commentary

Topological semimetals

A. A. Burkov

Topological semimetals and metals have emerged as a new frontier in the field of quantum materials. Novel macroscopic quantum phenomena they exhibit are not only of fundamental interest, but may hold some potential for technological applications.

The study of the electronic structure topology of crystalline materials has emerged in the past decade as a major new theme in modern condensed-matter physics. The starting impetus came from the remarkable discovery of topological insulators[1,2], but the focus has recently shifted towards topological semimetals and even metals. Although the idea that metals can have a topologically nontrivial electronic structure is not entirely new and some of the recent developments were anticipated in earlier work[3–5], this shift was precipitated by the theoretical discovery of Weyl[6–12] and later Dirac semimetals[13–15]. The experimental realization of both Weyl and Dirac semimetals[16–20] within the past couple of years has brought the field to the forefront of quantum condensed-matter research.

Weyl semimetals

Whereas ordinary metals owe most of their unique observable character to the existence of a Fermi surface, and insulators to its lack — together with the presence of a finite gap between the highest occupied and the lowest unoccupied state — the defining feature of topological semimetals is the appearance of band touching points or nodes at the Fermi energy, where two or more bands are exactly degenerate at particular values of the crystal momentum in the first Brillouin zone (BZ). Line nodes, where the bands are degenerate along closed lines in momentum space, may also exist in topological semimetals. The existence of such band touching nodes was recognized in the early days of solid-state physics[21], yet their importance was only appreciated recently.

Naively, a band touching point should be a very unlikely and very unstable feature, and can thus be hardly expected to be of any importance. Indeed, as we know from basic quantum mechanics, a degeneracy between energy levels is always lifted unless required by a symmetry. However, this naive viewpoint overlooks the possibility of an accidental degeneracy of a single pair of bands in a three-dimensional material. To see how this happens, suppose two bands touch at some point, $\mathbf{k}_0$, in the first BZ and at energy ε_0. In the vicinity of this point, the momentum-space Hamiltonian may be expanded in a Taylor series with respect to the deviation of the crystal momentum from $\mathbf{k}_0$. The expansion will generally have the form (ignoring possible anisotropies for simplicity)

$$H(\mathbf{k}) = \varepsilon_0\sigma_0 \pm \hbar v_{\mathrm{F}}(\mathbf{k} - \mathbf{k}_0)\cdot\boldsymbol{\sigma} \qquad (1)$$

where σ_0 is a 2×2 unit matrix, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the three Pauli matrices, $\mathbf{k}$ is the crystal momentum and v_{F} is the Fermi velocity. This represents nothing more than an expansion of a general 2×2 Hermitian matrix in terms of the unit matrix and the three Pauli matrices. What makes this interesting is that there is nothing we can do to equation (1) to get rid of the band touching point. Changing ε_0 or $\mathbf{k}_0$ can only change the location of the band touching point in energy or crystal momentum, whereas changing the parameter v_{F}, which has dimensions of velocity, only changes the slope of the band dispersion away from the point. The point itself is always there. Formally, this is related to the fact that there are three crystal momentum components and the same number of Pauli matrices. This means that for such an unremovable band touching to occur we need three spatial dimensions, which is not a problem, and non-degenerate (at a general value of the crystal momentum) bands. This second requirement cannot be satisfied in a material that possesses two fundamental symmetries: inversion, P (which means that the crystal structure has an inversion centre) and time reversal, Θ (meaning the material is nonmagnetic). The reason is that in such a material all bands must be at least doubly degenerate at every value of $\mathbf{k}$ due to the fundamental property of any system of fermions that $(P\Theta)^2 = -1$. Thus, we come to the conclusion that unremovable band touching points may only occur in non-centrosymmetric or magnetic materials.

Interestingly, if we set $\varepsilon_0 = 0$ in equation (1), which simply resets the energy zero-point, equation (1) takes the exact form (up to a trivial replacement of the speed of light by v_{F}) of a Weyl Hamiltonian, that is, the Hamiltonian of a massless relativistic particle of right-handed or left-handed chirality, which corresponds to the $\pm$ sign in the equation. Weyl fermions are fundamental building blocks, from which the standard model of particle physics is constructed. All the known elementary fermions in nature have masses, acquired due to complex interactions between the fundamental Weyl fermions. The fact that massless Weyl fermions may occur as quasiparticles in a crystal is very intriguing and may have deep implications for our understanding of the Universe[3]. Due to this analogy with the Weyl equation, the touching points of pairs of non-degenerate bands are called Weyl points or nodes.

Another important feature of the Weyl Hamiltonian is that it is a topologically nontrivial object. Indeed, the eigenstates of equation (1) may be labelled by helicity, which is the sign of the projection of $\boldsymbol{\sigma}$ onto the direction of the crystal momentum $\mathbf{k}$ (helicity coincides with chirality for a massless particle, but not in general). The expectation value of $\boldsymbol{\sigma}$ in an eigenstate of a given helicity forms a vector field in momentum space that wraps around the Weyl node location $\mathbf{k}_0$, forming a 'hedgehog', or a 'hairy ball' as shown in Fig. 1. Such a hedgehog may be characterized by a topological invariant, defined as a flux of a vector

$$\boldsymbol{\Omega}(\mathbf{k}) = \pm\frac{\mathbf{k}}{2|\mathbf{k}|^3} \qquad (2)$$

through any surface in momentum space, enclosing the Weyl node, where the sign in front is the chirality of the node. This flux, normalized by 2π, is equal to the chirality, and therefore this is a quantized integer-valued invariant. The vector $\boldsymbol{\Omega}(\mathbf{k})$ is called the Berry curvature and the Weyl nodes may thus be regarded as point 'charges', which are sources and sinks of the Berry curvature. This implies that the Weyl nodes must occur in pairs of opposite chirality,

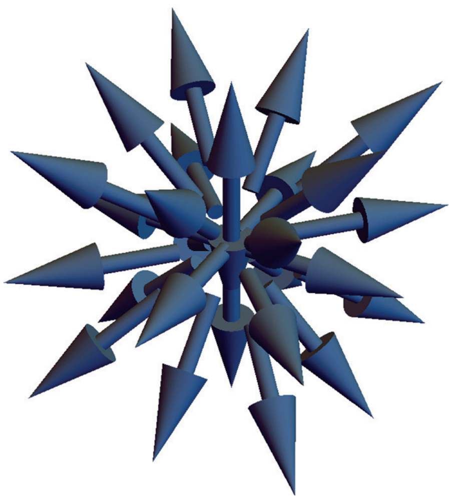

Figure 1 | Berry curvature field forming a 'hedgehog' around a Weyl node in momentum space.

since the field lines of the Berry curvature must begin and end somewhere within the BZ, and that the only way to eliminate the Weyl nodes is to annihilate them pairwise, which may only happen if they occur at the same point in momentum space. Thus the origin of the stability of the Weyl nodes is ultimately topological.

An important consequence of the fact that Weyl nodes are point sources and sinks of the Berry curvature field is that when the sample has edges in real space, parallel to the line connecting a pair of Weyl nodes in momentum space, there exist edge-localized states that are confined to the projection of the internode interval onto the first BZ of the sample surface. These edge states are called Fermi arcs[6].

So far we have deduced that any three-dimensional non-centrosymmetric or magnetic material will have Weyl band touching points somewhere in its electronic structure. Yet in order for these points to have a significant effect on observable properties of the material they must occur at or very close to the Fermi energy and there must not be any other states at the Fermi energy. This is what is nontrivial to achieve in general. What may help here is the fact that any non-degenerate band must have exactly the same number of states, equal to the number of unit cells in the crystal, and the Pauli principle, which prohibits double occupation of any state. This implies that, for any even integer number of electrons per unit cell, the bands are always either completely filled or completely empty, provided they do not overlap in energy. The problem is that they generally do overlap, but we may find a situation when they do not if we look among narrow-gap semiconductors, which in the absence of broken inversion or time reversal symmetry would be insulators with a very small gap. All currently known Weyl materials are of this sort.

Equation (1), as written, actually makes an implicit omission, which leads to a loss of an important piece of physics of Weyl semimetals. In general, the coefficient of the unit matrix σ_0 is a function of the crystal momentum, not just a constant ε_0, which is simply the zeroth-order term in the Taylor expansion of the function. An important question is whether this expansion has a term linear in $(\mathbf{k} - \mathbf{k}_0)$ or not. Only if $\mathbf{k}_0$ is a special time-reversal-invariant momentum (TRIM), which occurs only at the BZ centre or at special points at the BZ boundary, there is no linear term. In general, Weyl points occur away from a TRIM, which means that a linear term does exist:

$$H = \hbar\tilde{v}_F(\mathbf{k} - \mathbf{k}_0)\sigma_0 \pm \hbar v_F(\mathbf{k} - \mathbf{k}_0)\cdot\boldsymbol{\sigma} \qquad (3)$$

where we have set $\varepsilon_0 = 0$. Now, something new happens when $\tilde{v}_F > v_F$. The Weyl point is of course still there, but the two bands that touch at the Weyl point now overlap in energy, forming electron and hole pockets (Fig. 2). The Weyl point then becomes a point at which an electron and a hole pocket touch. Such Weyl semimetals are called type-II Weyl semimetals[22]. $MoTe_2$ is an example of a type-II Weyl semimetal material[23]. TaAs, the first Weyl semimetal material discovered[18,19], is type-I. An additional subclass of Weyl semimetal materials are the so-called multi-Weyl semimetals[24], in which Weyl nodes carry topological charges of higher magnitude (for example 2 or 3). Such higher-charge Weyl nodes may be stabilized by certain point-group crystal symmetries.

Dirac semimetals

As mentioned above, we may imagine getting a Weyl semimetal by breaking inversion or time reversal symmetry in a narrow-gap semiconductor. The simplest Hamiltonian that describes this situation has the form

$$H = \begin{pmatrix} \hbar v_F \boldsymbol{\sigma}\cdot\mathbf{k} & m \\ m & -\hbar v_F \boldsymbol{\sigma}\cdot\mathbf{k} \end{pmatrix} \qquad (4)$$

This equation describes two Weyl fermions of opposite chirality at the same point (TRIM) in the BZ. The off-diagonal entry m in equation (4) mixes the two Weyl fermions and opens up a gap of magnitude $2m$. This Hamiltonian in fact describes a transition between a topological and ordinary insulator in three dimensions. The transition point corresponds to setting $m = 0$ and is what is called a Dirac semimetal. At the Dirac point there is degeneracy between four bands instead of two, as at a Weyl point. Such a degeneracy requires either fine tuning or an extra symmetry to force $m = 0$. In the case when $m = 0$ by symmetry we get a stable Dirac semimetal phase, distinct from a Weyl semimetal. It can be shown that symmetry-protected Dirac points of the kind described above may only occur at a TRIM at the BZ edge and require a non-symmorphic symmetry, which means a rotation combined with a partial translation[13]. While proposals for specific material realizations have been made, no such Dirac semimetal has yet been realized experimentally.

However, another possibility exists. Instead of a single Dirac point at a TRIM, it is possible to create two Dirac points on a rotation axis, protected by a rotational symmetry about this axis. The Hamiltonian that describes this situation has the form

$$H = \hbar v_F(\tau_x\sigma_z k_x - \tau_y k_y) + m(k_z)\tau_z \qquad (5)$$

where $\boldsymbol{\tau}$ and $\boldsymbol{\sigma}$ are Pauli matrices, describing orbital and spin degrees of freedom respectively and $m(k_z) = -m_0 + m_1 k_z^2$. For a given eigenvalue of the spin operator $\sigma = \pm 1$, this Hamiltonian describes the simplest Weyl semimetal with two Weyl nodes at

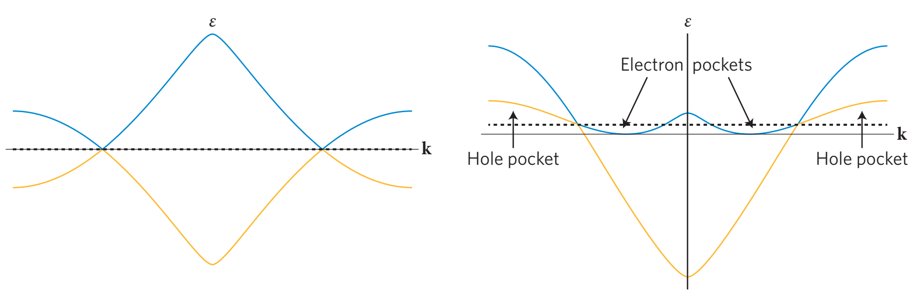

Figure 2 | Type-I and type-II Weyl nodes. In the type-I case (left), the Fermi energy, represented by the dashed line, only crosses the nodes themselves and is otherwise in the gap between the two touching bands. In the type-II case (right) there are electron and hole pockets that touch at the Weyl nodes between the two touching bands, denoted by the blue and yellow lines.

points $\mathbf{k}_\pm = (0, 0, \pm(m_0/m_1)^{1/2})$. The two Weyl semimetals, corresponding to $\sigma = \pm 1$, are related to each other by the time reversal operation. Thus each of the two band touching points $\mathbf{k}_\pm$ contain two Weyl points of opposite chirality and they are therefore two Dirac points. This kind of Dirac semimetal is distinct from the one discussed above, with a single Dirac point at a TRIM. Such a Dirac semimetal has been realized experimentally in two compounds, Na_3Bi and Cd_2As_3 (refs 14–17).

Topological response

Whereas the existence of special edge states, required by the electronic structure topology, is a feature of Weyl and Dirac semimetals that is similar to the other class of topological materials, the topological insulators, what makes the former potentially more interesting is that they exhibit special topology-related responses, in addition to special spectroscopic features like the edge states.

The best known example of a universal (namely detail-independent) topological transport property is the Hall conductivity of a two-dimensional quantum Hall liquid, which is quantized to an incredible precision, in spite of all the imperfections and dirt in the semiconductor heterostructure samples in which it is measured. What leads to this precision and universality is that a two-dimensional quantum Hall liquid is a (topological) insulator and it is the intrinsic insensitivity of an insulator to perturbations that is the source of it. One would not expect such a universal transport to exist in a gapless system.

Surprisingly, Weyl and Dirac semimetals do exhibit universal transport. Its origin is the chiral anomaly, a phenomenon that was discovered many years ago in the particle physics context. In this case chiral anomaly is an unexpected non-conservation of chiral charge in a system of massless relativistic fermions, coupled to electromagnetic field with collinear electric and magnetic components. It is unexpected because being massless for a relativistic particle, described by the Dirac equation, is equivalent to having a conserved chirality, or chiral symmetry. The Dirac equation also has negative energy state solutions, which must be filled with unobservable particles, forming the so-called Dirac sea. The seemingly obvious chiral symmetry disappears, however, once the infinite negative energy Dirac sea is filled with particles, and it is the infinity of the Dirac sea, which by itself is surely a mathematical artefact of the relativistic field theory, that ultimately leads to the anomaly. Astonishingly, the chiral anomaly

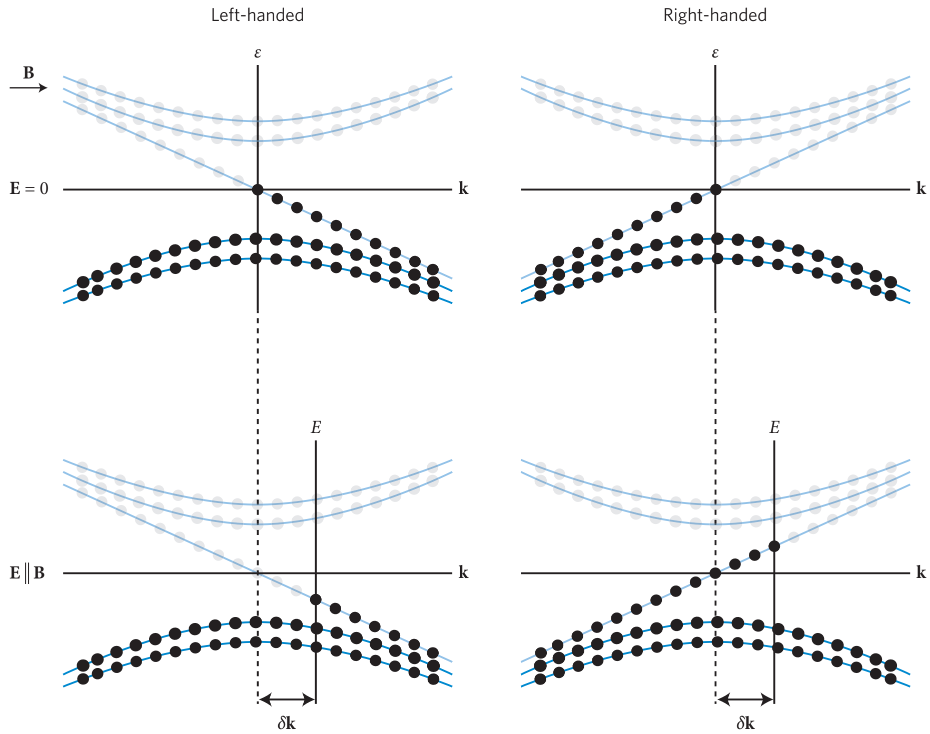

Figure 3 | Origin of the chiral anomaly. The schematic shows a simple explanation of the chiral anomaly, based on a Landau level spectrum of Weyl fermions in an external magnetic field. Top: the energy spectrum of the left-handed and the right-handed fermions in equilibrium in the presence of a magnetic field **B**. Filled states with negative energy are shown by black dots, while empty states with positive energy by grey dots. Bottom: the same spectrum as above, but in the presence of an additional electric field **E**, parallel to the magnetic field **B**. All states have been displaced in momentum space by an amount $\delta\mathbf{k} \sim -\mathbf{E}$ from their equilibrium locations. As a consequence, right-handed particles and left-handed antiparticles have been produced.

is nevertheless a real effect, with observable consequences, explaining for example the decay of a neutral pion into two photons. Mathematically, the chiral anomaly may be expressed most conveniently as an extra term in the action for the electromagnetic field induced by the fermions

$$S = \frac{e^2}{32\pi^2\hbar^2 c}\int dt d^3r\,\theta(\mathbf{r},t)\varepsilon_{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta} = \frac{e^2}{8\pi^2\hbar^2 c}\int dt d^3r\,\theta(\mathbf{r},t)\mathbf{E}\cdot\mathbf{B} \quad (6)$$

where $\theta(\mathbf{r}, t) = 2(\mathbf{b}\cdot\mathbf{r} - b_0 t)$ and $b_\mu = (b_0, \mathbf{b})$ are chiral gauge fields, which couple antisymmetrically to Weyl fermions of different chirality, while $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor and A_μ are the ordinary electromagnetic gauge fields, which do not care about chirality. In the Weyl and Dirac semimetal context, what leads to the analogous phenomena is the emergence of an effective relativistic structure in the Hamiltonian near Weyl or Dirac band touching points, as described above. In this case, chiral symmetry is really not there to begin with and there is thus nothing 'anomalous' in the chiral charge non-conservation. The astonishment is still there, however, but the reason for it is opposite: in spite of chirality not being strictly conserved, the anomaly term looks exactly like equation (6), as long as the Fermi energy is close to the nodes! In this context

$$\mathbf{b} = \frac{1}{2}\sum_i C_i\mathbf{k}_i,\ b_0 = \frac{1}{2}\sum_i C_i\varepsilon_i \quad (7)$$

where C_i is the topological charge (chirality) of the Weyl node, i is its location in momentum space and ε_i is its energy. This leads to universal transport phenomena, such as the anomalous Hall effect, whose magnitude is determined only by the location of the Weyl nodes in magnetic Weyl semimetals; and a large negative longitudinal magnetoresistance in both Weyl and Dirac semimetals[25–28]. The latter effect arises as a result of a two-step process: generation of a chiral chemical potential by charge current in the presence of an external magnetic field (Fig. 3), which in turn leads to an extra contribution to the charge current, known as the chiral magnetic effect[29–31]. As both steps involve the magnetic field, the resulting magnetoresistance depends quadratically on it. Chiral anomaly in Weyl and Dirac semimetals is a perfect example of an emergent phenomenon, characterized by simplicity and universality emerging at low energies, in this particular example due to nontrivial electronic structure topology. While we limited our discussion above to the electrical transport phenomena, chiral anomaly in fact affects thermoelectric response as well[32–34]. Many of the topological insulator and topological semimetal materials exhibit large thermopower, which by itself is of significant interest from the viewpoint of practical applications. Chiral anomaly, in addition, leads to a strong magnetic field dependence of the thermoelectric transport coefficients.

Outlook

Perhaps the most important current issue in the field of topological semimetals is finding more and better material realizations. As described in the previous section, universality of the topological transport properties of Weyl and Dirac semimetals is an emergent property, which becomes more and more precise as the Fermi energy is approaching the band touching points. Thus it is important to search for materials with the Fermi energy very close to the nodes and no other states at it. Of the currently known materials, Na_3Bi comes closest to this ideal, and it is not an accident that it exhibits a very strong negative longitudinal magnetoresistance due to the chiral anomaly[26–28].

Another important direction is to look for a magnetic Weyl semimetal material, which would exhibit another consequence of the chiral anomaly, the universal anomalous Hall effect. A material exhibiting a quantum anomalous Hall effect at room temperature would be of great technological importance, offering low-dissipation transport in the absence of external magnetic fields. This could hold some potential for helping us to avoid the quickly approaching technological brick wall: current semiconductor electronics is nearly at its limits for both miniaturization and speed. Unless a fundamentally new technology emerges, our computers will very soon, within a few years, stop getting smaller and faster, something that we have learned to take for granted, but which in reality is not guaranteed. Materials with nontrivial electronic structure topology, exhibiting low-dissipation transport at room temperature, may help us overcome this challenge by dramatically reducing power consumption and heating.

Finally, it is of interest, both from the fundamental and the practical standpoint, to explore the possible effects of electron–electron interactions in topological semimetals. At the simplest (in theory) level, superconductivity, resulting from attractive electron–electron interactions, will arise at low temperatures in almost any metal. In Weyl metals, especially in those with magnetic properties, superconductivity is predicted to be topologically nontrivial, leading to Majorana surface states[35,36], which may be of interest for topological quantum computing. Venturing into a more exotic area, one could speculate if sufficiently strong repulsive interactions could lead to fractionalized Weyl and Dirac semimetals, in analogy to fractional topological insulators[37,38]. ❒

A. A. Burkov is in the Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, and at ITMO University, Saint Petersburg 197101, Russia.
e-mail: aburkov@uwaterloo.ca

References

1. Hasan, M. Z. & Kane, C. L. *Rev. Mod. Phys.* **82,** 3045 (2010).
2. Qi, X.-L. & Zhang, S.-C. *Rev. Mod. Phys.* **83,** 1057 (2011).
3. Volovik, G. *The Universe in a Helium Droplet* (Oxford Univ. Press, 2003).
4. Haldane, F. D. M. *Phys. Rev. Lett.* **93,** 206602 (2004).
5. Murakami, S. *New J. Phys.* **9,** 356 (2007).
6. Wan, X., Turner, A. M., Vishwanath, A. & Savrasov, S. Y. *Phys. Rev. B* **83,** 205101 (2011).
7. Yang, K.-Y., Lu, Y.-M. & Ran, Y. *Phys. Rev. B* **84,** 075129 (2011).
8. Burkov, A. A. & Balents, L. *Phys. Rev. Lett.* **107,** 127205 (2011).
9. Xu, G., Weng, H., Wang, Z., Dai, X. & Fang, Z. *Phys. Rev. Lett.* **107,** 186806 (2011).
10. Witczak-Krempa, W. & Kim, Y. B. *Phys. Rev. B* **85,** 045124 (2012).
11. Weng, H., Fang, C., Fang, Z., Bernevig, B. A. & Dai, X. *Phys. Rev. X* **5,** 011029 (2015).
12. Huang, S.-M. *et al. Nat. Commun.* **6,** 7373 (2015).
13. Young, S. M. *et al. Phys. Rev. Lett.* **108,** 140405 (2012).
14. Wang, Z. *et al. Phys. Rev. B* **85,** 195320 (2012).
15. Wang, Z., Weng, H., Wu, Q., Dai, X. & Fang, Z. *Phys. Rev. B* **88,** 125427 (2013).
16. Liu, Z. K. *et al. Science* **343,** 864–867 (2014).
17. Neupane, M. *et al. Nat. Commun.* **5,** 3786 (2014).
18. Xu, S.-Y. *et al. Science* **349,** 613–617 (2015).
19. Lv, B. Q. *et al. Phys. Rev. X* **5,** 031013 (2015).
20. Lu, L. *et al. Science* **349,** 622–624 (2015).
21. Herring, C. *Phys. Rev.* **52,** 365 (1937).
22. Soluyanov, A. A. *et al. Nature* **527,** 495–498 (2015).
23. Huang, L. *et al.* Preprint at https://arxiv.org/abs/1603.06482 (2016).
24. Fang, C., Gilbert, M. J., Dai, X. & Bernevig, B. A. *Phys. Rev. Lett.* **108,** 266802 (2012).
25. Son, D. T. & Spivak, B. Z. *Phys. Rev. B* **88,** 104412 (2013).
26. Burkov, A. A. *Phys. Rev. B* **91,** 245157 (2015).
27. Xiong, J. *et al. Science* **350,** 413–416 (2015).
28. Burkov, A. A. & Kim, Y. B. *Phys. Rev. Lett.* **117,** 136602 (2016).
29. Fukushima, K., Kharzeev, D. E. & Warringa, H. J. *Phys. Rev. D* **78,** 074033 (2008).
30. Vazifeh, M. M. & Franz, M. *Phys. Rev. Lett.* **111,** 027201 (2013).
31. Chen, Y., Wu, S. & Burkov, A. A. *Phys. Rev. B* **88,** 125105 (2013).
32. Hirschberger, M. *et al. Nat. Mater.* http://dx.doi.org/10.1038/nmat4684 (2016).
33. Chen, Q. & Fiete, G. A. *Phys. Rev. B* **93,** 155125 (2016).
34. Spivak, B. Z. & Andreev, A. V. *Phys. Rev. B* **93,** 085107 (2016).
35. Bednik, G., Zyuzin, A. A. & Burkov, A. A. *Phys. Rev. B* **92,** 035153 (2015).
36. Li, Y. & Haldane, F. D. M. Preprint at https://arxiv.org/abs/1510.01730 (2015).
37. Maciejko, J. & Fiete, G. A. *Nat. Phys.* **11,** 385–388 (2015).
38. Meng, T., Grushin, A. G., Shtengel, K. & Bardarson, J. H. Preprint at https://arxiv.org/abs/1602.08856 (2016).

Acknowledgements

Financial support was provided by the Natural Sciences and Engineering Research Council of Canada.