakly interacting 2DCS can be assumed in a first approximation. In this limit, and at low temperatures, for $kT \ll E_F$, the 2D carrier gas exhibits Pauli paramagnetism, and thus χ_m is proportional to the density of states (DOS) at the Fermi level, $D(E_F)$ [9]:

$$\chi_m = g^2\mu_B^2 D(E_F)\,, \tag{1}$$

where g is the Landé factor and μ_B is Bohr's magneton. These experiments are in the tradition of earlier measurements on bulk Si that were based on the static [9] and the electron spin resonance (ESR) susceptibility [10], on the Knight shift [11], and on specific heat [12]. We obtain here qualitatively different results for a 2DCS due to the different DOS and smooth potential fluctuations.

For a weakly interacting electron gas, the Thomas-Fermi screening wave vector q_{TF} is also proportional to $D(E_F)$ [13]:

$$q_{\mathrm{TF}} = \frac{e^2}{2\epsilon\epsilon_0}\, D(E_F) \qquad (\text{for } kT \ll E_F) \tag{2}$$

with the permittivity $\epsilon\epsilon_0$ of the medium. In this limit, CESR thus allows also for an experimental evaluation of q_{TF}. Although this simple model is not expected to hold all the way to the MIT, we show that the carrier mobilities calculated from the experimental q_{TF} agree over a large r_s range surprisingly well with the measured mobilities justifying this approach *a posteriori*.

The experiments exploit the recently identified CESR signal of a 2D electron gas in the strained Si channel of a modulation-doped SiGe/Si/SiGe double heterostructure [14–16]. The 2DCS in this heterostructure has the same twofold valley degeneracy and effective transport mass as the inversion layers of the more widely studied metal-oxide-semiconductor field-effect transistor (MOSFETs) [1,3], but modulation doping allows for much higher mobilities [17]. Other important differences comprise the lack of oxide charges [6], the reduced interface roughness, and the presumably much lower density of interface charges at the crystalline Si/SiGe interfaces. The weak spin-orbit interaction and the lack of magnetic impurities lead to extremely long spin relaxation times of several μs [18,19], which provide exceptionally narrow ESR signals, and thus an unusually low detection limit of 3×10^9 spins/cm^2.

The ESR experiments were performed at temperatures down to 2 K in a standard X-band spectrometer (Bruker 200 SRC) with a TE102 cavity. The samples were grown by molecular beam epitaxy on 1000 Ω cm Si (001) substrates, which show complete carrier freeze-out below 30 K. A 20 nm thick Si channel with tensile in-plane strain was deposited on a strain-relaxed $Si_{0.75}Ge_{0.25}$ buffer layer, which consists of a 0.5 μm thick $Si_{0.75}Ge_{0.25}$ layer on top of a 2 μm thick $Si_{1-x}Ge_x$ layer with compositional grading [15,17]. The upper $Si_{0.75}Ge_{0.25}$ barrier was modulation doped with a 12.5 nm thick, nominally undoped spacer layer and capped with 5 nm of Si. Sample 1 has a low Sb doping concentration of 5×10^{17} cm^{-3} in a 30 nm wide supply layer. It shows strong persistent photoconductivity: After cooling in darkness to 2 K the carrier concentration n_s is less than 7×10^{10} cm^{-2}, but it increases with the dose of Si band gap illumination and saturates at $n_s = 3 \times 10^{11}$ cm^{-2}. Sample 2 from the same wafer had a palladium Schottky gate, and sample 3 was higher doped, leading to a constant $n_s = 7 \times 10^{11}$ cm^{-2}. Sample 3 has a mobility of about 200 000 cm^2/V s at 2 K, whereas the mobilities of the other samples are strongly n_s dependent (see Fig. 2 below).

All spectra were taken with the magnetic field perpendicular to the 2DCS which in this geometry yields a broad cyclotron resonance (CR) absorption and a very narrow CESR signal [15,19] (see inset in Fig. 1). In sample 1, n_s was adjusted by the illumination dose and monitored by fitting the CR signal shape [18]. In the gated sample 2, n_s was determined by gate-voltage-dependent Shubnikov–de Haas experiments.

Figure 1 shows the integral CESR absorption, and thus χ_m, normalized to its saturation value, which is the same for all samples, as a function of n_s. Good agreement is found between the gated and illumination-dosed samples. Following Eqs. (1) and (2), χ_m can be converted into an absolute value of q_{TF}, as given by the right-hand side ordinate in Fig. 1. The constant of proportionality was determined in the saturation range by assigning to it the constant value of the ideal 2DCS: $q_{\mathrm{TF}}^{\mathrm{2D}} = 2g_\nu/a_B = 1.27 \times 10^7$ cm^{-1}, with $g_\nu = 2$ in tensely strained Si. The striking feature in Fig. 1 is that q_{TF} reaches the intrinsic value of an ideal 2DCS only for $n_s > 3 \times 10^{11}$ cm^{-2}, whereas it decreases monotonously below that concentration and extrapolates to zero at a critical density of 7×10^{10} cm^{-2}. This compares well with the critical MIT carrier concentration of 9×10^{10} cm^{-2} reported for high-mobility Si MOSFETs [1].

The Thomas-Fermi approximation may become inadequate for describing transport close to the MIT, where r_s is 7 in our samples. It is therefore important to assess the n_s range over which q_{TF} derived from the CESR experiments is relevant for the transport properties. For that purpose we measured the Hall mobility of gated and illumination-dosed samples as a function of n_s (Fig. 2, solid squares). Below 4.5×10^{11} cm^{-2} the experimental

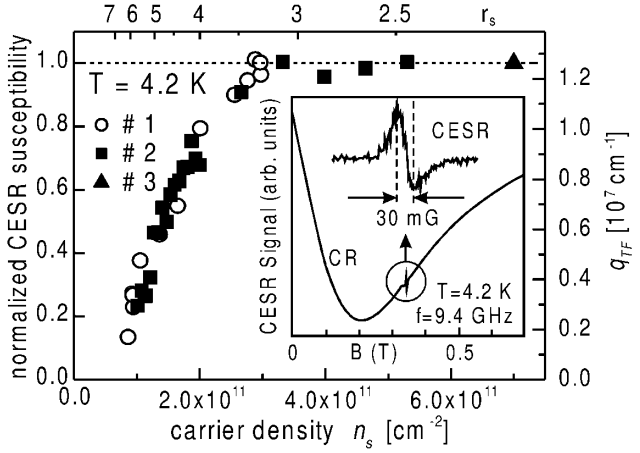

FIG. 1. Integral CESR absorption normalized to the saturation value that corresponds to the Pauli susceptibility of an ideal 2DCS. The right ordinate has been converted into the Thomas-Fermi wave vector q_{TF}. The inset shows a typical spectrum with B perpendicular to the 2DCS; CR labels the cyclotron resonance signal. The top scale gives r_s (see text).

mobility follows a power law in n_s with a fitted exponent of 2.34 (solid line). For comparison, we calculated the Thomas-Fermi transport scattering time [13] replacing the constant $q_{\mathrm{TF}}^{\mathrm{2D}}$ of the ideal 2D electron gas with the experimentally determined $q_{\mathrm{TF}}(n_s)$ from Fig. 1. For simplicity a Howard-Fang envelope function for the wave function perpendicular to the well was assumed [13]. The resulting mobilities are plotted as open symbols in Fig. 2.

The absolute values of the calculated and experimental mobilities differ by almost a factor of 2, but this is not too surprising, since only remote impurity scattering has been considered. Much more striking is the fact that the Thomas-Fermi mobility perfectly reproduces the experimentally found $n_s^{2.34}$ power law down to about 1.5×10^{11} cm^{-2}. For $n_s > 3 \times 10^{11}$ cm^{-2}, the calculated mobility bends over to the $n_s^{1.0}$ relation of an ideal 2DCS with constant

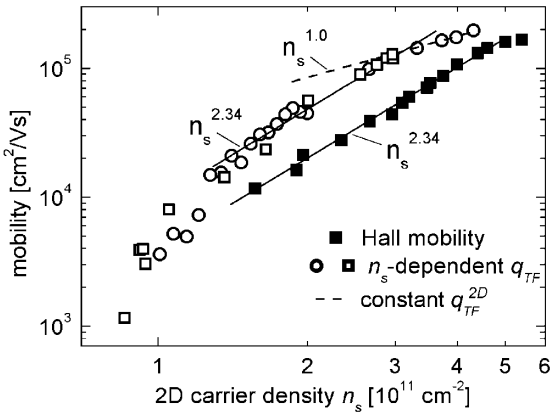

FIG. 2. Hall mobility measured at 0.36 K (solid symbols) and calculated mobility in Thomas-Fermi approximation (open symbols) versus n_s. The dashed line represents the ideal 2DCS. Solid lines are fits to the data and represent an $n_s^{2.34}$ power law.

$q_{TF}(E_F)$ (dashed line). The excellent agreement between the fitted power laws suggests that the n_s-dependent q_{TF} determined from CESR is meaningful for transport experiments down to about 1.5×10^{11} cm^{-2}.

In order to explain the observed decrease of χ_m and the associated breakdown in screening efficiency [$\chi_m \sim q_{TF}(E_F)$] as n_s approaches n_s^c we consider the effect of disorder. The latter introduces a low energy tail to the ideally sharp onset of the DOS which we model by a Gaussian distribution of the potential fluctuations that are superimposed on the band edge. The width of the distribution corresponds to the screened fluctuation amplitude δV. Since $D(E_F) \sim \chi_m$ is directly measured by the CESR signal according to Eq. (1), the two parameters δV and E_F can be determined as unambiguous functions of n_s by simultaneously fitting the experimental data of χ_m and n_s applying Fermi statistics (see Fig. 3, inset) [18].

Figures 3a and 3b show results of δV and E_F versus n_s, with E_F being measured relative to the band edge of the ideal 2D electron gas. For higher n_s, where screening is effective, $e\delta V$ remains smaller than E_F (i.e., $r_d < 1$). At lower n_s, δV increases drastically and finally diverges. Simultaneously, E_F becomes negative at about 1.5×10^{11} cm^{-2}; i.e., it drops below the ideal band edge and thus it enters the tail states. These results show that the drop of E_F, which was also observed in other experiments

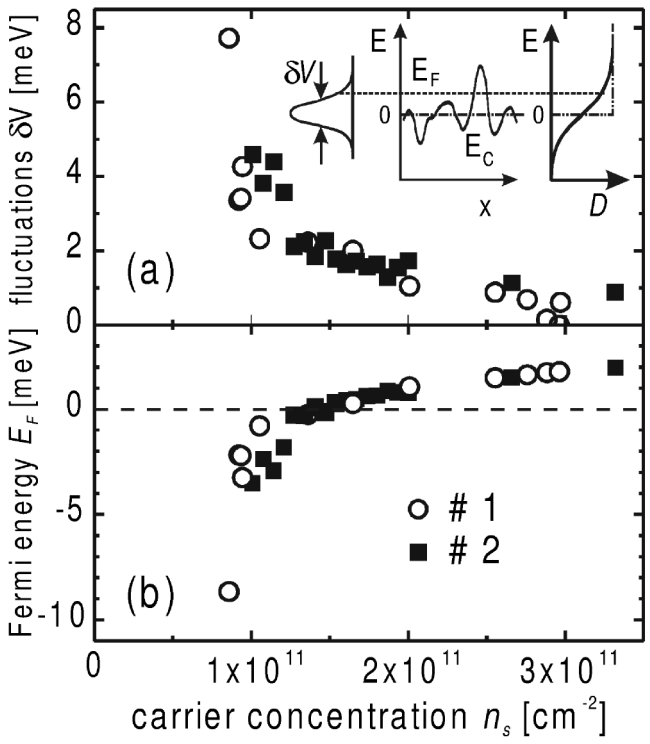

FIG. 3. Fluctuation amplitude δV (a), and Fermi energy E_F (b), as a function of the carrier density n_s for samples 1 and 2. The inset shows the Gaussian distribution of conduction band edge fluctuations (E_C) with respect to E_F (dotted line) and the resulting DOS (solid line) in comparison to the ideal DOS (dash-dotted line).

near the MIT [20,21], is caused by the increase of δV and the concomitant formation of tail states.

Close to n_s^c the calculated mobilities in Fig. 2 drop exponentially. We attribute this behavior to a positive feedback mechanism that is driven by the loss of screening and the concomitant increase of the potential fluctuations. Such mechanisms have been proposed theoretically [7,22,23], and our combination of susceptibility and transport measurements provides strong experimental evidence for the relevance of such an inherently nonlinear mechanism. In the nonlinear regime $e\delta V$ exceeds E_F (i.e., $r_d < 1$), which means that the 2DCS becomes localized.

The data presented so far do not provide the characteristic length scales l_V of the potential variations and l_c for carrier localization. For the latter a lower limit can be extracted from the following observation: The CESR linewidth at the lowest carrier concentrations investigated is about 30 mG [15,19]. This implies a lower limit for l_c, because the minimum CESR linewidth is given by the hyperfine interaction with nuclear spins in the probing area. With the relative abundance of 4.7% of ^{29}Si (the only stable Si isotope with nuclear spin) and the hyperfine constant [24] of P in Si, we estimate a minimum l_c on the order of 1 μm. The localization behavior of the system is then determined by the relation between l_V and l_c. For $l_V \ll l_c$ the 2DCS is expected to show Anderson localization, whereas $l_V \approx l_c$ would cause puddles of mobile carriers confined to shallow long-range potential minima [21,25].

We cannot clearly distinguish between these pictures on the base of our experiments. However, two general remarks can be made regarding the predominant sources of the potential fluctuations in metal-oxide-semiconductor- and modulation-doped structures. The former are dominated by interface charges, which lead to large angle scattering and Dingle ratios around 1. The density of scatterers, however, is typically an order of magnitude lower than n_s. It can therefore be expected that Anderson localization will occur in such a system [26]. In contrast, the donors in modulation-doped structures are separated from the channel by $\geq$10 nm, which leads to much smoother potential fluctuations. These cause small angle scattering, with Dingle ratios around 15 in our samples [17]. However, because of the adjacent surface depletion region the density of ionized remote dopants is typically an order of magnitude larger than n_s. In such a situation long-range potential fluctuations, e.g., due to the statistical variations of the doping concentration, can be expected to become important. Near the MIT the positive feedback between potential fluctuations and screening will strongly enhance these long-range fluctuations, and it appears plausible that this mechanism can finally cause a phase separation between mobile and localized carriers. Such a behavior is corroborated by spatially resolved compressibility measurements on modulation-doped GaAs hole channels [21] and could be interpreted as a disintegration of the 2DCS into weakly linked puddles as suggested in Ref. [25].

In summary, by combining CESR and transport experiments we have studied high-mobility 2DCSs in Si/SiGe heterostructures on their way from the metallic to the insulating state. We found that above $n_s = 1.5 \times 10^{11}$ cm^{-2} the 2DCS is reasonably well described by Thomas-Fermi screening in the presence of weak potential fluctuations. At lower n_s the potential fluctuations increase and diverge at a critical value of 7×10^{10} cm^{-2}. This process is driven by the mutual dependence of DOS, screening efficiency, and potential fluctuations in a self-consistent, self-amplifying way. Thus, even in our high-mobility modulation-doped samples scattering is obviously dominated by potential fluctuations and the loss of screening rather than by the *e-e* interaction.

Valuable discussions with S. Lyon, G. Abstreiter, G. Brunthaler, and A. Prinz are gratefully acknowledged. This work is supported by *FWF* (Grants No. P-12143 and No. 13497 PHY), GMe, and ÖAD (all Vienna), and by KBN Grant No. 2 P 03B 007 16 (Poland).

[1] S. V. Kravchenko, D. Simonian, M. P. Sarachik, W. Mason, and J. E. Furneaux, Phys. Rev. Lett. **77**, 4938 (1996).
[2] J. Yoon, C. C. Li, D. Shahar, D. C. Tsui, and M. Shayegan, Phys. Rev. Lett. **82**, 1744 (1999), and references therein.
[3] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, JETP Lett. **65**, 932 (1997); **68**, 442 (1998).
[4] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[5] Here we use the general definition of r_s [see B. Tanatar *et al.,* Phys. Rev. B **39**, 5005 (1989)]; note that the ratio $E_{e\text{-}e}/E_{\mathrm{kin}} = g_\nu \cdot r_s$, with a valley degeneracy $g_\nu = 2$ in our system.
[6] B. L. Altshuler and D. L. Maslov, Phys. Rev. Lett. **82**, 145 (1999); **83**, 2092 (1999).
[7] Q. Si and C. M. Varma, Phys. Rev. Lett. **81**, 4951 (1998).
[8] V. Dobrosavljevic, E. Abrahams, E. Miranda, and S. Chakravarty, Phys. Rev. Lett. **79**, 455 (1997).
[9] M. A. Paalanen and R. N. Bhatt, Physica (Amsterdam) **169B**, 223 (1991).
[10] J. D. Quirt and J. R. Marko, Phys. Rev. Lett. **26**, 318 (1971).
[11] W. W. Warren, Jr., J. Non-Cryst. Solids **4**, 168 (1970).
[12] N. Mott, *Metal-Insulator Transitions* (Taylor & Francis Ltd., London, 1974), and references therein.
[13] See, e.g., J. H. Davies, *The Physics of Low-Dimensional Semiconductors* (Cambridge University Press, New York, 1997), p. 353 ff.
[14] N. Nestle, G. Denninger, M. Vidal, C. Weinzierl, K. Brunner, K. Eberl, and K. von Klitzing, Phys. Rev. B **56**, R4359 (1997).
[15] W. Jantsch, Z. Wilamowski, N. Sandersfeld, and F. Schäffler, Phys. Status Solidi (b) **210**, 643 (1998).
[16] C. F. O. Graeff, M. S. Brandt, M. Stutzmann, M. Holzmann, G. Abstreiter, and F. Schäffler, Phys. Rev. B **59**, 13 242 (1999).
[17] For a review, see F. Schäffler, Semicond. Sci. Technol. **12**, 1515 (1997), and references therein.
[18] W. Jantsch, Z. Wilamowski, N. Sandersfeld, and F. Schäffler, Physica (Amsterdam) **6E**, 218 (2000).
[19] Z. Wilamowski, W. Jantsch, N. Sandersfeld, and F. Schäffler, Ann. Phys. (Leipzig) **8**, 507 (1999).
[20] J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **68**, 674 (1992).
[21] S. Ilani, A. Yacoby, D. Mahalu, and H. Shtrikman, Phys. Rev. Lett. **84**, 3133 (2000).
[22] S. Das Sarma, Phys. Rev. Lett. **50**, 211 (1983); S. Das Sarma and E. H. Hwang, Phys. Rev. Lett. **83**, 164 (1999).
[23] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[24] G. Feher, Phys. Rev. **114**, 1219 (1959).
[25] Y. Meir, Phys. Rev. Lett. **83**, 3506 (1999).
[26] N. Mott, M. Pepper, S. Pollitt, R. H. Wallis, and C. J. Adkins, Proc. R. Soc. London A **143**, 160 (1975).