Temperature-driven transition from a semiconductor to a topological insulator

Steffen Wiedmann,1,* Andreas Jost,1 Cornelius Thienel,2 Christoph Brüne,2 Philipp Leubner,2 Hartmut Buhmann,2 Laurens W. Molenkamp,2 J. C. Maan,1 and Uli Zeitler1
1High Field Magnet Laboratory and Institute for Molecules and Materials, Radboud University, Toernooiveld 7, 6525 ED Nijmegen, The Netherlands.
2Physikalisches Institut (EP3), Universität Würzburg, Am Hubland, D-97074, Würzburg, Germany.

(Received 19 November 2014; published 19 May 2015)

We report on a temperature-induced transition from a conventional semiconductor to a two-dimensional topological insulator investigated by means of magnetotransport experiments on HgTe/CdTe quantum well structures. At low temperatures, we are in the regime of the quantum spin Hall effect and observe an ambipolar quantized Hall resistance by tuning the Fermi energy through the bulk band gap. At room temperature, we find electron and hole conduction that can be described by a classical two-carrier model. Above the onset of quantized magnetotransport at low temperature, we observe a pronounced linear magnetoresistance that develops from a classical quadratic low-field magnetoresistance if electrons and holes coexist. Temperature-dependent bulk band structure calculations predict a transition from a conventional semiconductor to a topological insulator in the regime where the linear magnetoresistance occurs.

DOI: 10.1103/PhysRevB.91.205311 PACS number(s): 73.25.+i, 73.20.At, 73.43.–f

I. INTRODUCTION

Narrow-gap semiconductors possess conduction bands which are strongly non-parabolic and spin-orbit splittings that can be even larger than the fundamental band gap [1]. A particular interesting system is a type-III heterostructure composed of the semimetal HgTe and the wide-gap semiconductor HgCdTe with a low Hg content. These quantum well (QW) structures with an energy gap of several meV have already been experimentally investigated by means of optics and magnetotransport in the late 90s in order to obtain information about the band structure (BS) and the Landau level dispersion [2,3].

In 2006, Bernevig et al. predicted that the quantum spin Hall effect (QSHE) can be observed in inverted HgTe/CdTe QW structures if the layer thickness is larger than a critical value [4]. The system is then referred to as a two-dimensional (2D) topological insulator (TI) [5,6]. The hallmark of this state of matter is a quantized longitudinal conductance when the Fermi energy is in the bulk band gap and transport is governed by spin-polarized counter-propagating edge states. This quantized conductance has been found experimentally first in inverted HgTe QWs [7] and later on in InAs/GaSb heterostructures [8,9]. The existence of the helical states has been confirmed in inverted HgTe QWs by nonlocal measurements [10] and by verifying their spin polarization [11].

Bulk HgTe crystallizes in zincblende structure. When the semimetal HgTe with a negative energy gap of $E_v = -0.3$ eV is combined with the semiconductor HgCdTe, a type-III QW is formed. The band order in HgTe QWs with HgCdTe barriers depends strongly on the quantum confinement, i.e., the width $d$ of the QW. For $d < d_o$, the system is a conventional direct band-gap semiconductor with a $s$-type $\Gamma_8$ conduction band and $p$-type $\Gamma_8$ valence band. $d_o$ is the critical thickness of the QW where the system becomes a zero-gap semiconductor [12]. Calculations within the $8 \times 8 k \cdot p$ Kane model yield $d_o = 6.3$ nm for a QW on a Cd$_{0.96}$Zn$_{0.04}$Te substrate and $d_o = 6.7$ nm on a CdTe substrate in the zero-temperature limit. For $d > d_o$, the BS is inverted, i.e., the $\Gamma 1$ band is the conduction and the $E1$ band is the valence band, and the system is a 2D TI. Moreover, the system has an indirect band gap. If the well width is increased further, the confinement energy decreases and the system exhibits semimetallic behavior [13].

Apart from the variation of the QW thickness, temperature can induce a phase transition from a normal state ($T > T_c$) to a topologically nontrivial state of matter ($T < T_c$). This is caused by the strong temperature dependence of the $E1$ band, which is the topic of this paper. In order to provide experimental evidence for such a temperature-driven transition, we have performed magnetotransport experiments on four different HgTe quantum wells, all with a width exceeding the critical thickness $d_o$ at $T = 0$. We show that we are in the regime of the QSHE and observe an ambipolar quantized Hall resistance at low temperature when the Fermi energy is tuned through the bulk band gap. In contrast, at room temperature we find electron and hole conduction that can be described by a classical two-carrier model. In an intermediate temperature range ($100$ K $\leq T \leq 205$ K), where Shubnikov-de Haas oscillations and quantum Hall effect are absent, we observe a pronounced linear magnetoresistance (LMR) that develops from a classical quadratic low-field magnetoresistance (MR). Bulk band structure calculations using an eight-band $k \cdot p$ model demonstrate a transition from a conventional semiconductor to a topological insulator in the regime where the LMR is observed.

II. EXPERIMENTAL DETAILS AND SAMPLE CHARACTERIZATION AT $T = 4.2$ K

We have grown inverted HgTe QWs with (001) surface orientation by molecular beam epitaxy on a CdTe (sample S1, and S4) and on a Cd$_{0.96}$Zn$_{0.04}$Te substrate (sample S2 and S3). Details for the samples S1–S4 are given in Table T1 in the supplemental material [14]. Lithographically defined Hall-bar structures have been produced with the dimension ($L \times W$) = ($600 \times 200$) $\mu$m$^2$. All samples are equipped with a metallic Au
of $\mu = 5.6 \times 10^4 \text{ cm}^2/\text{V} \cdot \text{s}$ ($\mu = 4.6 \times 10^5 \text{ cm}^2/\text{V} \cdot \text{s}$). The fact that we can indeed tune the Fermi energy through the bulk band gap is demonstrated by measuring the Hall resistivity $\rho_{xy}$, shown in Figs. 1(c) and 1(d). Depending on $V_g$, we find a positive $\rho_{xy}$ caused by negatively charged electrons in (i) and (ii), and a negative $\rho_{xy}$ in (iii) indicating hole transport. At higher magnetic fields, we observe the quantum Hall effect for electrons, and the $\rho_{xy}$ for holes diverges. Quantization in $\rho_{xy}$ for holes occurs at lower temperatures, see Fig. 1(d) at 0.3 K owing to the higher effective mass for holes [2].

III. MAGNETOTRANSPORT AT ROOM TEMPERATURE

We now present magnetotransport at room temperature. At high temperatures we are limited to apply a high $|V_g|$ due to an increase in the leak current through the insulator. In Fig. 2, we show magnetotransport for sample S1 at $T = 300$ K. Applying a gate voltage $V_g$ at $B = 0$, we find that $\rho_{xx}$ increases with decreasing $V_g$, see inset of Fig. 2(a), indicating that we deplete electrons when decreasing the gate voltage. In Figs. 2(a) and 2(b), we illustrate $\rho_{xx}$ and $\rho_{xy}$ as a function of the magnetic field for various fixed $V_g$. For all gate voltages, $\rho_{xx}$ displays a pronounced MR and $\rho_{xy}$ shows a strong nonlinear behavior. For positive $V_g$, $\rho_{xx}$ increases quadratically as a function of $B$ but with decreasing $V_g$, we find that $\rho_{xx}$ deviates from the quadratic behavior and exhibits a small $MR$ seemingly saturating in higher magnetic fields.

Both observations point towards a system where electrons and holes coexist. Notably, the slope of $\rho_{xy}$ is first positive, indicating a dominant contribution of mobile electrons. With increasing magnetic field, the slope becomes negative due to holes with a higher concentration and lower mobility.

In Fig. 1, we present transport at $T = 4.2$ K for sample S1 and S2 with a well width of $d = 12$ nm. Figures 1(a) and 1(b) show the longitudinal resistance $R_{xx}$ at $T = 4.2$ K as a function of top-gate voltage $V_g$ for samples S1 and S2. According to band structure calculations, both QWs are inverted at 4.2 K. The $H1$ band is the conduction band and the $H2$ band is the valence band; see sketch of the QW in the inset of Fig. 1(b). The finite maximum in $R_{xx}$ is an indication for the QSHE [4]. Its value is higher than $h/2e^2$ which can be explained by inelastic scattering in large samples [7,8]. This resistance is by an order of magnitude smaller than in samples with comparable size and a 20 meV bulk band gap [7] (see band structure calculations in Fig. 4(c) and in the Supplemental Material [14]). At $V_g = 0$ (ii), both samples are $n$ conducting and sample S1 (S2) has a carrier concentration of $n = 3.7 \times 10^{11} \text{ cm}^{-2}$ ($n = 4.5 \times 10^{11} \text{ cm}^{-2}$) and a mobility top-gate, separated from the structure by an insulator made of a superlattice of Si$_3$N$_4$ and SiO$_2$, with a total thickness of 110 nm to tune the carrier concentration as a function of the applied gate voltage $V_g$. Four-probe measurements of longitudinal and transverse electrical resistances have been carried out using Stanford Research Systems SR830 Lock-In amplifiers with low constant voltage excitation. The samples were placed in an $\text{O}_2$ flow cryostat in a 33 T Bitter-type magnet.

In Fig. 1, we present transport at $T = 4.2$ K for sample S1 and S2 with a well width of $d = 12$ nm. Figures 1(a) and 1(b) show the longitudinal resistance $R_{xx}$ at $T = 4.2$ K as a function of top-gate voltage $V_g$ for samples S1 and S2. According to band structure calculations, both QWs are inverted at 4.2 K. The $H1$ band is the conduction band and the $H2$ band is the valence band; see sketch of the QW in the inset of Fig. 1(b). The finite maximum in $R_{xx}$ is an indication for the QSHE [4]. Its value is higher than $h/2e^2$ which can be explained by inelastic scattering in large samples [7,8]. This resistance is by an order of magnitude smaller than in samples with comparable size and a 20 meV bulk band gap [7] (see band structure calculations in Fig. 4(c) and in the Supplemental Material [14]). At $V_g = 0$ (ii), both samples are $n$ conducting and sample S1 (S2) has a carrier concentration of $n = 3.7 \times 10^{11} \text{ cm}^{-2}$ ($n = 4.5 \times 10^{11} \text{ cm}^{-2}$) and a mobility...
Temperature-driven transition from a ... Physical Review B 91, 205311 (2015)

Fig. 3. (Color online) Analysis of room-temperature magneto-transport with a two-carrier Drude model (symbols represent the experimental data, solid lines fits to the model): (a)–(d) MR and $\rho_{xy}$ up to $B = 30$ T for several chosen gate voltages. (e) Charge carrier concentrations and (f) charge carrier mobilities for electrons and holes at different $V_g$ as extracted from the fits. With decreasing gate voltage, the electron (hole) concentration and mobility decreases (increases).

We can extract quantitative information on the charge carrier properties using a semiclassical Drude-model with field-independent electron and hole concentrations and mobilities where we sum up the individual contributions of both electrons and holes to the conductivity tensor

$$\sigma_{xx} = \frac{n e \mu_e}{1 + \mu_e^2 B^2} + \frac{p e \mu_p}{1 + \mu_p^2 B^2}$$

$$\sigma_{xy} = \frac{n e \mu_e^2 B}{1 + \mu_e^2 B^2} - \frac{p e \mu_p^2 B}{1 + \mu_p^2 B^2}$$

and perform a fit to the experimentally measured resistivity tensor $\hat{\rho} = \hat{\sigma}^{-1}$.

The results of the simultaneous fitting of the relative magnetoresistance, defined as $MR = (\rho_{xx}(B) - \rho_{xx}(0))/\rho_{xx}(0)$, and the Hall resistance $\rho_{xy}(B)$ at four chosen fixed gate voltages are shown as the solid lines in Figs. 3(a)–3(d). The gate dependencies of the electron and hole concentrations $n$ and $p$, and the mobilities $\mu_e$ and $\mu_p$ extracted from the fits, are illustrated in Figs. 3(e) and 3(f). Therefore, transport is governed by an electron band with very mobile carriers, i.e., carriers with a small effective mass, coexisting with a hole band with a large amount of charge carriers with a large effective mass and low mobility.

IV. TEMPERATURE-DEPENDENT BAND STRUCTURE CALCULATIONS

So far, we have presented that our system is a 2D TI at low temperature whereas magnetotransport can be described within a classical two-carrier model for one electron and one hole band at room temperature as expected for a conventional semiconductor if the thermal energy $k_BT$ is larger than the band gap $E_g$. This transition can be elucidated by performing temperature-dependent bulk band structure calculations of the QW structures shown for sample S1 in Fig. 4. Additional BS
calculations for sample S2 are illustrated in the Supplemental Material \[14\]. Our calculations are based on an eight-band $\mathbf{k} \cdot \mathbf{p}$ model in the envelope function approach \[15\]. The $\mathbf{k} \cdot \mathbf{p}$ model takes into account the temperature dependence of all relevant parameters, in particular the change in the lattice constants of Ho$_{6}$-Cd$_{x}$Te and the elastic constants $C_{11}$, $C_{12}$ (bulk modulus) and $C_{44}$ with temperature \[16,17\]. The elastic constants $C_{ij}$ increase by a few % with decreasing $T$ but the ratios which enter the calculations remain constant. In Figs. 4(a)–4(d), we plot $E(k)$ for the 12 nm thick QW at different temperatures used in our experiment. At 300 K, the gap between the conduction band $E_{1}$ and the valence band $H1$ is $E_{g} \simeq 26$ meV. With decreasing temperature, the band gap $E_{g}$ considerably decreases and the 12 nm HgTe QW becomes a zero-gap SC at $T = 223$ K \[10\], see Fig. 4(b).

In this temperature range $E_{g} \ll k_{B}T$, and transport is still governed by thermally activated charge carriers within the $E1$ and $H1$ bands. Decreasing the temperature further, the BS becomes inverted and the system has an indirect bulk band gap. Thus, the BS calculations demonstrate a transition from a conventional semiconductor with normal band ordering to a 2D TI. At $T = 100$ K, see Fig. 4(c), transport is still dominated by thermally activated charge carriers but now in the $H1$ conduction and the $H2$ valence band since $E_{g} \ll k_{B}T$. At low temperatures, see Fig. 4(d) for $T = 4.2$ K, $E_{g} > k_{B}T$ and thermal activation becomes negligible, and we would observe an infinite resistance in the bulk band gap for a perfectly homogeneous gate potential if our system was not a 2D TI.

As can be seen from the calculation, the system undergoes a transition from a conventional semiconductor to a 2D TI due to the decrease of the $E1$ band in energy, highlighted in Fig. 4(f) where the temperature dependence of $E1$ and $H1$ is illustrated. For $T > 223$ K, our system is a conventional semiconductor with the conduction band $E1$ and the valence band $H1$ with a direct band gap [indirect band gap] for $k||1(1,0)$ $k||1(1,1)$.

V. REGIME OF LINEAR MAGNETORESISTANCE

Let us now draw our attention to the intermediate temperature regime where, according to the presented BS calculation, the band order is inverted but transport is still dominated by thermally activated charge carriers. In Fig. 5(a), we plot $\rho_{xx}(T)$ at $V_{g} = 0$ and observe that $\rho_{xx}$ first increases with decreasing temperature, which is characteristic for a semiconductor with thermally activated carriers. Around 130 K, $\rho_{xx}$ displays a maximum and then starts to decrease with further decreasing temperature before saturating for $T \leq 10$ K. This behavior is characteristic for a metallic system with a constant carrier concentration and an increasing mobility with decreasing temperature. In Fig. 5(b), we plot the resistivity $\rho_{xx}$ as a function of $V_{g}$ at 100, 150, and 204 K. For $T > 100$ K, $\rho_{xx}$ decreases monotonously with increasing $V_{g}$. For $T = 100$ K, we observe a maximum in $\rho_{xx}$, which we refer to as the region of charge neutrality.

A particular feature of this intermediate-temperature regime is the emergence of a strong LMR that develops from a classical quadratic low-field MR. In Figs. 5(c) and 5(d) we plot $\rho_{xx}$ and $\rho_{xy}$ at $V_{g} = 0$ as a function of $B$ for several temperatures. In the temperature range presented in Fig. 5(c), $E_{g} \ll k_{B}T$ at $V_{g} = 0$ and magnetotransport is governed by bulk electrons and holes. For high temperatures ($T > 200$ K), we observe a quadratic MR that can be perfectly modeled by a two-carrier Drude model (see also Fig. 3). For lower temperatures, however, we find two different regimes in $\rho_{xx}(B)$: a quadratic MR at low magnetic fields, as expected from the classical two-carrier Drude model, and a LMR at high $B$. Furthermore, the onset of LMR shifts continuously to lower $B$ with decreasing temperature. Interestingly, at $T = 150$ and 100 K, we observe a wide range of LMR, e.g., from 8 to 30 T at 150 K and 3 to 30 T at 100 K, respectively. The experimentally observed MR can not be described by the classical two-carrier model though the corresponding $\rho_{xy}$ traces indicate that both electrons and holes contribute.
holes still contribute to the transport. We estimate the electron concentrations from the linear increase of $\rho_{xx}$ at low $B$ and the hole concentration from the slope of $\rho_{xy}$ at high $B$, and the results are shown in Fig. 5(e). When we decrease the temperature from 300 to 100 K, $n$ and $p$ decrease by approximately a factor of two and 1.5, respectively, which can again be explained by the decrease of thermally exited carriers. We can also estimate the carrier mobilities, by limiting the two-carrier model to the classical regime at low magnetic fields. An example of a two-carrier fit is shown in Fig. 6(c) for the two-carrier model to the classical regime at low magnetic fields.

As an example for the peculiar MR below $T_c$, when the band structure is inverted, we plot $\rho_{xx}$ and $\rho_{xy}$ in Figs. 6(a) and 6(b) as a function of the magnetic field at different $V_g$, respectively. For $V_g \gtrsim 1$ V, $\rho_{xx}$ exhibits a strong positive MR at low fields that evolves into a strong linear MR with increasing $B$ as the field is increased, see also inset of Fig. 6(a), and the onset of LMR shifts to lower magnetic field with decreasing $V_g$. The linear dependence of $\rho_{xx}$ and its onset can be clearly illustrated in the first-order derivative $d\rho_{xx}/dB$, as plotted in Fig. 6(d) and (e) as a function of the magnetic field. We define a critical field $B_{crit}$ as the magnetic field corresponding to the maximum in $d\rho_{xx}/dB$, which marks the deviation from a squared dependence in the two-carrier model for low $B$. For all $V_g$, $\rho_{xx}$ is first positive due to mobile electrons and becomes negative with increasing $B$ due to the presence of holes.

VI. DISCUSSION

From the above bulk BS calculations we see that temperature induces a transition from a normal state to a topologically nontrivial state in HgTe QWs. With decreasing temperature ($T > T_c$), the gap closes, see Fig. 4(c). The conduction band exhibits a significant dependence on $k$, yielding a small effective mass $0.015 m_e < m^* < 0.04 m_e$ [2,15,16] for our QW. In contrast, the valence band is more flat pointing to a much higher effective mass ($m^* \approx 0.2 m_e$ [16]). As demonstrated in Fig. 4(f), the transition from normal to inverted band order occurs at $T_c = 223$ K. Since the thermal energy is larger than $E_g$, magnetotransport is dominated by bulk electrons and holes due to the small bulk band gap above and below $T_c$. Thus, this transition does not occur abruptly in magnetotransport as demonstrated by our experimental data in Figs. 5 and 6. However, as presented in Fig. 6(c), the classical two-carrier model fails to describe the observed MR for $B > B_{crit}$. Moreover, the peculiar MR occurs if electrons and holes with a considerable difference in carrier concentration and mobility coexist. The fact that our bulk band structure is inverted for $T < 223$ K implies that transport can also take place in helical edge states with a linear dispersion in the bulk band gap [4–7].

The observation of LMR has been reported in various systems such as bulk narrow-band gap semiconductors [18,19] and semimetals [20] as well as recently in TIs [21]. In fact, the occurrence of a strong LMR has been ascribed to surface states in three-dimensional TIs [22]. In a 2D TI (HgTe QW), LMR has been also found at low magnetic field and low temperature when the chemical potential moves through the bulk gap [23]. In contrast, since $k_B T > E_g$, LMR in our system is governed by mobile bulk electrons with low density and less mobile holes with high carrier concentration and helical edge states. From $V_g$-dependent measurements we found that the onset $B_{crit}$ of LMR shifts to lower $B$ with increasing carrier concentration of holes.

Theoretical models have also addressed the appearance of LMR [24,25]. The classical percolation model by Parish and Littlewood [24] for a nonsaturating LMR due to distorted current paths caused by disorder-induced inhomogeneities in the electron mobility cannot be applied to our system since our MBE grown samples do not show strong fluctuations in $\mu$, and it does not explain the transition from classical MR to LMR with increasing magnetic field. The quantum model, that has been proposed by Abrikosov [25], is valid for systems with a gapless linear dispersion spectrum when only the lowest Landau level remains occupied. Moreover, the energy difference between the lowest LL $E_0$ and the first LL $E_1$ should be much larger than $E_F$ and $k_B T$. We reach the quantum limit for one type of charge carriers, e.g., at $V_g = 0$ for $T = 100$ K, since $E_1 - E_0 > E_F > k_B T$, however, the LMR in our 2D system occurs in the presence of two types of charge carriers in the bulk and charge carriers in the helical edge states in
contrast to the three-dimensional model for one type of charge carrier proposed by Abrikosov [25].

Recently, MR has been theoretically investigated in two-component systems such as narrow-band semiconductors or semimetals at high temperatures [26]. For equal carrier concentrations of electrons and holes, a nonsaturating LMR has been predicted to occur in finite size at charge neutrality due to the interplay between bulk and edge contributions. At room temperature, our data in Fig. 2 shows qualitatively the expected behavior for $\rho_{xx}$ and $\rho_{xy}$ as proposed and illustrated in Ref. [26] for broken electron-hole symmetry. Yet we have shown that both $\rho_{xx}$ and $\rho_{xy}$ can also be explained within the classical two-carrier model without any contribution due to a quasiparticle density that develops near the sample edges. A satisfactory theoretical explanation of the origin of LMR for $T < T_e$ and $B > B_{c2}$, that also addresses the role of helical edge states at high temperature remains open and is certainly challenging for theoretical models in the future.

VII. CONCLUSION

We have demonstrated in bulk band structure calculations on HgTe QWs that temperature induces a transition from a semiconductor at room temperature to a TI at low temperature. Experimentally, we can distinguish between three regimes in magnetotransport: (i) transport of coexisting electrons and holes that can be described within a classical two-carrier model at room temperature, (ii) the appearance of a strong LMR for $B > B_{c1}$, and $T < T_e$, where electrons and holes still coexist, and (iii) the regime of quantized transport ($h\omega_c > k_B T$) at low temperature where we are also in the regime of the QSHE. We note that apart from inverted HgTe QWs, the only other system known to be a 2D TI is the InAs/GaSb hybrid system [27] that has been investigated at low temperature [8,9,28]. Temperature-dependent magnetotransport experiments could demonstrate whether the MR effects are unique in inverted HgTe QWs due to their bulk band structure or are a fundamental property of 2D topological insulators.

ACKNOWLEDGMENTS

This work has been performed at the HFML-RU/FOM member of the European Magnetic Field Laboratory (EMFL) and has been supported by EuroMagNET II under EU Contract No. 228043, by the DARPA MESO project through the Contract No. N66001-11-1-4105, by the German Research Foundation (DFG Grant No. HA5893/4-1 within SPP 1666, the Leibniz Program, and DFG-JST joint research project ‘Topological Electronics’), and the EU ERC-AG program (Project 3-TOP). S.W. is financially supported by a VENI grant of the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

REFERENCES


