ЖЭТФ, 2005, том 127, вып. 3, стр. 674–685

THERMAL CORRECTION TO RESISTIVITY IN DILUTE Si-MOSFET TWO-DIMENSIONAL SYSTEMS

M. V. Cheremisin[*]

Ioffe Institute for Physics and Technology
194021, St. Petersburg, Russia

Submitted 15 April 2003,
after revision 24 April 2004

Neglecting electron–electron interactions and quantum interference effects, we calculate the classical resistivity of a two-dimensional electron (hole) gas taking the degeneracy and the thermal correction due to the combined Peltier and Seebeck effects into account. The resistivity is found to be a universal function of the temperature, expressed in the units of $(h/e^2)(k_F l)^{-1}$. Analysis of the compressibility and thermopower points to the thermodynamic nature of the metal–insulator transition in two-dimensional systems. We reproduce the beating pattern of Shubnikov – de Haas oscillations in both the crossed field configuration and Si-MOSFET valley splitting cases. The consequences of the integer quantum Hall effect in a dilute Si-MOSFET two-dimensional electron gas are discussed. The giant parallel magnetoresistivity is argued to result from the magnetic-field-driven disorder.

PACS: 73.40.Qv, 71.30.+h, 73.20.Fz

1. INTRODUCTION

Recently, much interest has been focused on the anomalous transport behavior of a variety of low-density two-dimensional (2D) systems [1–5]. It has been found that below some critical density, the cooling causes an increase in resistivity, whereas in the opposite high-density case, the resistivity decreases. Another property of dilute 2D systems is their unusual response to the parallel magnetic field. At low temperatures, the magnetic field was found to suppress the metallic behavior and result in increasing the resistivity upon enhancement of the spin polarization degree [6, 7]. A strong perpendicular magnetic field, if applied simultaneously with the parallel one, results in suppression of the parallel magnetoresistivity [8]. Although numerous theories have been put forward to account for these effects, the origin of the above behavior is still the subject of a heated debate.

The ohmic measurements are known to be carried out at a low current ($I \to 0$) in order to prevent Joule heating. In contrast to the Joule heat, the Peltier and Thomson effects are linear in the current. As shown in Refs. [9–11], the Peltier effect results in a correction to the measured resistance. When the current is running, one of the sample contacts is heated, and the other is cooled because of the Peltier effect. The contact temperatures are different. The voltage drop across the circuit includes thermoelectromotive force, which is linear in the current. There exists a thermal correction $\Delta\rho$ to the ohmic resistivity ρ of the sample. For low-density 2D electron gas (2DEG), the correction may be comparable to the resistivity because $\Delta\rho/\rho \sim (kT/\mu)^2$, where μ is the Fermi energy. In the present paper, we report on a study of low-T transport in 2D systems, taking both the carrier degeneracy and the Peltier-effect-induced correction to resistivity into account.

2. GENERAL FORMALISM

For clarity, we consider the (100) Si-MOSFET 2DEG system. Within the strong inversion regime, we further neglect a depletion layer charge in the semiconductor bulk. At a fixed gate voltage, the quasi-Fermi level μ in the semiconductor is shifted with respect to that in the metal gate. The number of occupied states below the quasi-Fermi level determines the density of electrons assumed to occupy the first

[*]E-mail: maksim.vip1@pop.ioffe.rssi.ru

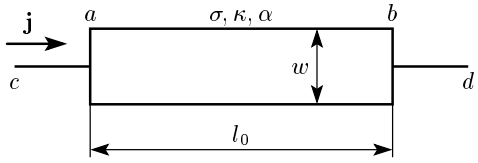

Fig. 1. The experimental setup

quantum-well subband with the isotropic energy spectrum $\varepsilon(\mathbf{k}) = \hbar^2 k^2/2m$.

We consider a sample connected to the current source by means of two identical leads (Fig. 1). Both contacts are ohmic. The voltage is measured between the open ends c and d kept at the temperature of the external thermal reservoir. The sample is placed in a chamber with the mean temperature T_0. According to our basic assumption, the contacts a and b may have different respective temperatures, T_a and T_b. Including the temperature gradient term, the current density $\mathbf{j}$ and the energy flux density $\mathbf{q}$ are given by

$$\mathbf{j} = \sigma(\mathbf{E} - \alpha\nabla T), \quad \mathbf{q} = (\alpha T - \zeta/e)\mathbf{j} - \kappa\nabla T, \qquad (1)$$

where $\mathbf{E} = \nabla\zeta/e$ is the electric field, $\zeta = \mu - e\varphi$ is the electrochemical potential, α is the thermopower, $\sigma = Ne\mu_0$ is the conductivity, $\mu_0 = e\tau/m$ is the mobility, τ is the momentum relaxation time, $\kappa = LT\sigma$ is the thermal conductivity, and $L = \pi^2 k^2/3e^2$ is the Lorentz number.

In general, one can solve Eq. (1) and then find the difference of contact temperatures, $\Delta T = T_a - T_b$, for an arbitrary circuit cooling. But below approximately 1 K, the electron–phonon coupling is known to be weak [12]. In the actual case where $I \to 0$, we can then omit the Joule heating. We therefore consider a simple case of adiabatic cooling, with the 2DEG thermally insulated from the environment. We emphasize that under the above conditions, the sample is not heated. Indeed, at small currents, we have $T_a \approx T_b \approx T_0$. Hence, the amount of the Peltier heat $Q_a = I\Delta\alpha T_0$ evolved at contact a and that absorbed at contact b are equal. Here, $\Delta\alpha$ is the difference of the 2DEG and metal conductor thermopowers. We recall that the energy flux is continuous at each contact,

$$-\kappa\nabla T|_{a,b} = \mathbf{j}\Delta\alpha T_{a,b},$$

and therefore the temperature gradient is constant downstream the current. The difference of the contact temperatures is then given by [9, 10]

$$\Delta T = \frac{\Delta\alpha l_0}{L\sigma w} I,$$

where l_0 and w are respectively the sample length and width. For example, for a 2×2-mm sample, the typical current $I = 1$ nA, the 2D resistivity of the order of h/e^2, and $\alpha \sim k^2 T/e\mu \sim 0.01k/e$, the contact temperature difference is $\Delta T = 10$ mK $\ll T_0$, and therefore our approach is well justified. From Eq. (1), the voltage drop between ends c and d is given by

$$U = RI + \Delta\alpha\Delta T,$$

where R is the ohmic resistance of the circuit. The second term is the conventional Seebeck thermoelectromotive force. Because $\Delta T \propto I$, we finally obtain the total 2DEG resistivity as

$$\rho^{tot} = \rho(1 + \alpha^2/L), \qquad (2)$$

where we assume that $\Delta\alpha \approx -\alpha$. We note that within the adiabatic approach, Eq. (2) can also be applied for the 2D hole gas and in the case of four-point contact measurements. In the Appendix, we discuss the case of 2DEG realistic cooling in more detail.

3. RESULTS

3.1. 2D density and thermopower

Using Gibbs statistics, we find that the 2DEG density $N = -(\partial\Omega/\partial\mu)_T$ is given by

$$N = N_0 \xi F_0(1/\xi), \qquad (3)$$

where

$$\Omega = -kT \sum_{\mathbf{k}} \ln\left[1 + \exp\left(\frac{\mu - \varepsilon(\mathbf{k})}{kT}\right)\right]$$

is the thermodynamic potential, $\xi = kT/\mu = T/T_F$ is the dimensionless temperature, T_F is the Fermi temperature, and $F_n(z)$ is the Fermi integral. At the moment, we disregard the valley splitting, reported to be of the order of 1 K in the (100) Si-MOSFET 2DEG system [13]. Below we discuss the importance of a nonzero valley splitting in the context of low-field Shubnikov–de Haas oscillations. Next, we introduce the density of states $D = 2m/\pi\hbar^2$, where m is the effective mass. For the density of strongly degenerate 2DEG, we have $N_0 = D\mu$. In what follows, we consider both the classical Boltzman ($\mu < 0$) and Fermi ($\mu > 0$) cases, and therefore use the dimensionless concentration $n = N/|N_0|$ (Fig. 2*a*). In the classical Boltzman limit ($\mu < 0$, $|\xi| \ll 1$), the 2D electron density is thermally activated:

$$n = |\xi| \exp(-1/|\xi|).$$

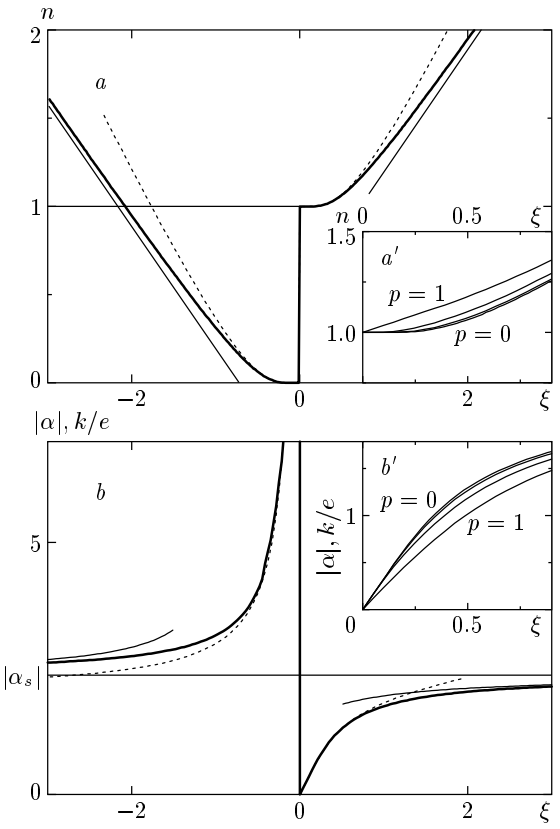

Fig. 2. The zero-field 2DEG density (a) and thermopower (b) given by Eqs. (3) and (4) respectively vs the dimensionless temperature ξ. Asymptotes shown by dotted lines correspond to $|\xi| \ll 1$ and those shown by thin lines to $|\xi| \gg 1$. Insets: 2DEG density (a') and thermopower (b') for the spin polarization degree $p = 0,\ 0.3,\ 0.6,\ 1$

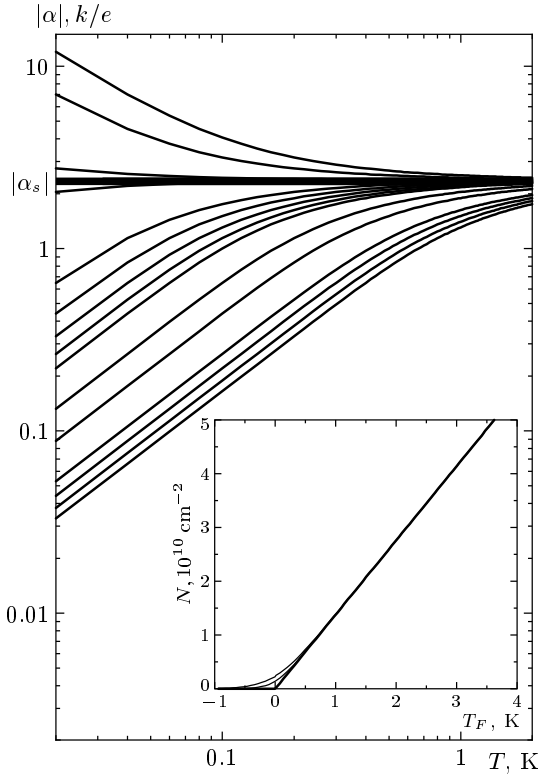

Fig. 3. Temperature dependence of the thermopower given by Eq. (4) for T_F [K] $= 2 - 0.25$ (step 0.25), $0.2 - 0.05$ (the step 0.05), 0.01, 0 (bold line), -0.1, -0.2. Inset: density vs Fermi energy at the fixed temperature T [K] $= 0$ (piecewise bold line), 0.15, 0.25 in Si-MOSFET system

For strongly degenerate electrons ($\xi \ll 1$), we obtain

$$n = 1 + \xi \exp(-1/\xi).$$

Then, at elevated temperatures ($\xi \geq 1$), the density

$$n = 1/2 + \xi \ln 2$$

becomes linear in the temperature. We note that at a fixed temperature, the 2DEG density always exceeds the zero-temperature value, i. e., $N > N_0$ (see Fig. 3, inset). Experimentally, the concentration extracted from the period of the Shubnikov–de Haas quantum oscillations [14, 15] determines the density of strongly degenerate 2DEG, i. e., $N_{SdH} = N_0$. In contrast, the classical low-field Hall measurements [4, 14] allow finding the total carrier density $N_{Hall} = N$, which coincides with the density of strongly degenerate electrons for $\xi \ll 1$. We argue that in dilute 2D systems, the accuracy provided by both methods becomes questionable, which seems to be the reason for the sample and temperature-dependent deviation $N_{Hall} - N_{SdH}$ observed in Si-MOSFETs [14].

Following the conventional Boltzman equation formalism, the explicit formula for the 2DEG thermopower can be written as

$$\alpha = -\frac{k}{e}\left[\frac{2F_1(1/\xi)}{F_0(1/\xi)} - \frac{1}{\xi}\right]. \qquad (4)$$

For simplicity, we assume that the electron scattering is characterized by the energy-independent momentum relaxation time. In the classical limit ($\mu < 0$, $|\xi| \ll 1$), the thermopower is given by the conventional formula

$$\alpha = -\frac{k}{e}\left(2 - \frac{1}{\xi}\right).$$

For strongly degenerate 2DEG ($\xi \ll 1$), we obtain the

temperature dependence of the thermopower (Fig. 2*b*) as

$$\alpha = -\frac{k}{e}\left[\frac{\pi^2\xi}{3} - (1+3\xi)\exp\left(-\frac{1}{\xi}\right)\right].$$

At elevated temperatures ($\xi > 1$), the thermopower first grows with the temperature and then approaches the universal value

$$\alpha_s = -\frac{k}{e}\frac{2F_1(0)}{F_0(0)} = -\frac{k}{e}\frac{\pi^2}{6\ln 2}.$$

The above behavior is confirmed by low-temperature thermopower measurement data [16], found to diverge at a certain value near $0.6k/e$, which is of the order of α_s (see the bold line in Fig. 3).

3.2. Zero-field resistivity

We now calculate the total 2DEG resistivity given by Eq. (2). Figure 4 represents the T-dependent resistivity at fixed Fermi temperatures that correspond to a certain 2DEG density range (see Fig. 3, inset). For a fixed disorder strength, we represent the data found at different densities (or T_F) in a single plot (Fig. 4). In real units, increasing the disorder results in the upshift of resistivity curves. The temperature dependence of the resistivity (see, e.g., the curve at $T_F = 0.25$ K in Fig. 4) exhibits the metallic behavior (i. e. $d\rho/dT > 0$) for $T \le T_F$ and then becomes insulating ($d\rho/dT < 0$) at $T \ge T_F$. Within the low-temperature metallic region, the 2DEG resistivity can be approximated (see the dotted line in Fig. 4) by

$$\rho^{tot} = \rho_0(1 + \pi^2\xi^2/3),$$

where

$$\rho_0 = \frac{1}{N_0 e\mu_0} = \frac{h}{2e^2}\frac{1}{k_F l}$$

is the resistivity at $T \to 0$, $k_F = \sqrt{2m\mu}/\hbar$ is the Fermi vector, and $l = \hbar k_F \tau/m$ is the mean free path. For the high-temperature insulating region, we then obtain the asymptote

$$\rho^{tot} = \rho_0\frac{1+\alpha_s^2/L}{\xi\ln 2} \propto \frac{1}{T},$$

depicted in Fig. 4 by the thin line. The metallic-to-insulating behavior crossover occurs at $T \approx 0.8T_F$. In fact, the low-temperature metallic resistivity is explained within our model in terms of the thermal correction given by Eq. (2), while the high-temperature insulating behavior results from a decrease of the 2D degeneracy. It is to be specially noted that in the Boltzman limit (see curves at $\mu < 0$ in Fig. 4), the resistivity can be scaled in units of the disorder parameter $k_F l$, where the substitution $\mu \to |\mu|$ must be made.

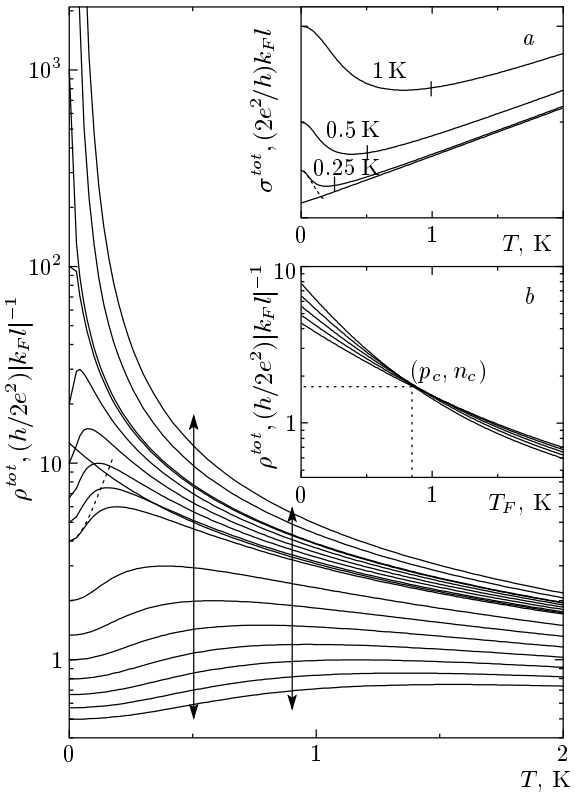

Fig. 4. Zero-field temperature dependence of resistivity given by Eq. (2) for Fermi temperatures depicted in the main panel of Fig. 3. The dotted line corresponds to the asymptote $\xi \ll 1$, the thin line to $\xi > 1$ for fixed $T_F = 0.25$ K. The arrows depict the region of temperatures explored in inset *b*. Inset *a*: temperature dependence of the inverse resistivity σ^{tot} at T_F[K] = 0.25, 0.5, 1.0 (marked by vertical bars). Inset *b*: density dependence of the resistivity depicted in the main panel in the temperature range $T = 0.5$–0.9 K

The resistivity data (Fig. 4, inset *b*), being represented as a function of the 2D density (or T_F), exhibit a well-pronounced transition point. The critical resistivity ρ_c is roughly inversely proportional to the critical density n_c. We note that the same experimental range $(0.1\text{–}10)h/e^2$ expected to eliminate the metal–insulator transition in Si-MOSFETs with different mobilities provides the higher temperature range, critical density n_c, and hence lower ρ_c for more disordered samples. This result is confirmed by experimental observations [17].

Recent experiments [4, 18–20] confirm our predictions and demonstrate that the metallic-region data obey a scaling where the disorder parameter $k_F l$ (not the ratio of the Coulomb interaction energy to the Fermi energy [20]) and the dimensionless temperature T/T_F appear explicitly. These experimental findings

were argued to rule out the electron–electron interactions [4], the shape of the potential well [18], spin–orbit effects, and quantum interference effects [19, 20] as possible origins of the metallic behavior mechanism. In addition, our concept of the high-temperature insulating behavior is qualitatively confirmed by the experimental data [21, 22] within the insulating side of the metal–insulator transition exhibiting the nonhopping $1/T$ dependence. As an example, for the p-GaAs/AlGaAs two-dimensional hole gas [22] with the peak mobility $\mu_0 = 2 \cdot 10^5$ cm^2/V$\cdot$s, we obtain the linear dependence (see the thin-line asymptote in Fig. 4, inset *a*) of the inverse resistivity, $\sigma^{tot}[e^2/h] = 1/\rho^{tot} = 1.4T$ [K], which is consistent with the experimental value $3.3T$ [K]. It is to be noted that the conventional theory [23, 24] used to explain the 2D metallic behavior [7, 21, 25] fails to account for both $T \to 0$ and $T \geq T_F$ cases.

We emphasize that Eq. (2) provides the actually measured effective 2D mobility and yields

$$\mu_{eff} = \frac{\mu_0}{1 + \alpha^2/L}.$$

Experimentally, at a fixed temperature, the mobility data can be unambiguously extracted using independent measurements of the 2D resistivity and the low-field Hall density $N_{Hall} \approx N$. With the help of Eq. (3), we plot the density dependence of μ_{eff} in Fig. 5. Upon depletion of 2DEG, the dependence $\mu_{eff}(N)$ falls down at low densities near 10^9 cm^{-2} as $\xi \to 0$. In the high-density case, we predict $\mu_{eff} \approx \mu_0$. We argue that in real experiment, the above behavior can be masked by impurity-assisted (Si–SiO_2 roughness-associated) suppression of the momentum scattering time in the respective cases of low (high) densities [26].

3.3. 2DEG magnetoresistivity

In contrast to the conventional Shubnikov – de Haas formalism extensively used to reproduce low-field data, we use the alternative approach [27] that seems to be aimed at resolving the magnetotransport problem within both Shubnikov – de Haas and integer-quantum-Hall-effect regimes.

The Si-MOSFET energy spectrum modified with respect to the valley and spin splitting is given by

$$\varepsilon_n = \hbar\omega_c \left(n_L + \frac{1}{2}\right) \pm \frac{\Delta_s}{2} \pm \frac{\Delta_v}{2}, \qquad (5)$$

where $n_L = 0, 1, \ldots$ is the Landau level number, $\omega_c = eB_\perp/mc$ is the cyclotron frequency, $\Delta_s = g^*\mu_B B$ is the Zeeman splitting, g^* is the effective g-factor, and $B = \sqrt{B_\perp^2 + B_\parallel^2}$ is the total magnetic field. Next,

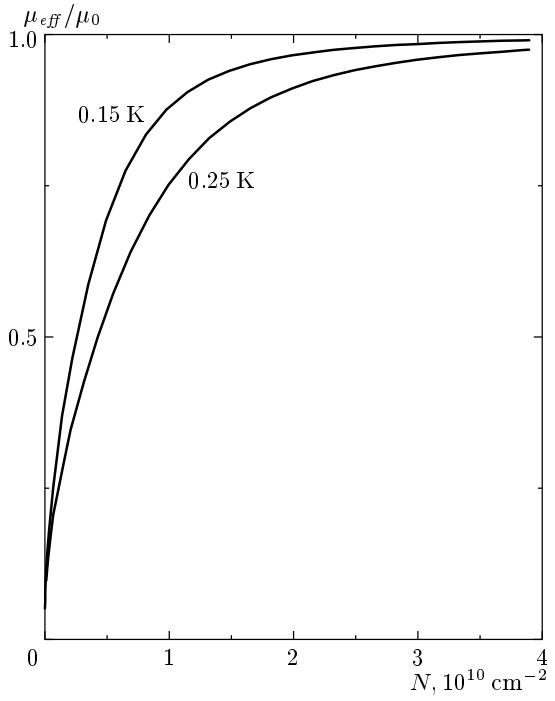

Fig. 5. The dimensionless effective mobility specified in the text vs the 2D density for the Si-MOSFET system at $T = 0.15$ K and $T = 0.25$ K

Δ_v[K]$= \Delta_v^0 + 0.6B_\perp$[T] is the density-independent [28] valley splitting. In contrast to the valley splitting, the spin susceptibility $\chi = g^*m/2m_0$ (where m_0 is the free electron mass) is known to exhibit strong enhancement upon 2D carrier depletion. This result is confirmed independently by magnetotransport measurements in a tilted magnetic field [6, 29], the perpendicular field [30], and by the beating pattern of the Shubnikov – de Haas oscillations [31] in crossed fields.

We recall that in strong magnetic fields ($\hbar\omega_c \gg kT, \hbar/\tau$), the electrons can be considered dissipationless, and therefore $\sigma_{xx}, \rho_{xx} \approx 0$. Under current carrying conditions, the only reason for a finite longitudinal resistivity seems to be the thermal correction mechanism discussed earlier [27]. Following Ref. [27], we obtain

$$\rho = \rho_{yx}\alpha^2/L, \qquad (6)$$

where α is the thermopower, $\rho_{yx}^{-1} = Nec/B_\perp$ is the Hall resistivity, $N = -(\partial\Omega/\partial\mu)_T$ is the 2D density,

$$\Omega = -kT\Gamma \sum_n \ln\left[1 + \exp\left(\frac{\mu - \varepsilon_n}{kT}\right)\right]$$

is the thermodynamic potential modified with respect to the energy spectrum mentioned above, and

$\Gamma = eB_\perp/hc$ is the zero-width of the Landau-level density of states. In strong magnetic fields, the $2D$ thermopower is a universal quantity [32], proportional to the entropy per electron, $\alpha = -S/eN$, where $S = -(\partial\Omega/\partial T)_\mu$ is the entropy. Both S and N, and hence α and ρ are universal functions of ξ and the dimensionless magnetic field $\hbar\omega_c/\mu = 4/\nu$, where $\nu = N_0/\Gamma$ is the conventional filling factor.

Using the Lifshitz–Kosevich formalism, we can easily derive asymptotic formulas for N and S, and hence for ρ_{yx} and ρ, valid at low temperatures $\xi < 1$ and weak magnetic fields $\nu^{-1} < 1$:

$$N = N_0\xi F_0\left(\frac{1}{\xi}\right) + \\ + 2\pi\xi N_0 \sum_{b=1}^{\infty} \frac{(-1)^b \sin(\pi b\nu/2)}{\operatorname{sh} r_b} R(\nu), \quad (7)$$

$$S = S_0 - 2\pi^2\xi k N_0 \sum_{b=1}^{\infty} (-1)^b \Phi(r_b) \cos\left(\frac{\pi b\nu}{2}\right) R(\nu),$$

where

$$S_0 = kN_0 \frac{d}{d\xi}\left[\xi^2 F_1\left(\frac{1}{\xi}\right)\right]$$

is the entropy at $B_\perp = 0$, $F_n(z)$ is the Fermi integral,

$$\Phi(z) = \frac{1 - z\operatorname{cth} z}{z\operatorname{sh} z},$$

and $r_b = \pi^2\xi\nu b/2$ is the dimensionless parameter. Then $R(\nu) = \cos(\pi bs)\cos(\pi bv)$ is the form factor, $s = \Delta_s/\hbar\omega_c = \chi B/B_\perp$ is the dimensionless Zeeman spin splitting, and $v = \Delta_v/\hbar\omega_c = \Delta_v^0\nu/4\mu + 0.12$ is the dimensionless valley splitting.

We first consider the zero-$B_\parallel$ case, where the Zeeman spin splitting is reduced to a field-independent constant, i. e., $s = \chi$. In the low-$T, B_\perp$ limit, the valley splitting Δ_v^0 is then known to be resolved [13], and therefore leads to beating of the Shubnikov–de Haas oscillations. In the actual first-harmonic case (i. e., $b = 1$), the beating nodes can be observed when $\cos(\pi v) = 0$, or

$$\nu_i^v = \frac{4\mu(i/2 - 0.12)}{\Delta_v^0},$$

where $i = 1, 3, \ldots$ is the beating node index. For 2DEG parameters reported in Ref. [13] (Fig. 6), we estimate $\nu_1^v = 101$, and therefore $\Delta_v^0 = 0.92$ K. The second node is expected to appear at $\nu_3^v = 368$. However, the Shubnikov–de Haas oscillations are in fact resolved when $\nu \le 1/\xi = 203$, and therefore the second beating node was not observed in experiment [13]. Moreover, the disappearance of the first beating node upon 2D carrier depletion $N < 3 \cdot 10^{11}$ cm^{-2} is governed by the same condition because in this case $\nu \le 1/\xi = 73$ is of the order of the first beating node. We note, however, that suppression of the beating nodes at higher densities ($N > 9 \cdot 10^{11}$ cm^{-2}) reported in Ref. [13] is unexpected within our simple scenario.

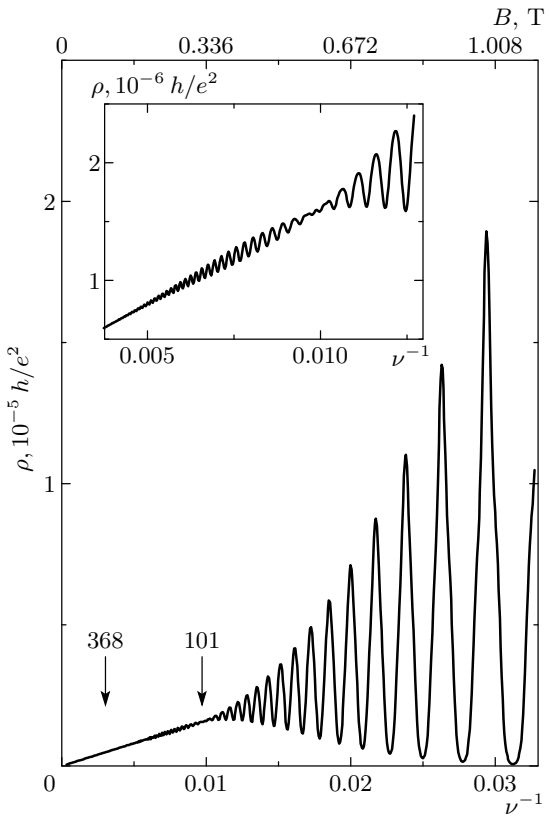

Fig. 6. The Shubnikov–de Haas oscillations at $T = 0.3$ K for a Si-MOSFET sample [13]: $N_0 = 8.39 \cdot 10^{11}$ cm^{-2}, the spin susceptibility $\chi = 0.305$, and the valley splitting Δ_v[K]$=\Delta_v^0 + 0.6B_\perp$ [T]. The zero-field valley splitting $\Delta_v^0 = 0.92$ K is a fitting parameter. The arrows depict the beating nodes at $i = 1, 3$. Inset: an enlarged plot of the beating node from the main panel

We now analyze the case of a low-density 2D system in a strong magnetic field with only the lowest Landau levels occupied. For an extremely dilute 2DEG ($N \approx 10^{11}$ cm^{-2}), the energy spectrum (see the inset in Fig. 7) is known to be strongly affected by enhanced spin susceptibility. In contrast to the high-density case with cyclotron minima occurring at $\nu = 4, 8, 12, \ldots$, only the spin minima ($\nu = 2, 6, 10, \ldots$) are observed in dilute 2DEG [33]. As expected, the spin (cyclotron) minimum fillings are proportional to odd (even) num-

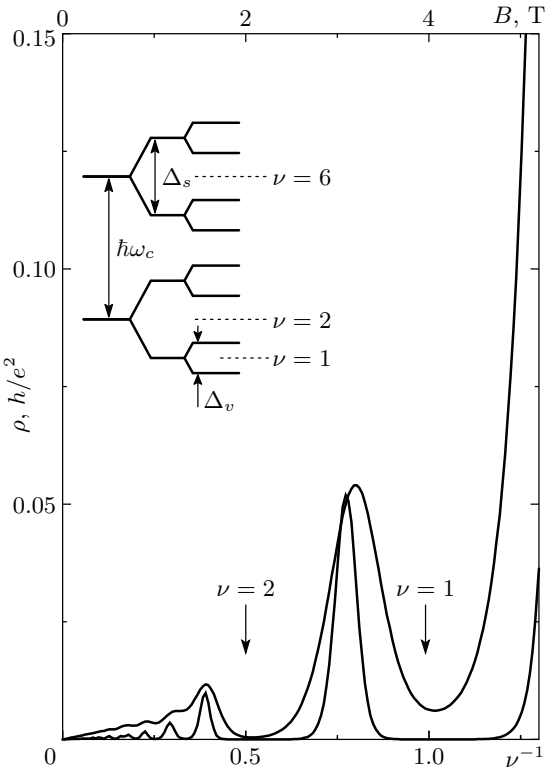

Fig. 7. Magnetoresistivity at $T = 0.36$ K (upper curve) and $T = 0.18$ K for dilute 2DEG Si-MOSFET [33]: $N_0 = 10^{11}$ cm^{-2}, the spin susceptibility $\chi = 0.5$, and the valley splitting specified in the capture Fig. 6. Inset: energy spectrum given by Eq. (5) for two lowest Landau levels

bers multiplied by a factor of two due to the valley degeneracy. In stronger fields, the magnetoresistivity data exhibit a $\nu = 1$ minimum associated with valley splitting. Using the energy spectrum implied by Eq. (5), we can easily find that the last minimum occurs when the Fermi level lies between the lowest valley-split Landau levels, i. e., at $\mu = \hbar\omega_c(1-\chi)/2$. The sequence of minima at $B = 4, 2, 0.66$ T reported in Ref. [33] provides an independent test for spin susceptibility in the high-$B_\perp$ limit. In Fig. 7, we represent the magnetoresistivity specified by Eq. (6) and then use $\chi = 0.5$ in order to fit the observed sequence of minima. Surprisingly, the value of spin susceptibility is lower than the value $\chi = 0.86$ extracted from the crossed-field low-field Shubnikov–de Haas beating pattern analysis [31]. We attribute this discrepancy, for example, to a possible magnetic-field dependence of spin susceptibility.

Finally, we focus on the magnetotransport problem in the crossed magnetic field configuration. Following experiments [13], we further neglect the zero-field valley splitting in the actual high-density case ($N > 9 \cdot 10^{11}$ cm^{-2}). At a fixed parallel magnetic field, the dimensionless Zeeman splitting is given by

$$s = \chi\sqrt{1 + \nu^2/\nu_\parallel^2},$$

where we introduce the auxiliary «filling factor» $\nu_\parallel = hcN_0/eB_\parallel$ associated with the parallel field. In the low-$B_\perp$ limit, the spin splitting induced by the parallel field also results in the beating of the Shubnikov–de Haas oscillations. We can easily derive the condition for the Shubnikov–de Haas beating nodes as $\cos(\pi s) = 0$ or

$$\nu_j^s = \nu_\parallel\sqrt{(j/2\chi)^2 - 1}, \quad j = 1, 3, \ldots$$

The sequence of the beating nodes observed in Ref. [31] allowed the authors to deduce the density dependence of the spin susceptibility. As an example, in Fig. 8, we reproduce the magnetoresistivity implied by Eqs. (6) and (7) for 2DEG parameters [31]. The phase of the Shubnikov–de Haas oscillations remains the same between the adjacent beating nodes and changes by π through the node in agreement with experiments.

We now consider the 2DEG magnetotransport in a tilted configuration with the sample rotated in a constant magnetic field [6, 7, 29]. In this case, the Shubnikov–de Haas beating pattern is known to depend on the spin polarization degree $p = \Delta_s/2\mu = 2\chi/\nu_{tot}$, where we introduce the auxiliary «filling factor» $\nu_{tot} = hcN_0/eB$ associated with the total magnetic field. Conventionally, the spin polarization degree is related to the parallel field B_c required for the complete spin polarization, and therefore $p = B/B_c$. Performing a minor modification in Eq. (6), namely that $s = \chi\nu/\nu_{tot}$, in Fig. 9 we represent the magnetoresistivity as a function of the filling factor for a 2DEG plane rotated with respect to the constant magnetic field $B = 18$ T [7]. For simplicity, we omit the zero-field valley splitting. Then, assuming that the Landau-level broadening is neglected within our simple approach, we use a somewhat higher temperature compared to that in experiment [7]. For a spin-polarized system, the Shubnikov–de Haas oscillations ($p = 1.01$ in Fig. 9*a*) are caused by the only lowest valley-degenerate spin-up subband. At low temperatures, the valley splitting at $\nu = 3$ is resolved. With the energy spectrum specified by Eq. (5), the high-filling maxima occur at $4(N + 1/2)/(1 + p) \sim 2N + 1$ and therefore have the period $\Delta\nu = 2$. In contrast, the partially polarized high-density 2DEG case ($p = 0.29$) depicted in Fig. 9*b* demonstrates a rather complicated beating pattern caused by both spin-up and spin-down subbands.

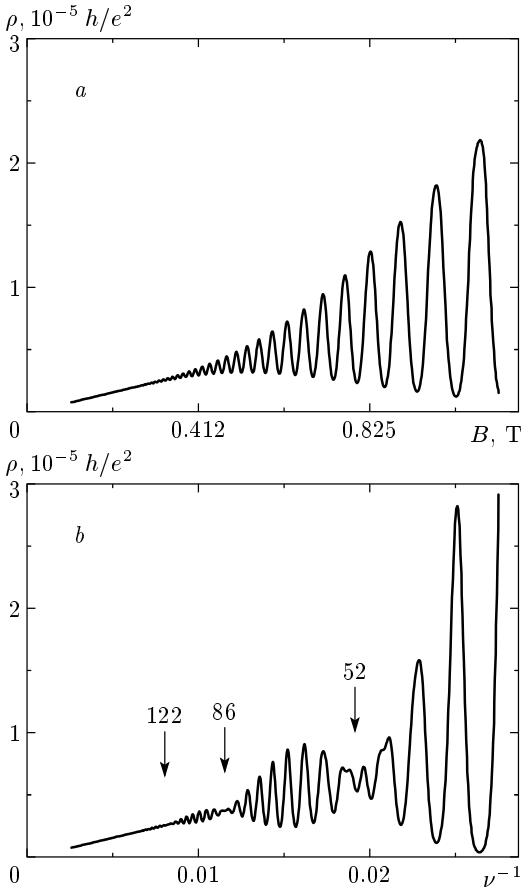

Fig. 8. The Shubnikov – de Haas beating pattern oscillations at $T = 0.35$ K for a Si-MOSFET sample [31]: $N_0 = 10.6 \cdot 10^{11}$ cm^{-2}, the spin susceptibility $\chi = 0.27$, $\Delta_v^0 = 0$ and $B_\parallel = 0$ (a), $B_\parallel = 4.5$ T (b), $\nu_\parallel = 9.25$. The arrows depict the beating nodes at $j = 3,\ 5,\ 7$

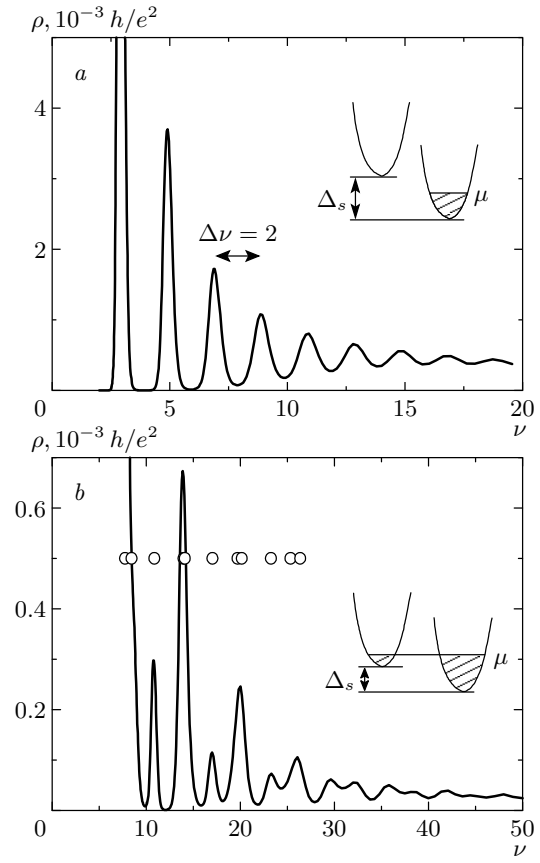

Fig. 9. The small-angle Shubnikov – de Haas oscillations at $T = 1.35$ K for a Si-MOSFET system [7]: a) spin polarized electrons ($p = 1.01$) at $N_0 = 3.72 \cdot 10^{11}$ cm^{-2}, the spin susceptibility $\chi = 0.42$ [31], the «effective filling factor» $\nu_{tot} = 0.83$, and b) the partially polarized case ($p = 0.29$) at $N_0 = 9.28 \cdot 10^{11}$ cm^{-2}, the spin susceptibility $\chi = 0.30$ [31], and $\nu_{tot} = 2.06$. The positions of maxima are represented by empty dots. Insets: schematic band diagrams at $B = B_\parallel$

It can be easily demonstrated that high-filling maxima occur at $4(N + 1/2)/(1 \pm p)$ (dots in Fig. 9b) and hence depend on the spin polarization degree. The ratio of oscillation frequencies for the two spin subbands is

$$\frac{f_\downarrow}{f_\uparrow} = \frac{1-p}{1+p},$$

consistently with experiment [7]. At the moment, however, we cannot explain the puzzling behavior of low-filling magnetoresistivity data known to be insensitive to the parallel field component [6, 33].

We emphasize that the data represented in Figs. 6–9 differ from those provided by the conventional formalism in the following aspects: i) the low-field ($\omega_c\tau \le 1$) quantum interference and classical negative magnetoresistivity background is excluded within our approach and ii) in contrast to the conventional Shubnikov – de Haas analysis, our approach determines (at $\omega_c\tau \gg 1$) the absolute value of magnetoresistivity and, moreover, provides a continuous transition from the Shubnikov – de Haas regime to the quantum Hall effect ($\hbar\omega_c \gg kT$). A minor point is that our approach predicts a somewhat lower Shubnikov – de Haas oscillation amplitude compared to that in experiment. However, in the integer-quantum Hall effect regime, the magnetoresistivity magnitude is well comparable with experimental values [27].

3.4. Parallel-field magnetoresistivity

One of the most intriguing features of the Si-MOSFET 2D system is its enormous response to the

magnetic field applied in the plane of the electrons. At a fixed temperature, the parallel-field resistivity is known to exhibit a dramatic increase at both sides of the zero-field metal–insulator transition. On the metallic side, the resistivity increases by more than an order of magnitude and then saturates above a certain value of the parallel magnetic field. The saturation field corresponds to the complete spin polarization [6], when $p = 1$. On the insulating metal–insulator transition side, the saturation of the magnetoresistivity is not observed [6]. We now give a qualitative argument in favor of the magnetic-field-driven disorder origin of the observed magnetoresistivity data.

At a fixed parallel magnetic field, behavior of the T-dependent resistivity is reported [17] to be similar to that in the zero-field case (see Fig. 4). Moreover, the same data plotted as a function of density also exhibit a well-pronounced transition point as in the case of the zero-field metal–insulator transition (see Fig. 4, inset b). Both the critical resistivity ρ_c^B and the density n_c^B depend on the magnetic field strength. Surprisingly, the critical diagram ρ_c^B vs n_c^B was found [17] to coincide with that obtained in the case of the zero-field metal–insulator transition for different mobility Si-MOSFET samples. Assuming that the thermal correction mechanism is also valid in the presence of the parallel field, we attribute the observed magnetoresistivity behavior to the field-driven disorder enhancement, i. e., $\tau(p) < \tau(0)$. Indeed, with the energy spectrum specified by Eq. (5), the explicit formulas for the 2DEG density and thermopower are

$$N = \frac{N_0\xi}{2}\sum_i F_0\left(\frac{1-\varepsilon_i}{\xi}\right),$$
$$\alpha = -\frac{k}{e}\times$$
$$\times\left[\frac{\sum\limits_i\left[2F_1\left(\frac{1-\varepsilon_i}{\xi}\right)+\frac{\varepsilon_i}{\xi}F_0\left(\frac{1-\varepsilon_i}{\xi}\right)\right]}{\sum\limits_i F_0\left(\frac{1-\varepsilon_i}{\xi}\right)}-\frac{1}{\xi}\right], \quad (8)$$

where $\varepsilon_i = \pm p$ is the dimensionless energy deficit between the bottom of spin subbands and that of the ground state. For simplicity, we here neglect the zero-field valley splitting. Both the 2D density and thermopower exhibit (see Fig. 2, insets a' and b') only a minor perturbation upon parallel field enhancement within $0 < p < 1$. We therefore conclude that the field-driven disorder enhancement can be responsible for the observed magnetoresistivity behavior. The detailed analysis of the prevailing $\tau(p)$ mechanism (see, e.g., Ref. [34]) is beyond the scope of the present paper.

3.5. 2D compressibility

Hereafter, we refer the reader to the experimental data mostly obtained for the n-GaAs/AlGaAs 2DEG system, and therefore we should substitute $D \to D/2$ in what follows. In general, of particular interest is the 2DEG compressibility

$$K = \frac{dN}{d\mu} = -\frac{d^2\Omega}{d\mu^2},$$

known to be a fundamental quantity, generally more amenable to theoretical and experimental analysis [15, 35]. For noninteracting electrons, Eq. (3) yields

$$K(\xi) = DF_0'(1/\xi),$$

where $F_n'(z) = dF_n(z)/dz$ is the derivative of the Fermi integral. Figure 10 represents the dependence of the actually measured inverse compressibility $d(\mu) = \varepsilon/Ke^2$. For strongly degenerate electrons ($\xi \ll 1$), we obtain a constant value $d_0 = \varepsilon/De^2$, consistent with the conventional capacitance measurements [36]. But as the 2DEG degeneracy decreases, the AC electric field penetration data [15, 35] demonstrates diminution and, furthermore, the negative inverse compressibility also decreases compared to d_0.

Conventionally, this behavior is explained [15] in terms of a Hartree–Fock exchange, which is omitted in our simple approach. In contrast, for extremely depleted 2DEG, the inverse compressibility data always exhibit an abrupt upturn, which cannot be explained within the Hartree–Fock scenario [35]. We assume that the above feature has a natural explanation within our model (see the dotted line in Fig. 10) because $d = d_0\exp(-1/|\xi|)$ at $\mu < 0$, $|\xi| \ll 1$, and hence the inverse compressibility exhibits the T-activated behavior. For example, upon depletion, the inverse compressibility [15] strongly increases at $T_F = 0.63$ K ($N = 2\cdot 10^9$ cm^{-2}), being of the order of the bath temperature $T = 0.3$ K.

In the general case of 2DEG placed in the perpendicular magnetic field, the compressibility is

$$K = -\left(\frac{\partial\Omega}{\partial\mu}\right)_T = \frac{D}{\xi\nu}\sum_n \mathrm{ch}^{-2}\left(\frac{\varepsilon_n-\mu}{2kT}\right),$$

or

$$K \approx D\left[F_0'\left(\frac{1}{\xi}\right)+\pi^2\xi\nu\sum_b\frac{(-1)^b b\cos(\pi b\nu/2)}{\mathrm{sh}\, r_b}\right], \quad (9)$$

where we use the thermodynamic potential modified with respect to the single-valley spin-unresolved Landau level energy spectrum. According to Eq. (9), at the

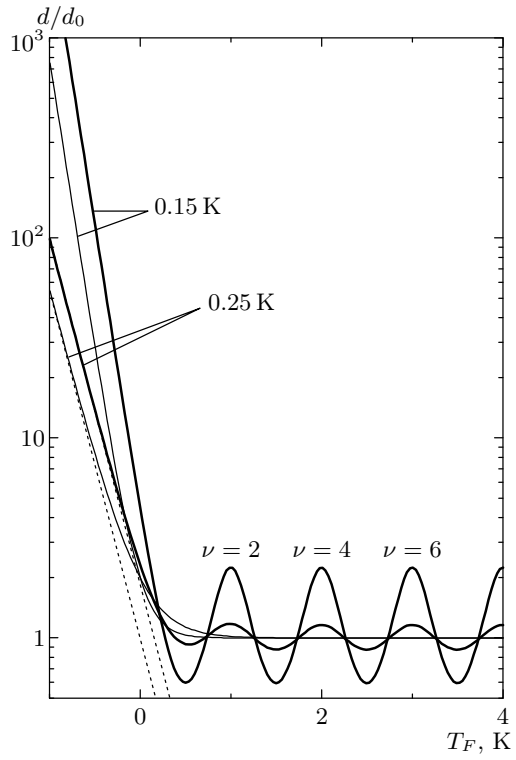

Fig. 10. The dimensionless inverse compressibility vs the Fermi temperature at zero magnetic field (thin lines) and $\hbar\omega_c = 1$ K (bold lines) at fixed temperatures $T = 0.15$ K and $T = 0.25$ K. Dotted lines depict asymptotes at $\xi < 0, |\xi| \ll 1$

fixed magnetic field and temperature, the dependence $d(\mu)$ can be viewed (see Fig. 10) as a superposition of the zero-field dependence and the Landau-level-related oscillations. For a typical GaAs/AlGaAs system [15], we represent the data for $B = 0.5$ T ($\hbar\omega_c = 1$ K $> T$) in Fig. 10. The Landau-level-assisted oscillations at $\nu = 2, 4, 6, \ldots$ are well resolved. We note that for a dilute 2DEG system at $\xi < 0, |\xi| \ll 1$ in the presence of a strong magnetic field, $\hbar\omega_c \gg kT$ (i. e., $\xi\nu \ll 1$), we obtain the T-activated behavior as

$$d = \frac{d_0\xi\nu}{2} \exp\left(\frac{1}{\xi\nu} - \frac{1}{|\xi|}\right),$$

similarly to the zero-field case (see dotted lines in Fig. 10).

4. CONCLUSIONS

In conclusion, the total resistivity of a dilute 2D system in Si-MOSFET with the thermal correction included is found to be a universal function of the temperature, expressed in units $(h/e^2)(k_F l)^{-1}$. We have demonstrated the relevance of the approach suggested in Ref. [27] to the low-field beating pattern of the Shubnikov–de Haas oscillations in both crossed and tilted magnetic field configurations. The features concerning the integer quantum Hall effect in dilute Si-MOSFET systems are discussed. The strong increase of the parallel magnetiresistivity was argued to result from spin-dependent disorder.

This work was supported by the RFBR (grant № 03-02-17588) and LSF (HPRI-CT-2001-00114, Weizmann Institute).

APPENDIX

Real cooling of the 2D system

We consider the more realistic situation of electron cooling caused by a finite strength of the electron–phonon coupling. The phonon-to-mixing-chamber cooling could then predominately occur over the sample surface. The power balance equations linearized with respect to small temperature perturbations are

$$\begin{aligned} &\mathrm{div}(\kappa\nabla T) + \frac{j^2}{\sigma} - jT\nabla\alpha - \beta(T - T_p) = 0, \\ &\mathrm{div}(\kappa_p\nabla T_p) - \gamma(T_p - T) - \beta(T_p - T_0) = 0, \end{aligned} \tag{10}$$

where κ_p is the phonon thermal conductivity, T_p is the local phonon temperature, and β and γ are the respective electron–phonon and sample-to-mixing-chamber cooling strengths. With the phonon diffusion assumed weak in the sample bulk, the phonon temperature

$$T_p = \frac{\gamma T + \beta T_0}{\gamma + \beta}$$

coincides with the electron (bath) temperature upon predominant cooling. In general, $T_0 < T_p < T$. The electron–phonon coupling term in Eq. (??), rewritten in terms of the bath temperature, is

$$\beta^*(T - T_0) = \frac{\gamma\beta}{\gamma + \beta}(T - T_0),$$

which depends on both coupling constants. As expected, a weak heat path channel provides thermal cooling of the 2DEG system.

With $T_0^* = T_0 + j^2/\sigma\beta^*$ being the Joule heat enhanced temperature, Eq. (??) yields

$$\Delta\Theta^2 - u\nabla\Theta^2 - 2(\Theta - 1) = 0, \tag{11}$$

where $\theta = T/T_0^*$ is the dimensionless electron temperature, $\eta = x/\lambda$ is the dimensionless coordinate,

$\lambda = (L\sigma T_0^*/\beta^*)^{1/2}$ is the thermal diffusion length scale, and $u(\Theta) = (\lambda j/L\sigma)d\alpha/dT$ is the dimensionless parameter. Because $T \approx T_0$, the 2D thermopower can be considered constant, and we therefore omit the second term in Eq. (10). Then the energy flux continuity at both ends of the sample provides symmetric boundary conditions with the temperature gradients $\nabla\Theta|_{a,b} = -j\lambda\Delta\alpha/L\sigma T_0^*$. Under these conditions, solving Eq. (10) is straightforward [10]. The temperature profile downstream the sample is governed by the sample-to-thermal diffusion length ratio l_0/λ. Our approach of the adiabatic cooling is justified when $l_0 \ll \lambda$. In the opposite case of strong cooling ($l_0 \gg \lambda$), the electron temperature exhibits sharp deviation with respect to T_0^* near the contacts and then coincides with T_0^* in the sample bulk.

Considering that the use of interior potential probes gives uniform resistivity data, Prus et al. [37] suggested these data as a precursor of strong cooling in real Si-MOSFETs. With the T-dependent resistivity of 2DEG used as a thermometer, the electron–phonon coupling constant was extracted [37] from the simplified energy balance condition $T = T_0^*$ valid in the sample bulk when $l_0 \gg \lambda$. We stress that the above procedure is only justified when the electron–phonon coupling, and hence the thermal diffusion length are known *a priori*. Indeed, the weak coupling, if present, provides a constant temperature gradient, and hence a uniform resistivity as well. Nevertheless, the simple balance condition used in [37] becomes useless because $T \neq T_0^*$ throughout the sample. It turns out that the weak electron–phonon coupling constant cannot be extracted in the conventional manner. One can estimate the critical electron–phonon coupling at which our adiabatic approach is valid, i. e., $l_0 < \lambda$ or $\beta^* < L\sigma T_0^*/l_0^2$. For $l_0 = 3$ mm, $\sigma \sim e^2/h$, and $T_0 = 100$ mK, we obtain $\beta^* < 1.1 \cdot 10^{-10}$ W/K$\cdot$cm^2.

It is to be noted that the Peltier effect correction to resistivity becomes strongly damped at higher frequencies because of the thermal inertial effects [9]. Our dc approach is valid below some critical frequency $f_{cr} \approx \hbar/ml_0^2 = 0.3$ kHz, and therefore the spectral dependence of the 2D resistivity can be used to estimate the thermal correction.

REFERENCES

1. S. V. Kravchenko, G. V. Kravchenko, J. E. Furneaux et al., Phys. Rev. B **50**, 8039 (1994).
2. D. Popovic, A. B. Fowler, and S. Washburn, Phys. Rev. Lett. **79**, 1543 (1997).
3. P. T. Coleridge, R. L. Williams, Y. Feng, and P. Zawadzki, Phys. Rev. B **56**, R12764 (1997).
4. M. Y. Simmons, A. R. Hamilton, M. Pepper et al., Phys. Rev. Lett. **80**, 1292 (1998); Physica E **11**, 161 (2001).
5. Y. Hanein, U. Meirav, D. Shahar et al., Phys. Rev. Lett. **80**, 1288 (1998).
6. T. Okamoto, K. Hosoya, S. Kawaji, and A. Yagi, Phys. Rev. Lett. **82**, 3875 (1999).
7. S. A. Vitkalov, H. Zheng, K. M. Mertes et al., Phys. Rev. Lett. **85**, 2164 (2000); E-print archives, cond-mat/0101196.
8. D. Simonian, S. V. Kravchenko, M. P. Sarachik, and V. M. Pudalov, Phys. Rev. Lett. **79**, 2304 (1997).
9. C. G. M. Kirby and M. J. Laubitz, Metrologia **9**, 103 (1973).
10. M. V. Cheremisin, Zh. Eksp. Teor. Fiz. **119**, 409 (2001).
11. M. V. Cheremisin, in *Proceedings on ICPS'26*, D113, Edinburgh, UK (2002).
12. A. Mittal, M. W. Keller, R. G. Wheeler, and D. E. Prober, Physica B **194–196**, 167 (1994).
13. V. M. Pudalov, A. Punnoose, G. Brunthaler et al., E-print archives, cond-mat/0104347.
14. V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, Pis'ma v ZhETF **70**, 48 (1999).
15. J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **68**, 674 (1992); Phys. Rev. B **50**, 1760 (1994).
16. R. Fletcher, V. M. Pudalov, A. D. B. Radcliffe, and C. Possanzini, Semicond. Sci. Technol. **16**, 386 (2001).
17. V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, E-print archives, cond-mat/0103087.
18. A. R. Hamilton, M. Y. Simmons, M. Pepper et al., Phys. Rev. Lett. **87**, 126802 (2001).
19. G. Brunthaler, A. Prinz, G. Bauer, and V. M. Pudalov, Phys. Rev. Lett. **87**, 096802 (2001).
20. A. Lewalle, M. Pepper, C. J. B. Ford, and E. H. Hwang, Phys. Rev. B **66**, 075324 (2002).
21. H. Noh, M. P. Lilly, D. C. Tsui, and J. A. Simmons, Phys. Rev. B **68**, 165308 (2003); E-print archives, cond-mat/0301301.

22. Xuan P. A. Gao, A. P. Mills Jr., A. P. Ramirez, and L. N. Pfeiffer, Phys. Rev. Lett. **88**, 166803 (2002); E-print archives, cond-mat/0308003.

23. A. Gold and V. T. Dolgopolov, Phys. Rev. B **33**, 1076 (1986).

24. G. Zala, B. N. Narozhny, and I. L. Aleiner, Phys. Rev. B **65**, 020201 (2001).

25. V. M. Pudalov, M. E. Gershenson, H. Kojima, and G. Brunthaler, Phys. Rev. Lett. **91**, 126403 (2003).

26. A. K. Savchenko, private communication.

27. M. V. Cheremisin, in *Proceedings on NATO ASI*, Windsor, UK (2001); E-print archives, cond-mat/0102153.

28. V. M. Pudalov, S. G. Semenchinskii, and V. S. Edelman, Zh. Eksp. Teor. Fiz. **89**, 1870 (1985).

29. F. F. Fang and P. J. Stiles, Phys. Rev. **174**, 823 (1968).

30. S. V. Kravchenko, A. A. Shashkin, D. A. Bloore, and T. M. Klapwijk, Sol. St. Comm. **116**, 495 (2000).

31. V. M. Pudalov, M. E. Gershenson, H. Kojima, and N. Butch, Phys. Rev. Lett. **88**, 196404 (2002).

32. S. M. Girvin and M. Jonson, J. Phys. C **15**, L1147 (1982).

33. S. V. Kravchenko, D. Simonian, M. P. Sarachik, and A. D. Kent, Phys. Rev. B **58**, 3553 (1998).

34. V. T. Dolgopolov and A. Gold, Pis'ma v Zh. Eksp. Teor. Fiz. **71**, 42 (2000).

35. S. C. Dultz and H. W. Jiang, Phys. Rev. Lett. **84**, 4689 (2000).

36. S. I. Dorozhkin, J. H. Smet, and K. von Klitzing, Phys. Rev. B **63**, 121301(R) (2001).

37. O. Prus, M. Reznikov, U. Sivan, and V. M. Pudalov, Phys. Rev. Lett. **88**, 016801 (2002).