nature COMMUNICATIONS

ARTICLE

Received 12 Jul 2013 | Accepted 1 Oct 2013 | Published 1 Nov 2013

DOI: 10.1038/ncomms3696

Evidence for massive bulk Dirac fermions in $Pb_{1-x}Sn_xSe$ from Nernst and thermopower experiments

Tian Liang[1], Quinn Gibson[2], Jun Xiong[1], Max Hirschberger[1], Sunanda P. Koduvayur[1], R.J. Cava[2] & N.P. Ong[1]

Topological surface states protected by mirror symmetry are of interest for spintronic applications. Such states were predicted to exist in the rocksalt IV–VI semiconductors, and several groups have observed the surface states in (Pb,Sn)Te, (Pb,Sn)Se and SnTe using photoemission. An underlying assumption in the theory is that the surface states arise from bulk states describable as massive Dirac states, but this assumption is untested. Here we show that the thermoelectric response of the bulk states displays features specific to the Dirac spectrum. By relating the carrier density to the peaks in the quantum oscillations, we show that the first ($N=0$) Landau level is non-degenerate. This finding provides robust evidence that the bulk states are indeed massive Dirac states. In the lowest Landau level, S_{xx} displays a striking linear increase versus magnetic field characteristic of massive Dirac fermions. In addition, the Nernst signal displays a sign anomaly in the gap-inverted phase at low temperatures.

[1] Department of Physics, Princeton University, Jadwin Hall, PO Box 708, Princeton, New Jersey 08544, USA. [2] Department of Chemistry, Princeton University, Frick Chemistry, Princeton, New Jersey 08544, USA. Correspondence and requests for materials should be addressed to N.P.O. (email: npo@princeton.edu).

The rocksalt IV–VI semiconductors have been identified by Fu *et al.*[1,2] as a novel class of insulators—the topological crystalline insulators—which display surface states that are protected by crystalline symmetry. The topological surface states in topological crystalline insulators are to be contrasted with those in the widely investigated Z2 invariant topological insulators, which are protected by time-reversal invariance[3,4]. Angle-resolved photoemission spectroscopy (ARPES) experiments have obtained evidence for the surface states in $Pb_{1-x}Sn_xSe$ (ref. 5), SnTe (ref. 6) and $Pb_{1-x}Sn_xTe$ (ref. 7).

In the alloys $Pb_{1-x}Sn_xTe$ and $Pb_{1-x}Sn_xSe$, the bulk electrons occupy four small Fermi surface (FS) pockets located at the L points in **k** space (inset, Fig. 1). The conduction band is predominantly derived from the cation Pb $6p$ orbitals, whereas the uppermost valence band is predominantly anion $4p$ (or $5p$) orbitals (ordering similar to the atomic limit)[8]. As the Sn content x increases, the system undergoes gap inversion when x exceeds a critical value x_c (refs 9–12). In samples with $x \geq x_c$, gap inversion occurs when the temperature T is lowered below the gap-inversion temperature T_{inv}. The ARPES experiments[5–7] confirm that the predicted topological surface states appear in the gap-inverted phase.

The new topological ideas invite a fresh look at the bulk states of the IV–VI semiconductors. To date, the gap inversion appears to have no discernible effect on transport properties (the resistivity, Hall coefficient and thermopower vary smoothly through T_{inv}). This is surprising given that transport probes the states at the Fermi level. Moreover, a long-standing prediction[8,13] is that the bulk electrons occupy states described by the massive Dirac Hamiltonian. This assumption underlies the starting Hamiltonian of Hsieh *et al.*[2] However, no experimental test distinguishing the massive Dirac from the Schrödinger Hamiltonian has appeared to our knowledge.

We have grown crystals of $Pb_{1-x}Sn_xSe$ ($x = 0.23$) in which the n-type carriers have high mobilities ($\mu = 114{,}000\ cm^2\,V^{-1}\,s^{-1}$ at 4 K). The low electron density ($3.46 \times 10^{17}\ cm^{-3}$) enables the quantum limit to be reached at 7.7 T (measurements reveal that holes are absent). In addition to resistivity, we have used both thermopower and the Nernst effect to probe the states in fields up to 34 T. Surprisingly, the Nernst signal is observed to change its sign at T_{inv}. To date, this appears to be the only transport or thermodynamic quantity that is strongly affected by gap inversion.

In a thermal gradient $-\nabla T \parallel \hat{\mathbf{x}}$ and an applied magnetic field $\mathbf{B} \parallel \hat{\mathbf{z}}$, the diffusion of carriers produces an electric field **E**, which is expressed as the thermopower signal $S_{xx} = -E_x/|\nabla T|$ and the Nernst signal $S_{xy} = E_y/|\nabla T|$. In the semiclassical regime, the Mott relation[14] simplifies S_{xx} and S_{xy} to the form (see Methods section)

$$S_{xx}(B) = \mathcal{A}\left(\frac{\sigma^2}{\sigma^2+\sigma_{xy}^2}\mathcal{D} + \frac{\sigma_{xy}^2}{\sigma^2+\sigma_{xy}^2}\mathcal{D}_H\right) \quad (1)$$

$$S_{xy}(B) = \mathcal{A}\frac{\sigma\sigma_{xy}}{\sigma^2+\sigma_{xy}^2}(\mathcal{D}_H - \mathcal{D}), \quad (2)$$

where $\mathcal{A} = \pi^2 k_B^2 T/3e$. The dependence on B appears only in the conductivity matrix elements $\sigma_{ij}(B)$ (for brevity, we write $\sigma \equiv \sigma_{xx}$). The parameters $\mathcal{D} = \partial \ln\sigma/\partial\zeta$ and $\mathcal{D}_H = \partial \ln\sigma_{xy}/\partial\zeta$ are

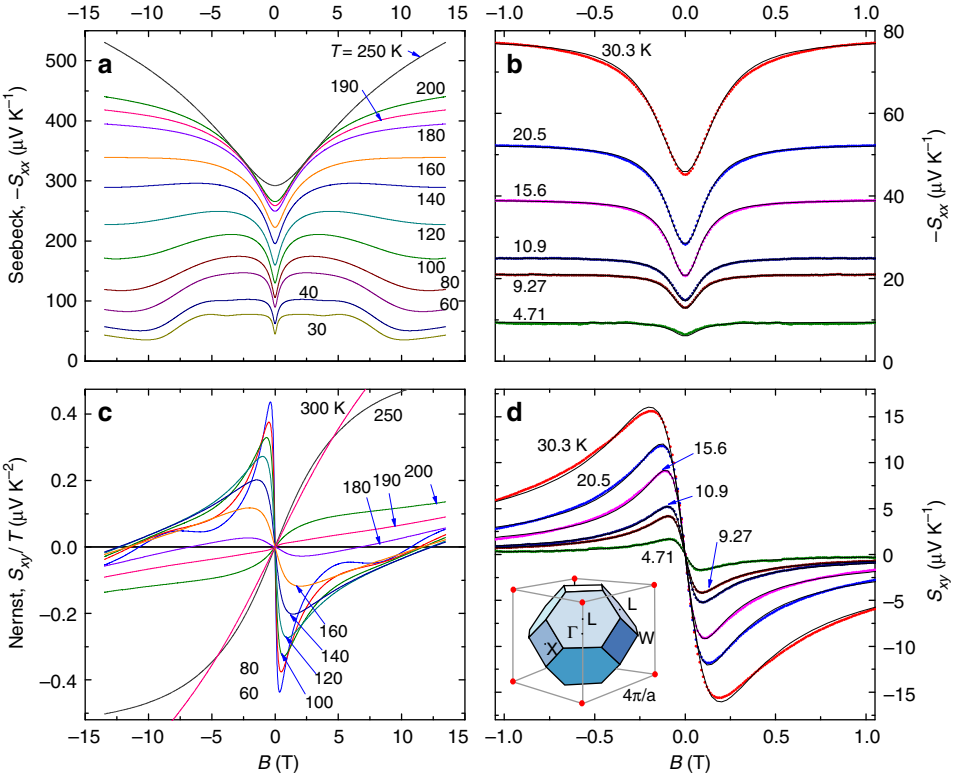

Figure 1 | Thermopower and Nernst effect in a topological crystalline insulator. The panels provide an overview of how the thermopower S_{xx} and Nernst signal S_{xy} vary with field B in a high-mobility crystal (sample 1) of $Pb_{1-x}Sn_xSe$ with $x = 0.23$. (**a**, **b**) Curves of S_{xx} versus B at selected T (sample 1). At each T, the V-profile bracketing $B = 0$ reflects the rapid crossover from small μB to large μB regime. (**c**) The Nernst signal S_{xy}/T from 60 to 300 K. The sharp peaks reflect the semiclassical response. An anomalous sign change occurs at $T_{inv} = 180$ K. (**b**,**d**) Fits equations 1 and 2 (thin curves) to S_{xx} and S_{xy} at low B. For S_{xy} (**d**), we have had to invert the sign. At 30.3 K, the best-fit values of μ, $\mathcal{D}$ and $\mathcal{D}_H$ are $51{,}404\ cm^2\,V^{-1}\,s^{-1}$, $61.5\ eV^{-1}$ and $104.6\ eV^{-1}$, respectively. At 4.71 K, the corresponding values are $113{,}250 cm^2\,V^{-1}\,s^{-1}$, $52.3\ eV^{-1}$ and $81.3\ eV^{-1}$. The inset shows the L (111) points on the hexagonal faces of the Brillouin Zone.

independent of the mobility μ (ζ is the chemical potential). Equation 1 describes the crossover in S_{xx} from $\mathcal{AD}$ (at $B=0$) to $\mathcal{AD}_H$ when $\mu B >> 1$. Correspondingly, S_{xy} increases linearly from 0 to peak at the value $1/2\mathcal{A}(\mathcal{D}_H - \mathcal{D})$ at $B = 1/\mu$ before falling as $1/B$ when $\mu B >> 1$. For n-type carriers, both $\mathcal{A}$ and σ_{xy} are negative, and $S_{xx} < 0$. From equation 2, the Nernst signal S_{xy} is positive if $\mathcal{D}_H > \mathcal{D}$ (we discuss the sign convention in the Methods section). In terms of the exponents β and β_H defined by $\sigma(E) \sim E^{\beta}$ and $\sigma_{xy} \sim E^{\beta_H}$, we have $\mathcal{D} = \beta/E_F$ and $\mathcal{D}_H = \beta_H/E_F$.

Even for one-band systems, equations 1 and 2 have not received much experimental attention, possibly because real materials having only a single band of carriers with a low density (and high mobility) are rare. The analysis of S_{ij} is complicated by the extreme anisotropy of the FS pockets in many semi-metals. For recent Nernst measurements on Bi and Bi_2Se_3, see refs 15,16. For results on S_{ij} in graphene, see refs 17–19. The angular variation of the SdH period in Bi is investigated in refs 20,21. In the annealed crystals of $Pb_{1-x}Sn_xSe$ (with $x \simeq x_c$) investigated here, our measurements reveal only n-type carriers (holes are absent). The FS pockets are very small (the quantum limit is reached at 7.7 T) and nearly isotropic (as shown by field-tilt measurements). We find that equations 1 and 2 provide a very good fit to S_{xx} and S_{xy} in the semiclassical regime.

Here we provide evidence that the bulk electronic states at the L points are in fact massive Dirac states. A characteristic feature of the massive Dirac spectrum is that in a magnetic field, the lowest Landau level (LL) is non-degenerate with respect to the spin degrees, whereas all higher LLs are doubly degenerate. Knowing the carrier density, we show that the field at which the chemical potential jumps to the lowest LL accurately determines its spin degeneracy to be 1. This confirms the starting assumption of ref. 2. The unusual thermoelectric response is also investigated deep in the quantum limit. In addition, we show that the sign of the Nernst signal is anomalous (relative to standard Boltzmann theory) within the gap-inverted phase.

Results

Semiclassical regime. Figure 1a and 1b plots curves of the thermopower S_{xx} versus B for selected T. From 250 to 160 K, the dominant feature is the rapid increase in weak B followed by saturation to a B-independent plateau at large B. As noted, the Nernst signal (shown as S_{xy}/T in Fig. 1c) changes from positive to negative as T is decreased below 180 K (identified with T_{inv}).

As shown in Fig. 1b, the curves of S_{xx} versus B fit very well to equation 1 in the semiclassical regime ($|B| < 1$T). Likewise, below 100 K the curves of S_{xy} also fit well to equation 2 up to an overall sign (Fig. 1d). Although the fit parameters (μ, $\mathcal{D}$, $\mathcal{D}_H$) for S_{xx} are independent of those for S_{xy}, we find that they agree with each other (at the level of $\pm$2%) below 60 K (see Methods section). At each T, the two curves, $S_{xx}(B)$ and $S_{xy}(B)$, are described by just three parameters. This provides a potent self-consistency check of equations 1 and 2. As a further test, we have also fitted the measured conductivity tensor $\sigma_{ij}(B)$ and obtained similar values for μ below 100 K (Methods). By and large, the close fits to both tensors S_{ij} and σ_{ij} demonstrate that we have one band of carriers below 100 K.

The semiclassical expressions are no longer valid when quantum oscillations appear at higher B. In particular, the giant step at 7.7 T in the curves at 30 and 40 K (Fig. 1a) is a relic of the quantum regime that remains resolvable up to 100 K. The step has a key role in the discussion later. In the opposite extreme above 100 K, the two sets of fit parameters begin to deviate. The disagreement is especially acute near 180 K, where S_{xy} changes sign. We reason that the one-band assumption breaks down because of strong thermal activation of holes as the gap closes and re-opens across T_{inv}. The evidence comes from the T dependence of the Hall density $n_H = B/e\rho_{yx}$ (solid circles in Fig. 2a). Whereas n_H is nearly T-independent below 80 K, in agreement with the one-band model, it deviates upwards above 180 K. Thermal activation of a large population of holes leads to partial cancellation of the Hall E-field and a reduction in $|\rho_{yx}|$.

In Fig. 2a, we have also plotted the zero-H thermopower $S \equiv S_{xx}(0)$ to bring out its nominally T-linear variation below 100 K (bold curve). The large value of the slope $S(T)/T = 1.41$ μV K^{-2} implies an unusually small E_F.

As discussed, the Nernst signal changes sign at $T_{inv} = 180$ K. The T dependence of its initial slope dS_{xy}/dB ($B \to 0$) is displayed in Fig. 2b. From the fits to equations 1 and 2, we may address the interesting question whether the sign anomaly occurs in the gap-inverted phase ($T < T_{inv}$) or in the uninverted phase. On both sides of T_{inv}, the fits of S_{xx} imply $\mathcal{D}_H > \mathcal{D}$ (that is, $|S_{xx}|$ always increases as μB goes from 0 to values $>> 1$). As $S_{xy} \sim (\mathcal{D}_H - \mathcal{D})$, we should observe a positive S_{xy}. Hence, the sign anomaly occurs in the gap-inverted phase (in Fig. 1d, we multiplied the curves by an overall minus sign). The sign of the Nernst signal below T_{inv} disagrees with that inferred from equations 1 and 2, despite the close fit. Further discussion of the sign anomaly is given below (see Discussion section). However, we note that the sign of S_{xy} is

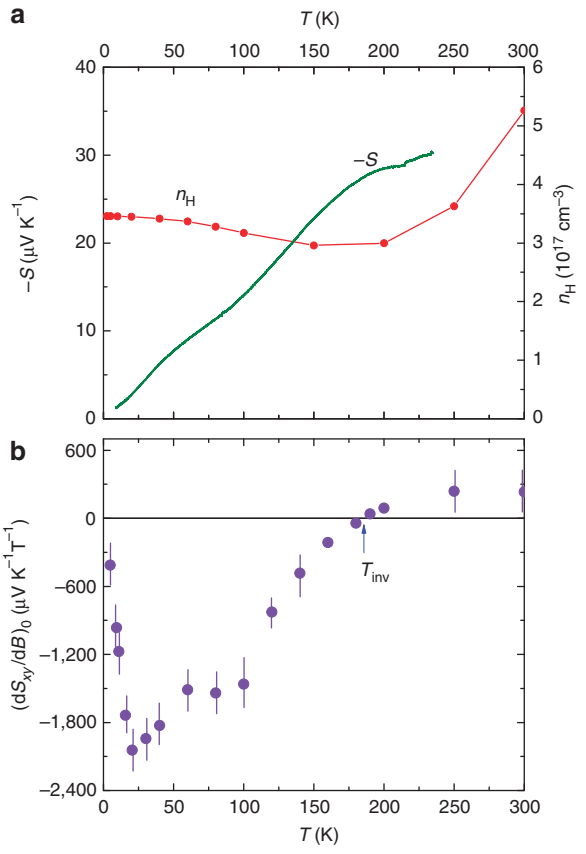

Figure 2 | Temperature dependence of Hall density, thermopower and Nernst slope. (**a**) The T dependence of the Hall density $n_H = B/\rho_{yx}e$ inferred from the Hall resistivity ρ_{yx} and the zero-B thermopower $S(T) \equiv S_{xx}(T, B=0)$ in $Pb_{1-x}Sn_xSe$ ($x = 0.23$). The Hall signal is n-type at all T. Below 20 K, n_H equals 3.46×10^{17} cm^{-3} (sample 1). n_H increases significantly above 200 K signalling thermal activation of holes across the band gap. (**b**) The T dependence of the initial slope of the Nernst signal dS_{xy}/dB ($B \to 0$) to show the sign change at T_{inv}. The error bars correspond to 1 s.d. in determining dS_{xy}/dB in the limit $B \to 0$.

independent of the carrier sign. As seen in Fig. 2a, both S and n_H vary smoothly through T_{inv} without a sign change.

Quantum oscillations. As shown in Fig. 3, oscillations in S_{xx} and S_{xy} grow rapidly below 60 K to dominate the weak-B semiclassical profile. The most prominent feature in S_{xx} is the large step decrease at the field $B_1 = 7.7$ T (at which the chemical potential ζ jumps from the $N = 1$ LL to the $N = 0$ LL). In the Nernst curves, plotted as S_{xy}/T in Fig. 3b, the quantum oscillations are more sharply resolved. As S_{xy} is the off-diagonal term of the tensor S_{ij}, its maxima (or minima) are shifted by 1/4 period relative to the extrema of the diagonal S_{xx} (analogous to the shift of σ_{xy} relative to σ). This shift is confirmed in Fig. 4a, which plots the traces of S_{xx} and S_{xy} versus $1/B$. For the analysis below, we ignore the weak spin splitting, which is resolved in the $N = 1$ LL (and barely in $N = 2$).

Figure 4b shows the index plot of $1/B_n$ (inferred from the maxima in $|S_{xx}|$ and S_{xy}) plotted versus the integers n. From the slope of the line, we derive the FS section $S_F = 5.95\,\mathrm{T} = 5.67 \times 10^{16}\,\mathrm{m}^{-2}$. Assuming a circular cross-section, we have $k_F = 0.0134\,\text{Å}^{-1}$. The electron concentration per FS pocket is then $n_e = k_F^3/3\pi^2 = 8.2 \times 10^{16}\,\mathrm{cm}^{-3}$. As there are four pockets, the total carrier density is $4n_e = 3.28 \times 10^{17}\,\mathrm{cm}^{-3}$, in good agreement with the Hall density n_H at 4 K ($3.46 \times 10^{17}\,\mathrm{cm}^{-3}$).

Using sample 2, we have tracked the variation of the SdH period versus the tilt angle θ of **B**. Figure 5a plots the fields B_1 and B_2 versus θ (**B** is rotated in the y–z plane). Here, B_1 and B_2 are the fields at which ζ jumps from $N = 1 \to 0$ and from $N = 2 \to 1$, respectively. To our resolution, the SdH period is nearly isotropic. The fields B_1 and B_2 are also independent of tilt angle when **B** is rotated in the x–y plane. This justifies treating the FS pockets as nominally spherical.

The $N = 0$ Landau level. We next address the question whether the bulk states in the inverted phase are Dirac fermions or Schrödinger electrons. The two cases differ by a distinctive feature in their LL spectrum that is robust against small perturbations. In the quantum limit, the massive Dirac Hamiltonian exhibits an interesting twofold difference in degeneracy between the $N = 0$ and $N = 1$ levels. Wolff[22] considered a three-dimensional (3D) massive Dirac Hamiltonian with spin–orbit interaction but no Zeeman energy term. More recently, Serajedh *et al.*[23] included the

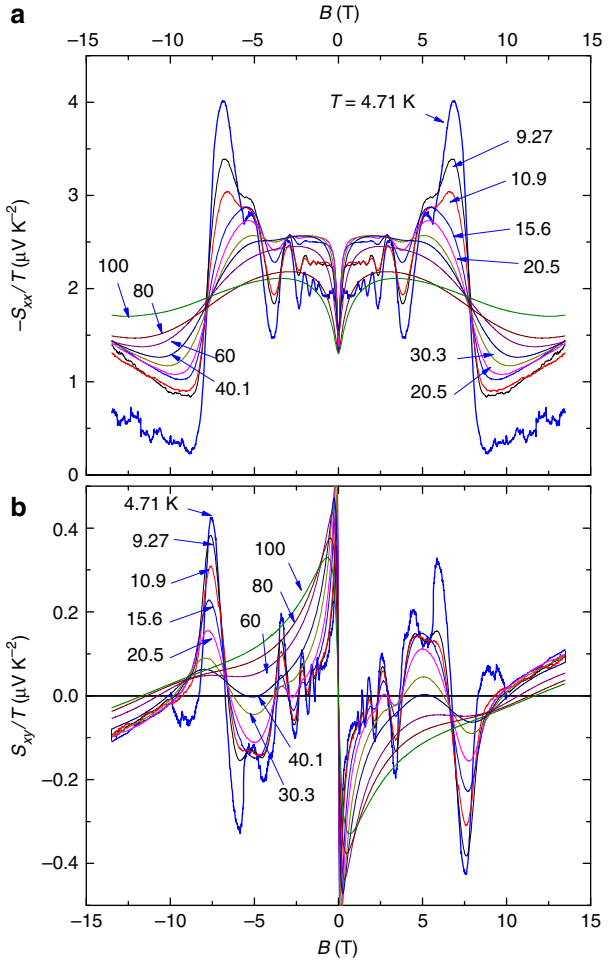

Figure 3 | Quantum oscillations in the thermopower and Nernst signal of PbSnSe. (**a**) Curves of S_{xx}/T versus B at $T = 4.71$ to 100 K of $Pb_{1-x}Sn_xSe$ ($x = 0.23$, sample 1). $|S_{xx}/T|$ displays a maximum when ζ is at the DOS maximum in each LL. At 4.71 K, the $N = 1$ LL (5–7 T) displays a weak spin splitting. The giant step at 7.7 T occurs when ζ enters the $N = 0$ LL. (**b**) Curves of S_{xy}/T for the same T. The sharp resonance-like peaks at low fields are the semiclassical response of large μ electrons. Below 30 K, they are eclipsed by strong quantum oscillations.

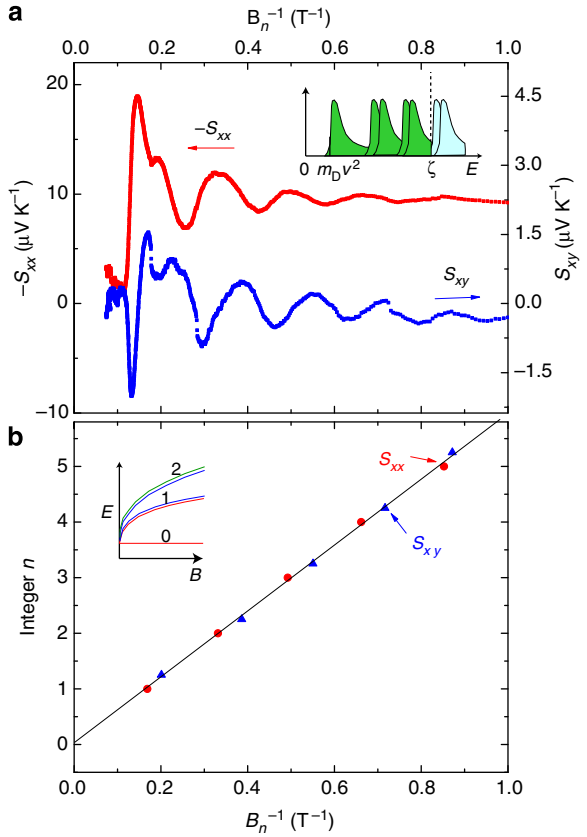

Figure 4 | Analysis of the quantum oscillations. (**a**) Comparison of curves of S_{xx} and S_{xy} versus $1/B$ at 4.71 K in $Pb_{1-x}Sn_xSe$ ($x = 0.23$). The maxima in S_{xy} are shifted by a $\frac{1}{4}$ period relative to the maxima in S_{xx}. The $N = 1$ LL shows a weak spin splitting. The sketch (inset) shows the peaks in the DOS of each LL for 3D massive Dirac fermions. (**b**) Index plot of B_n corresponding to the maxima in $|S_{xx}|$ (solid circles) and S_{xy} (triangles) versus the integers n. The straight line is the relation $\mathcal{S}_F = 2\pi\,(n+\gamma)/l_B^2$ where $\mathcal{S}_F$ is the FS section and γ the Onsager phase. The maxima in S_{xy} are shifted by $\frac{1}{4}$ in n. From the slope, we infer the Fermi wavevector $k_F = 0.0134\,\text{Å}^{-1}$ and $n_e = 8.20 \times 10^{16}\,\mathrm{cm}^{-3}$ (per valley). The inset shows the LL energy E versus B in the massive Dirac spectrum for $k_z = 0$ (ref. 22).

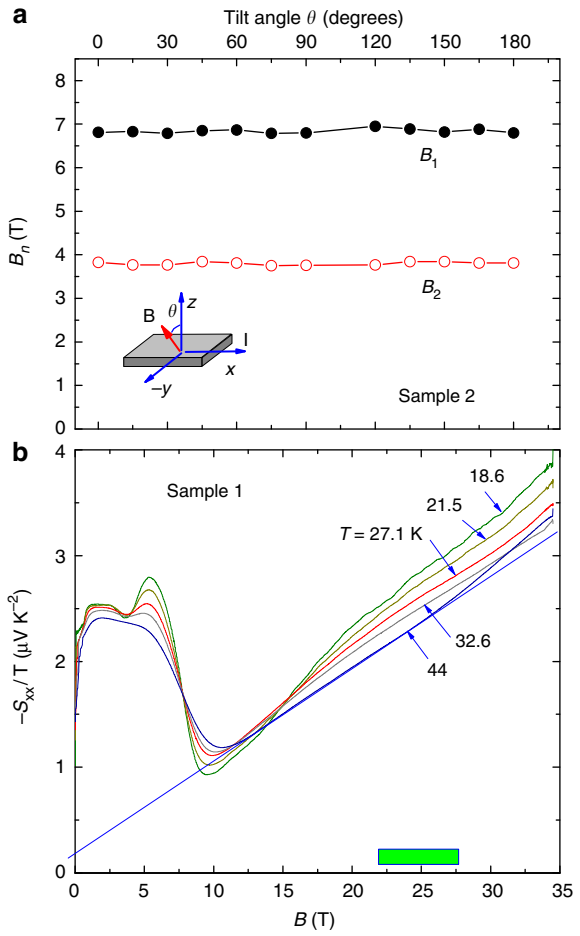

Figure 5 | Dependence of SdH period on field-tilt angle and high-field thermopower. (**a**) The dependence of the transition fields B_1 and B_2 versus tilt angle θ of **B** in sample 2 ($x=0.23$) inferred from magnetoresistance (B_1 and B_2 are the fields at which ζ jumps from LL with $N=1\rightarrow0$ and $N=2\rightarrow1$, respectively). Within our resolution, no angular dependence of B_1 and B_2 is observed. **B** is rotated in the *y*–*z* plane (sketch in inset). (**b**) High-field measurements of S_{xx}/T to 34 T at several T in sample 1. The *B*-linear dependence smoothly extends through the region 22–28 T where the transition (0, −)→(0, +) should have appeared. The interval 22–28 T is indicated by the horizontal green bar.

Zeeman energy term as well as a Rashba term in the massive two-dimensional (2D) Dirac Hamiltonian. Other 3D massive Dirac cases are discussed by Bernevig and Hughes[24]. All these authors find that $N=0$ LL is non-degenerate with respect to spin degrees, whereas the LLs with $N\neq0$ are doubly spin degenerate (we discuss in Methods section a pedagogical example, which shows that this anomaly is related to the conservation of states). By contrast, for the Schrödinger case, all LLs are doubly degenerate.

In $Pb_{1-x}Sn_xSe$, the ability to measure accurately both n_e and the 'jump' field B_1 provides a crisp confirmation of this prediction. The energy of the N^{th} LL is $E_{N,k_z}=\sqrt{(m_Dv^2)^2+(\sqrt{2N}\hbar v/\ell_B)^2+(\hbar vk_z)^2}$, where m_D is the Dirac mass and $\ell_B=\sqrt{\hbar/eB}$ the magnetic length[24]. At B_1, E_F lies just below the bottom of the $N=1$ LL so that all the electrons are accomodated in the $N=0$ LL. Integrating the density of states (DOS) for one spin polarization in the $N=0$ LL from m_Dv^2 to E_F, we find (see Methods section)

$$n_{e\uparrow}=\sqrt{2}/(2\pi^2\ell_B^3). \quad (3)$$

Ignoring the small spin splitting, we equate B_1 with 7.7 T. Equation 3 then gives $n_{e\uparrow}=9.0\times10^{16}\,cm^{-3}$, which agrees within 10% with the measured n_e (the agreement is improved if we correct for spin splitting). All the electrons are accomodated by an $N=0$ LL that is non-degenerate, in agreement with the prediction for massive Dirac states[22–24], but disagreeing with the Schrödinger case by a factor of 2. As the singular spin degeneracy of the $N=0$ LL cannot be converted to a double degeneracy, the experiment uncovers a topological feature of the bulk states that is robust. As predicted in refs 22–24, the $N=0$ LL has only one spin state (0, +); the spin-down partner (0, −) is absent.

To check this further, we extended measurements of S_{xx} to 34 T to search for the transition from the sublevel (0, −) to (0, +) (which should occur if $N=0$ LL were doubly degenerate). From extrapolation of the spin split $N=1$ and $N=2$ LLs, we estimate that the transition (0, −)→(0, +) should appear in the interval 22–28 T. As shown in Fig. 5b, the measured curves show no evidence for this transition to fields up to 34 T.

Finally, we note an interesting thermopower feature in the quantum limit. At fields above B_1, S_{xx} displays a *B*-linear profile that extends to 34 T (Fig. 5). The *B*-linear behaviour is most evident in the curve at 44 K. As T is decreased to 18.6 K, we resolve a slight downwards deviation from the linear profile in the field interval 10–20 T. The *B*-linear profile appears to be a characteristic property of massive Dirac fermions in the quantum limit. We discuss below a heuristic, semiclassical approach that reproduces the observed profile.

Discussion

We summarize the electronic parameters inferred from our experiment and relate them to ARPES measurements.

As noted, the FS section derived from the index plot (Fig. 4b) corresponds to a total electron density $4n_e=3.28\times10^{17}\,cm^{-3}$, in good agreement with the Hall density n_H at 4 K ($3.46\times10^{17}\,cm^{-3}$).

We may estimate E_F from the slope of the thermopower $S(T)/T=1.41\,\mu V\,K^{-2}$. Using the Mott expression $S(T)=(\pi^2/3)(k_B/e)(k_BT/E_F)\beta$, we find for the Fermi energy $E_F=17.0\,\beta$ meV. For the massive Dirac dispersion, we have $n_e\sim k_F^3$, which implies that β has the minimum value 3 (if the mobility increases with E, β is larger). Using the lower bound, $\beta=3$, S/T gives $E_F=51$ meV.

These numbers may be compared with ARPES measurements. We estimate the Fermi velocity from the expression $v\simeq E_F/\hbar k_F$ (valid when $E_F\gg m_Dv^2$ with m_D the Dirac mass). Using our values of E_F and k_F, we find $v=5.74\times10^5\,m\,s^{-1}$ as the lower bound. Although ARPES experiment cannot resolve v in the conduction band, the ARPES estimate[5] for the hole band velocity is $5.6\times10^5\,m\,s^{-1}$, in good agreement with our lower bound. It is likely to be that the conduction band has a higher velocity (which would then require $\beta>3$).

One of our findings is that gap inversion changes the sign of the Nernst signal. As the energies of states involved in gap inversion are very small, the resulting dispersion can be hard to resolve by ARPES measurement. Transport quantities would appear to be more sensitive to these changes. As noted, however, most transport quantities are either unaffected or only mildly perturbed. The Hall effect and thermopower are unchanged in sign across T_{inv} (Fig. 2a). Although n_H shows a gradual increase, this is largely attributed to thermal activation of holes across a reduced gap for $T>T_{inv}$. Hence, the dramatic sign change observed in S_{xy} stands out prominently; its qualitative nature may provide a vital clue.

It has long been known[10] that in the lead rocksalt IV–VI semiconductors, the energy gap E_g undergoes inversion as the Sn content x increases from 0. Moreover, within a narrow range of x,

gap inversion is also driven by cooling a sample (the critical temperature is x dependent within this interval). Strauss[10] performed early optical transmission measurements of E_g in a series of single-crystal films of $Pb_{1-x}Sn_xSe$, with x ranging from 0 to 0.35. For $x=0.25$, he reported that E_g closes at 195 K. A slight interpolation of his data shows that at our doping $x=0.23$, E_g should vanish at 179 K, remarkably close to our $T_{inv}=180$ K. The recent ARPES measurements of Dziawa *et al.*[5] is consistent with E_g closing between 100 and 200 K. Given the ARPES resolution, these results are all consistent with our inference that our T_{inv} corresponds to the gap inversion temperature. Hence, we reason that the Nernst signal changes sign either at, or very close to, the gap inversion temperature. The inverted sign of S_{xy} below 180 K in Fig. 1c,d occurs in the gap-inverted phase. We refrain from making the larger claim that this is also the topological transition because we are unable to resolve the surface states in our experiments.

The fits of S_{xy} to equation 2 (Fig. 1d) shows that the curves below 100 K are well described by the Boltzmann–Mott expression assuming a single band of carriers, but there is an overall sign disagreement. Despite the sign problem, the analysis singles out the physical factors that fix the sign and delineates the scope of the problem. For example, reversing the sign of both β and β_H inverts the sign of S_{xy}, but also that of S_{xx}. Alternately, one might try reversing the signs of β and β_H, and e simultaneously. This will invert the sign of S_{xy} but leave S_{xx} unchanged. However, ρ_{yx} is forced to change sign.

The analysis assumes that in the gap-inverted phase, the FS is simply connected. This may not be valid. Gap inverion may lead to the existence of a small pocket surrounded by a larger FS sheet (topologically similar to the FS of the 'giant Rashba' material BiTeI[25]). As the small pocket dominates the thermoelectric response, the Nernst effect may be detecting this novel situation. These issues will be left for future experiments.

We may attempt to understand the striking B-linear profile of S_{xx}/T in Fig. 5 using a semiclassical approach. In $N=0$ LL, the long-lived quasiparticles complete a large number of cyclotron orbits between scattering events (for example, from $\mu B\sim 220$, we estimate this number is $\sim$35 at 20 T). The scattering results in the drift of the orbit centres $\mathbf{X}$ in a direction transverse to the applied $-\nabla T$. Ignoring the fast cyclotron motion, we may apply the Boltzmann equation to $\mathbf{X}$. The thermopower is then given by the high-B limit of equation 1, $S_{xx}(T,H)\rightarrow \mathcal{A}\beta_H'/E_F$, where E_F is now measured from the bottom of $N=0$ LL and β_H' differs from the weak-field β_H. In this picture, the B dependence of S_{xx} arises solely from how E_F changes with B.

For $B>B_1$, only $N=0$ LL is occupied. From equation 14 (Methods), we have the relation between E_F, n_e and B, viz.

$$E_F^2=(m_D v^2)^2+\frac{\mathcal{P}^2}{B^2},\quad \left(\mathcal{P}=\frac{2\pi^2\hbar^2 v n_e}{g_s e}\right). \tag{4}$$

In the limit $E_F\gg m_D v^2$, we obtain the relation $E_F\sim 1/B$. This immediately implies that S_{xx}/T increases linearly with B as observed. Setting $g_s=1$, we derive from equation 4 the rate of increase

$$\frac{\partial S_{xx}/T}{\partial(B)}=\frac{k_B^2}{6\hbar^2}\frac{\beta_H'}{v n_e} \tag{5}$$

Repeating this calculation for the Schrödinger case, we get instead $S_{xx}/T\sim B^2$.

From Fig. 5, the thermopower slope $\partial(S_{xx}/T)/\partial(B)=8.71\times 10^{-8}\,\mathrm{V\,K^{-2}\,T^{-1}}$. Using the above values of v and n_e in equation 5, we find $(\partial S_{xx}/T)/\partial(B)=6.1\,\beta_H'\times 10^{-8}\,\mathrm{V\,K^{-2}\,T^{-1}}$. The value of β_H' is not known. Comparison of the calculated slope with experiment suggests $\beta'_H\sim 1.5$. Hence, this back-of-the-envelop estimate can account for the rate at which S_{xx}/T increases with B.

Methods

Semiclassical fits to S_{xx} and S_{xy}. In the presence of a magnetic field $\mathbf{B}$, an electric field $\mathbf{E}$ and a temperature gradient $-\nabla T$ (in an infinite medium), the total current density is given by[14] $\mathbf{J}=\boldsymbol{\sigma}\cdot\mathbf{E}+\boldsymbol{\alpha}\cdot(-\nabla T)$. Here, σ_{ij} is the conductivity tensor and α_{ij} is the thermoelectric tensor. Setting $\mathbf{J}=0$ (for a finite sample) and solving for $\mathbf{E}$, we have $\mathbf{E}=-\boldsymbol{\rho}\cdot\boldsymbol{\alpha}\cdot(-\nabla T)$, with $\boldsymbol{\rho}=\boldsymbol{\sigma}^{-1}$ the resistivity tensor.

In the geometry with $\mathbf{B}\parallel\hat{z}$ and $-\nabla T\parallel\hat{x}$, the components of the E-field (for an isotropic system) are

$$E_x/\mid\nabla T\mid=-(\rho_{xx}\alpha_{xx}+\rho_{yx}\alpha_{xy}) \tag{6}$$

$$E_y/\mid\nabla T\mid=\rho_{xx}\alpha_{xy}-\rho_{yx}\alpha_{xx}. \tag{7}$$

The thermoelectric tensor S_{ij} is given by $E_i=S_{ij}\partial_j T$ ($S_{xx}>0$ for hole carriers and $S_{xy}>0$ if $E_y>0$ when $H_z>0$).

The Mott relation[14],

$$\alpha_{ij}=\mathcal{A}\left[\frac{\partial\sigma_{ij}}{\partial\varepsilon}\right]_\zeta,\quad \left(\mathcal{A}=\frac{\pi^2}{3}\frac{k_B^2 T}{e}\right), \tag{8}$$

where k_B is Boltzmann's constant, e is the elemental charge and ζ the chemical potential, has been shown to hold under general conditions, for example, in the quantum Hall Effect (QHE)[26,27]. Using equation 8, equations 6 and 7 reduce to equations 1 and 2, respectively.

The fits of S_{ij} to these equations displayed in Fig. 1d were carried out using the one-band, Boltzmann–Drude expressions for the conductivity tensor, viz.

$$\sigma_{xx}(B)=N_e e\mu/(1+\mu^2B^2), \tag{9}$$

$$\sigma_{xy}(B)=N_e e\mu^2B/(1+\mu^2B^2), \tag{10}$$

where the total carrier density N_e is $4n_e$ (n_e is the density in each of the FS pocket at the L points).

In the geometry with $\mathbf{B}\parallel\hat{z}$ and $-\nabla T\parallel\hat{x}$, we define the sign of the Nernst signal to be that of the y component of the E-field E_y. More generally, if $\mathbf{E}_N$ is the E-field produced by the Nernst effect, the sign of the Nersnt signal is that of the triple product $\mathbf{E}_N\cdot\mathbf{B}\times(-\nabla T)$. This agrees with the old convention based on 'Amperean current'[28] and with the one adopted for vortex flow in superconductors[29].

At each T, we have fitted the measured curves of S_{xx} and S_{xy} versus B to equations 1 and 2 using equations 9 and 10 for the conductivity tensor. The separate fits of S_{xx} and S_{xy} yield two sets of the parameters μ, $\mathcal{D}$ and $\mathcal{D}_H$, which are displayed in Fig. 6 (solid triangles and open circles, respectively). The three-parameter fit places strong constraints on the curves of S_{xx} and S_{xy}. Disagreement between the two sets signals that the one-band model is inadequate.

Below 100 K, the two sets agree well, whereas closer to T_{inv} they begin to deviate. The reason is that equation 2 cannot account for the change of sign in the Nernst signal given the relative magnitudes of $\mathcal{D}$ and $\mathcal{D}_H$ fixed by the curves of S_{xx}. Above 200 K, the two sets are inconsistent because thermal excitations of holes across the small band gap is important at elevated T and the one-band assumption becomes inadequate. This is evident in the onset above 200 K of significant T dependence in the Hall density n_H (see Fig. 2a).

We remark that $S_{xx}=V_x/\delta T$ is directly obtained from the observed voltage difference V_x and the temperature difference δT between longitudinal electrical contacts (their spatial separation L_x is immaterial). However, for the Nernst signal, we have $S_{xy}=(V_y/\delta T)(L_x/L_y)$, where L_y is the spatial separation between the transverse contacts. Hence, the aspect ratio L_y/L_x is needed to convert the observed Nernst voltage V_y to S_{xy}. The ratio L_y/L_x is measured to be 4 ± 0.4. The fits are improved significantly if this value is refined to 4.20, which we adopt for the curves at all T.

Fits to equation 9 and 10 of the conductivity tensor measured in the same sample are shown in Fig. 7 for weak B at selected T from 5 to 150 K. The fits yield values of the mobility μ similar to those shown in Fig. 6a. The inferred carrier density N_e is also similar to the measured Hall density n_H.

Indexing the quantum oscillations. For the 3D systems, one identifies the index field B_n as the field at which the DOS displays a sharp maximum (diverging as $[E-(n+\frac{1}{2})\hbar\omega_c]^{-\frac{1}{2}}$ in the absence of disorder). From the quantization rule for areas in k-space, B_n is related to the FS cross-section $\mathcal{S}_F$ as

$$\mathcal{S}_F=\frac{2\pi}{\ell_B^2}(n+\gamma), \tag{11}$$

where $\ell_B=\sqrt{\hbar/eB}$ and γ (the Onsager phase) is $\frac{1}{2}$ for Schödinger electrons. The plot in Fig. 3b follows equation 11. From its slope, we obtain $\mathcal{S}_F$. The intercept γ is close to zero in Fig. 3b. We will discuss γ elsewhere.

We note that, in the 2D systems in the QHE regime, the index field is the field at which the chemical potential ζ falls between adjacent LLs, where the DOS vanishes, and the Hall conductance displays a plateau. The difference between 2D and 3D systems arises because the integer n counts the number of edge states in the QHE

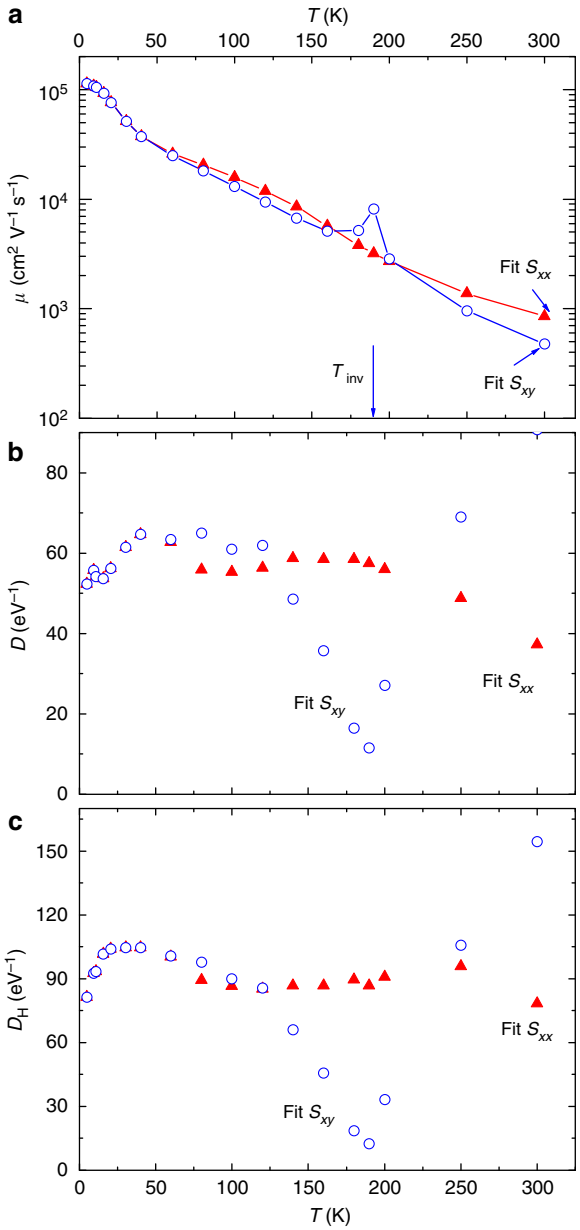

Figure 6 | Variation of fit parameters versus temperature. The T dependence of the parameters μ (**a**), $\mathcal{D}$ (**b**) and $\mathcal{D}_H$ (**c**) obtained from best fits to the curves of S_{xx} (solid triangles) and S_{xy} (open circles). The fits are most reliable below 100 K where the values obtained from fitting S_{xx} and S_{xy} are in agreement. Above 100 K, disagreement between the two sets is significant, especially close to $T_{inv} = 180$ K. Above 200 K, the one-band model is no longer valid.

case, whereas n indexes the DOS peaks in the 3D case. One needs to keep this in mind in interpreting γ.

We have verified that the slope in Fig. 4b is insensitive to the tilt angle θ of **B** relative to the crystalline axes. As shown in Fig. 5a, the SdH period is virtually independent of θ within the experimental uncertainties, consistent with negligible anisotropy in the small FS pockets. The good agreement between $\mathcal{S}_F$ and n_H (Hall density) at 5 K is also evidence for a negligible anisotropy.

Spin degeneracy in $N = 0$ LL. Knowledge of the field B_1 (the transition from $N = 1$ to the $N = 0$ LL) and the electron density per valley n_e suffices to determine the spin degeneracy of the $N = 0$ LL.

For the 3D Dirac case[24], the energy in the Nth LL is

$$E_{N,k_z} = \sqrt{(m_D v^2)^2 + \left(\frac{\sqrt{2N}\hbar v}{\ell_B}\right)^2 + (\hbar v k_z)^2}, \quad (12)$$

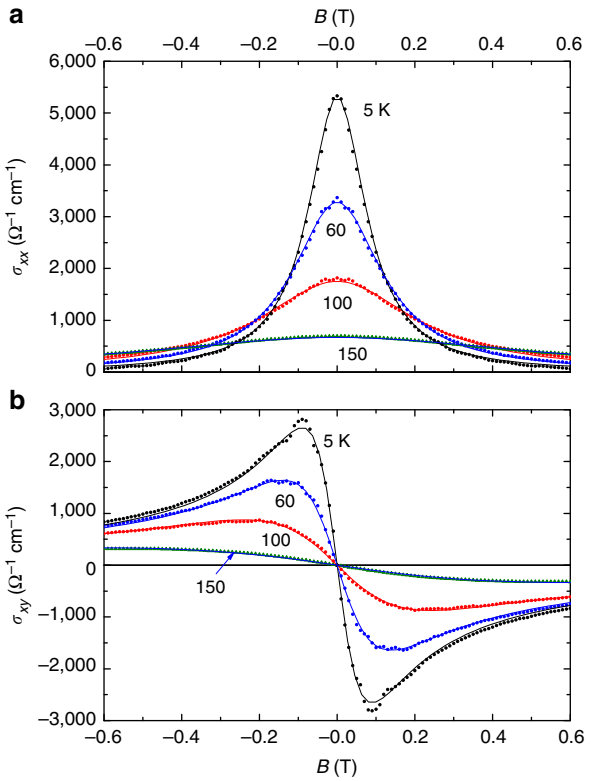

Figure 7 | Fits of the conductivity tensor. The measured weak-B conductivity σ_{xx} (**a**) and Hall conductivity σ_{xy} (**b**) of $Pb_{1-x}Sn_xSe$ ($x = 0.23$) versus B at selected T from 5 to 150 K, together with the fits to equation 9 and 10 (thin solid curves).

with m_D the Dirac mass, k_z the component of **k** along **B** and $\ell_B = \sqrt{\hbar/eB}$ the magnetic length.

For $N = 0$ LL, we solve for $k_z(E)$

$$k_z(E) = \sqrt{E^2 - E_{00}^2}/\hbar v, \quad (13)$$

where $E_{00} = m_D v^2$.

Let us assume that only $N = 0$ LL is occupied. To obtain the relation linking E_F, B and n_e, we integrate the 3D DOS $\mathcal{D}(E)dE = (g_L g_s/\pi)dk_z$, with g_s the spin degeneracy and $g_L = 1/2\pi\ell_B^2$ the 2D LL degeneracy per spin. Using equation 13, we have

$$n_e = \int_{E_{00}}^{E_F} \mathcal{D}(E)dE = \frac{g_L g_s}{\pi\hbar v}\sqrt{E_F^2 - E_{00}^2}. \quad (14)$$

This equation is valid until B is reduced to the jump field B_1, whereafter electrons enter the $N = 1$ LL. At the jump field, E_F lies just below the bottom of $N = 1$ LL, that is, $E_F^2 = E_{10}^2 = (m_D v^2)^2 + (\sqrt{2}\hbar v/\ell_B)^2$. Using this in equation 14, we have

$$n_e = \frac{\sqrt{2}g_s}{2\pi^2\ell_B^3} \quad (B = B_1). \quad (15)$$

In relation to equation 3, we showed that equation 15 gives a value equal (within 10%) to the total electron density per valley if $g_s = 1$, that is, when $B > B_1$, all the electrons can be accomodated by the $N = 0$ LL with only one spin polarization. This is direct evidence for the non-degeneracy of the $N = 0$ LL.

Interestingly, equation 15 is identical for the isotropic Schrödinger case, for which

$$E_{N,k_z} = (N + \frac{1}{2})\hbar\omega_c + \frac{\hbar^2 k_z^2}{2m}, \quad (16)$$

where $\omega_c = eB/m$ and m is the mass. However, for $N = 0$ LL of the Schrödinger spectrum, we must have $g_s = 2$, so it can be excluded.

A simple example of massive Dirac spectrum. An example illustrating the non-degeneracy of $N = 0$ LL is the spinless fermion on the 2D hexagonal lattice (valley degeneracy replaces spin degeneracy in this example). The sublattices A and B have

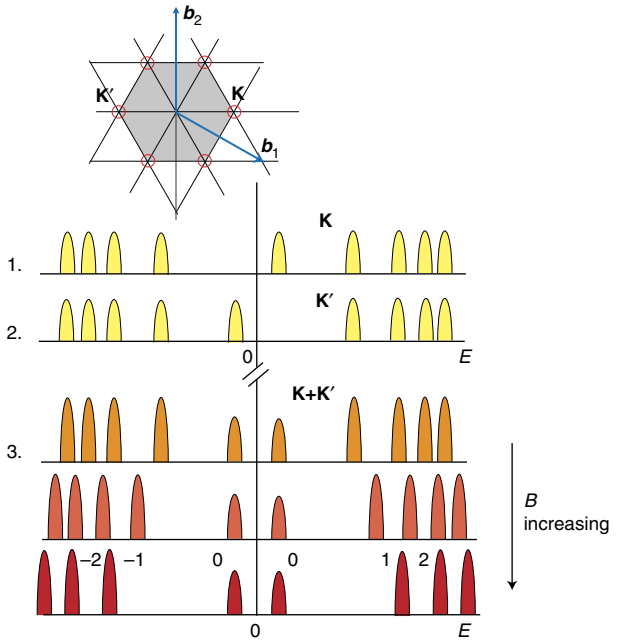

Figure 8 | The massive Dirac spectrum in the 2D hexagonal lattice. We assume distinct on-site energies on sublattices A and B (as in boron nitride). The Dirac cones sit at the high-symmetry points **K** and **K′** on the edge of the Brillouin Zone (upper inset). In a magnetic field **B**, the $N=0$ LL at **K** moves to a value above $E=0$ (curve 1), whereas the $N=0$ LL at **K′** moves below $E=0$ (curve 2). The sum of the two spectra (**K** + **K′**) is symmetric about $E=0$ (curves 3). However, the $N=0$ LLs are non-degenerate, whereas all other LLs ($N\neq 0$) have a valley degeneracy of 2. As B increases, all LLs fan out except for the $N=0$ levels (curves 3).

distinct on-site energies ϵ_A and ϵ_B as in BN. The Dirac cones remain centred at the inequivalent 'valleys' **K** and **K′** in **k**-space (inset, Fig. 8). Both valleys acquire a mass gap.

For states close to the valley at **K**, the 2D massive Dirac Hamiltonian is

$$\mathcal{H}_{2D}=\begin{bmatrix} m & k_x-ik_y \\ k_x+k_y & -m \end{bmatrix}, \qquad (17)$$

in the basis $(1,0)^T$ (pseudospin up) and $(0,1)^T$ (pseudospin down), where **k** is measured from **K** and $m>0$ represents the gap parameter proportional to $\epsilon_A-\epsilon_B$ (we set the velocity v to 1). In a field **B**, we replace **k** by $\pi=\mathbf{k}-e\mathbf{A}$ with the vector gauge $A=(0,Bx,0)$. Introducing the operators

$$a^\dagger=(\ell_B/\sqrt{2})\pi_-, \quad a=(\ell_B/\sqrt{2})\pi_+, \qquad (18)$$

with $\pi_\pm=\pi_x\pm i\pi_y$, and eigenstates $|N\rangle$ satisfying

$$a^\dagger\,|\,N\rangle=\sqrt{N+1}\,|\,N+1\rangle, \quad a\,|\,N\rangle=\sqrt{N}\,|\,N-1\rangle, \qquad (19)$$

we diagonalize the Hamiltonian to get eigenenergies E_N given by

$$E_N^2=m^2+(2N/\ell_B^2) \qquad (20)$$

(for brevity, we will write E for E_N).

For positive E, the (unrenormalized) two-spinor eigenstates are (for $N=0,1,\cdots$)

$$|\,\Psi_{N,+}\rangle=\begin{pmatrix} |\,N\rangle \\ \frac{1}{E+m}\frac{\sqrt{2N}}{\ell_B}\,|\,N-1\rangle \end{pmatrix}, \quad (E>0). \qquad (21)$$

For the negative energy states, the corresponding eigenvectors are ($N=1,2,\cdots$)

$$|\,\Psi_{N,-}\rangle=\begin{pmatrix} |\,N\rangle \\ -\frac{1}{|E|-m}\frac{\sqrt{2N}}{\ell_B}\,|\,N-1\rangle \end{pmatrix}, \quad (E<0). \qquad (22)$$

Setting $N=0$ in equation 21, we find that the positive-energy state $|\Psi_{0,+}\rangle=(|0\rangle,0)^T$ (pseudospin up). For $E<0$, however, the lower entry in equation 22 is non-determinate (0/0). This implies that the state $N=0$ does not exist for $E<0$. Thus, for the valley at **K**, there is only one LL with $N=0$. It has positive energy $E_0=|m|$; the corresponding LL at $-|m|$ is absent (the spectrum of K is sketched as curve 1 in Fig. 8).

Repeating the calculation for **K′**, we find the opposite situation (the Hamiltonian is the conjugate of equation 17). Now $N=0$ LL has energy $E_0=-|m|$, but $N=0$ LL is absent in the positive spectrum (the spectrum of K′ is the curve 2 in Fig. 8).

A transport experiment detects the sum of the two spectra (curves of K + K′ at different B are collectively labelled as 3 in Fig. 8). In the total spectrum, the two $N=0$ LLs are non-degenerate, whereas all LLs with $N\neq 0$ have a valley degeneracy of 2. The difference simply reflects the conservation of states. In the limit $m\to 0$, we recover the spectrum of graphene. If, at finite m, each of the $N=0$ LLs had a valley degeneracy of 2, we would end up with an $N=0$ LL in graphene with fourfold valley degeneracy.

The authors in refs 22–24 and others have shown that the non-degeneracy of the $N=0$ LL also holds in massive Dirac systems even when a Rashba term and a Zeeman energy term are included.

References

1. Fu, L Topological crystalline insulators. *Phys. Rev. Lett.* **106,** 106802 (2011).
2. Hsieh, T. H., Lin, H., Liu, J., Duan, W., Bansil, A. & Fu, L. Topological crystalline insulators in the SnTe material class. *Nat. Commun.* **3,** 982 (2012).
3. Hasan, M. Z. & Kane, C. L. Colloquium: topological insulators. *Rev. Mod. Phys.* **82,** 3045–3067 (2010).
4. Qi, X. -L. & Zhang, S. -C. Topological insulators and superconductors. *Rev. Mod. Phys.* **83,** 1057–1110 (2011).
5. Dziawa, P. *et al.* Topological crystalline insulator states in $Pb_{1-x}Sn_xSe$. *Nat. Mater.* **11,** 1023–1027 (2012).
6. Tanaka, Y. *et al.* Experimental realization of a topological crystalline insulator in SnTe. *Nat. Phys.* **8,** 800–803 (2012).
7. Xu, S. Y. *et al.* Observation of a topological crystalline insulator phase and topological phase transition in $Pb_{1-x}Sn_xTe$. *Nat. Commun.* **3,** 1192 (2012).
8. Mitchell, D. L. & Wallis, R. F. Theoretical energy-band parameters for the lead salts. *Phys. Rev.* **151,** 581–595 (1966).
9. Dimmock, J. O., Melngailis, I. & Strauss, A. J. Band structure and laser action in $Pb_xSn_{1-x}Te$. *Phys. Rev. Lett.* **26,** 1193–1196 (1966).
10. Strauss, A. J. Inversion of conduction and valence bands in $Pb_{1-x}Sn_xSe$ alloys. *Phys. Rev.* **157,** 608–611 (1967).
11. Lee, S. B. & Dow, J. D. Electronic structure of $Pb_xSn_{1-x}Te$ semiconductor alloys. *Phys. Rev. B* **36,** 5968–5973 (1987).
12. Gao, X. & Daw, M. S. Investigation of band inversion in (Pb,Sn)Te alloys using *ab initio* calculations. *Phys. Rev. B* **77,** 033103 (2008).
13. Svane, A. *et al.* Quasiparticle self-consistent GW calculations for PbS, PbSe, and PbTe: Band structure and pressure coefficients. *Phys. Rev. B* **81,** 245120 (2010).
14. Ziman, J. M. *Electrons and Phonons* (Oxford Clarendon Press, Oxford, 1960) p 500.
15. Behnia, K., Méasson, M. A. & Kopelevich, Y. Oscillating Nernst-Ettingshausen effect in bismuth across the quantum limit. *Phys. Rev. Lett.* **98,** 166602 (2007).
16. Fauqu1, B. *et al.* Magnetothermoelectric properties of Bi_2Se_3. *Phys. Rev. B* **87,** 035133 (2013).
17. Zuev, Y. M., Chang, W. & Kim, P. Thermoelectric and magnetothermoelectric transport measurements of graphene. *Phys. Rev. Lett.* **102,** 096807 (2009).
18. Wei, P., Bao, W., Pu, Y., Lau, C. N. & Shi, J. Anomalous thermoelectric transport of Dirac particles in graphene. *Phys. Rev. Lett.* **102,** 166808 (2009).
19. Checkelsky, Joseph G. & Ong, N. P. Thermopower and Nernst effect in graphene in a magnetic field. *Phys. Rev. B* **80,** 081413 (2009).
20. Li, L. u. *et al.* Phase transitions of Dirac electrons in bismuth. *Science* **321,** 547–550 (2008).
21. Zhu, Z., Fauqu, B., Fuseya, Y. & Behnia, K. Angle-resolved Landau spectrum of electrons and holes in bismuth. *Phys. Rev. B* **84,** 115137 (2011).
22. Wolff, P. A. Matrix elements and selection rules for the two-band model of bismuth. *J. Phys. Chem. Solids* **25,** 1057–1068 (1964).
23. Seradjeh, B., Wu, J. & Phillips, P. Signatures of surface states in bismuth at high magnetic fields. *Phys. Rev. Lett.* **103,** 136803 (2009).
24. Bernevig, B. A. & Taylor, L. H. *Topological Insulators and Topological Superconductors* (Princeton University Press, 2013).
25. Ishizaka, K. *et al.* Giant Rashba-type spin splitting in bulk BiTeI. *Nat. Mater.* **10,** 521–526 (2011).
26. Girvin, S. M. & Jonson, M. Inversion layer thermopower in high magnetic field. *J. Phys. C Solid State Phys.* **15,** L1147–L1151 (1984).
27. Jonson, M. & Girvin, S. M. Thermoelectric effect in a weakly disordered inversion layer subject to a quantizing magnetic field. *Phys. Rev. B* **29,** 1939–1946 (1984).
28. Bridgman, P. W. The connnections between the four transverse galvano-magnetic and thermomagnetic phenomena. *Phys. Rev.* **24,** 644–651 (1924).
29. Wang, Y., Li, L. & Ong, N. P. Nernst effect in high-Tc superconductors. *Phys. Rev. B* **73,** 024510 (2006).

Acknowledgements

We acknowledge helpful discussions with B.A. Bernevig, F.D.M. Haldane and M.Z. Hasan. N.P.O., T.L. and S.P.K. acknowledge support by the Army Research Office (ARO W911NF-11-1-0379). R.J.C., Q.G. and J. X. were supported by a MURI grant on Topological Insulators (ARO W911NF-12-1-0461) and the US National Science Foundation (grant number DMR 0819860). T.L acknowledges scholarship support from the Japan Student Services Organization. High-field measurements were performed at the

National High Magnetic Field Laboratory, which is supported by NSF (Award DMR-084173), by the State of Florida, and by the Department of Energy.

Author contributions

T.L., Q.G., R.J.C. and N.P.O. planned and carried out the experiment. T.L. and N.P.O. analysed the data and wrote the manuscript. J.X., M.H. and S.P.K. assisted with the measurements and analyses. Q.G. and R.J.C. grew the crystals. All authors contributed to editing the manuscript.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Liang, T. *et al.* Evidence for massive bulk Dirac fermions in $Pb_{1-x}Sn_xSe$ from Nernst and thermopower experiments. *Nat. Commun.* 4:2696 doi: 10.1038/ncomms3696 (2013).