Collective spin modes in Fermi liquids with spin-orbit coupling

Dmitrii L. Maslov*
Department of Physics, University of Florida, Gainesville, FL 32611-8440, USA

Abhishek Kumar
Center for Materials Theory, Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA

Saurabh Maiti
Department of Physics and Centre for Research in Molecular Modeling, Concordia University, Montreal, QC H4B 1R6, Canada

(Dated: November 8, 2022)

A combination of spin-orbit coupling and electron-electron interaction gives rise to a new type of collective spin modes, which correspond to oscillations of magnetization even in the absence of the external magnetic field. We review recent progress in theoretical understanding and experimental observation of such modes, focusing on three examples of real-life systems: a two-dimensional electron gas with Rashba and/or Dresselhaus spin-orbit coupling, graphene with proximity-induced spin-orbit coupling, and the Dirac state on the surface of a three-dimensional topological insulator. This paper is dedicated to the 95th birthday of Professor Emmanuel I. Rashba.

I. INTRODUCTION

Spin-orbit coupling (SOC) plays an important and, sometimes decisive, role in many condensed matter systems, including two-dimensional (2D) electron and hole gases in semiconductor heterostructures,[1, 2] non-centrosymmetric normal metals [3] and superconductors,[4, 5] bismuth tellurohalides [6], a variety of iridates and vanadates, [7], surface/edge states of three-dimensional (3D)/2D topological insulators,[8–12] conducting states at oxide interfaces,[13] 2D transition metal dichalcogenides (TMD),[14, 15] graphene on TMD substrates,[16] atomic Bose[17, 18] and Fermi[19, 20] gases in simulated non-Abelian magnetic fields, etc. Coupling between electron spins and momenta leads to a number of fascinating consequences, such as the electric-dipole spin resonance (EDSR), [21–23] current-induced spin polarization, [24–26] persistent spin helices, [27–29] quantum spin [30–32] and anomalous Hall effects,[33–35] to name just a few. An interesting and still largely open question is the interplay between spin-orbit and electron-electron interactions. Such interplay gives rise to new phases of matter, e.g., topological Mott insulator,[36, 37] gyrotropic and multipolar orders in normal metals,[38] helical Fermi liquid (FL),[39] Gor'kov-Rashba superconductor with mixed singlet-triplet order parameter,[40] topological Kondo insulators, [41] etc. It also affects in a non-trivial way many physical phenomena, e.g., optical conductivity,[42, 43] plasmon spectra,[44–46] RKKY interaction,[47–49] non-analytic behavior of the spin susceptibility,[50–52] etc., and gives rise to spin-dependent electron-electron interaction.[53]

In this paper, we review recent progress in theoretical understanding and experimental observation of a new type of collective spin modes in 2D FLs with SOC. Such modes are perhaps the most direct manifestation of an interplay between spin-orbit and electron-electron interactions, as their existence hinges on both components being present. Unlike the conventional Silin mode in a partially spin-polarized FL,[54] these modes exist even in the absence of an external magnetic field; in addition, they modify in a characteristic way the Silin mode if both SOC and magnetic field are present. As long as SOC is weak, the new modes correspond to oscillations of the magnetization which are decoupled from the oscillations of charge. The origin of the new modes can be traced to the effective spin-orbit magnetic field, which depends on the orientation and magnitude of the electron momentum, and also on the position of electron valley in the Brillouin zone (for multi-valley systems, such as graphene with proximity-induced SOC). Some of these modes have already been observed experimentally in $Cd_{1-x}Mn_xTe$ quantum wells (in the presence of the magnetic field)[55–60] and in the surface state of a three-dimensional (3D) topological insulator (TI) Bi_2Se_3 (in zero magnetic field);[61] however, many more predictions are still awaiting their experimental confirmation.

The rest of the paper is organized as follows. In Sec. II A, we introduce three systems considered in the rest of the article: 1) a 2D electron gas (2DEG) with Rashba and/or Dresselhaus SOC, 2) graphene with proximity-induced SOC, and 3) a Dirac helical state on the surface of a 3D TI. In Sec. II B, we describe the single- and two-valley FL theories,

* maslov@ufl.edu

which will be applied to study the collective spin modes in 2DEGs and graphene, respectively, given that SOC and/or magnetic field are weak. In Sec. II C, we explain why a FL theory cannot be applied to the cases of arbitrarily strong SOC and/or magnetic field. Sec. III serves as a short reminder of collective modes in a FL without SOC, in general, and of the Silin modes in a partially spin-polarized FL, in particular. In Sec. IV, we discuss collective spin modes in a 2DEG. Sec. IV A describes the FL theory for the case of Rashba SOC. In Sec. IV B, we show that the FL kinetic equation for a 2DEG with Rashba and/or Dresselhaus SOC and in the presence of the magnetic field can be mapped onto an effective tight-binding model for an artificial one-dimensional (1D) lattice, whose sites are labeled by the projections of the angular momentum. Within this mapping, the role of FL interaction is to produce "defects", both of the on-site and bond types, and the collective modes arise as bound states due to such defects. In Sec. IV C, we illustrate how this mapping works for the case of a 2DEG with Rashba SOC and in the presence of the magnetic field using the s-wave approximation for the Landau function. Sec. V deals with collective spin modes in Dirac systems. In Sec. V A we apply a two-valley version of the FL theory to graphene with proximity-induced SOC. In Sec. V B, we derive the spectrum of inter-band spin excitations in a Dirac surface state within the ladder approximation. In Sec. VI, we discuss the spatial dispersion of collective spin modes. Sec. VII is devoted to damping due to both disorder and electron-electron interaction. In Sec. VIII, we discuss both the current and future experiment. Sec. VIII A summarizes the results of a series of Raman experiments on $Cd_{1-x}Mn_xTe$. In Sec. VIII B, we provide a summary of recent Raman spectroscopy of a collective spin mode on the surface of Bi_2Se_3. Sec. VIII C contains the theoretical predictions for electron spin resonance (ESR) and EDSR experiments on graphene with proximity-induced SOC, both in zero and strong (compared to SOC) magnetic field. Our conclusions are given in Sec. IX.

II. MODEL SINGLE-PARTICLE HAMILTONIANS AND ELECTRON-ELECTRON INTERACTION

A. Model single-particle Hamiltonians for spin-orbit coupling

Despite the variety of real-life systems, the effects of SOC on the electron spectrum can be described by just a few low-energy Hamiltonians, constructed by using the symmetry arguments. In this paper, we will consider three examples of two-dimensional electron systems: a two-dimensional electron gas (2DEG) sandwiched between two dissimilar semiconductors, monolayer graphene with substrate-induced SOC, and the surface state of a three-dimensional (3D) topological insulator (TI).

The effect of SOC on a 2DEG sandwiched between two dissimilar centrosymmetric semiconductors is described by the venerable Rashba Hamiltonian[62, 63]

$$\hat{H}_{\rm R} = \frac{k^2}{2m} + \alpha(\hat{\boldsymbol{\sigma}} \times \mathbf{k}) \cdot \hat{\mathbf{z}} = \frac{k^2}{2m} + \alpha \left(k_y \hat{\sigma}_x - k_x \hat{\sigma}_y\right), \tag{1}$$

where $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ is a vector of Pauli matrices which describe electron spin, $\hat{\mathbf{z}}$ is a unit vector along the normal to the 2DEG plane, and α is a phenomenological parameter (with units of velocity). The simplest way to arrive at this Hamiltonian is to notice that a combination $(\hat{\boldsymbol{\sigma}} \times \mathbf{k}) \cdot \hat{\mathbf{z}}$ is the only scalar which can be formed out of an axial vector ($\hat{\boldsymbol{\sigma}}$) and two polar vectors ($\mathbf{k}$ and $\hat{\mathbf{z}}$).

In a bulk non-centrosymmetric semiconductor, e.g., of the A3B5 family (GaAs, CdTe, etc.), symmetry allows for a cubic (Dresselhaus) coupling between momentum and spin.[64] A quantum well on the surface of such a semiconductor is described by a 2D Dresselhaus Hamiltonian with linear coupling between spin and momentum obtained by projecting the bulk Dresselhaus term onto the quantum-well plane.[65, 66] A particular form of the 2D Dresselhaus Hamiltonian depends on the orientation of the quantum-well plane with respect to crystallographic axes. We will consider the most common case of a quantum well grown along the (001) direction. With the x-axis along the (100) direction and y-axis along the (010) the direction, the Dresselhaus Hamiltonian reads

$$\hat{H}_{\rm D} = \frac{k^2}{2m} + \beta \left(k_x \hat{\sigma}_x - k_y \hat{\sigma}_y\right). \tag{2}$$

In heterostructures made from non-centrosymmetric semiconductors the Rashba and Dresselhaus types of SOC usually occur simultaneously, and the total spin-orbit part of the Hamiltonian is the sum of the Rashba and Dresselhaus terms. Without loss of generality, we assume that $\alpha, \beta > 0$.

Another popular system is graphene adsorbed on a transition-metal-dichalcogenide (TMD) substrate, such as WS_2, WSe_2 and MoS_2.[67–80] In this case, the induced SOC is expected to be a mixture of two types:[68, 81, 82] of Rashba SOC and valley-Zeeman (VZ) or Ising SOC, which acts as an out-of-plane magnetic field whose direction alternates

between the K and K' valleys of graphene. We focus on the case of monolayer graphene with proximity-induced SOC, described by the following low-energy Hamiltonian:

$$\begin{aligned}\hat{H}_{\rm GR} &= v_{\rm D}(\tau_z\hat{\nu}_x k_x + \hat{\nu}_y k_y) + \Delta\hat{\nu}_z + \frac{\lambda_{\rm R}}{2}(\tau_z\hat{\sigma}_y\hat{\nu}_x - \hat{\sigma}_x\hat{\nu}_y) \\ &+ \frac{\lambda_{\rm VZ}}{2}\tau_z\hat{\sigma}_z, \end{aligned} \tag{3}$$

where $v_{\rm D}$ is the Dirac velocity, $\mathbf{k}$ is the electron momentum measured either from the K or K' point of the graphene Brillouin zone, $\lambda_{\rm R}$ and $\lambda_{\rm VZ}$ are the coupling constants of the Rashba and VZ spin-orbit interactions, respectively, Δ is the gap due to substrate-induced asymmetry between the A and B sites of the honeycomb lattice, $\hat{\nu}_i$ $(i = x, y)$ are the Pauli matrices in the sublattice (pseudospin) subspace, and $\tau_z = \pm 1$ labels the K and K' points. (An expression of the type $\hat{\sigma}_i\hat{\nu}_j$ is to be understood as a tensor product of two matrices. A single matrix in one of the subspaces implies that it is tensor product with a unity matrix in the other subspace.) In Eq. (3), we neglected the intrinsic, Kane-Mele type of SOC,[83] which is much weaker than the induced ones.[1]

Finally, the surface of a 3D TI, e.g., Bi_2Se_3, harbors a Dirac helical state. If the hexagonal warping of the energy contours due to underlying crystal lattice can be neglected, this state is described by the Rashba-like Hamiltonian without the parabolic term[89]

$$\hat{H}_{\rm TI} = v_{\rm D}(\hat{\boldsymbol{\sigma}} \times \mathbf{k}) \cdot \hat{\mathbf{z}}. \tag{4}$$

In all cases presented above, the effect of an in-plane magnetic field of magnitude B and at angle γ to the x-axis is accounted for by adding the usual Zeeman term

$$\hat{H}_{\rm Z} = \frac{1}{2}\Delta_{\rm Z}\left(\cos\gamma\,\hat{\sigma}_x + \sin\gamma\,\hat{\sigma}_y\right) \tag{5}$$

to the corresponding Hamiltonian. Here, $\Delta_{\rm Z} = g\mu_B B$ is the Zeeman splitting, g is the effective Landé factor (assumed to be isotropic) and μ_B is the Bohr magneton.

The properties of Hamiltonians (1)-(4) are well-understood by now and we will not reproduce the known results here. It suffices to say that Rashba and Dresselhaus types of SOC lead to spin textures in the momentum space, which can be interpreted as the effect of an effective magnetic field, while the VZ type of SOC in (3) acts as an out-of-plane magnetic field whose orientation alternates between the K and K' valleys.

B. Models of electron-electron interaction

Since the focus of this article is on the collective modes, we need to invoke the electron-electron interaction as it is essential to induce collective behaviour. In all cases, we assume that our system is doped, such that the Fermi energy (E_F) is significantly larger than the spin-orbit and/or Zeeman splitting of the electron spectrum. An exception is the surface state of a TI, where SOC forms the spectrum rather than modifies the already existing one. With this exception, SOC can be treated as a perturbation imposed on a single-valley (2DEG) or two-valley (graphene) Fermi liquid. Modulo renormalizations by the interaction, the spin-orbit and Zeeman energy scales determine the frequencies of the corresponding collective modes. Therefore, to leading order in these energy scales, one can neglect the effect of SOC and magnetic field on the Landau interaction function. For the single-valley case, the latter is given by the usual form [90–92]

$$N_F^*\hat{f}(\mathbf{k},\mathbf{k}') = F^s(\vartheta) + \hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{\sigma}}' F^a(\vartheta), \tag{6}$$

where N_F^* is the total (including the spin degeneracy), renormalized density of states at the Fermi energy, ϑ is the angle between the momenta $\mathbf{k}$ and $\mathbf{k}'$ of two quasiparticles on the Fermi surface ($|\mathbf{k}| = |\mathbf{k}'| = k_F$), functions F^s and F^a describe the interaction in the symmetric (direct) and asymmetric (exchange) channels, and unprimed/primed σ-matrices refer to the spin state of the first/second quasiparticle.

The valley degree of freedom in graphene allows for more interaction channels. In addition to direct and exchange interactions between electrons within a single valley, one now also has an exchange interaction between the valleys, and a mixed spin-valley exchange interaction. If doping is low enough to neglect the trigonal warping of the Fermi contours, the Landau function of a two-valley FL can be written as[93]

---

[1] Theoretical estimates place the Kane-Mele coupling in the range from 1 μeV [84, 85] to 25-50 μeV,[86, 87] while a recent ESR experiment reports the value of 42.2 μeV [88]

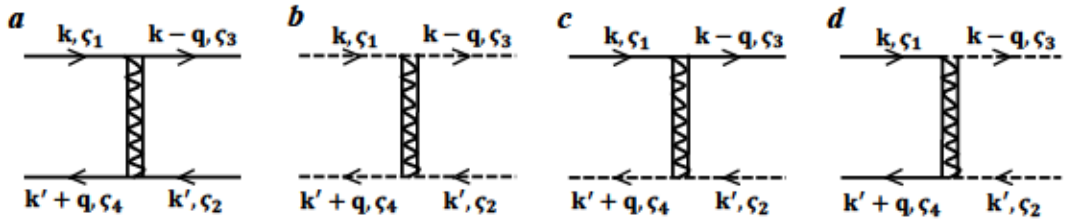


Figure 1. Interaction vertices for intra-valley (a and b) and inter-valley (c and d) scattering processes. The solid (dashed) lines refer to electrons in the K (K') valley. Diagrams a and b also have exchange partners with outgoing states swapped (not shown). Diagram d involves a large momentum transfer $\approx |\mathbf{K}-\mathbf{K}'|$. Reproduced with permission from Ref. 94. Copyright 2021 of the American Physical Society.

$$N_F^*\hat{f}(\mathbf{k},\mathbf{k}') = F^s(\vartheta)+\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{\sigma}}'F^a(\vartheta)+\hat{\boldsymbol{\tau}}_{||}\cdot\hat{\boldsymbol{\tau}}'_{||}G^{a||}(\vartheta)+\tau_z\tau'_zG^{a\perp}(\vartheta)+(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{\sigma}}')\left[\hat{\boldsymbol{\tau}}_{||}\cdot\hat{\boldsymbol{\tau}}'_{||}H^{||}(\vartheta)+\hat{\tau}_z\hat{\tau}'_zH^{\perp}(\vartheta)\right], \quad (7)$$

where where N_F^* is the total (including the spin and valley degeneracies), renormalized density of states, $\hat{\boldsymbol{\tau}}_{||}=\hat{x}\tau_x+\hat{y}\tau_y$ and unprimed/primed τ-matrices refer to the valley state of the first/second quasiparticle. Scattering processes in a two-valley FL are depicted diagrammatically in Fig. 1, where the solid (dashed) lines refer to electrons in the K (K') valley and $\varsigma_{1,2}$ ($\varsigma_{3,4}$) label the spin indices of the initial (final) states. Diagrams a and b describe intra-valley processes, while diagram c describes an inter-valley process with small momentum transfers, in which electrons stay in their respective valleys. Diagram d corresponds to an inter-valley process with large ($\approx |\mathbf{K}-\mathbf{K}'|$) momentum transfer, in which electrons are swapped between the valleys. For Coulomb interaction, the corresponding matrix element is smaller than the matrix elements in diagrams $a-c$. If the inter-valley matrix element is neglected, then the rotational symmetry in the valley space is restored,[93] i.e., $G^{a||}(\vartheta)=G^{a\perp}(\vartheta)$ and $H^{a||}(\vartheta)=H^{a\perp}(\vartheta)$.

Collective modes of a FL in the presence of SOC and/or Zeeman magnetic field (i.e., a field that affects only electron spins but not their orbital motion) can be found from the self-consistent FL kinetic equation for the (matrix) occupation number $\hat{n}(\mathbf{k},\mathbf{r},t)$

$$\frac{\partial\hat{n}}{\partial t}+i[\hat{\varepsilon},\hat{n}]+\frac{1}{2}\{\boldsymbol{\nabla}_{\mathbf{k}}\hat{\varepsilon},\boldsymbol{\nabla}_{\mathbf{r}}\hat{n}\}-\frac{1}{2}\{\boldsymbol{\nabla}_{\mathbf{r}}\hat{\varepsilon},\boldsymbol{\nabla}_{\mathbf{k}}\hat{n}\}=0, \quad (8)$$

where $[\hat{a}_1,\hat{a}_2]=\hat{a}_1\hat{a}_2-\hat{a}_2\hat{a}_1$, $\{\hat{\mathbf{a}}_1,\hat{\mathbf{a}}_2\}=\hat{\mathbf{a}}_1\cdot\hat{\mathbf{a}}_2+\hat{\mathbf{a}}_2\cdot\hat{\mathbf{a}}_1$, and $\hat{\varepsilon}(\mathbf{k},\mathbf{r},t)$ is the functional of quasiparticle energy. In the most general case, $\hat{\varepsilon}(\mathbf{k},\mathbf{r},t)$ can be written as

$$\hat{\varepsilon}(\mathbf{k},\mathbf{r},t)=\hat{\varepsilon}_0(\mathbf{k})+\hat{\varepsilon}_{\mathrm{SO}}(\mathbf{k})+\hat{\varepsilon}_{\mathrm{Z}}+\hat{\varepsilon}_{\mathrm{FL}}(\mathbf{k},\mathbf{r},t), \quad (9)$$

where $\hat{\varepsilon}_0$ is the equilibrium quasiparticle energy in the absence of SOC and Zeeman field, $\hat{\varepsilon}_{\mathrm{SO/Z}}$ are the changes in the energy due to SOC/Zeeman field, and

$$\hat{\varepsilon}_{\mathrm{FL}}(\mathbf{k},\mathbf{r},t)=\mathrm{Tr}'\int\frac{d^Dk'}{(2\pi)^D}\hat{f}(\mathbf{k},\mathbf{k}')\hat{n}'(\mathbf{k}',\mathbf{r},t) \quad (10)$$

accounts for the interaction of a given quasiparticle with the rest. The effect of oscillatory magnetic and electric fields, applied in the ESR and EDSR measurements, are accounted for by adding the corresponding terms to the right-hand side (RHS) of Eq. (9).

A number of comments are in order.

i) In the case of graphene with broken A/B symmetry, the kinetic equation (8) needs to be modified to include the effect of a non-Abelian Berry curvature, which arises due to the combined effect of broken A/B symmetry and Rashba SOC.[95] In this case, the term $\boldsymbol{\nabla}_{\mathbf{k}}\hat{n}$ is replaced by the covariant derivative $\mathcal{D}\hat{n}=\boldsymbol{\nabla}_{\mathbf{k}}\hat{n}-i[\mathcal{A},\hat{n}]$, where $\mathcal{A}$ is the non-Abelian Berry connection.[96–98] This leads to an effective orbital magnetic field which, in its turn, gives rise to an EDSR peak in the Hall conductivity, see Sec. V A.

ii) Equation 8 neglects scattering of quasiparticles either by external sources (disorder, phonons) or by other quasiparticles, which lead to damping of collective modes. These effects will be discussed separately in Sec. VII.

iii) The FL theory cannot describe a collective spin excitation that condenses out of the continuum of spin-flip transitions between the lower and upper cones of the Dirac surface state described by Eq. (4). The reason is that the energy of such an excitation is comparable to $2E_F$, while the FL theory can only describe physics at energies much smaller than E_F. In this case, one has to use microscopic, rather than phenomenological methods to describe the electron-electron interaction, see Sec. VIII B.

iv) If the reader is content with our assumption that the effects of SOC and Zeeman field on a FL can be treated perturbatively, they can skip the next section and go directly to Sec. III. A more demanding reader is invited to read the next section, which explains why the FL theory cannot deal with strong SOC and/or strong Zeeman field.

C. Does the Fermi-liquid theory work for arbitrarily strong spin-orbit coupling and/or magnetic field?

At first glance, the answer to the question in the title of this section is in the affirmative. All one needs to do is to construct a new Landau interaction function, accounting for broken rotational invariance in the spin space and, in the case of SOC, for coupling between momentum and spin. A modification of the Landau function in Eq. (6) for the case of a ferromagnetic metal was proposed long time ago in Ref. 99. For the case of Rashba SOC, the modified Landau function was composed in Ref. 100, but it is too long to be displayed here. It suffices to say that it contains eight instead of two components which, in contrast to Eq. (6), depend not only on the angle between $\mathbf{k}$ and $\mathbf{k}'$ but also on the magnitudes of these momenta. And this is the first sign of a problem with the FL theory. To make this problem more evident, we consider the case of non-interacting electrons with Rashba spectrum (cf. Fig. 2a):

$$\varepsilon_k^{\pm} = \frac{k^2}{2m} \pm \alpha k. \tag{11}$$

If both Rashba branches of the spectrum are occupied, the Fermi surface consists of two concentric circles with radii $k_F^{\pm} = \mp m\alpha + \sqrt{(m\alpha)^2 + 2mE_F}$, as shown in Fig. 2b. Let us calculate the spin susceptibility. In the diagrammatic language, the ij component of the spin susceptibility is given by a polarization bubble with the corresponding Pauli matrices at the vertices. In the absence of SOC, the spin susceptibility comes from the states in the immediate vicinity of the Fermi surface. This is not so in the presence of SOC. The simplest case is when the Zeeman magnetic field is applied along the normal to the 2DEG plane, i.e., $i = j = z$, see Fig. 2c. In this case the $T = 0$ static susceptibility arises solely from transitions between the two Rashba branches:

$$\begin{aligned} \chi_{zz} &= \frac{g^2\mu_B^2}{2} \int_0^{\infty} \frac{dkk}{2\pi} \frac{n_F(\varepsilon_k^-) - n_F(\varepsilon_k^+)}{\varepsilon_k^+ - \varepsilon_k^-} \\ &= \frac{g^2\mu_B^2}{2} \int_{k_F^+}^{k_F^-} \frac{dk}{4\pi\alpha} = \frac{g^2\mu_B^2}{2} \frac{m}{2\pi}, \end{aligned} \tag{12}$$

where $n_F(\varepsilon)$ is the Fermi function. We see that the integral over k does not come from the vicinity of either Fermi circle, but rather from the entire interval between the Ferm circles of width $\Delta k_F = k_F^- - k_F^+ = 2m\alpha$.[2] It is not a problem for a non-interacting case, because the spectrum is known for arbitrary k.[3] But it does become a problem for the interacting system. Indeed, the concept of quasiparticles is applicable only to long-lived states in the vicinity of each of the Fermi circle, see Fig. 2d. But states away from each of the Fermi circles (shown by red shaded regions) are just some complicated many-body states, which the FL theory cannot describe.

Suppose that one ignores this warning and goes ahead with calculating the renormalized g-factor, using the Landau function modified by Rashba SOC from Ref. 100. The only modification which matters here is that, because SOC breaks spin-rotational invariance, the $\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\sigma}}' F^a$ term in Eq. (6) is replaced by a combination $\hat{\boldsymbol{\sigma}}_{||} \cdot \hat{\boldsymbol{\sigma}}'_{||} f^{a||}(\mathbf{k}, \mathbf{k}') + \hat{\sigma}_z \hat{\sigma}'_z f^{a\perp}(\mathbf{k}, \mathbf{k}')$, where $\hat{\boldsymbol{\sigma}}_{||} = \hat{\sigma}_x \hat{\mathbf{x}} + \hat{\sigma}_y \hat{\mathbf{y}}$. With this modification, one follows the same procedure as for a usual FL [90, 91] and arrives at the integral equation, for e.g., the out-of-plane g-factor

$$g_{zz}^*(k) = g - \int_{k_F^{*+}}^{k_F^{*-}} \frac{dk'k'}{\pi} \frac{f_0^{a\perp}(k,k')}{\varepsilon_{k'}^{*+} - \varepsilon_{k'}^{*-}} g_{zz}^*(k'), \tag{13}$$

---

[2] The in-plane components, $\chi_{xx} = \chi_{yy}$, consist of two parts: one is the Ferm-surface contribution, as in the absence of SOC, and another one is the contribution from the entire interval Δk_F.

[3] Note that the integral in Eq. (12) can be also solved by rewriting it as $\chi_{zz} = (1/4\pi\alpha) \int_0^{\infty} dk \left[n_F(\varepsilon_k^+) - n_F(\varepsilon_k^-) \right]$ and then integrating by parts. In this way, one obtains the same result as the sum of two contributions from the Fermi circles. This means that spin susceptibility is an anomalous quantity in the field-theoretical sense, i.e., it can be viewed equivalently either as low-energy or high-energy contribution.[101] But this equivalence is lost in the presence of interactions.

where * indicates a renormalized quantity. (Note that the Luttinger theorem guarantees only that the total area of the Fermi surface is not renormalized, i.e., that $k_F^{+*2} + k_F^{-*2} = k_F^{+2} + k_F^{-2}$, which does exclude renormalizations of $k_F^{\pm}$ individually.)

A brief inspection of Eq. (13) shows that, in general, it does not make sense. Indeed, the integral on its RHS involves energies of quasiparticles at an arbitrary point between two Fermi circles. However, as we explained before, quasiparticle states are not well-defined away from either of the Fermi circles. Therefore, the renormalization of the g-factor cannot be calculated within the FL theory. Only if SOC is weak and thus $\Delta k_F \ll k_F$, one can replace $\varepsilon_k^{*\pm}$ by their quasiparticle forms. But in this case the width of the integration interval cancels out with the energy splitting due to SOC, and we are back to the usual result for the renormalized g-factor: $g^* = g/(1+F_0^a)$. (For weak SOC, the difference between the in-plane and out-of-plane components of the exchange interaction becomes negligible, and we replaced $F_0^{a,\perp}$ by F_0^a at the last step.)

A similar conclusion had eventually been reached in regard to a FL in the Zeeman magnetic field: if the Zeeman splitting is comparable to the Fermi energy, the FL theory does not work.[102–105] Interestingly, problems with applying the FL theory to the cases of strong spin polarization, ferromagnetism, and SOC had been foreseen by Conway Herring as early as in 1966.[106] The problem discussed above is a generic feature of systems with broken $SU(2)$ symmetry.

What was said above does not imply that an interacting system of electrons with strong SOC belong to the category of non-Fermi liquids. Indeed, momentum-resolved probes, such as angular-resolved photoemission, would find well-defined quasiparticles near each of the Fermi circles. What it means is that the thermodynamic quantities characterizing the spin sector of such a FL cannot be described by a (small) set of Landau parameters, but must involve the information about states away from the Fermi surfaces.

This is not such an unusual situation. For example, renormalization of the effective mass of a Galilean-invariant FL is described by a single Landau parameter: $m^* = m(1+F_1^s)$. However, if the system is not Galilean-invariant (but still isotropic), the last formula changes to $m^* = mQ(1+F_1^s)$, where Q involves integrals of the interaction vertices over the entire momentum space, [92, 107] and thus the effective mass in this case cannot be described by any finite number of Landau parameters.

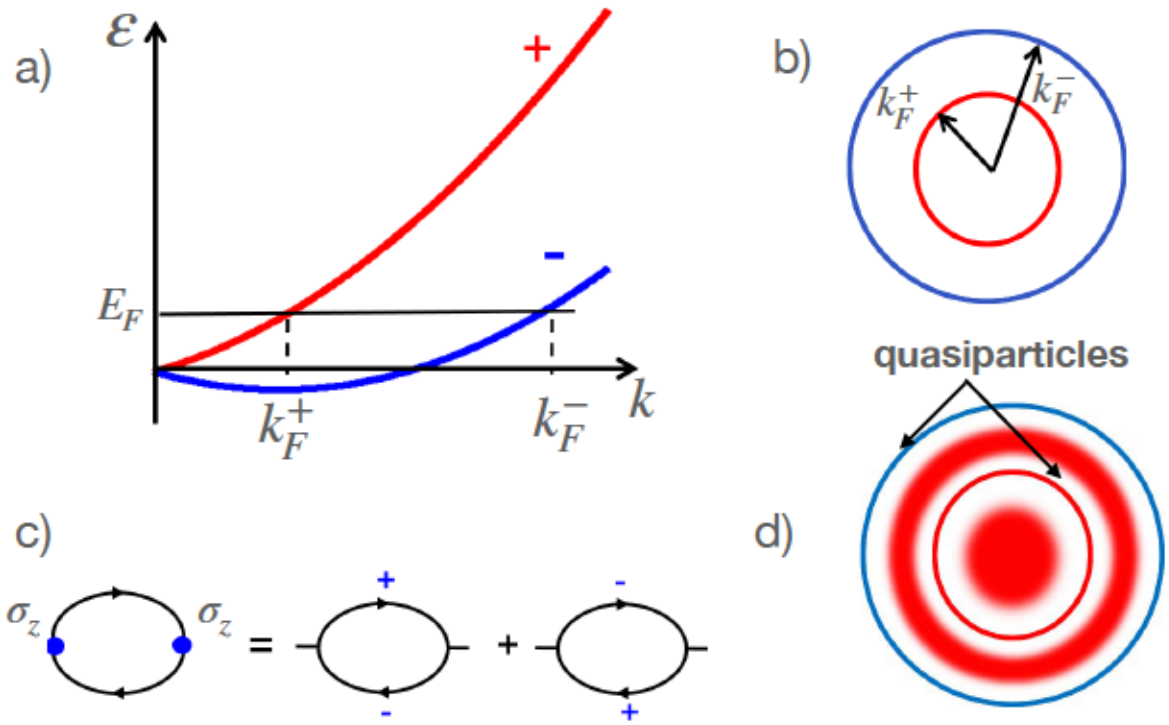


Figure 2. a) Spectrum of the Rashba Hamiltonian (1). b) The Fermi surface of the Rashba-split spectrum for non-interacting electrons. c) Diagrammatic representation of the out-of-plane spin susceptibility, χ_{zz}. $\pm$ label the Green's functions of the Rashba subbands. d) Same as in b) for interacting electrons. Red shaded regions denote states away from the Fermi circles which cannot be described by the Fermi-liquid theory.

III. PREAMBLE: SILIN MODES IN A PARTIALLY SPIN-POLARIZED FERMI LIQUID

Switching to the subject of collective modes, it is instructive to begin with a short reminder about the collective modes in a single-valley FL without SOC.[54, 91, 92] A neutral FL, e.g., normal ^{3}He, supports long-lived collective excitation in the (charge) density sector: zero-sound waves with acoustic spectrum (line cZS in Fig. 3*a*). In a charged FL, the $l = 0$ branch of zero-sound waves a replaced by plasmons, which are gapped in 3D (curve P_{3D}) and gapless in 2D (curve P_{2D}). The spin zero-sound mode does not fare that well. There is no general theorem that such a mode cannot be long-lived. However, as long as the interaction between fermions is repulsive (which corresponds to the attractive exchange interaction) and under reasonable assumptions about the harmonics of the Landau function $F^a(\vartheta)$ in Eq. (6), i.e., $F_0^a < F_1^a < 0$, etc., the mode is subsumed by the particle-hole continuum and thus overdamped (line sZS in Fig. 3*a*).

The situation changes in the presence of the Zeeman magnetic field, which gaps out the continuum of particle-hole excitations accompanied by spin-flips, see Fig. 3*b*. At $q = 0$ the gap in the continuum is equal to the renormalized Zeeman energy $\Delta_{\rm Z}^* = \Delta_{\rm Z}/(1+F_0^a)$, then it decreases as q increases, and vanishes eventually at $q \approx \Delta_{\rm Z}/v_F^*$, where v_F^* is the renormalized Fermi velocity. The region below the lower boundary of the continuum is now free of particle-hole excitations and can support long-lived collective modes.

The frequencies of these modes at $q = 0$ are found by writing the occupation number as $\hat{n}(\mathbf{k},t) = n_F + n_F' \left[\Delta_{\rm Z}^*/2 + \mathbf{u}(\mathbf{k},t)\cdot\hat{\boldsymbol{\sigma}}\right]$ and the quasiparticle energy as $\hat{\varepsilon}(\mathbf{k},t) = \varepsilon_0 + \Delta_{\rm Z}^* \hat{\mathbf{b}}\cdot\hat{\boldsymbol{\sigma}}/2 + \hat{\varepsilon}_{\rm FL}(\mathbf{k},t)$, where $\hat{\mathbf{b}}$ is the unit vector in the direction of the magnetic field and $n_F' = \partial n_F(\varepsilon)/\partial\varepsilon$, and using Eq. (8) without the gradient term.[91, 92] This gives an equation of motion for the vector $\mathbf{u}$

$$i\partial_t \mathbf{u}(\mathbf{k}_F,t) = \Delta_{\rm Z}^* \hat{\mathbf{b}} \times \left[\mathbf{u}(\mathbf{k}_F,t) + \int \frac{d\mathcal{O}'}{\mathcal{O}_D} \mathbf{u}(\mathbf{k}_F',t) F^a(\vartheta)\right], \tag{14}$$

where $\mathbf{k}_F = k_F\mathbf{k}/k$ and $\mathbf{k}_F' = k_F\mathbf{k}'/k'$, $d\mathcal{O}'$ is the element of the solid angle subtended by the vector $\mathbf{k}_F'$, and $\mathcal{O}_D$ is the full solid angle in D dimensions. Equation (14), which describes precession of the vector $\mathbf{u}$ around the magnetic field, is an eigenmode equation for the Silin collective modes.[54] The frequencies of these modes depend on the angular momentum, ℓ, in 3D or its projection onto the normal to the plane of motion, m, in 2D [92]

$$\Omega_n = \Delta_{\rm Z}^*(1+F_n^a) = \Delta_{\rm Z}\frac{1+F_n^a}{1+F_0^a}, \tag{15}$$

with $n = \ell$ in 3D and $n = m$ in 2D. Here, F_n^a are the harmonics of the Landau function defined by

$$F^a(\vartheta) = \sum_\ell (2\ell+1) F_\ell^a \mathcal{P}_\ell(\vartheta) \tag{16}$$

and

$$F^a(\vartheta) = \sum_m F_m^a e^{im\vartheta} \tag{17}$$

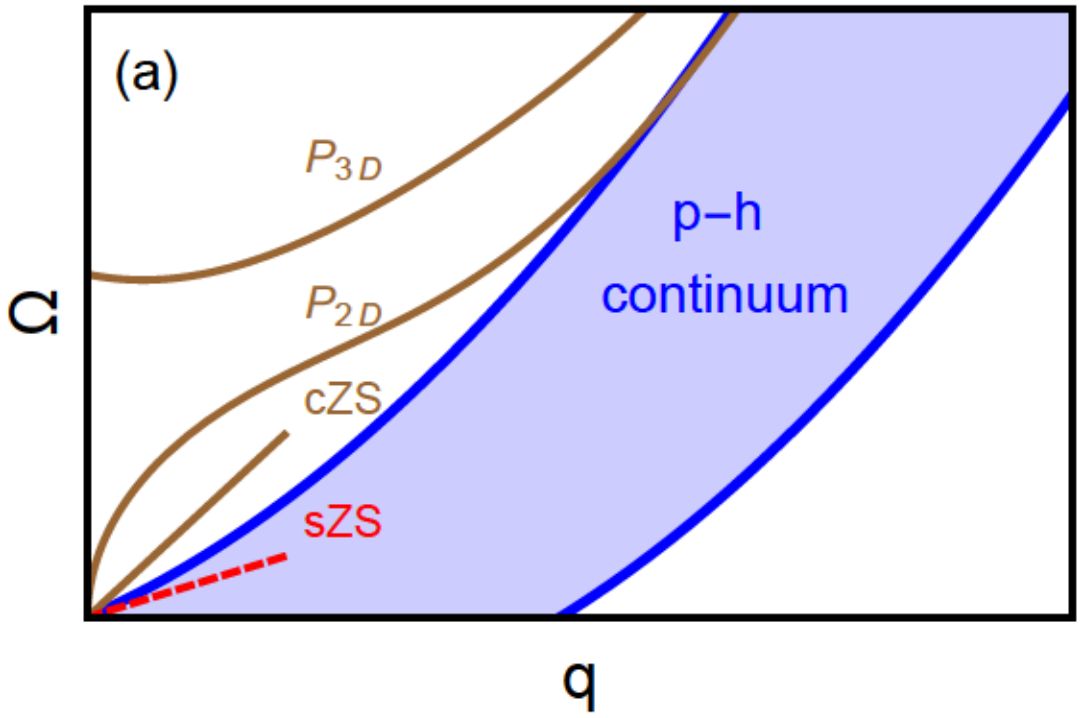
in 3D and 2D, respectively, and $\mathcal{P}_\ell(\vartheta)$ are the Legendre polynomials. Although there is, in principle, an infinite number of Silin modes, only the $n = 0$ mode couples to an oscillatory magnetic field, applied in an ESR or nuclear magnetic resonance (NMR) measurements. Indeed, the magnetization is expressed solely through the zeroth harmonic of the function $\mathbf{u}(\mathbf{k}_F,t)$: $\boldsymbol{\mathcal{S}} = -(1/2)g^*\mu_B N_F^* \int d\mathcal{O}\mathbf{u}(\mathbf{k}_F,t)/\mathcal{O}_D = (1/2)g^*\mu_B N_F^* \mathbf{u}^0(t)$.

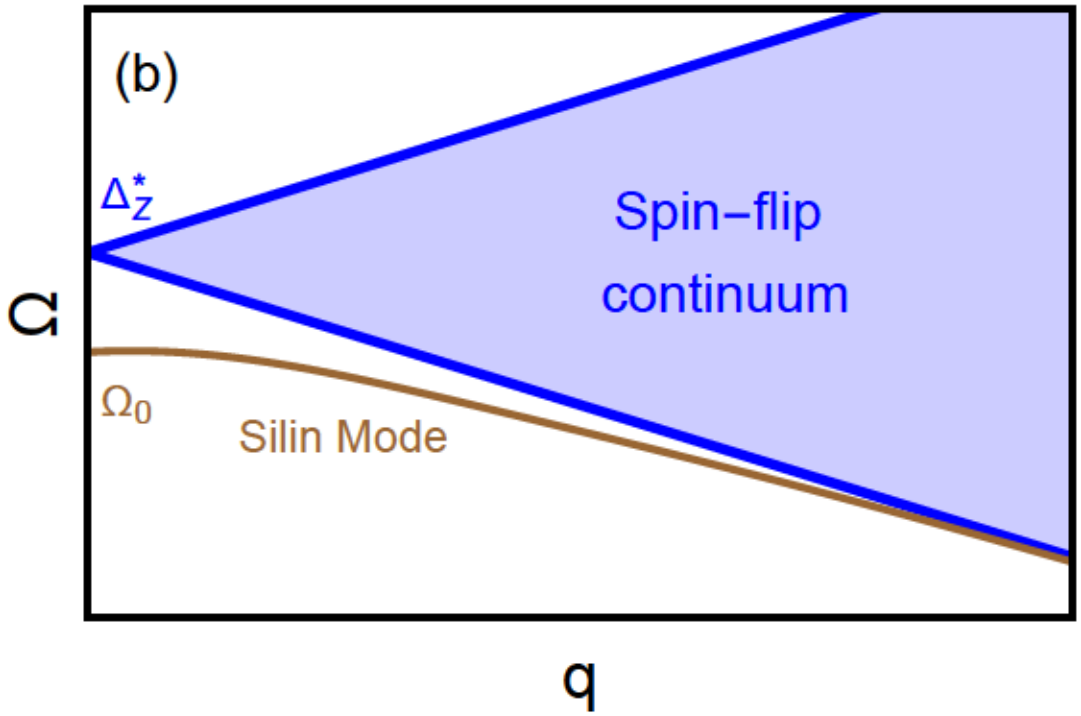
As we see from Eq. (15), the frequency of the $n = 0$ mode is not renormalized by the interaction and coincides with the Larmor frequency for free fermions. This is a general property of Hamiltonians with interactions that conserve spin and do not depend on velocities.[108] Diagrammatically, the results comes about as a cancellation between the self-energy and vertex corrections the spin susceptibility.[108]

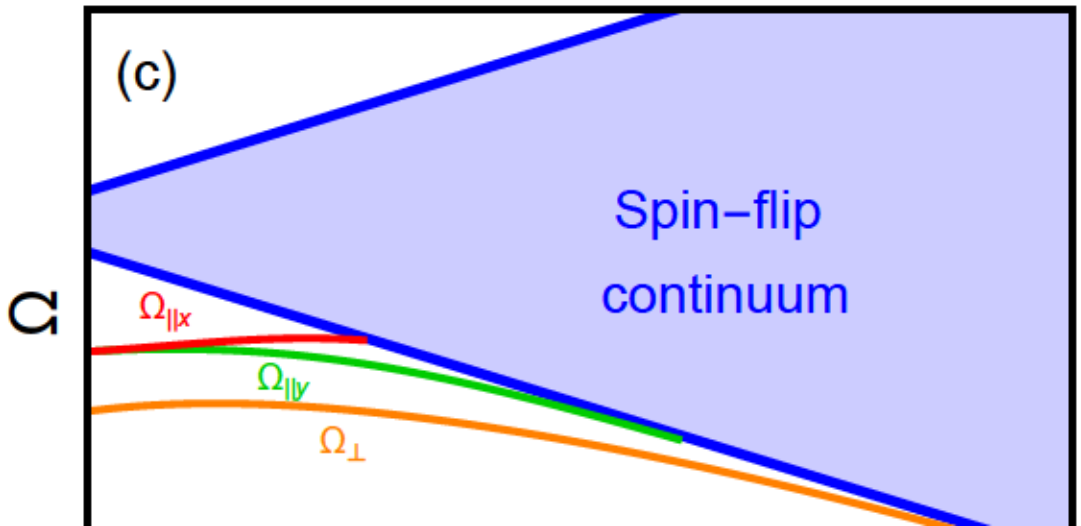
The difference between spin precession of free fermions and FL quasiparticles is illustrated in Fig. 4. Although the spins of free fermions (on the left) and FL quasiparticles (on the right) precess with the same frequencies, the phases of the former are not correlated but the phases of the latter are locked.

The $n = 0$ Silin mode disperses down with q, at first quadratically, and then grazes the continuum of spin-flip excitations, as shown in Fig. 3*b*. The downward sign of the dispersion can be explained by attraction in the exchange channel of the interaction.[55] Unlike the frequency at $q = 0$, which is not renormalized by the interaction, the functional form of the dispersion encapsulates all harmonics of F^a. In particular, the quadratic part contains harmonics F_0^a (which can be also extracted from the spin susceptibility, if the effective mass is measured independently

from the specific heat) and F_1^a (which cannot). The dispersion of the Silin mode in normal ^{3}He was studied extensively by NMR experiments in a spatially varying magnetic field which produced confined waves, see Ref. 109 and references therein. In solid-state systems, the Silin modes where measured in the late 60s by ESR on alkali metals. [110, 111] More recently, the dispersion of the Silin mode was measured by high-precision, finite-q Raman spectroscopy in a $Cd_{1-x}Mn_xTe$ quantum wells.[55–60] In this case, however, the Silin mode bears clear fingerprints of both Rashba and Dresselhaus SOCs, which are discussed in Sec. VIII A. A tantalizing proposal to observe Silin modes in a quantum spin liquid with spinon Fermi surface has recently been put forward in Ref. 112.


(a)
P_{3D}
P_{2D}
cZS
sZS
p–h
continuum
ω
q


(b)
Δ_z^*
Spin–flip
continuum
Ω_0
Silin Mode
ω
q


(c)
Spin–flip
continuum
$\Omega_{\|x}$
$\Omega_{\|y}$
$\Omega_\perp$
ω

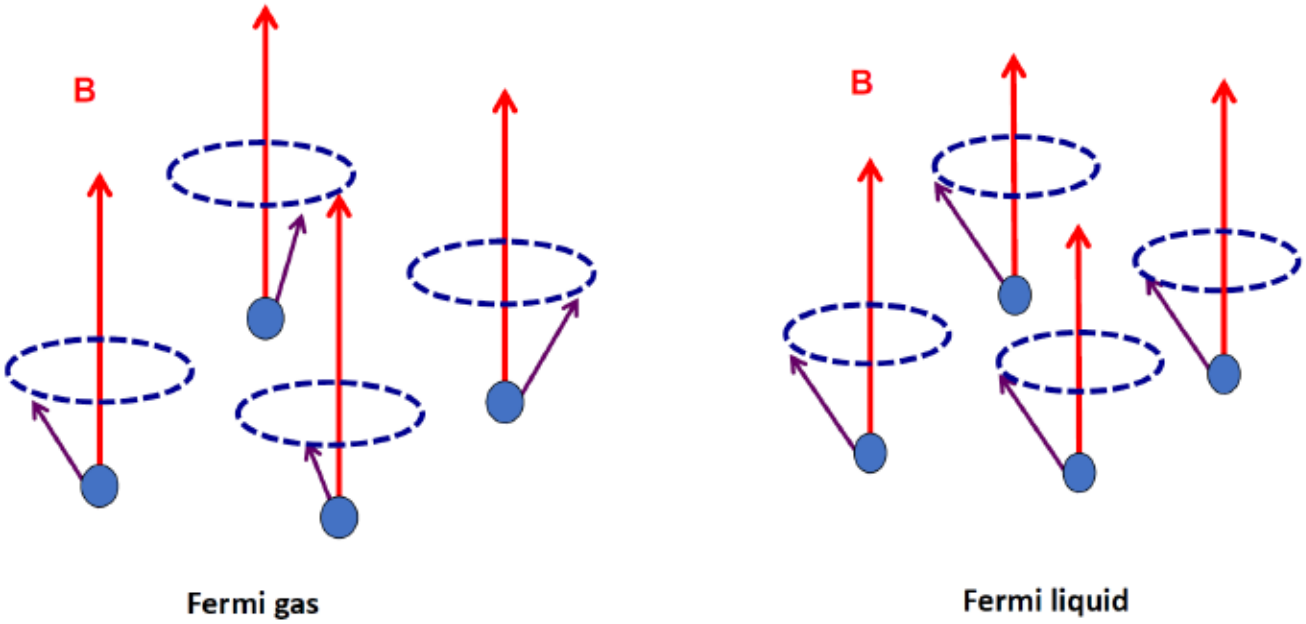


Figure 4. Precession of spins in a Fermi gas (left) and Fermi liquid (right). Although the frequencies of the $n = 0$ mode are the same in both cases, the phases of precessing spins are uncorrelated in a Fermi gas and locked in a Fermi liquid.

IV. CHIRAL SPIN WAVES IN TWO-DIMENSIONAL ELECTRON GASES

A. Chiral spin waves with Rashba spin-orbit coupling

We now proceed with discussing the main subject of this paper–collective spin modes in the presence of SOC–starting from the simplest case of a single-valley 2DEG with Rashba SOC, studied first in Ref. 113. The single-particle Hamiltonian is given by Eq. (1) and the Landau function is given by Eq. (6). It is convenient to introduce a set of rotated Pauli matrices which depend on the electron momentum[113]

$$\begin{aligned}\hat{\Sigma}_x(\mathbf{k}) &= -\hat{\sigma}_z,\ \hat{\Sigma}_y(\mathbf{k}) = \cos\phi_{\mathbf{k}}\hat{\sigma}_x + \sin\phi_{\mathbf{k}}\hat{\sigma}_y,\\ \hat{\Sigma}_z(\mathbf{k}) &= \sin\phi_{\mathbf{k}}\hat{\sigma}_x - \cos\phi_{\mathbf{k}}\hat{\sigma}_y,\end{aligned} \tag{18}$$

where $\phi_{\mathbf{k}}$ is the azimuthal angle of $\mathbf{k}$. When taken at the same momentum, $\hat{\Sigma}$'s obey the usual algebra $\left[\hat{\Sigma}_\alpha(\mathbf{k}), \hat{\Sigma}_\beta(\mathbf{k})\right] = i\epsilon^{\alpha\beta\gamma}\hat{\Sigma}_\gamma(\mathbf{k})$, where $\epsilon^{\alpha\beta\gamma}$ is the Levi-Civita tensor, while $\left[\hat{\Sigma}_y(\mathbf{k}), \hat{\Sigma}_z(\mathbf{k}')\right] = i\hat{\Sigma}_x(\mathbf{k})\cos(\phi_{\mathbf{k}} - \phi_{\mathbf{k}'})$. In this basis,

$$\hat{\varepsilon}_{\rm SO} = \frac{1}{2}\Delta_{\rm R}\hat{\Sigma}_z(\mathbf{k}), \tag{19}$$

where $\Delta_{\rm R} = 2\alpha k_F$. Treating $\hat{\varepsilon}_{\rm SO}$ as a static perturbation, we obtain the renormalized value of the Rashba splitting[113, 114]

$$\Delta^*_{\rm R} = \Delta_{\rm R}/(1 + F^a_1). \tag{20}$$

This energy marks the end point of the continuum of particle-hole excitations which involve spin-flip transitions between the branches of the Rashba spectrum, see Fig. 3*c*. In a non-interacting system, the continuum at $q = 0$ occupies a finite region of width $\Delta\Omega = 4m\alpha^2$. However, the FL theory is valid only to first order in α, therefore, within this theory the continuum shrinks to a single point at $\Omega = \Delta^*_{\rm R}$.

Now we consider a time-dependent perturbation of the occupation number

$$\delta\hat{n}(\mathbf{k}, t) = n'_F\left[\frac{1}{2}\Delta^*_{\rm R}\hat{\Sigma}_z(\mathbf{k}) + \mathbf{u}(\mathbf{k}, t)\cdot\hat{\mathbf{\Sigma}}(\mathbf{k})\right]. \tag{21}$$

Substituting this form into Eq. (8) (without the gradient term) and linearizing with respect to $\mathbf{u}$, we obtain the following equations of motion for angular harmonics of $\mathbf{u}$:

$$\partial_t u_x^m = -\Delta_{\mathrm{R}}^* u_y^m \left[1 + \frac{1}{2}\left(F_{m-1}^a + F_{m+1}^a\right)\right], \tag{22a}$$

$$\partial_t u_y^m = \Delta_{\mathrm{R}}^* u_y^m \left(1 + F_m^a\right), \tag{22b}$$

$$\partial_t u_3^m = 0. \tag{22c}$$

The last equation implies that $u_3^m = \text{const}$ and can thus be ignored. The eigenvalues of this system are the frequencies of a new type of collective modes: chiral spin resonances.[113] Their frequencies are

$$\Omega_m = \Delta_{\mathrm{R}}^* \sqrt{\left[1 + \frac{1}{2}(F_{m+1}^a + F_{m-1}^a)\right](1 + F_m^a)}. \tag{23}$$

(Note that $\Omega_{-m} = \Omega_m$ because $F_{-m}^a = F_m^a$.) As in the case of the Silin mode, there is an infinite number of chiral-spin resonances. Now, however, only two modes with $m = 0$ and $m = 1$ couple to the external magnetic field. This is because the magnetization

$$\boldsymbol{\mathcal{S}} = -\frac{1}{4} g^* \mu_B N_F^* \mathrm{Tr} \int \frac{d\phi_{\mathbf{k}}}{2\pi} \left[\mathbf{u}(\mathbf{k}_F, t) \cdot \hat{\boldsymbol{\Sigma}}(\mathbf{k}_F)\right] \hat{\boldsymbol{\sigma}} \tag{24}$$

contains not only the 0^{th} but also $\pm 1^{\text{st}}$ harmonics of $\mathbf{u}$ due to the angular dependence of matrices $\hat{\Sigma}_\alpha(\mathbf{k})$. The $m = 0$ mode with frequency

$$\Omega_0 \equiv \Omega_\perp = \Delta_{\mathrm{R}}^* \sqrt{(1 + F_0^a)(1 + F_1^a)} \tag{25}$$

corresponds to oscillations of the z-component of $\boldsymbol{\mathcal{S}}$, while the doubly-degenerate $m = \pm 1$ mode with frequency

$$\Omega_{\pm 1} \equiv \Omega_{||} = \Delta_{\mathrm{R}}^* \sqrt{\left[1 + \frac{1}{2}\left(F_0^a + F_2^a\right)\right](1 + F_1^a)} \tag{26}$$

corresponds to oscillations of the in-plane components of $\boldsymbol{\mathcal{S}}$. Provided that $F_0^a < F_2^a$, we have $\Omega_\perp < \Omega_{||}$, which is the case shown in Fig. 3*c*. Since spin is not a conserved quantity anymore, the frequencies of all modes are renormalized away from Δ_{R}.

The chiral spin resonances are the $q = 0$ end points of the dispersive modes: chiral spin waves (CSW).[46, 115] Once the direction of $\mathbf{q}$ in the plane is chosen, the degeneracy of the two in-plane mode is lifted and there are now two modes with frequencies $\Omega_{||x}(\mathbf{q})$ and $\Omega_{||y}(\mathbf{q})$, polarized along and perpendicular to $\mathbf{q}$, respectively. The dispersions of these modes are shown schematically in Fig. 3*c*. The in-plane modes run into the continuum at some values of q and disappear from the spectrum. The out-of-plane mode behaves similarly to the Silin mode: it disperses down with q and grazes the continuum of particle-hole excitations. We will discuss the spatial dispersion of CSW in more detail in Sec. VI.

The equations of motion (22a)-(22c) allow for a simple physical interpretation.[116] Namely, one can think of vector $\mathbf{u}^m(t)$ as a classical spin on site m of a linear spin chain, aligned along the y-axis as in Fig. 5. Spins do not interact with each other but are subject to an effective magnetic field due to Rashba SOC, directed along the z axis and of magnitude $\Delta_{\mathrm{R}}^*/2$. The effective Landé factor of these spins is anisotropic in the (x, y) plane with components $\gamma_x^m = 2 + F_{m+1}^a + F_{m-1}^a$ and $\gamma_y^m = 2(1 + F_m^a)$. The spin chain is *non-uniform* because γ_1^m and γ_2^m depend on the lattice site. Both anisotropy and site-dependence of the g-factor arise from the FL interaction. Because $u_3^m = \text{const}$, the spins precess around the Rashba field. The effective-lattice interpretation will be even more instructive in the case of both Rashba SOC and Zeeman field being present, which is discussed in the next section.

B. Effective lattice model

Now we add a Zeeman field aligned along the x-axis. The total Hamiltonian is the sum of Eqs. (1) and (5), and the additional term in the quasiparticle energy in Eq. (9) is

$$\hat{\varepsilon}_{\mathrm{Z}} = \frac{\Delta_{\mathrm{Z}}^*}{2} \hat{\sigma}_x = \frac{\Delta_{\mathrm{Z}}^*}{2} \left[\cos\phi_{\mathbf{k}} \hat{\Sigma}_x(\mathbf{k}) + \sin\phi_{\mathbf{k}} \hat{\Sigma}_y(\mathbf{k})\right]. \tag{27}$$

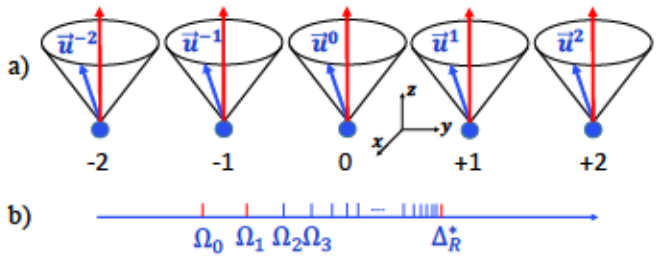


Figure 5. a) An interpretation of collective modes in a Fermi liquid with Rashba spin-orbit coupling in terms of a linear spin chain. b) The spectrum of the system consists of an infinite number of discrete levels, converging towards the continuum at $\Delta_{\rm R}^*$. Reprinted with permission from Ref. 116. Copyright 2017 by the American Physical Society.

Correspondingly, the equations of motion for vector $\mathbf{u}$ become

$$\begin{aligned}
\partial_t u_x^m &= -\Delta_{\rm R}^* u_y^m \left[1+\frac{1}{2}\left(F_{m-1}^a+F_{m+1}^a\right)\right] - \frac{1}{2i}\Delta_{\rm Z}^*\left(u_y^{m-1}-u_y^{m+1}\right)\left(1+F_m^a\right) \\
&\quad +\frac{1}{2i}\Delta_{\rm R}^* u_z^m\left(F_{m-1}^a-F_{m+1}^a\right)+\frac{1}{2}\Delta_{\rm Z}^*\left(u_z^{m-1}+u_z^{m+1}\right)\left(1+F_m^a\right), \qquad (28a)\\
\partial_t u_y^m &= \Delta_{\rm R}^* u_x^m\left(1+F_m^a\right)+\frac{1}{2i}\Delta_{\rm Z}^*\left[u_x^{m-1}\left(1+F_{m-1}^a\right)-u_x^{m+1}\left(1+F_{m+1}^a\right)\right], \qquad (28b)\\
\partial_t u_3^m &= -\frac{1}{2}\Delta_{\rm Z}^*\left[u_x^{m-1}\left(1+F_{m-1}^a\right)+u_1^{m+1}\left(1+F_{m+1}^a\right)\right]. \qquad (28c)
\end{aligned}$$

Note that the angular dependence of the Zeeman energy in the Σ-basis in Eq. (27) leads to a non-locality in the m-space: the Zeeman terms in the equation of motion shift the angular momentum by ± 1. Now the equations of motion resemble those for the tight-binding (TB) model with three orbitals per site, in which the Rashba and Zeeman terms play the roles of on-site and hopping energies, respectively. To make this analogy more transparent, we eliminate components u_y^m and u_z^m and introduce the "Bloch wavefunction"

$$\psi_m = i^{-m}u_x^m. \qquad (29)$$

Assuming the oscillatory time dependence with frequency Ω, the equation for ψ_m is reduced to the TB form:[116]

$$\begin{aligned}
\Omega^2\psi_m &= \left[\Delta_{\rm R}^{*2}\left(1+F_m^a\right)\left(1+\frac{F_{m+1}^a+F_{m-1}^a}{2}\right)+\Delta_{\rm Z}^{*2}\left(1+F_m^a\right)^2\right]\psi_m \\
&\quad -\Delta_{\rm R}^*\Delta_{\rm Z}^*\left[\left(1+\frac{F_m^a+F_{m+1}^a}{2}\right)\left(1+F_{m+1}^a\right)\psi_{m+1}+\left(1+\frac{F_{m-1}^a+F_m^a}{2}\right)\left(1+F_{m-1}^a\right)\psi_{m-1}\right]. \qquad (30)
\end{aligned}$$

In the absence of interaction, Eq. (30) is simplified to

$$\Omega^2\psi_m = \left(\Delta_{\rm R}^2+\Delta_{\rm Z}^2\right)\psi_m - \Delta_{\rm R}\Delta_{\rm Z}\left(\psi_{m+1}+\psi_{m-1}\right). \qquad (31)$$

The eigenvalue of this equation

$$\Omega(\phi_{\mathbf{k}}) = \left[\Delta_{\rm R}^2+\Delta_{\rm Z}^2-2\Delta_{\rm R}\Delta_{\rm Z}\cos\phi_{\mathbf{k}}\right]^{1/2} \qquad (32)$$

is nothing but the difference between the energies of the Rashba subbands at a given direction of $\mathbf{k}$ (modulo a phase shift inflicted by the transformation (29)). $\Omega(\phi_{\mathbf{k}})$ disperses with $\phi_{\mathbf{k}} \in (0, 2\pi)$, which is a conjugate variable to m. Therefore, $\phi_{\mathbf{k}}$ plays the role of "quasimomentum" confined to the first Brillouin zone $(0, 2\pi)$. The minimum and maximum values of $\Omega(\phi_{\mathbf{k}})$, $\Omega_{\rm max} = \Delta_{\rm R}+\Delta_{\rm Z}$ and $\Omega_{\rm min} = |\Delta_{\rm R}-\Delta_{\rm Z}|$, mark the edges of the "conduction band", which is nothing but the particle-hole continuum. At $\Delta_{\rm Z} = \Delta_{\rm R}$ the gap in the continuum collapses to zero. At this value of the magnetic field, the Fermi contours of the Rashba branches touch at one point and thus a transition between the two branches costs no energy.

We now come back to the interacting version of Eq. (30), which allows for a simple physical interpretation. Imagine a 1D lattice with sites labeled by index $m = 0, \pm 1, \ldots$. The Bloch wavefunction ψ_m resides on these sites. The first

Table I. Mapping of the Fermi-liquid kinetic equation onto an effective 1D tight-binding model.

Fermi-liquid kinetic equation	**1D tight-binding model**
angular momentum m	lattice site m
azimuthal angle of momentum $\mathbf{k}$ ($\phi_{\mathbf{k}}$)	quasimomentum
m^{th} harmonic of the spin part of the occupation number u_x^m	Bloch wavefunction on site m
Rashba spin-orbit coupling	on-site energy
Zeeman splitting	nearest-neighbor hopping
Dresselhaus spin-orbit coupling	next-to-nearest neighbor hopping
harmonics of the Landau function	on-site and bond defects
continuum of spin-flip particle-hole excitations	conduction band
collective modes	bound states

term on the RHS of Eq. (30) is the energy on site m, the last two terms describe hopping between site m and its the nearest neighbors, $m \pm 1$. If we artificially remove all F_m^a's from the equation but keep the renormalized values of Δ_{R} and Δ_{Z}, we will obtain the renormalized band (continuum) with boundaries $\Omega^*_{\text{max}} = \Delta^*_{\text{R}} + \Delta^*_{\text{Z}}$ and $\Omega^*_{\text{min}} = |\Delta^*_{\text{R}} - \Delta^*_{\text{Z}}|$. The role of F_m^a's is to renormalize both on-site energies and hopping matrix elements. Because F_m^a depends on m, both the on-site energies and hopping matrix elements vary along the lattice. In other words, F_m^a's introduce both one-site and bond defects.

In principle, F_m^a is non-zero for any m. Therefore, each site of our lattice is different from the others, and we can view our lattice as an ordered alloy composed from an infinite variety of chemically distinct atoms. But if we keep only a few first harmonics of $F^a(\vartheta)$ and neglect the rest, the central region of the lattice will contain defects, while the outer regions will be defect-free. Lattice defects produce bound states which split off the band. In this language, therefore, the collective modes of a FL are the bound states of an effective 1D lattice. Studying a much more transparent problem of bound states, one can understand a more complicated case of a FL with SOC and in the presence of the magnetic field.

The case of Dresselhaus SOC can be analyzed in a similar way. According to Eq. (2), the corresponding term in the quasiparticle energy reads

$$\begin{aligned} \hat{\varepsilon}_{\text{SO}} &= \frac{\Delta^*_{\text{D}}}{2} \left(\cos\phi_{\mathbf{k}} \hat{\sigma}_x - \sin\phi_{\mathbf{k}} \hat{\sigma}_y\right) \\ &= \Delta^*_{\text{D}} \left[\cos 2\phi_{\mathbf{k}} \hat{\Sigma}_y(\mathbf{k}) + \sin 2\phi_{\mathbf{k}} \hat{\Sigma}_z(\mathbf{k})\right], \end{aligned} \tag{33}$$

where $\Delta^*_{\text{D}} = \Delta_{\text{D}}/(1 + F_1^a)$. A double angle in this equation implies that the corresponding terms in the equations of motion shift the angular momentum by ± 2. In the lattice interpretation, this corresponds to hopping between next-to-nearest neighbors. A complete dictionary of mapping between the FL kinetic equation and TB model is given in Table I.

We note that the zero-sound modes of a 2D FL in the absence of SOC and Zeeman field can also be understood as bound states of a 1D lattice. In this case, hopping between sites arises due the gradient term in the kinetic equation, which also shifts the harmonic index by ± 1. For example, the shear zero-sound waves of a 2D FL were analyzed in this way in Ref. 117.

C. Effective lattice model for Rashba spin-orbit coupling and Zeeman magnetic field

We now illustrate how the lattice model works for the case when both Rashba SOC and Zeeman magnetic field are present. In addition, we assume that the Landau function is isotropic and thus contains only the zeroth harmonic: $F_m^a = \delta_{m,0} F_0^a$ (the s-wave approximation). In a dimensionless form and in the s-wave approximation, Eq. (30) can be written as

$$\omega^2 \psi_m = W\psi_m - J\left(\psi_{m+1} + \psi_{m-1}\right) + U_0 \delta_{m,0} \psi_m + U_1 \left(\delta_{m,1} + \delta_{m,-1}\right)\psi_m - \delta J \delta_{m,0}\left(\psi_1 + \psi_{-1}\right) - \delta J' \left(\delta_{m,1} + \delta_{m,-1}\right)\psi_0, \tag{34}$$

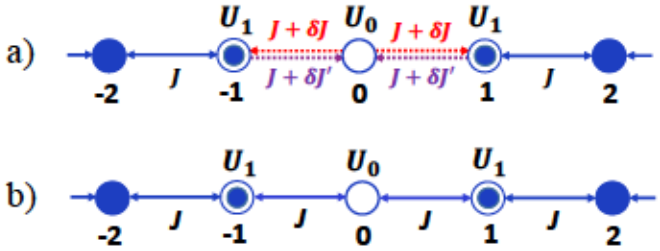


Figure 6. a) Effective lattice for the FL kinetic equation (34) with Rashba spin-orbit coupling and in-plane magnetic field, and in the s-wave approximation for the Landau function. Filled circles represent sites of ideal lattice with on-site energies W, connected by bonds with hopping amplitudes J. The $m = 0, \pm 1$ sites are occupied by "impurities" with on-site energies U_0 and $U_{\pm 1}$, respectively, connected by "defective" bonds with hopping amplitudes $J \pm \delta J$. b) A simplified version of the effective lattice with on-site disorder only. Reprinted with permission from Ref. 116. Copyright 2017 by the American Physical Society.

where

$$\omega = \Omega/\Delta_{\rm R},\ J = \Delta_{\rm Z}^*/\Delta_{\rm R},\ W = 1 + J^2, \tag{35a}$$

$$U_0 = F_0^a \left[1 + J^2 \left(2 + F_0^a\right)\right],\ U_1 = \frac{F_0^a}{2} \tag{35b}$$

$$\delta J = J\frac{F_0^a}{2},\ \delta J' = \frac{F_0^a}{2}\left(3 + F_0^a\right). \tag{35c}$$

[Note that, according to Eq. (20), $\Delta_{\rm R}$ is not renormalized in the s-wave approximation.] The corresponding lattice is shown pictorially in Fig. 6*a*. The first two terms on the RHS of Eq. (34) describe an ideal lattice with "on-site energies" W and "hopping amplitudes" between the nearest neighbors J. The rest of the terms represent "defects". The third and fourth term correspond to three "impurities" on sites $m = 0$ and $m = \pm 1$ with on-site energies U_0 and U_1, respectively. The last two terms describe defective bonds between the $0^{\rm th}$ and $\pm 1^{\rm st}$ sites. The amplitudes are equal to $J + \delta J$ for hopping from 0 to ± 1 and $J + \delta J'$ for hopping in the opposite direction, i.e., the bond defects are *chiral*. This means that the effective TB Hamiltonian is non-Hermitian. This does not present any difficulties, however, because the eigenvalues of Eq. (34) are real.

To solve Eq. (34), we choose wavefunctions ψ_0 and $\psi_{\pm 1}$ as independent variables, and assume that, starting from sites $m = \pm 2$, the wavefunctions of the bound states decreases exponentially with m :

$$\psi_{\pm(|m|+2)} = e^{-(|m|+1)\lambda}\psi_{\pm 1} \tag{36}$$

with ${\rm Re}\lambda > 0$. This yields a transcendental equation for λ, whose solutions are presented in Fig. 7. The left panel is for magnetic fields below the gap closing point ($\Delta_{\rm Z}^* < \Delta_{\rm R}$). There are two collective modes in zero field–these are the same modes as given by Eqs. (25) and (26). An in-plane magnetic field lifts the double degeneracy of the $\Omega_{||}$ mode (similar to the case of finite q in Fig. 3*c*) and, at finite but small $\Delta_{\rm Z}^*$ there are three modes. The in-plane modes run into the continuum at some critical values of $\Delta_{\rm Z}^*$, but the out-of-plane mode continues to graze the continuum down the gap-closing point. For $\Delta_{\rm Z}^* > \Delta_{\rm R}$, the single mode appears again, see Fig. 7, right. Its frequency increases with the magnetic field and, in the limit of $\Delta_{\rm Z}^* \gg \Delta_{\rm R}$, the mode evolves into the Silin mode with frequency $\Delta_{\rm Z}$.

A rather complex spectrum shown in Fig. 7 can be understood in terms of simplified versions of the TB model. Namely, the merging of the two in-plane modes with the continuum can be understood qualitatively by ignoring bond defects, i.e., by setting $\delta J = \delta J' = 0$. In this case, we have a TB model with identical bonds between all sites and three impurities on sites $m = 0, \pm 1$, see Fig. 6*b*. For an even simpler case of a single impurity in a 1D lattice, it is well-known that there always exists a bound state located either below (for an attractive impurity) or above (for a repulsive impurity) the conduction band. For realistic values $-1 < F_0^a < 0$, our "impurities" are attractive, and thus the bound state is below the band. Given that a single impurity has at least one bound state, it is natural to expect that a complex of three impurities will have up to three bound states, which is indeed confirmed by an explicit solution of the TB model (see Appendix A.1b in Ref. 116). Qualitatively, this can be understood in the continuum limit, in which a three-impurity complex is replaced by a 1D potential well of finite width (a) and depth (U). Such a well has at least one bound state but may also have two, three, etc. states, if the product Ua exceeds some critical values.

Our original problem corresponds to a TB model with parameters given by Eqs. (35a-35c). In the limit $\Delta_{\rm Z}^* \ll \Delta_{\rm R}$ and not too weak interaction $|F_0^a| \sim 1$, the potential energies of the impurity sites are of the order of 1, which is much larger than the bandwidth $2J = 2\Delta_{\rm Z}^*/\Delta_{\rm R} \ll 1$. Thus we have three strong impurities with the maximum number of bound states, equal to three. As $\Delta_{\rm Z}^*$ increases, the bandwidth increases linearly whereas the potential energies

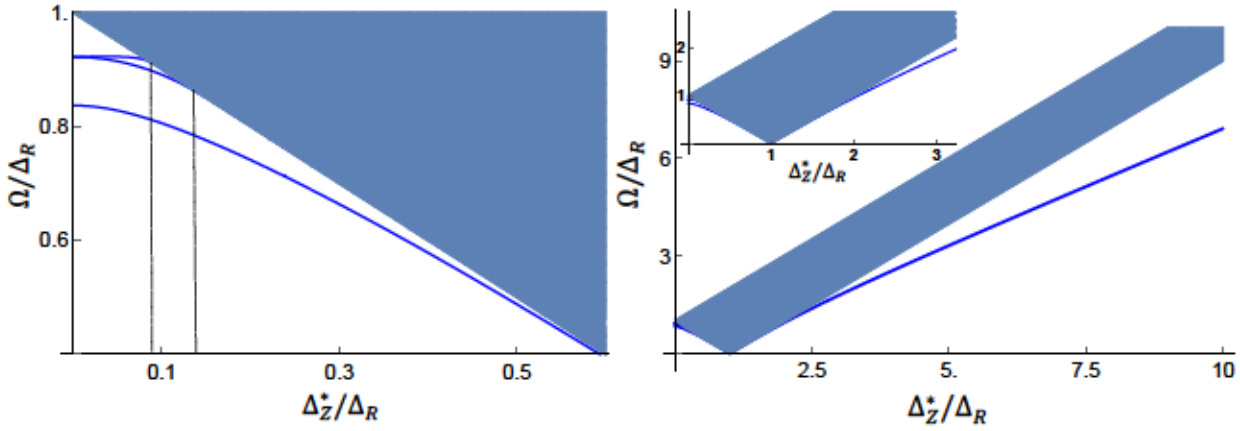


Figure 7. Collective modes of a FL with Rashba SOC and in the presence of the in-plane magnetic field. The Landau function is taken in the s-wave approximation: $F^a(\theta) = F_0^a = -0.3$. Δ_Z^* is the (renormalized) Zeeman energy and Δ_R is the Rashba energy splitting. Left: $\Delta_Z^* < \Delta_R$. Right: $\Delta_Z^* > \Delta_R$. Inset: Same as in main panel for a wider range of fields. Reprinted with permission from Ref. 116. Copyright 2017 by the American Physical Society.

increase only as Δ_Z^{*2}. Therefore, the impurities get relatively weaker (compared to the bandwidth), and we lose first the highest and then next-to-highest bound state. The lowest bound state also disappears but only at the gap-closing point. This seems to contradict the fact that there is at least one bound state in a 1D problem regardless to its parameters. However, this statement is true only for the continuum Schroedinger equation, which does not have the notion of bonds. In our full problem, two bonds are defective and there is an interesting competition between the on-site and bond defects, which does allow the bound state to disappear precisely at $\Delta_Z^* = \Delta_R$.[116]

As the field increases beyond the gap-closing point ($\Delta_Z^* > \Delta_R$), the $m = 0$ impurity becomes stronger while the $m = \pm 1$ impurities remain the same. Therefore, we are back to a single-impurity problem with only one bound state or, equivalently, one collective mode.

From the solution of the kinetic equation one can also deduce the polarization of the collective modes. Figure 8 shows how the polarization of chiral spin modes changes with the magnetic field.[118] In zero field, there are linearly polarized modes. At finite $B < B_c$, where B_c is the field at which the gap in the continuum closes, along the x-axis, the longitudinal mode (with the magnetization along the field) remains linearly polarized, while two transverse mode (with the magnetizations perpendicular to the field) are elliptically polarized (as long as they are located outside the continuum). When the single wave emerges on the other side of the gap-closing point ($B > B_c$), it is elliptically polarized, with the magnetization vector precessing around the magnetic field. As the field becomes much larger than B_c, the wave transforms gradually into the Silin mode with circular polarization.

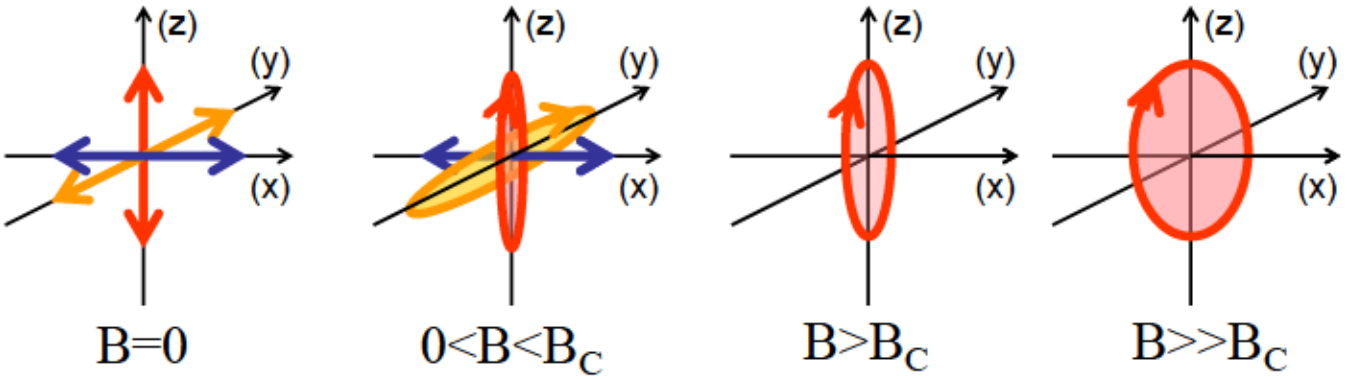


Figure 8. Polarizations of chiral spin modes at different values of the magnetic fields applied along the x-axis. B_c is the gap-closing field, at which $\Delta_Z^* = \Delta_R^*$. Reprinted with permission from Ref. 118. Copyright 2016 by the American Physical Society.

If the Landau function contains a large or infinite numbers of harmonics, one has to resort to numerical diagonal-

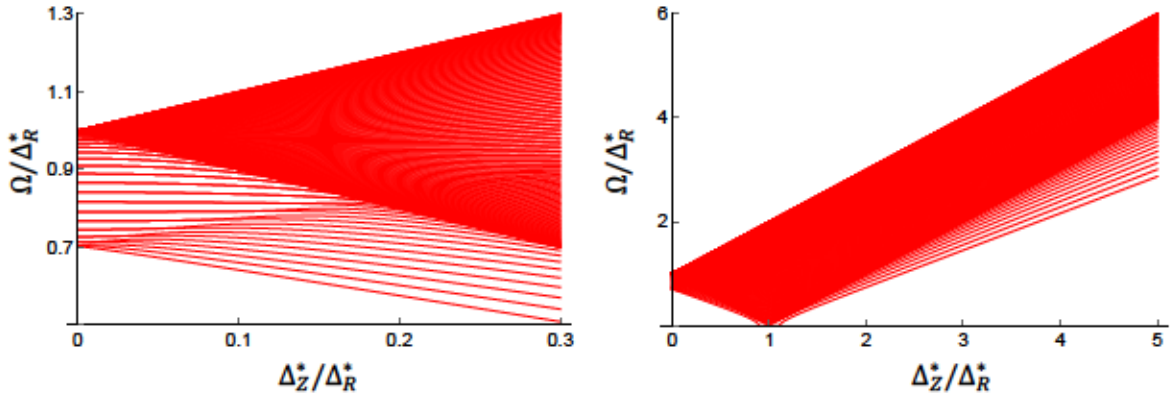


Figure 9. Collective modes of a FL with Rashba SOC (left) and in the presence of the in-plane magnetic field (right) for a model form of the Landau function: $F_m^a = F_0^a \exp(-m^2/m_0^2)$ with $F_0^a = -0.3$ and $m_0 = 10$. Reprinted with permission from Ref. 116. Copyright 2017 by the American Physical Society.

ization of Eq. (30). Figure 9 shows the result of such diagonalization for a model form of the Landau function, given by $F_m^a = F_0^a \exp(-m^2/m_0^2)$. In this case, the modes are very densely spaced and form a quasi-continuum.

V. COLLECTIVE SPIN MODES IN DIRAC SYSTEMS

A. Graphene with proximity-induced spin-orbit coupling

The case of graphene with proximity-induced SOC differs from the one considered in the previous section in two important aspects. First, at the single-particle level, the Hamiltonian (3) contains the valley-Zeeman (VZ) term, in addition to the Rashba term. Second, at the many-body level, the Landau function contains exchange interaction between the valleys. In this section, we ignore the asymmetry gap as it has no interesting consequences for the spectrum of the collective mode. Furthermore, we will neglect a process in which electrons are swapped between the valleys (see Fig. 1*d*). Then the Landau function becomes isotropic not only in the spin but also in valley subspaces:

$$N_F^* \hat{f}(\mathbf{k},\mathbf{k}') = F^s(\vartheta) + \hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{\sigma}}' F^a(\vartheta) + \hat{\boldsymbol{\tau}}\cdot\hat{\boldsymbol{\tau}}' G^a(\vartheta) + (\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{\sigma}}')(\hat{\boldsymbol{\tau}}\cdot\hat{\boldsymbol{\tau}}') H(\vartheta). \tag{37}$$

Since we are interested in collective modes with frequencies $\ll E_F$, the valence band can be projected out by using Löwdin method,[119] and the effective single-particle Hamiltonian for the conduction band becomes[94]

$$\hat{H}_c = v_F k + \frac{\lambda_{\rm R}}{2}(\hat{\mathbf{k}}\times\hat{\boldsymbol{\sigma}})\cdot\hat{\mathbf{z}} + \frac{\lambda_{\rm VZ}}{2}\tau_z\hat{\sigma}_z. \tag{38}$$

Accordingly, the change in the quasiparticle energy due to SOC is now given by [4]

$$\begin{aligned}\hat{\varepsilon}_{\rm SO}(\mathbf{k}) &= \frac{\lambda_{\rm R}^*}{2}(\hat{\mathbf{k}}\times\hat{\boldsymbol{\sigma}})\cdot\hat{\mathbf{z}} + \frac{\lambda_{\rm VZ}^*}{2}\hat{\tau}_z\hat{\sigma}_z, \\ &= \frac{\lambda_{\rm R}^*}{2}\hat{\Sigma}_z - \frac{\lambda_{\rm VZ}^*}{2}\hat{\tau}_z\hat{\Sigma}_x,\end{aligned} \tag{39}$$

while the non-equilibrium part of the occupation number can be written as

$$\begin{aligned}\delta\hat{n}(\mathbf{k},t) = n_F' \Big[&\mathbf{u}(\mathbf{k},t)\cdot\hat{\boldsymbol{\Sigma}}(\mathbf{k}) + \mathbf{w}(\mathbf{k},t)\cdot\hat{\boldsymbol{\tau}} \\ &+ M_{\alpha\beta}(\mathbf{k},t)\hat{\tau}_\alpha\hat{\Sigma}_\beta(\mathbf{k})\Big],\end{aligned} \tag{40}$$

[4] Note that in τ_z is just ±1 in Eq. (38, but it is a matrix $\hat{\tau}_z$ in Eq. (39).

where the rotated set of Pauli matrices is defined by Eq. (18).Vector $\mathbf{u}$ and tensor $M_{\alpha\beta}$ describe oscillations of the uniform

$$\boldsymbol{\mathcal{S}} = -\frac{g^*\mu_B}{8}\int \frac{d^2k}{(2\pi)^2}\mathrm{Tr}\left[\delta\hat{n}(\mathbf{k},t)\,\hat{\boldsymbol{\sigma}}\right] \tag{41}$$

and valley-staggered magnetization

$$\boldsymbol{\mathcal{M}} = -\frac{g^*\mu_B}{8}\int \frac{d^2k}{(2\pi)^2}\mathrm{Tr}\left[\delta\hat{n}(\mathbf{k},t)\,\hat{\tau}_z\hat{\boldsymbol{\sigma}}\right], \tag{42}$$

respectively. The vector $\mathbf{w}$ describes valley polarization, which is decoupled from the spin sector and will not be considered below.

Collective modes of our system correspond to coupled oscillations of the uniform and valley-staggered magnetizations. In the absence of VZ such oscillations are decoupled. The spectrum of collective excitations consists of a doublet of in-plane modes and a single out-of plane mode in each of the sectors, to a total of six modes. VZ SOC mixes the $\boldsymbol{\mathcal{S}}$ and $\boldsymbol{\mathcal{M}}$ sectors. Solving the coupled system of equations of motion, one obtains the following expressions for the frequencies of the in-plane modes[94]

$$\Omega^2_{||,\pm} = \lambda^{*2}_{\mathrm{R}}\left(\frac{ff_+ + hh_+}{2}\right) + \lambda^{*2}_{\mathrm{VZ}}\left(f_+h_+ + f_-h_-\right) \pm \Omega_0^2, \tag{43a}$$

where

$$\Omega_0^2 = \left[\lambda^{*4}_{\mathrm{R}}\left(\frac{ff_+ - hh_+}{2}\right)^2 + \lambda^{*4}_{\mathrm{VZ}}(f_-h_+ + h_-f_+)^2 + \lambda^{*2}_{\mathrm{R}}\lambda^{*2}_{\mathrm{VZ}}(ff_- + hh_-)(f_-h_+ + h_-f_+)\right]^{1/2}, \tag{43b}$$

$$\begin{aligned} f &= 1 + F_1^a,\ f_+ = 1 + \frac{F_0^a + F_2^a}{2},\ f_- = \frac{F_0^a - F_2^a}{2}, \\ h &= 1 + H_1,\ h_+ = 1 + \frac{H_0 + H_2}{2},\ h_- = \frac{H_0 - H_2}{2}, \end{aligned} \tag{44}$$

and $\lambda^*_{\mathrm{R}} = \lambda_{\mathrm{R}}/(1+F_1^a)$ and $\lambda^*_{\mathrm{VZ}} = \lambda_{\mathrm{VZ}}/(1+H_0)$ are the renormalized spin-orbit coupling constants. The frequencies of the out-of-plane modes are given by

$$\begin{aligned} \Omega^2_{\perp,+} &= h\left[(h_+ + h_-)\lambda^{*2}_{\mathrm{R}} + f\lambda^{*2}_{\mathrm{VZ}}\right] \\ \Omega^2_{\perp,-} &= f\left[(f_+ + f_-)\lambda^{*2}_{\mathrm{R}} + h\lambda^{*2}_{\mathrm{VZ}}\right] \end{aligned} \tag{45}$$

Finite q and/or Zeeman magnetic field lift the degeneracy of the in-plane modes. The spectrum of the collective modes in these cases is a subject of the on-going study.[120]

B. Dirac surface state of a 3D topological insulator

The Dirac state on the surface of a 3D topological insulator, described by the Hamiltonian (4), is characterized by locking of the spin and charge degrees of freedom. Indeed, Eq. (4) implies the charge current and spin densities are related by an operator identity[45]

$$\hat{\mathbf{j}} = ev_{\mathrm{D}}\hat{\boldsymbol{\sigma}} \times \hat{\mathbf{z}}. \tag{46}$$

Due to this coupling, the system supports a new kind of collective modes: spin-plasmons.[45] Although the dispersion of spin-plasmons at small q is similar to that of usual plasmons in 2DEGs ($\omega \propto \sqrt{q}$), the nature of the two modes is quite different because a spin-plasmon corresponds to coupled oscillations of spin and charge densities. For example, the weight of the spin part at small q is much larger than that of the charge part.

Low-energy excitations of a doped surface state can be described by the effective theory of helical FL, obtained by projecting out the occupied Dirac cone.[39] While the charge sector of such a FL can be described by a few Landau

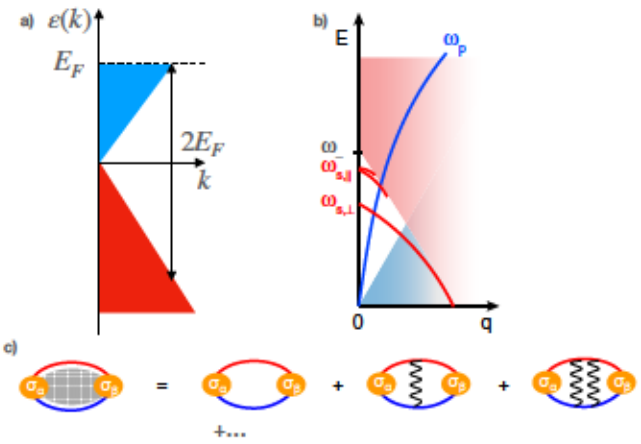


Figure 10. a) Spectrum of the Dirac state on the surface of a 3D topological insulator. b) Spectrum of the collective modes of the surface state. ω_p denotes the spin-plasmon [45], $\omega_{s,||}$ and $\omega_{s,\perp}$ denote the out- and in-plane collective spin modes, and $\omega_- = 2E_F$ is the lower boundary of the continuum of spin-flip particle-hole excitations. c) Ladder diagrams for the spin susceptibility $\chi_{\alpha\beta}$ $(\alpha, \beta = x, y, z)$. Panel b) is reprinted with permission from Ref. 61. Copyright 2017 by the American Physical Society.

parameters, description of the spin sector runs into difficulties similar to those discussed in Sec. II C. Namely, the total spin susceptibility contains contribution from high-energy states which cannot be accounted for within the FL theory. Nevertheless, the FS contribution to the spin susceptibility can be expressed through the Landau parameters of a helical FL.[39]

In what follows, we will be interested in a collective mode of the spin-excitonic type observed in Ref. 61; see Sec. VIII B for more details on the experiment. This mode condenses out of the continuum of inter-band particle-hole excitations. In a non-interacting system with Dirac spectrum such a continuum starts at the Pauli threshold $\omega = 2E_F$, see Fig. 10*a*, and disperses with q as shown by the shaded region in Fig. 10*b*. The blue shaded region depicts the continuum of gapless intra-band particle-hole excitations. With interactions, one expects to see the spin-plasmon mode (blue curve) and inter-band collective spin modes (red curves). By the same symmetry arguments as given in Sec. IV A, there are one out-of-plane mode and two in-plane modes at finite q; the latter become degenerate at $q = 0$. Inter-band excitations with energies comparable to E_F cannot be described within the FL theory, and one has to treat electron-electron interaction explicitly. The simplest method is the ladder approximation, shown diagrammatically in Fig. 10*c*. For a Hubbard-like interaction and at $q = 0$, one obtains the spin susceptibility as [46]

$$\chi_{\alpha\beta}(\Omega) = -\delta_{\alpha\beta}\frac{g^2\mu_B^2}{2}\frac{\Pi_{\alpha\beta}(\Omega)}{1+\frac{U}{2}\Pi_{\alpha\beta}(\Omega)}, \tag{47}$$

where the components of the polarization bubble $\hat{\Pi}$ are obtained by analytic continuation of the corresponding Matsubara expressions

$$\Pi_{\alpha\beta}(i\Omega_n) = T\sum_m \int \frac{d^2k}{(2\pi)^2}\mathrm{Tr}\left[\hat{\sigma}_\alpha \hat{G}_0(\mathbf{k}, i\varepsilon_m + i\Omega_n)\hat{\sigma}_\beta \hat{G}_0(\mathbf{k}, i\varepsilon_m)\right]. \tag{48}$$

with $\alpha, \beta = x, y, z$ and $G_0(\mathbf{k}, i\epsilon_m) = \left(i\epsilon_m - \hat{H}_{\mathrm{HTI}}\right)^{-1}$ is the Green's function of the Hamiltonian (4).

For Dirac spectrum, the momentum integrals for $\Pi_{\alpha\alpha}$ diverge in the ultraviolet and need to be cut off at some momentum Λ, chosen from the condition that the cubic term in the dispersion due to hexagonal warping[89] becomes comparable to the linear one. With such a cutoff, we obtain

$$\mathrm{Re}\Pi_{xx}(\Omega) = \mathrm{Re}\Pi_{yy}(\Omega) = \frac{1}{2}\mathrm{Re}\Pi_{zz}(\Omega) = -\frac{1}{4\pi v_{\mathrm{D}}}\left(\Lambda + \frac{\Omega}{4v_{\mathrm{D}}}\ln\left|\frac{\Omega+2E_F}{\Omega-2E_F}\right|\right). \tag{49}$$

The poles of Eq. (47) give the frequencies of the collective modes. At weak coupling, the modes are exponentially close to the boundary of the continuum:

$$\Omega_{||} = 2E_F - Ve^{-8/u},\ \Omega_{\perp} = 2E_F - Ve^{-4/u}, \tag{50}$$

where $u = UE_F/2\pi v_D^2$ is the dimensionless coupling constant and $V = 4E_F \exp(\Lambda/2k_F)$.

A reader familiar with the semiconductor literature of the 60s may notice that the derivation presented above is very close to that of "Mahan excitons" in degenerate semiconductors.[121] Indeed, because of the relation (46) the in-plane components of the spin susceptibility and conductivity are proportional to each other. [45] The same analysis, therefore, would predict that the optical conductivity of a helical state should have a peak at $\Omega = \Omega_{||}$. This is exactly what Mahan obtained by (effectively) resumming the ladder series for the optical conductivity of a degenerate semiconductor; the only difference is that he considered Coulomb rather than Hubbard interaction but this difference is not significant. Indeed, a similar analysis for the Coulomb case shows that the collective modes are exponentially close to the continuum boundary as well, namely[122]

$$\begin{aligned}
\Omega_{||} &= 2E_F \left\{1 - \frac{\alpha}{2} \exp\left[-\frac{\pi}{2\ln \alpha^{-1}} \left(\frac{1}{\alpha} - \ln \frac{\Lambda}{k_F}\right)\right]\right\}, \\
\Omega_{\perp} &= 2E_F \left\{1 - \frac{\alpha}{2} \exp\left[-\frac{\pi}{2\ln \alpha^{-1}} \left(\frac{1}{2\alpha} - \ln \frac{\Lambda}{k_F}\right)\right]\right\},
\end{aligned} \tag{51}$$

where $\alpha = e^2/v_D$ is the dimensionless coupling constant of the Coulomb interaction.

Subsequent analysis showed, however, that the ladder approximation is not adequate for this problem:[123, 124] because electrons and holes interact not in vacuum (as it is the case for an undoped semiconductor) but in the presence of the Fermi sea, the excitonic state is overdamped due to Auger-like electron-hole interactions processes, which start at the indirect threshold of $\Omega = E_F$. Moreover, damping due to electron-electron and electron-hole interactions starts already at $\Omega \ll E_F$, and is still operational at $\Omega \sim E_F$.[125, 126] Both types of damping give rise to the linewidth proportional to u^2 (or α^2), whereas the binding energy is exponentially small in $1/u$ (or $1/\alpha$). Therefore, collective modes (and Mahan excitons) must be totally washed out at weak coupling.

Nevertheless, Eq. (47) describes very well a spin-resolved collective mode observed by Raman spectroscopy in Bi_2Se_3, which effectively probes the zz-component of the spin susceptibility, see Ref. 61 and Sec. VIII B. The frequency of the observed mode is about E_F, which implies that the system is not in the weak-coupling regime. On the other hand, an order-of magnitude estimate for the linewidth due to all types of electron- and electron-hole interactions is also of order E_F for $\Omega \sim E_F$ and $u \sim 1$ (or $\alpha \sim 1$), so one would expect to see only a very broad peak at best. Nevertheless, the observed peak is quite sharp: its linewidth is only about 1/20 of its position. Moreover, the linewidth does not vary with temperature as the latter is raised up to 300 K, which indicates that the broadening is due to impurity scattering. Therefore, damping due to electron- and electron-hole interactions is significantly weaker than a naive, order-of-magnitude estimate would suggest. We believe that the reason is purely numerical: as analytical and numerical results show, the linewidth due to both processes is only $5 \times 10^{-3} E_F$ for a 2D helical state.[122, 126] Therefore, the ladder approximation works much better than it might have been expected to.

VI. SPATIAL DISPERSION OF COLLECTIVE SPIN MODES

The dispersion of the transverse $n = 0$ Silin mode is quadratic in q for $q \ll \Delta_Z^*/v_F^*$:[54, 105, 108, 127]

$$\Omega(q) = \Delta_Z + a_2(\{F^a\}) \frac{v_F^{*2} q^2}{\Delta_Z}, \tag{52}$$

where $a_2(\{F^a\})$ depends on the angular harmonics of the Landau function; explicitly, $a_2 = (1+F_0^a)^2(1+F_1^a)/(F_0^a - F_1^a)$. Note that $a_2 < 0$ for $-1 < F_0^a < F_1^a < 0$ and thus dispersion is downward, as shown in Fig. 3. A quadratic scaling of the dispersion q^2 follows from the invariance of the spin subspace with respect to rotations about the magnetic field, which requires the dispersion to be isotropic, and from analyticity, which requires that an expansion in q starts from the quadratic term.

If only Rashba SOC is present, the group symmetry of the system is $C_{\infty v}$, i.e., the system is invariant with respect to rotations by an arbitrary angle about the normal to the plane. Therefore, as in the Silin's case, the dispersion is isotropic and, by analyticity, starts with a q^2 term. For the lowest three modes,[46, 115, 128, 129]

$$\Omega_\alpha(q) = \Omega_\alpha + \tilde{a}_2^\alpha(\{F^a\}) \frac{v_F^{*2} q^2}{\Delta_R}, \tag{53}$$

where $\alpha \in \{||x, ||y, \perp\}$, the $q = 0$ frequencies are given by Eqs. 25 and Eqs. 26, and $\tilde{a}_2^\alpha(\{F^a\})$ some functions of $F_0^a, F_1^a \ldots$ If only Dresselhaus SOC is present, the Hamiltonian (2) can be transformed back to the Rashba one (1) by by reflecting the spatial coordinates about a mirror plane that contains the (110) axis, upon which $k_x \to k_y$ and $k_y \to k_x$. Therefore, the spectrum of the collective modes is the same as in Eq. (53) with Δ_R being replaced by Δ_D.

If both Rashba and Dresselhaus SOCs are present, the symmetry is lowered to D_{2d}, which only allows rotations by π and mirror reflections about the two diagonals. The linear term in the dispersion is absent because by symmetry it can only be $\sqrt{q_x^2 + q_y^2}$, which is not allowed by analyticity, while a bilinear term is of the $c_1(q_x^2 + q_y^2) + c_2 q_x q_y$ form, where $c_{1,2}$ are constants.

An in-plane magnetic field breaks the rotational symmetry. Now, linear-in-q terms in the dispersion are allowed and, indeed, they have been observed experimentally.[55–60] The structure of such terms can be deduced just from the symmetries of $C_{\infty v}$ and D_{2d} groups.[57, 58, 129] In both cases, we need to form a scalar out of a polar vector $\mathbf{q}$ and axial vector $\mathbf{B}$. In the $C_{\infty v}$ group, this is only possible by forming the Rashba invariant $B_x q_y - B_y q_x = Bq \sin(\phi_\mathbf{q} - \phi_\mathbf{B})$, were $\phi_\mathbf{q}$ and $\phi_\mathbf{B}$ are the azimuthal angles of $\mathbf{q}$ and $\mathbf{B}$, respectively. Likewise, the only scalar that can be constructed in the D_{2d} group is the Dresselhaus invariant $B_x q_x - B_y q_y = Bq \cos(\phi_\mathbf{q} + \phi_\mathbf{B})$.

In addition to linear-in-q terms, an in-plane magnetic field also gives rise to the dependence of the mode frequency at $q = 0$ on the direction of the magnetic field. However, this effect needs both Rashba and Dresselhaus SOCs to be present. Indeed, since Rashba SOC has continuous rotational symmetry, the direction of the magnetic field is irrelevant in this case, while the case of pure Dresselhaus SOC is reduced back to pure Rashba one. If both types of SOC are present, the D_{2d} symmetry implies that mode frequency depends on the direction of $\mathbf{B}$ as $\Omega_\alpha(q = 0) \propto \sin 2\phi_\mathbf{B}$.[59, 129]

To be specific, from this point onwards we will focus on the case when $\Delta_\mathrm{Z} \gg \Delta_\mathrm{R}, \Delta_\mathrm{D} \neq 0$, which is relevant for Raman experiments on $Cd_{1-x}Mn_xTe$ quantum wells.[55–59] In this case, the mode frequency at $q = 0$ is given primarily by the (bare) Zeeman energy. The correction due to SOC must be quadratic in both Δ_R and Δ_D and symmetric on $\Delta_\mathrm{R} \leftrightarrow \Delta_\mathrm{D}$, i.e., to be of the form $\Delta_\mathrm{R}^2 + \Delta_\mathrm{D}^2$. The anisotropic term at the $q = 0$ exists only of both Δ_R and Δ_D are non-zero which, in the limit considered, implies that the corresponding term must be proportional to $\Delta_\mathrm{R}\Delta_\mathrm{D}$. The linear-in-$q$ terms due to Rashba and Dresselhaus SOC must be proportional to Δ_R and Δ_D, respectively. Finally, the bilinear-in-q is almost the same as for the Silin mode modulo an anisotropic correction, which arises again only due to the combined effect of Rashba and Dresselhaus SOCs, and thus is also proportional to $\Delta_\mathrm{R}\Delta_\mathrm{D}$. Combining all the arguments given above, we arrive at the following form of the dispersion[58, 59, 129]

$$\Omega(\mathbf{q}, \mathbf{B}) = \Omega_0(\mathbf{B}) + w(\mathbf{q}, \mathbf{B})q + S(\mathbf{q})\frac{v_F^{*2} q^2}{\Delta_\mathrm{Z}}, \tag{54}$$

where

$$\begin{aligned}
\Omega_0(\mathbf{B}) &= \Delta_\mathrm{Z} + a_0(\{F^a\})\frac{\Delta_\mathrm{R}^2 + \Delta_\mathrm{D}^2}{\Delta_\mathrm{Z}} \\
&+ \tilde{a}_0(\{F^a\})\frac{\Delta_\mathrm{R}\Delta_\mathrm{D}}{\Delta_\mathrm{Z}} \sin 2\phi_\mathbf{B}, \\
w(\mathbf{q}, \mathbf{B}) &= v_F a_1(\{F^a\}) \left[\frac{\Delta_\mathrm{R}}{\Delta_\mathrm{Z}} \sin(\phi_\mathbf{q} - \phi_\mathbf{B}) \right. \\
&\left. + \frac{\Delta_\mathrm{D}}{\Delta_\mathrm{Z}} \cos(\phi_\mathbf{q} + \phi_\mathbf{B})\right] \\
S(\mathbf{q}) &= a_2(\{F^a\}) + \tilde{a}_2(\{F^a\})\frac{\Delta_\mathrm{R}\Delta_\mathrm{D}}{\Delta_\mathrm{Z}} \sin 2\phi_\mathbf{q},
\end{aligned} \tag{55}$$

and the coefficient $a_2(\{F^a\})$ is the same as for a pure Silin mode, Eq. (52). In the equation above we omitted isotropic q^2 terms proportional to Δ_R^2 and Δ_D^2. The coefficients a_0, $\tilde{a}_0$, and a_1 were calculated in Ref. 129 in the s-wave approximation for the Landau function, see Eq. (20) in there. It was argued in Refs. 58 and 59 that, to linear order in SOC, the dispersion can be obtained by a canonical transformation of the Hamiltonian, which amounts to replacing $\mathbf{q}$ in the dispersion of the Silin mode, Eq. (52), by $\mathbf{q} + \mathbf{q}_0$, where $\mathbf{q}_0$ is proportional to Δ_R and Δ_D. This would imply that the coefficients of the q and q^2 terms in Eq. (55) are related as $|a_1| = 2|a_2|$. However, a microscopic calculation[118] shows that such relation is satisfied only in the weak-coupling limit ($|F_0^a| \ll 1$).

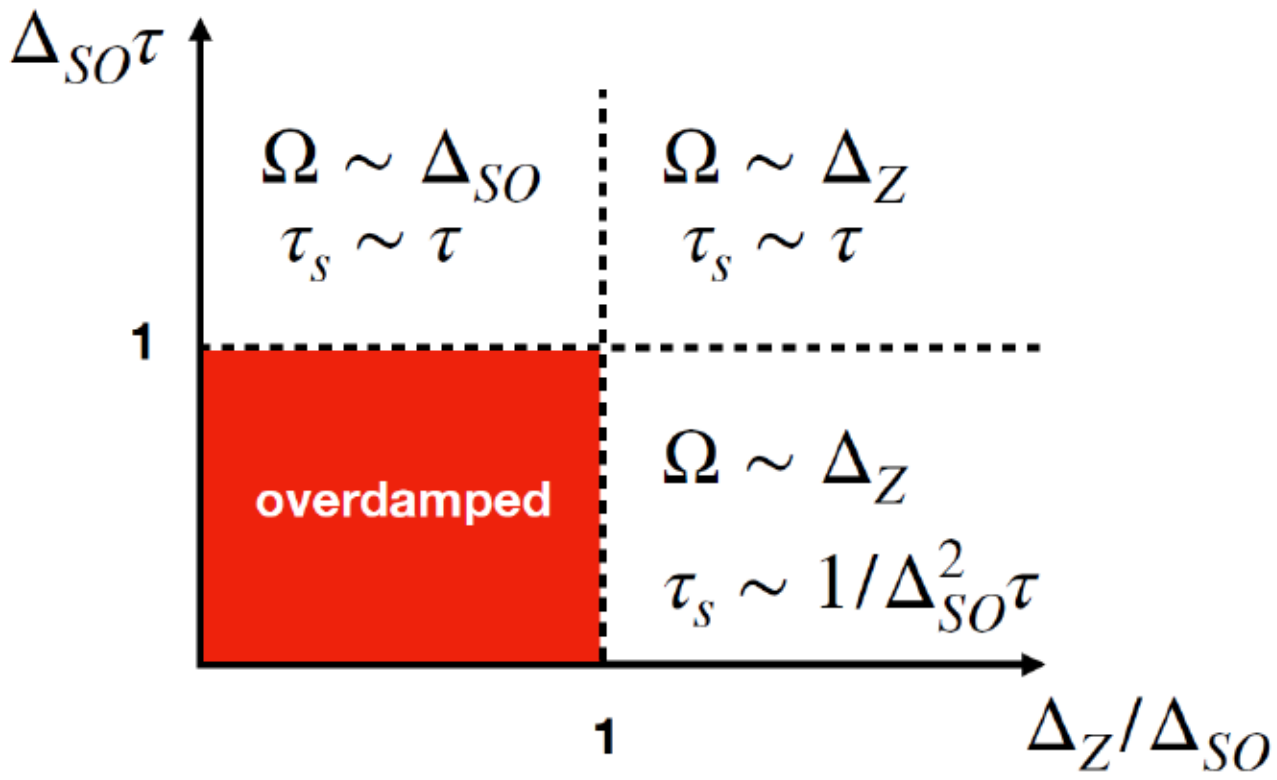


Figure 11. Damping of collective spin modes by non-magnetic disorder via D'yakonov-Perel' mechanism. Δ_{SO} and Δ are the splittings of the energy spectrum due to SOC and magnetic field, respectively, τ is the transport mean free time due to scattering by disorder. In the region outside the red square the spin collective modes are underdamped. The frequency of the collective mode (Ω) and spin relaxation time (τ_s) are indicated for each region. Within the red square, the modes are overdamped and thus cannot be observed.

VII. DAMPING OF COLLECTIVE SPIN MODES

Spin-orbit coupling is the reason for the collective spin modes, described in this paper, to exist. At the same time, however, it couples spin and momentum and thus enables damping of spin excitations via momentum relaxation due to scattering by non-magnetic degrees of freedom: non-magnetic disorder, phonons, etc.

Non-magnetic disorder in combination with SOC leads to spin relaxation via the Elliott-Yaffet and D'yakonov-Perel' (DP) mechanisms.[66] The former is present in both centro- and non-centrosymmetric systems, while the latter is specific for non-centrosymmetric systems, considered in this article, and in this case the DP mechanism is usually the dominant one. If τ is a characteristic time of momentum relaxation by disorder (to be defined more precisely later) and $\Delta_{\rm SO}$ is the energy splitting due to SOC, then the spin dynamics is ballistic with the DP spin relaxation time $\tau_s \sim \tau$ for $\Delta_{\rm SO}\tau \gg 1$, and diffusive with the DP spin relaxation time $\tau_s \sim 1/\Delta_{\rm SO}^2\tau$ for $\Delta_{\rm SO}\tau \ll 1$. On the other hand, the frequency of the collective mode is set by the largest of two energy scales: $\Delta_{\rm SO}$ and the Zeeman energy $\Delta_{\rm Z}$. The spin collective mode can be resolved only when it is underdamped, which corresponds to the region outside the red square in Fig. 11. This requires clean samples and strong SOC/magnetic fields.

One concrete example is a single-valley 2D FL with disorder and Rashba SOC. In this case, the width of the chiral spin resonance at $q = 0$ is controlled by two relaxation times, renormalized by the FL interaction. For $\Delta_{\rm R}^*\tau \gg 1$, the width of the resonance is[113]

$$\Gamma_{\rm i} = \frac{1}{2x_1\tau_1} + \frac{1}{2x_2(1+x_2/x_0)\tau_2}, \tag{56}$$

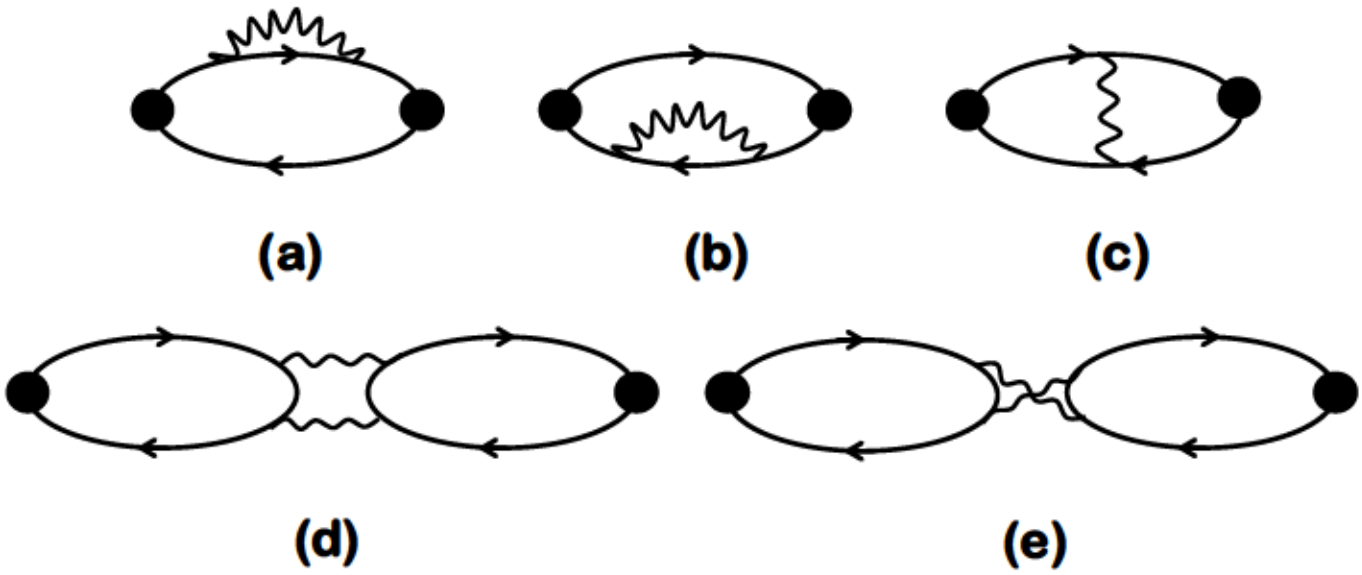


Figure 12. Diagrams contributing to damping of the collective modes by dynamically screened Coulomb interaction (wavy line). The filled circles denote the corresponding vertices. For the case of spin collective modes, the vertices are Pauli matrices.

where $x_m = 1/(1+F_m^a)$,

$$\frac{1}{\tau_n} = \pi N_F^* n_\mathrm{i} \int \frac{d\vartheta}{2\pi}(1-\cos n\vartheta)\left|u\left(2k_F \sin\frac{\vartheta}{2}\right)\right|^2, \tag{57}$$

n_i is the number density of impurities, and $u(Q)$ is the Fourier transform of the single-impurity potential. Note that τ_1 is the transport time that enters the mobility.

Another source of damping is electron-electron interaction. Because the collective modes lie outside the particle-hole continuum, they are not affected by Landau damping, which involves excitation of a single particle-hole pair. However, excitations involving multiple pairs are possible due to the residual interaction between FL quasiparticles, and such excitations give finite width even to modes outside the continuum. For example, plasmons, [125, 130–134] Silin mode in a partially spin-polarized FL,[105, 108, 127] and magnons in a ferromagnetic FL[105, 127, 135] are all damped via this mechanism. Figure 12 shows diagrams that contribute to damping to lowest order in a dynamically screened Coulomb interaction.[5]

Out of the examples listed above, the Silin mode is the closest one to collective spin modes due to SOC. However, there is an important difference. Namely, conservation of the spin component along the field ensures that the Silin mode is not damped at $q \to 0$.[136] On the contrary, spin is not a conserved quantity in a FL with SOC. Therefore, collective modes in this case are damped even at $q = 0$. Evaluation of diagrams in Fig. 12 leads to an intuitively clear result: at $T = 0$, the width of the resonance is given by[137]

$$\Gamma_\mathrm{ee} \sim \frac{\Delta_\mathrm{SO}^2}{E_F}\lambda^2 \ln\lambda^{-1}, \tag{58}$$

where $\lambda = e^2/v_F$ is the dimensionless coupling constant of the Coulomb interaction. The quadratic dependence on Δ_SO is an expected scaling of a relaxation rate in a FL.[6]

Modes with finite q are damped by the electron-electron interaction even in the absence of SOC. The same arguments of rotational invariance and analyticity that we used in Sec. VI to determine the q-dependence of the dispersion, can be applied to the damping rate. Namely, in the presence of rotational symmetry, the damping rate is proportional to q^2.[105, 108] If both types of SOC and an in-plane magnetic field are present, the damping rate contains a linear-in-q term, whose structure is the same as of the linear term in the dispersion.[59, 60]

[5] Although the two last (Aslamazov-Larkin) diagrams appear to be higher order in the Coulomb interaction than the first three, they actually contribute to the same order in the dimensionless coupling constant $e^2/\hbar v_F$.

[6] A single-particle relaxation rate in a 2D FL has an additional logarithmic factor, but it is canceled out between the diagrams for the spin relaxation rate, as guaranteed by the gauge invariance.

VIII. EXPERIMENT: OBSERVATIONS AND PREDICTIONS

A. Collective spin waves in $Cd_{1-x}Mn_xTe$ quantum wells

The dispersion of collective spin waves in $Cd_{1-x}Mn_xTe$ quantum wells was measured in a series of detailed Raman experiments.[55–59] This system has both Rashba and Dresselhaus SOCs and thus should have chiral spin waves even in the zero magnetic field. However, the corresponding energy scales are below the resolution of Raman spectroscopy, and one has to apply an in-plane magnetic field to increase the energies of spin-flip excitations. Even a moderate field of 2 T leads to a significant spin polarization due the exchange interaction between magnetic moments of Mn dopants and conduction electron spins, and the effective Zeeman energy is larger than the spin-orbit one. To get some sense of the numbers, for the reported values of $\alpha^* \approx 1.8\,\text{meV}\cdot\text{Å}$ and $\beta^* \approx 3.8\,\text{meV}\cdot\text{Å}$ (Ref. 58),[7] the combined spin-orbit splitting is $\Delta^*_{\rm SO} = \sqrt{\Delta^{*2}_{\rm R} + \Delta^{*2}_{\rm D}} \approx 0.1\,\text{meV}$ at $n = 2.7 \times 10^{11}\text{cm}^{-2}$, while the measured frequency of the collective mode varies from ~ 0.4 meV at $q = 0$ in samples with lower Mn fraction($x = 0.013$, Ref. 58) to $\sim 3\,$mev in samples with higher Mn fraction ($x = 0.8$, Ref. 55). Therefore, the effective Zeeman splitting is 4-30 times larger than the spin-orbit one, and the experimental situation corresponds to the right panel of Fig. 7 for $\Delta^*_{\rm Z} > \Delta^*_{\rm R}$ (with $\Delta^*_{\rm R}$ replaced by $\Delta^*_{\rm SO}$), i.e., there is a single collective mode, which evolves into a pure Silin mode in the limit of $\Delta^*_{\rm Z} \to \infty$. At lower Mn fractions, however, the effects of SOC on the dispersion and damping of this mode are quite pronounced.

Panels *a*-*d* in Fig. 13 show experimental data for a [001] $Cd_{87}Mn_{0.13}Te$ quantum well, reproduced from Refs. 57 and Refs. [59]. In the experiment, the in-plane magnetic field and vector $\mathbf{q}$ were kept at 90° to each other, while the pair of vectors was rotated by angle φ measured from the [100] direction, as shown in panel *a*. The Raman signal in panel *b* exhibits a well-resolved peak which disperses with q. Interestingly, the dispersion is not purely quadratic but has a sizable linear term, which is revealed by flipping the direction of $\mathbf{q}$, as shown in panel *c*. Panels *e*-*g*, reproduced from Ref. 59, show the angular dependence of the mode frequency (E_0), spin-wave velocity (E_1), and spin-wave stiffness (E_2).

Now we compare the experimental results with the theory presented in Sec. VI. Under the condition $\phi_{\mathbf{B}} - \phi_{\mathbf{q}} = \pm\pi/2$, parameters of the dispersion in Eq. (55) are reduced to

$$\begin{aligned}
\Omega_0(\mathbf{B}) &= \Delta_{\rm Z} + a_0(\{F^a\})\frac{\Delta_{\rm R}^2 + \Delta_{\rm D}^2}{\Delta_{\rm Z}} \\
&+ \tilde{a}_0(\{F^a\})\frac{\Delta_{\rm R}\Delta_{\rm D}}{\Delta_{\rm Z}}\sin 2\phi_{\mathbf{B}}, \\
w(\mathbf{B}) &= \pm v_F a_1(\{F^a\})\frac{\Delta_{\rm R} - \Delta_{\rm D}\sin 2\phi_{\mathbf{B}}}{\Delta_{\rm Z}} \\
S(\mathbf{B}) &= a_2(\{F^a\}) - \tilde{a}_2(\{F^a\})\frac{\Delta_{\rm R}\Delta_{\rm D}}{\Delta_{\rm Z}}\sin 2\phi_{\mathbf{B}}.
\end{aligned} \tag{59}$$

Figure 14 shows the theoretical results for $\Omega_0(\mathbf{B})$ (inset in panel *a*), $w(\mathbf{B})$ (panel *a*), and the q- dependence of the dispersion for two opposite orientations of the magnetic field, at $\pi/4$ and $-\pi/4$ (panel *b*). The value of $\Delta_{\rm Z} = 0.4\,\text{meV}$ at $B = 2\,\text{T}$ was taken from the experiment, while the Rashba and Dresselhaus coupling constants, and F_0^a were used as fitting parameters. The fitted values of α and β (1.9 mev Å and 3.8 mev Å) are very close to that reported in Ref. 58 ($1.83 \pm 0.08\,\text{meV Å}$ and $3.79 \pm 0.11\,\text{meV Å}$, respectively), while the fitted value $F_0^a = -0.41$ is quite close to the Hartree-Fock estimate $F_0^a = -0.3$ for a CdTe quantum well with $n = 2.7 \times 10^{11}\text{cm}^{-2}$.[129] We see that the spin-wave velocity (w) is indeed π-periodic and very close in magnitude the experimental result in panel *f* of Fig. 13. The π-modulation of $\Omega_0(\mathbf{B})$, shown in the inset of panel *a*, is much smaller than that of w because the former effect is second order in SOC.[59, 129] This is consistent with panel *e* of Fig. 13. The linear term in the dispersion is evident from panel *b* of Fig. 14 and consistent with panel *c* of Fig. 13. Finally, the experiment also observes small π-modulation of the spin stiffness (panel *g*), which is consistent with this effect being also second order in SOC.

B. Collective spin mode on the surface of Bi_2Se_3

Figure 15 summarizes the results of polarization-resolved Raman spectroscopy of the surface state of a topological insulator Bi_2Se_3.[61] To enhance the signal from the surface states, the frequency of incident light was tuned to the

[7] We assigned stars to α and β because the experiment measures only the renormalized values of the spin-orbit parameters.

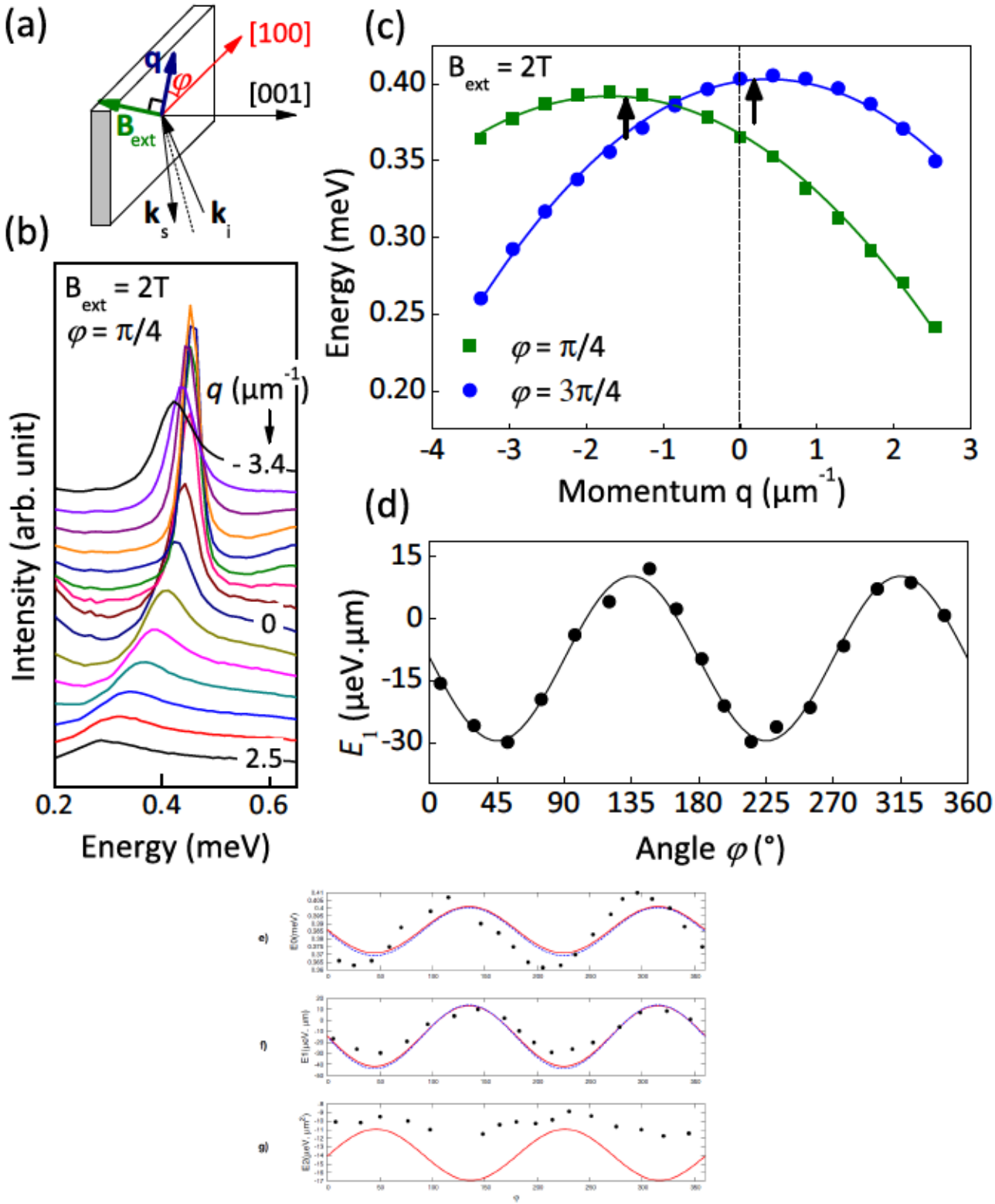


Figure 13. Experimental results for collective spin waves in $Cd_{1-x}Mn_xTe$ quantum wells. a) Geometry of the Raman experiment. b) Raman signal. c) q-dependences of the spin-wave frequency. d) Angular dependence of the spin-wave velocity [denoted by $w(\mathbf{B})$ in the main text]. Panels *e-g*: angular dependence of the mode frequency at $q = 0$ [denoted by $\Omega_0(\mathbf{B})$ in the main text], and spin-wave stiffness [denoted by $S(\mathbf{B})$ in the main text], respectively. Solid and dashed lines represent the numerical and analytical results of Ref. 59. Panels *a-d* and *e-f* are reproduced with permission from Ref. 57 and 59, respectively. Copyright 2015, 2017 of the American Physical Society.

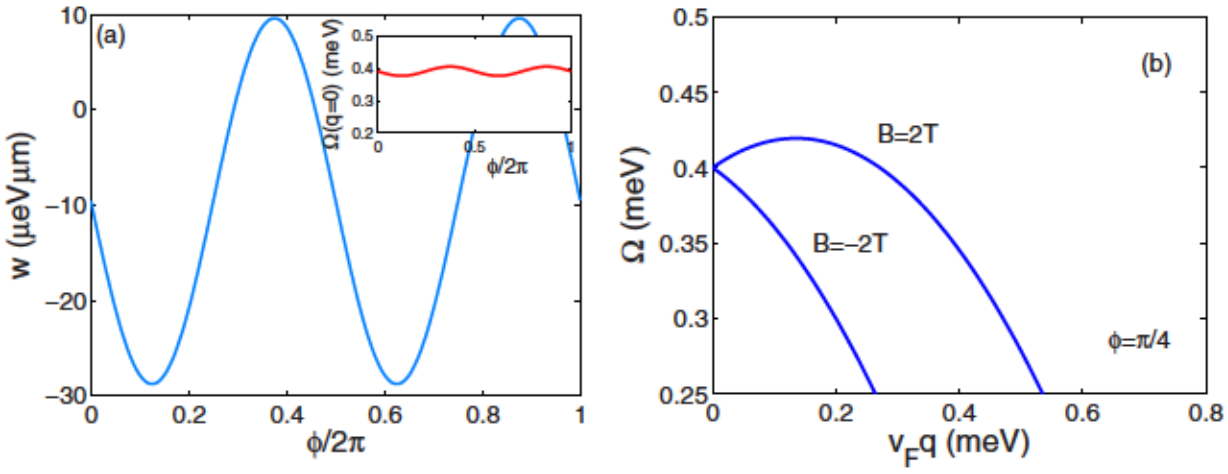


Figure 14. Theoretical results for the parameters of the spin-wave dispersion presented in Eq. (59). Reprinted with permission from Ref. 129. Copyright 2017 by the American Physical Society.

transition between two surface Dirac points: near the Fermi energy (SS1) and about $\Delta_S = 1.8$ eV above it (SS2), see panel I in Fig. 15. The Raman response shows a well-resolved peak at ≈ 150 meV, while the Pauli threshold for this sample is at $2E_F \approx 260$ meV. The peak is most pronounced for the incident frequency of 1.83 eV (panel II*c*), which is the closest to Δ_S. This proves that the signal is indeed due to the surface rather than the bulk states. (Broader peaks at higher frequency, interpreted as the result of excitonic photoluminescence,[138] were eliminated by subtracting the hatched areas from the data.) Furthermore, polarization-resolved experiments reveal a magnetic nature of the 150 meV excitation. As shown in panel II*d*, the signal is much stronger in the RR channel, when right-polarized photons are scattered into right-polarized ones, i.e, when the angular momenta of the incident and scattered photons differ by $2\hbar$, than in the RL channel, when right-polarized photons are scattered into left-polarized ones, i.e., without a change in the angular momentum. For linearly polarized light, the signal is the strongest in the cross-polarization channel (XY), when the polarization axis is rotated by $\pi/2$. Decomposing the signal into components corresponding to irreducible representations of the C_{6v} group, one finds that the 150 meV excitation belongs primarily to the A_2 representation, which is the pseudovector representation of C_{6v}, see panels III *a*-*c*. Noticeably, the excitation is very robust–it is observed up to 300 K, see panel III*c*.

Given the findings summarized above, one is prompted to interpret the 150 meV excitation as the $q = 0$ collective spin mode discussed in Sec. V B, namely, as the out-of-plane mode with frequency $\Omega_\perp$, because in the experiment both the incident and scattered beams were along the normal to the surface. This interpretation is confirmed by the theoretical analysis, which shows that the Raman intensity is proportional to the zz-component of the spin susceptibility[61]

$$\mathcal{R}(\Omega, T) \propto \chi_{zz}(\omega, T) \frac{\gamma_{\rm r}}{(\Omega_L - \Delta_S)^2 + \gamma_{\rm r}^2}, \tag{60}$$

where Ω_L is the frequency of incident light and $\gamma_{\rm r}$ is the linewidth of the resonance. $\chi_{zz}(\omega, T)$ was calculated within the ladder approximation for a realistic spectrum of the surface state and at finite temperature, and without using the weak-coupling assumption employed in Sec. V B. [For reasons discussed in Sec. V B, the ladder approximation works better than it might have been expected to, thanks to (accidental) numerical smallness of damping due to electron- and electron-hole interactions.] Since the linewidth of the observed peak is independent of temperature, the primary source of damping must be due to disorder. To mimic the effect of damping, the calculated $\chi_{zz}(\omega, T)$ was artificially broadened to produce the observed linewidth of ≈ 8 meV. In addition to the linewidth, the coupling constant of a Hubbard-like interaction was treated as a fitting parameter. The theoretical results, presented in panel III*d*, reproduce very well not only the profile of the peak as a function of frequency but also its temperature dependence. In particular, the theory reproduces the pronounced decrease of the resonance frequency with increasing temperature. This happens because the continuum broadens as T increases, which pushes the resonance peak down to lower frequency. The inset of panel III*d* shows a zoom on the interval between the resonance peak and continuum boundary, which is supposed to be at $2E_F \approx 260$ meV. We see, however, that the continuum is barely discernible because most of its spectral weight is transferred to the collective mode. The best fit was obtained for the Hubbard coupling $u \approx 0.6$. This is consistent with the estimate for the screened Coulomb interaction between electrons on the surface of Bi_2Se_3.

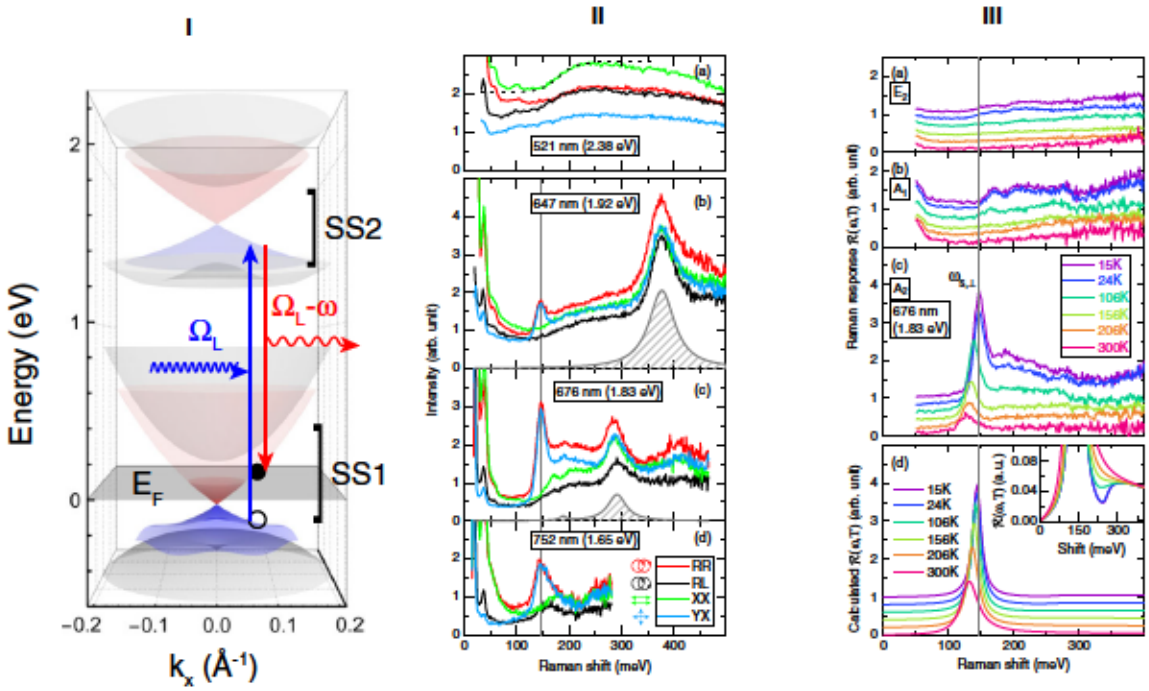


Figure 15. I) Bandstructure of Bi_2Se_3 around the Brillouin zone center. The two groups of Dirac States (SS1 and SS2, respectively) are separated by $\Delta_S \approx 1.8$ eV. The frequency of incident photons Ω_L is tuned to a resonance transition between the SS1 and SS2 states. II) Raman scattering data for different polarization geometries of the incoming and scattered photons, and different incident photon energies. The collective mode at the Raman shift of 150 meV is resonantly enhanced by the 1.83 eV photon. III (a)-(c): Raman scattering data decomposed into different channels, corresponding to E_2, A_1, and A_2 irreducible representations of the C_{6v} group, respectively. III (d): Theoretical results for the Raman response within the ladder approximation. Reprinted with permission from Ref. 61. Copyright 2017 by the American Physical Society.

In summary, the results discussed above present strong evidence for a new type of the collective spin mode, arising from the combined effect of SOC and Coulomb interaction. Further experiments measuring the spatial dispersion of this mode would be highly desirable.

C. Predictions for future experiments

1. *ESR and EDSR in 2DEGs*

The rate of absorption of electromagnetic wave at normal incidence by a 2DEG with Rashba and/or Dresselhaus SOC can be written as

$$Q = \frac{1}{2}\left[\mu_B^2 \Omega \chi_{||}''(\Omega) + \sigma'(\Omega)\right] E_{\rm em}^2, \tag{61}$$

where $E_{\rm em}$ is the amplitude of the electric field of the wave, $\sigma(\Omega)$ is the conductivity of a 2DEG and $\chi_{||}(\Omega)$ is the in-plane component of its spin susceptibility. The first term represents absorption due to ESR, which can be observed even in the absence of SOC. However, its magnitude is proportional to $1/c^2$ (in the Gaussian unit system) and thus small. In 2D systems instead of ESR one typically measures electrically-detected spin resonance, observed as a peak in the longitudinal resistivity under microwave radiation in the regime of integer Hall effect. [88, 139–142]

The second term in Eq. (61) represents absorption due to EDSR.[143–145] Its origin is an effective magnetic field acting on electron spins due to SOC and with magnitude proportional to k. The driving electric field (either from a *dc* current or electromagnetic wave) gives rise to a flow of electrons with a non-zero drift velocity, and hence the

electron system as a whole experiences an effective magnetic field due to SOC. This effect gives rise to a range of spectacular phenomena, e.g., a strong enhancement of microwave absorption in a geometry when the driving electric field is in the plane of a 2DEG [21] and a shift of the spin resonance frequency by *dc* current. [22, 23]

Collective spin modes described in the previous parts of the paper correspond to oscillations of the electron magnetization even in the absence of the external magnetic field. Therefore, they should be detectable both via ESR and EDSR which, in contrast to the conventional setup, should be present even in zero magnetic field. The structure of the signal can be understood qualitatively from Fig. 7, where one just has to replace $\Delta_R \to \sqrt{\Delta_R^2 + \Delta_D^2}$ to account for Dresselhaus SOC. In the absence of the magnetic field ($\Delta_Z^* = 0$), the signal consists of two peaks, at frequencies $\Omega_\perp$ and $\Omega_{||}$. At finite field, the peak at $\Omega_{||}$ splits into two. Upon further increase of the field, the $\Omega_{||}$ peaks merge with the continuum and die out, while the $\Omega_\perp$ peak continues to be present all the way till the gap closing point ($\Delta_Z^* = \sqrt{\Delta_R^2 + \Delta_D^2}$), and then emerges again at fields above this point.

To estimate the relative strength of the ESR and EDSR signals, we note that the conductivity in Eq. (61) is the sum of the Drude and spin-orbit parts: $\sigma'(\Omega) = \sigma'_D(\Omega) + \sigma'_{SO}(\Omega)$. Since for the Hamiltonians (1) and (2) the electric current is proportional to magnetization, the spin-orbit part of the conductivity and spin susceptibility are related by $\sigma'_{SO}(\Omega) \sim e^2 \max\{\alpha^2, \beta^2\}\chi''(\Omega)/\Omega$. Provided that the Drude part at the resonance frequency Ω_r can be neglected, the ratio of the EDSR to ESR signals can be estimated as[113]

$$\begin{aligned} \frac{Q_{EDSR}}{Q_{ESR}} &\sim \left(\frac{\max\{\alpha,\beta\}mc}{\Omega_r}\right)^2 \\ &\sim \left(\frac{\lambda_F}{\lambda_C}\right)^2 \begin{cases} 1, \text{ for } \Delta_Z \ll \max\{\Delta_R, \Delta_D\}, \\ \left(\frac{\max\{\Delta_R,\Delta_D\}}{\Delta_Z}\right)^2, \end{cases} \end{aligned} \tag{62}$$

where $\lambda_F = 2\pi/k_F$ is the Fermi wavelength and $\lambda_C = 2\pi/mc = 2.4 \times 10^{-10}$ cm is the Compton wavelength. For electron number densities in the interval $n = 10^{11} - 10^{12}\,\mathrm{cm}^{-2}$, the factor $(\lambda_F/\lambda_C)^2 \sim 10^8 - 10^9$, and the EDSR signal is stronger than the ESR one by many orders of magnitude, even if SOC is weak.

However, there is a caveat in this estimate, namely, it is valid provided that the Drude part of the conductivity is much smaller then the spin-orbit part, which imposes rather stringent conditions on the strength SOC and sample quality. Near the resonance, the spin-orbit part of the conductivity can be estimated as[113, 137]

$$\sigma'_{SO}(\Omega) \sim \frac{e^2}{h} m^* \max\{\alpha^2, \beta^2\} \frac{\Gamma}{(\Omega - \Omega_r)^2 + \Gamma^2}. \tag{63}$$

Assuming that the linewidth of the resonance is due to D'yakonov-Perel' mechanism in the ballistic regime (cf. Sec. VII), i.e, $\Gamma \sim 1/\tau$, and that the Drude conductivity is controlled by the same scattering mechanism, i.e., $\sigma'_D(\Omega_r) \sim (e^2/h)E_F/\Omega_r^2\tau$, we obtain for the ratio of the two parts of the conductivity right at the resonance:

$$K = \frac{\sigma'_{SO}(\Omega_r)}{\sigma'_D(\Omega_r)} \sim \frac{m^* \max\{\alpha^2, \beta^2\}}{E_F} (\Omega_r \tau)^2 . \tag{64}$$

Therefore, even if the resonance is underdamped, i.e., $\Omega_r\tau \gg 1$, it can be still masked by the Drude part if the first factor on the RHS of Eq. (64) is sufficiently small.

Estimates[137] show that the resonance in zero magnetic field, when $\Omega_r \sim \max\{\Delta_R, \Delta_D\}$, would be completely masked in a GaAs/GaAlAs heterostructure even with a record-hight mobility of $10^7\,\mathrm{cm}^2/\mathrm{V\,s}$ because, due to a relatively weak SOC in this system ($\alpha \sim \beta \sim 1\,\mathrm{mev\,\AA}$), K is only ~ 0.1. The problem is further exacerbated by the fact that, in the presence of both Rashba and Dresselhaus SOCs, the lower edge of the continuum is located at $\Omega_- = |\alpha - \beta|k_F$, which pushes the energies of the collective modes further down.

A better candidate is an InGaAs/InAlAs quantum well, where SOC is much stronger, i.e., $\alpha \sim 100$ meV Å (Ref. 146), which helps to compensate for smaller mobilities typical for these structures; the highest mobilities reported for InGaAs/InAlAs samples are in the range $\mu = (2-5) \times 10^5\ \mathrm{cm}^2/\mathrm{V\,s}$ (Refs. 147 and 148). Also, SOC in these structures is predominantly of the Rashba type,[146, 149] which alleviates the problem with a competition between the Rashba and Dresselhaus mechanisms. For a high-mobility InGaAs/InAlAs quantum well, $K \sim 1$ and a detailed calculation confirms that the zero-filed EDSR peak should be visible against the Drude background.[137]

2. *ESR and EDSR in graphene with proximity-induced spin-orbit coupling*

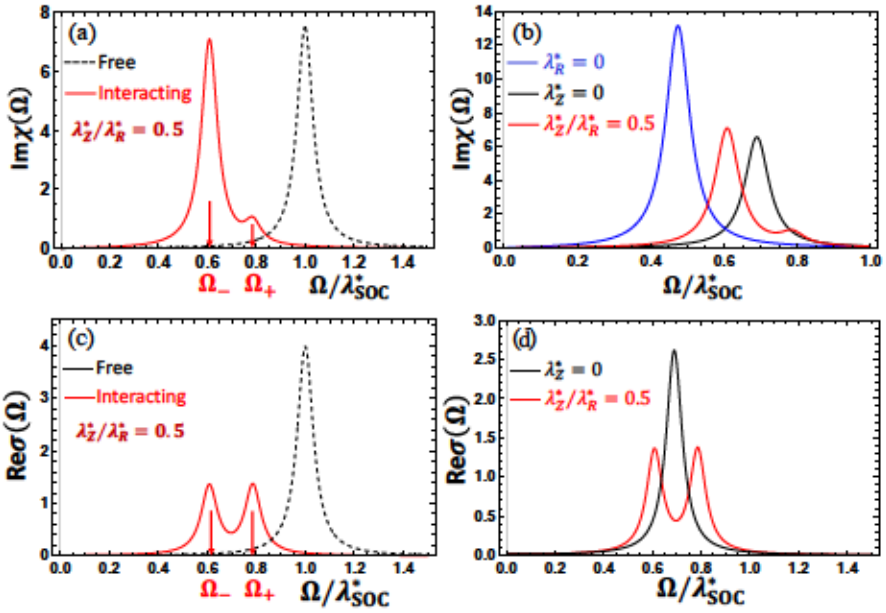


Figure 16. Theoretical predictions for the zero-field electron spin resonance (ESR) and electric-dipole spin resonance (EDSR) in graphene with proximity-induced spin-orbit coupling (SOC). (a) ESR signal. Vertical axis: the imaginary part of the dynamical spin susceptibility. The frequency on the horizontal axis is scaled with $\lambda_{\rm SOC}^* = \sqrt{\lambda_{\rm R}^{*2} + \lambda_{\rm VZ}^{*2}}$, where $\lambda_{\rm R}^*$ and $\lambda_{\rm VZ}^*$ are (renormalized) couplings of the Rashba and valley-Zeeman (VZ) types of SOC, respectively. $\Omega_\pm$ are the resonance frequencies, given by Eqs. (43a) and (43b). Dashed line: non-interacting system. Red solid line: a two-valley Fermi liquid (FL) with parameters $F_0^a = -0.5500$, $F_1^a = -0.2750$, $F_2^a = -0.1375$, $H_0 = -0.5000$, $H_1 = -0.2500$, and $H_2 = -0.1250$. The ratio $\lambda_{\rm VZ}^*/\lambda_{\rm R}^* = 0.5$. The choice of FL parameters is the same for all panels of the figure. (b) ESR signal in a FL for several values of $\lambda_{\rm VZ}^*/\lambda_{\rm R}^*$, as indicated in the legend. (c) EDSR signal. Vertical axis: the real part of the optical conductivity. Dashed line: non-interacting system. Solid line: FL. (d) EDSR signal in a FL for two values of $\lambda_{\rm VZ}^*/\lambda_{\rm R}^*$, as indicated in the legend. To account for smearing of the resonances by disorder, we added a damping term, $-\delta\hat{n}(\mathbf{k},t)/\tau_s$, to the right-hand side of the kinetic equation (8). In all panels of Fig. 16, $1/\tau_s = 0.04\lambda_{\rm SOC}^*$, where $\lambda_{\rm SOC}^* = \sqrt{\lambda_{\rm R}^{*2} + \lambda_{\rm VZ}^{*2}}$. For $\lambda_{\rm R}^* = 15.0\,$meV and $\lambda_{\rm VZ}^* = 7.5\,$meV, the spin relaxation time is $1/\tau_s = 1\,$ps. Reprinted with permission from Ref. 94. Copyright 2021 by the American Physical Society.

a. Zero magnetic field. The zero-field ESR and EDSR in graphene with proximity-induced spin-orbit coupling are predicted to have some interesting features.[94] In the presence of electron-electron interactions and both types of SOC (Rashba and valley-Zeeman), both ESR and EDSR signals consist of two peaks, centered at the frequencies of coupled oscillations of the uniform and valley-staggered magnetizations, see panels *a* and *c* in Fig. 16. Splitting of the resonance occurs as long as the Landau function in Eq. (7) has more than just the $m = 0$ harmonic in the spin-exchange and spin-valley-exchange channels, which is always the case for graphene. Next, even if the (interacting) system has only one type of SOC, there are still two resonance modes, Ω_+ and Ω_-, but one of them is both ESR- and EDSR-silent, because the spectral weights of the corresponding resonances vanish. This effect is illustrated in panels *b* and *d* of Fig. 16. Another interesting feature is that the two ESR (EDSR) peaks have different (comparable) magnitudes. Therefore, EDSR is better way to probe the two-peak structure of the resonance.

It is worth pointing out that the relative strengths of the Rashba and VZ components of SOC in graphene on TMD substrates is currently an open issue. While weak antilocalization experiments on monolayer graphene find VZ SOC to be much stronger than the Rashba one,[74, 75, 78] the opposite conclusion is reached in, e.g., Refs. 68–70, and 76. On the other hand, strong evidence for Rashba SOC being the dominant type in bilayer graphene on WSe_2 follows from the dependence of the splitting of the ShdH frequencies on the carrier number density.[69] Without getting deeper into this discussion, we note that the ESR and EDSR experiments can be used as an independent test for the dominant type of SOC. Indeed, the coupling between the electric field and electron spins is possible only due to Rashba SOC. Therefore, if the experiment shows no EDSR signal, while the ESR signal contains only a single peak, this would be a clear indication that VZ SOC is the dominant mechanism. On the contrary, if single peaks (at the same frequency) are observed both by EDSR and ESR, this would indicate that Rashba SOC is the dominant mechanism. Finally, if both ESR and EDSR signals are split into two peaks, this would indicate that the Rashba and VZ types of SOC are of comparable strength. A quantitative analysis of the signal shape would allow one not only to obtain the spin-orbit coupling constants, but also to extract up to six FL parameters in the $m = 0, 1, 2$ angular momentum channels, which

are hard, if at all possible, to be deduced from other types of measurements.

b. Strong magnetic field. The opposite case of a strong (compared to SOC) in-plane magnetic field was analyzed in Ref. 95 at $q = 0$ and at finite q in Ref. 150. If the effect of SOC on the spectrum of collective modes is neglected, the latter consists of the Silin modes, corresponding to oscillations of the uniform magnetization with frequencies as in Eq. (15), and of an additional set of modes, corresponding to oscillations of the valley-staggered magnetization with frequencies

$$\tilde{\Omega}_m = \frac{1 + H_m^\perp}{1 + F_0^a} \Delta_{\rm Z}, \tag{65}$$

where $H_m^\perp$ is the $m^{\rm th}$ harmonic of the function $H^\perp(\vartheta)$ in Eq. (7). In the absence of SOC, an external magnetic field couples only to the $m = 0$ Silin mode while the electric field does not couple to either of the modes. If both Rashba and valley-Zeeman types of SOC are present and, in addition, the Dirac point is gapped due to the breaking of the A-B symmetry of the honeycomb lattice by the substrate, the external electric field couples to the $m = 0$ and $m = 2$ Silin modes, and to the $m = 0$ and $m = 1$ valley-staggered modes. Therefore, the EDSR spectrum consists, in general, of four peaks. In addition, the resonances occurs not only in the longitudinal conductivity, but also in the transverse (Hall) one, although the external magnetic field does not affect the electron orbits. This last effect occurs due to the Berry curvature of the gapped Dirac point, and its mechanism can be understood already for non-interacting electrons as follows.[95]

Initially, all particle spins are polarized along the external magnetic field, which we take to be along the $\hat{\mathbf{x}}$ axis. Upon application of an external field $\mathbf{E}_{\rm em}(t)$ the particle spins feel an effective Rashba magnetic field $\mathbf{B}_{\rm R} \propto \hat{\mathbf{z}} \times \mathbf{j}$, where $\mathbf{j}$ is the electric current density, and therefore experience a spin torque $\mathbf{T} \propto \hat{\mathbf{x}} \times \mathbf{B}_{\rm R} \propto \hat{\mathbf{z}} j_x$. The x component of the current j_x is composed of regular and anomalous pieces, shown in the left of Fig. 17,

$$j_x(\omega) = -\frac{e^2 n E_{{\rm em},x}}{i\omega m^*} - e^2 N \mathcal{B}_z(k) E_{{\rm em},y}, \tag{66}$$

where n is the number density, $m^* = k_F/\epsilon_k'|_{k=k_F}$ is the effective mass, and $\boldsymbol{\mathcal{B}}(k) = \mp v_{\rm D}^2 \Delta/2(v_{\rm D}^2 k^2 + \Delta^2)^{3/2}\hat{\mathbf{z}}$ is the Berry curvature of the gapped Dirac points in the K (K') valleys. The first term in Eq. (66) creates identical torques in both valleys, while the second one, being proportional to the Berry curvature, yields valley-staggered torques depicted in the right panel of Fig. 17. The component of $\mathbf{E}_{\rm em}$ along $\mathbf{B}$ causes a valley-uniform torque on the spin, exciting the Silin mode spin, while the component of $\mathbf{E}_{\rm em}$ transverse to $\mathbf{B}$ causes a valley-staggered torque, and thus excites the valley-staggered spin mode. Because the charge-to-spin conversion in both cases is proportional to the Rashba coupling, this leads to a term in the conductivity proportional to α_R^2. Furthermore, the Silin mode contributes to σ_{xx}, while the valley-staggered mode contributes to σ_{xy}.

IX. CONCLUSIONS

In this review, we summarized recent progress in theoretical understanding and experimental observation of a new type of collective spin modes, arising from an interplay between spin-orbit coupling (SOC) and electron-electron interaction. We focused on three types of real systems: i) a two-dimensional (2D) electron gas (2DEG) with Rashba and/or Dresselhaus SOC, ii) graphene with proximity-induced SOC, and iii) the Dirac state on the surface of a three-dimensional topological insulator. Provided that SOC and/or external magnetic field are weak, i.e., the corresponding energy scales are much smaller than the Fermi energy, collective modes in systems i) and ii) can be described within the single-valley or two-valley versions of the Fermi-liquid (FL) theory, respectively. A transparent physical picture of such collective modes arises due to mapping of a kinetic equation for a 2D FL onto an effective tight-binding model for an artificial one-dimensional lattice, whose sites are labeled by the projections of angular momentum on the normal to the 2DEG plane (m). Rashba SOC plays the role of on-site energies, while Zeeman and Dresselhaus terms correspond to hopping between the nearest and next-to-nearest neighbors, respectively, whereas the m-dependent components of the Landau interaction function create "defects" of both on-site and bond types. Within this mapping, the continuum of particle-hole excitations plays a role of the conduction band, while collective modes are the bound states produced by defects.

We discussed the results of recent Raman experiments on $Cd_{1-x}Mn_xTe$ quantum wells[55–60] and the Dirac state on the surface of Bi_2Se_3,[61] in which some of the predicted collective modes have been observed, and formulated predictions for future electron spin resonance (ESR) and electric-dipole spin resonance (EDSR) experiments on graphene with proximity-induced SOC.

The new type of collective modes, discussed in this paper, may have potential applications in spintronics, magnonics, optoelectronics, and quantum sensing. Indeed, such modes can be thought of as massive "particles", with masses

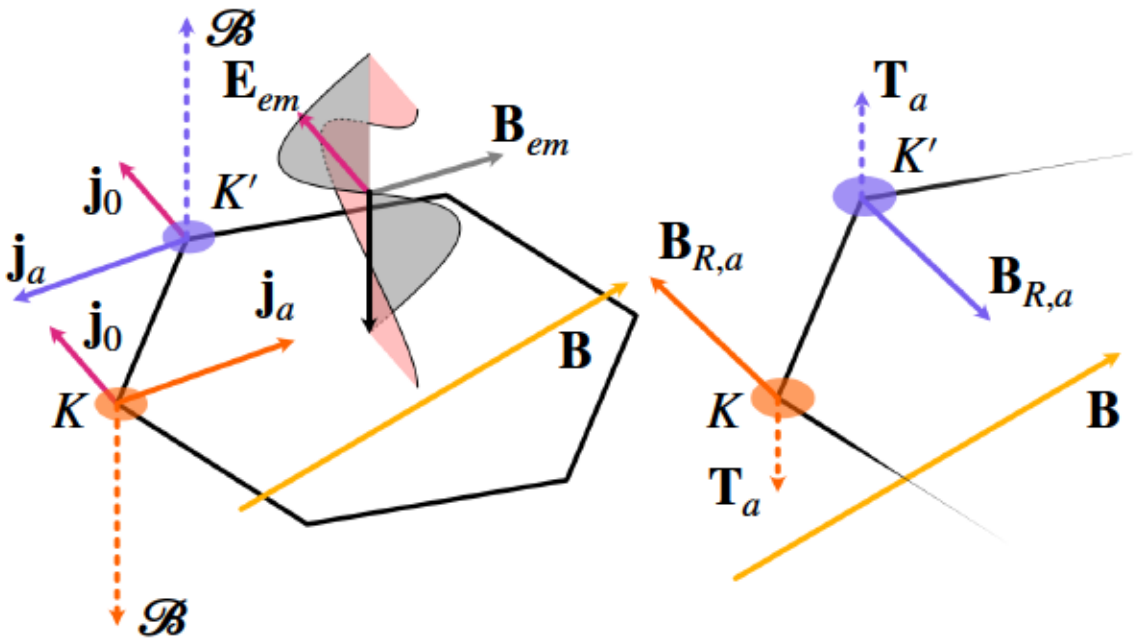


Figure 17. *Left*: Electric-dipole spin resonance in graphene with proximity-induced spin-orbit coupling in the presence of strong in-plane magnetic field. The regular, $\mathbf{j}_0 \propto \mathbf{E}_{\rm em}$, and anomalous, $\mathbf{j}_a \propto \boldsymbol{\mathcal{B}} \times \mathbf{E}_{\rm em}$, currents at the K (orange) and K' (purple) points in the Brillouin zone, induced by the electric field $\mathbf{E}_{\rm em}$ of the incident electromagnetic wave. Here, $\boldsymbol{\mathcal{B}}$ is the valley-staggered Berry curvature. *Right*: Anomalous fields and torques. Spins are initially polarized along the static magnetic field $\mathbf{B}$. The anomalous current-induced effective Rashba fields, $\mathbf{B}_{{\rm R},a} \propto \lambda_{\rm R}\,(\hat{\mathbf{z}} \times \mathbf{j}_a)$, produce valley-specific torques $\mathbf{T}_a \propto \mathbf{B} \times \mathbf{B}_{{\rm R},a}$, thus exciting the valley-staggered spin mode with an intensity proportional to $|\mathbf{E}_{\rm em} \times \mathbf{B}|$. Reprinted with permission from Ref. 95. Copyright 2022 by the American Physical Society.

fixed by the FL interaction, moving in a potential profile produced by SOC.[115] By modulating the strength of SOC along the plane of motion, e.g., by gating, one can confine the modes to waveguides and use them to transmit signals. Despite inherent disorder and other sources of damping, the so far observed modes of this type are quite sharp and robust; for example, the collective mode on the surface of Bi_2Se_3 is observed up to 300 K.[61]

ACKNOWLEDGMENTS

It is our great pleasure and honor to dedicate this paper to the 95th birthday of Professor Emmanuel Iosifovich Rashba, whose seminal work on spin-orbit effects in solids laid the foundation of modern spintronics. One of us (D.L.M.) has had a pleasure of knowing Emmanuel Iosifovich for a long time and have been fortunate to collaborate with him on Ref. 100, which was one of the most gratifying collaborations in his (D.L.M.'s) career. Emmanuel Iosifovich impresses anyone whom he interacts with even briefly by the clarity of his thoughts, encyclopedic knowledge of physics, acute attention to details, and outward friendliness.

We are grateful to our co-authors A. Ashrafi (Magine), G. Blumberg, L. Glazman, A. Goyal, S. Kung, Z. Raines, E. Rashba, and P. Sharma for collaborating with us on the projects discussed in this paper, and to H. Bouchiat, C. Bowers, A. Chubukov, S. Guéron, F. Perez, C. Ullrich, E. Sherman, O. Starykh, and T. Wakamura for stimulating discussions. This work was supported by the National Science Foundation under Grant No. NSF DMR-1720816 (D.L.M.), the Department of Energy (DOE) Basic Energy Sciences under Grant No. DE-SC0020353 (A.K.), and the Natural Sciences and Engineering Research Council of Canada (NSERC) Grant No. RGPIN-2019-05486 (S.M.).

[1] I. Žutić, J. Fabian, and S. Das Sarma, Rev. Mod. Phys. **76**, 323 (2004).
[2] R. Winkler, *Spin–Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer Berlin / Heidelberg, 2003).
[3] K.V. Samokhin, Ann. Phys. **324**, 2385 (2009).
[4] M. Sigrist and K. Ueda, Rev. Mod. Phys. **63**, 239 (1991).
[5] V. P. Mineev and M. Sigrist, in *Non-Centrosymmetric Superconductors*, Lecture Notes in Physics, edited by E. Bauer and M. Sigrist (Springer Berlin / Heidelberg, 2012) pp. 129–154.
[6] M. S. Bahramy, R. Arita, and N. Nagaosa, Phys. Rev. B **84**, 041202 (2011).
[7] W. Witczak-Krempa, G. Chen, Y. B. Kim, and L. Balents, Annual Review of Condensed Matter Physics **5**, 57 (2014).
[8] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[9] M. Z. Hasan and J. E. Moore, Annu. Rev. Conden. Matt. Phys. **2**, 55 (2011).
[10] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
[11] J. Alicea, Rep. Prog. Phys. **75**, 076501 (2012).
[12] J. H. Bardarson and J. E. Moore, Rep. Prog. Phys. **76**, 056501 (2013).
[13] C. Cen, S. Thiel, J. Mannhart, and J. Levy, Science **323**, 1026 ((2009).).
[14] S. Manzeli, D. Ovchinnikov, D. Pasquier, O. V. Yazyev, and A. Kis, Nature Reviews Materials **2**, 17033 (2017).
[15] G. Wang, A. Chernikov, M. M. Glazov, T. F. Heinz, X. Marie, T. Amand, and B. Urbaszek, Rev. Mod. Phys. **90**, 021001 (2018).
[16] Z. Wang, D.-K. Ki, H. Chen, H. Berger, A. H. MacDonald, and A. F. Morpurgo, Nature Communications **6**, 8339 EP (2015).
[17] Y.-J. Lin, R. L. Compton, A. R. Perry, W. D. Phillips, J. V. Porto, and I. B. Spielman, Phys. Rev. Lett. **102**, 130401 (2009).
[18] Y. J. Lin, K. Jiménez-García, and I. B. Spielman, Nature **471**, 83 (2011).
[19] P. Wang, Z.-Q. Yu, Z. Fu, J. Miao, L. Huang, S. Chai, H. Zhai, and J. Zhang, Phys. Rev. Lett. **109**, 095301 (2012).
[20] L. W. Cheuk, A. T. Sommer, Z. Hadzibabic, T. Yefsah, W. S. Bakr, and M. W. Zwierlein, Phys. Rev. Lett. **109**, 095302 (2012).
[21] M. Schulte, J. G. S. Lok, G. Denninger, and W. Dietsche, Phys. Rev. Lett. **94**, 137601 (2005).
[22] Z. Wilamowski, H. Malissa, F. Schäffler, and W. Jantsch, Phys. Rev. Lett. **98**, 187203 (2007).
[23] Z. Wilamowski, W. Ungier, and W. Jantsch, Phys. Rev. B **78**, 174423 (2008).
[24] Y. K. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Science **306**, 1910 (2004).
[25] Sih, V. and Myers, R. C. and Kato, Y. K. and Lau, W. H. and Gossard, A. C. and Awschalom, D. D., Nat. Phys **1**, 31 ((2005).).
[26] J. Wunderlich, B. Kaestner, J. Sinova, and T. Jungwirth, Phys. Rev. Lett. **94**, 047204 ((2005).).
[27] J. Schliemann, J. C. Egues, and D. Loss, Phys. Rev. Lett. **90**, 146801 ((2003).).
[28] B. A. Bernevig, J. Orenstein, and S.-C. Zhang, Phys. Rev. Lett. **97**, 236601 (2006).
[29] J. D. Koralek, C. P. Weber, J. Orenstein, B. A. Bernevig, S.-C. Zhang, S. Mack, and D. D. Awschalom, Nature **458**, 610 ((2009).).
[30] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).
[31] B. A. Bernevig and S.-C. Zhang, Phys. Rev. Lett. **96**, 106802 (2006).
[32] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
[33] Cui-Zu Zhang et al., Science **340**, 167.
[34] J. G. Checkelsky, R. Yoshimi, A. Tsukazaki, K. S. Takahashi, Y. Kozuka, J. Falson, M. Kawasaki, and Y. Tokura, Nat. Phys. **10**, 731 (2014).
[35] W. Chang, S. M. Albrecht, T. S. Jespersen, F. Kuemmeth, P. Krogstrup, J. Nygård, and C. M. Marcus, Nature Nanotechnology **10**, 232 EP (2015).
[36] D. Pesin and L. Balents, Nat. Phys. **6**, 376 ((2010).).
[37] W. Witczak-Krempa, G. Chen, Y. B. Kim, and L. Balents, arXiv:1305.2193.
[38] L. Fu, Phys. Rev. Lett. **115**, 026401 (2015).
[39] R. Lundgren and J. Maciejko, Phys. Rev. Lett. **115**, 066401 (2015).
[40] L. P. Gor'kov and E. I. Rashba, Phys. Rev. Lett. **87**, 037004 (2001).
[41] M. Dzero, K. Sun, V. Galitski, and P. Coleman, Phys. Rev. Lett. **104**, 106408 (2010).
[42] A.-K. Farid and E. G. Mishchenko, Phys. Rev. Lett. **97**, 096604 (2006).
[43] A. Agarwal, S. Chesi, T. Jungwirth, J. Sinova, G. Vignale, and M. Polini, Phys. Rev. B **83**, 115135 (2011).
[44] S. M. Badalyan, A. Matos-Abiague, G. Vignale, and J. Fabian, Phys. Rev. B **79**, 205305 (2009).
[45] S. Raghu, S. B. Chung, X.-L. Qi, and S.-C. Zhang, Phys. Rev. Lett. **104**, 116401 ((2010).).
[46] S. Maiti, V. Zyuzin, and D. L. Maslov, Phys. Rev. B **91**, 035106 (2015).
[47] P. Simon and D. Loss, Phys. Rev. Lett. **98**, 156401 (2007).
[48] S. Chesi and D. Loss, Phys. Rev. B **82**, 165303 (2010).
[49] S. M. Badalyan, A. Matos-Abiague, G. Vignale, and J. Fabian, Phys. Rev. B **81**, 205314 (2010).
[50] R. A. Żak, D. L. Maslov, and D. Loss, Phys. Rev. B **82**, 115415 (2010).

[51] R. A. Żak, D. L. Maslov, and D. Loss, Phys. Rev. B **85**, 115424 (2012).
[52] D. Miserev, J. Klinovaja, and D. Loss, Phys. Rev. B **103**, 075104 (2021).
[53] Y. Gindikin and V. A. Sablikov, ArXiv2206.12586 (2022).
[54] V. P. Silin, Sov. Phys. JETP **6**, 945 (1958).
[55] F. Perez, C. Aku-leh, D. Richards, B. Jusserand, L. C. Smith, D. Wolverson, and G. Karczewski, Phys. Rev. Lett. **99**, 026403 (2007).
[56] F. Baboux, F. Perez, C. A. Ullrich, I. D'Amico, G. Karczewski, and T. Wojtowicz, Phys. Rev. B **87**, 121303 (2013).
[57] F. Baboux, F. Perez, C. A. Ullrich, G. Karczewski, and T. Wojtowicz, Phys. Rev. B **92**, 125307 (2015).
[58] F. Perez, F. Baboux, C. A. Ullrich, I. D'Amico, G. Vignale, G. Karczewski, and T. Wojtowicz, Phys. Rev. Lett. **117**, 137204 (2016).
[59] S. Karimi, F. Baboux, F. Perez, C. A. Ullrich, G. Karczewski, and T. Wojtowicz, Phys. Rev. B **96**, 045301 (2017).
[60] I. D'Amico, F. Perez, and C. A. Ullrich, J. Phys. D: Applied Physics **52**, 203001 (2019).
[61] H.-H. Kung, S. Maiti, X. Wang, S.-W. Cheong, D. L. Maslov, and G. Blumberg, Phys. Rev. Lett. **119**, 136802 (2017).
[62] E. I. Rashba and V. I. Sheka, Fiz. Tverd. Tela: Collected Papers **2**, 162 (1959).
[63] Y. A. Bychkov and E. I. Rashba, JETP Lett. **39**, 78 (1984).
[64] G. Dresselhaus, Phys. Rev. **100**, 580 (1955).
[65] M. I. Dyakonov and V. Y. Kachorovskii, Sov. Phys. Semicond. **20**, 110 (1986).
[66] M. I. Dyakonov, in *Spin Physics in Semiconductors*, Vol. 157, edited by M. I. Dyakonov (Springer, 2017) p. 1.
[67] A. Avsar, J. Y. Tan, T. Taychatanapat, J. Balakrishnan, G. K. W. Koon, Y. Yeo, J. Lahiri, A. Carvalho, A. S. Rodin, E. C. T. O'Farrell, G. Eda, A. H. Castro Neto, and B. Özyilmaz, Nature Communications **5**, 4875 (2014).
[68] Z. Wang, D.-K. Ki, H. Chen, H. Berger, A. H. MacDonald, and A. F. Morpurgo, Nature Communications **6**, 8339 EP (2015).
[69] Z. Wang, D.-K. Ki, J. Y. Khoo, D. Mauro, H. Berger, L. S. Levitov, and A. F. Morpurgo, Phys. Rev. X **6**, 041020 (2016).
[70] B. Yang, M.-F. Tu, J. Kim, Y. Wu, H. Wang, J. Alicea, R. Wu, M. Bockrath, and J. Shi, 2D Materials **3**, 031012 (2016).
[71] B. Yang, M. Lohmann, D. Barroso, I. Liao, Z. Lin, Y. Liu, L. Bartels, K. Watanabe, T. Taniguchi, and J. Shi, Phys. Rev. B **96**, 041409 (2017).
[72] A. Dankert and S. P. Dash, Nature Communications **8**, 16093 (2017).
[73] T. S. Ghiasi, J. Ingla-Aynés, A. A. Kaverzin, and B. J. van Wees, Nano Letters **17**, 7528 (2017).
[74] T. Wakamura, F. Reale, P. Palczynski, S. Guéron, C. Mattevi, and H. Bouchiat, Phys. Rev. Lett. **120**, 106802 (2018).
[75] S. Zihlmann, A. W. Cummings, J. H. Garcia, M. Kedves, K. Watanabe, T. Taniguchi, C. Schönenberger, and P. Makk, Phys. Rev. B **97**, 075434 (2018).
[76] S. Omar and B. J. van Wees, Phys. Rev. B **97**, 045414 (2018).
[77] L. A. Benítez, J. F. Sierra, W. Savero Torres, A. Arrighi, F. Bonell, M. V. Costache, and S. O. Valenzuela, Nature Physics **14**, 303 (2018).
[78] T. Wakamura, F. Reale, P. Palczynski, M. Q. Zhao, A. T. C. Johnson, S. Guéron, C. Mattevi, A. Ouerghi, and H. Bouchiat, Phys. Rev. B **99**, 245402 (2019).
[79] H.-F. Wang, C. Tang, and Q. Zhang, Nano Today **25**, 27 (2019).
[80] J. O. Island, X. Cui, C. Lewandowski, J. Y. Khoo, E. M. Spanton, H. Zhou, D. Rhodes, J. C. Hone, T. Taniguchi, K. Watanabe, L. S. Levitov, M. P. Zaletel, and A. F. Young, Nature **571**, 85 (2019).
[81] A. W. Cummings, J. H. Garcia, J. Fabian, and S. Roche, Phys. Rev. Lett. **119**, 206601 (2017).
[82] J. H. Garcia, M. Vila, A. W. Cummings, and S. Roche, Chem. Soc. Rev. **47**, 3359 (2018).
[83] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[84] H. Min, J. E. Hill, N. A. Sinitsyn, B. R. Sahu, L. Kleinman, and A. H. MacDonald, Phys. Rev. B **74**, 165310 (2006).
[85] Y. Yao, F. Ye, X.-L. Qi, S.-C. Zhang, and Z. Fang, Phys. Rev. B **75**, 041401 (2007).
[86] J. C. Boettger and S. B. Trickey, Phys. Rev. B **75**, 121402 (2007).
[87] S. Konschuh, M. Gmitra, and J. Fabian, Phys. Rev. B **82**, 245412 (2010).
[88] J. Sichau, M. Prada, T. Anlauf, T. J. Lyon, B. Bosnjak, L. Tiemann, and R. H. Blick, Phys. Rev. Lett. **122**, 046403 (2019).
[89] L. Fu, Phys. Rev. Lett. **103**, 266801 (2009).
[90] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover, New York, 1963).
[91] E. M. Lifshitz and L. P. Pitaevski, *Statistical Physics, Part II, Course of Theoretical Physics, v. IX* (Pergamon Press, New York, 1980).
[92] G. Baym and C. J. Pethick, *Landau Fermi-Liquid Theory: Concepts and Applications* (John Wiley and Sons, 1991).
[93] Z. M. Raines, V. I. Fal'ko, and L. I. Glazman, Phys. Rev. B **103**, 075422 (2021).
[94] A. Kumar, S. Maiti, and D. L. Maslov, Phys. Rev. B **104**, 155138 (2021).
[95] Z. M. Raines, D. L. Maslov, and L. I. Glazman, Phys. Rev. B **105**, L201201 (2022).
[96] D. Culcer, Y. Yao, and Q. Niu, Phys. Rev. B **72**, 085110 (2005).
[97] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. **82**, 1959 (2010).
[98] E. Bettelheim, J. Phys. Math. Theor. **50**, 415303 (2017).
[99] A. A. Abrikosov and I. E. Dzyaloshinskii, Sov. Phys. JETP **8**, 535 (1958).
[100] A. Ashrafi, E. I. Rashba, and D. L. Maslov, Phys. Rev. B **88**, 075115 (2013).
[101] S. B. Treiman, R. Jackiw, and D. J. Gross, *Lectures on Current Algebra and Its Applications* (Princeton University Press,

1972).
[102] A. E. Meyerovich and K. A. Musaelian, J. Low Temp. Phys. **89**, 781 (1992).
[103] A. E. Meyerovich and K. A. Musaelian, J. Low Temp. Phys. **94**, 249 (1994).
[104] A. E. Meyerovich and K. A. Musaelian, J. Low Temp. Phys. **95**, 789 (1994).
[105] V. P. Mineev, Phys. Rev. B **69**, 144429 (2004).
[106] C. Herring, in *Magnetism*, Vol. IV, edited by G. Rado and H. Suhl (Academic Press, 1966).
[107] A. V. Chubukov, A. Klein, and D. L. Maslov, JETP **127**, 826 (2018).
[108] S.-k. Ma, M. T. Béal-Monod, and D. R. Fredkin, Phys. Rev. **174**, 227 (1968).
[109] D. Candela, N. Masuhara, D. S. Sherrill, and D. O. Edwards, J. Low Temp. Phys. **63**, 369 (1986).
[110] S. Schultz and G. Dunifer, Phys. Rev. Lett. **18**, 283 (1967).
[111] P. M. Platzman and W. M. Walsh, Phys. Rev. Lett. **19**, 514 (1967).
[112] L. Balents and O. A. Starykh, Phys. Rev. B **101**, 020401 (2020).
[113] A. Shekhter, M. Khodas, and A. M. Finkel'stein, Phys. Rev. B **71**, 165329 (2005).
[114] G.-H. Chen and M. E. Raikh, Phys. Rev. B **60**, 4826 (1999).
[115] A. Ashrafi and D. L. Maslov, Phys. Rev. Lett. **109**, 227201 (2012).
[116] A. Kumar and D. L. Maslov, Phys. Rev. B **95**, 165140 (2017).
[117] J. Y. Khoo and I. S. Villadiego, Phys. Rev. B **99**, 075434 (2019).
[118] S. Maiti, M. Imran, and D. L. Maslov, Phys. Rev. B **93**, 045134 (2016).
[119] P. Löwdin, J. Chem. Phys. **19**, 1396 (1951).
[120] S. Ganguly, A. Kumar, S. Maiti, and D. L. Maslov, unpublished.
[121] G. D. Mahan, Phys. Rev. **153**, 882 (1967).
[122] A. P. Goyal, *Electron-hole interactions in materials with and without a Fermi surface*, Ph.D. thesis, University of Florida (2022).
[123] J. Gavoret, P. Nozières, B. Roulet, and M. Combescot, Journal de Physique **30**, 987 (1969).
[124] D. Pimenov, J. von Delft, L. Glazman, and M. Goldstein, Phys. Rev. B **96**, 155310 (2017).
[125] P. Sharma, A. Principi, and D. L. Maslov, Phys. Rev. B **104**, 045142 (2021).
[126] A. P. Goyal, P. Sharma, and D. L. Maslov, unpublished.
[127] V. P. Mineev, Phys. Rev. B **72**, 144418 (2005).
[128] S.-S. Zhang, X.-L. Yu, J. Ye, and W.-M. Liu, Phys. Rev. A **87**, 063623 (2013).
[129] S. Maiti and D. L. Maslov, Phys. Rev. B **95**, 134425 (2017).
[130] D. F. DuBois and M. G. Kivelson, Phys. Rev. **186**, 409 (1969).
[131] E. G. Mishchenko, M. Y. Reizer, and L. I. Glazman, Phys. Rev. B **69**, 195302 (2004).
[132] M. Pustilnik, M. Khodas, A. Kamenev, and L. I. Glazman, Phys. Rev. Lett. **96**, 196405 (2006).
[133] A. Imambekov, T. L. Schmidt, and L. I. Glazman, Rev. Mod. Phys. **84**, 1253 (2012).
[134] A. Principi, G. Vignale, M. Carrega, and M. Polini, Phys. Rev. B **88**, 195405 (2013).
[135] I. E. Dzyaloshinskii and P. Kondratenko, Sov. Phys. JETP **43**, 1036 (1976).
[136] P. S. Kondratenko, JETP **20**, 1032 (1965).
[137] S. Maiti and D. L. Maslov, Phys. Rev. Lett. **114**, 156803 (2015).
[138] H.-H. Kung, A. P. Goyal, D. L. Maslov, X. Wang, A. Lee, A. F. Kemper, S.-W. Cheong, and G. Blumberg, Proc. Natl. Acad. Sci. USA **116**, 4006 (2019).
[139] D. Stein, K. v. Klitzing, and G. Weimann, Phys. Rev. Lett. **51**, 130 (1983).
[140] C. F. O. Graeff, M. S. Brandt, M. Stutzmann, M. Holzmann, G. Abstreiter, and F. Schäffler, Phys. Rev. B **59**, 13242 (1999).
[141] S. A. Vitkalov, C. R. Bowers, J. A. Simmons, and J. L. Reno, Phys. Rev. B **61**, 5447 (2000).
[142] E. Olshanetsky, J. D. Caldwell, M. Pilla, S.-c. Liu, C. R. Bowers, J. A. Simmons, and J. L. Reno, Phys. Rev. B **67**, 165325 (2003).
[143] E. I. Rashba and V. I. Sheka, in *Landau Level Spectroscopy*, edited by G. Landwehr and E. I. Rashba (North-Holland, Amsterdam, 1991) p. 131.
[144] E. I. Rashba and A. L. Efros, Applied Physics Letters **83**, 5295 (2003).
[145] A. L. Efros and E. I. Rashba, Phys. Rev. B **73**, 165325 (2006).
[146] J. Nitta, T. Akazaki, H. Takayanagi, and T. Enoki, Phys. Rev. Lett. **78**, 1335 (1997).
[147] Y. Sato, T. Kita, S. Gozu, and S. Yamada, Journal of Applied Physics **89**, 8017 (2001).
[148] S. Yamada, Science and Technology of Advanced Materials **4**, 77 (2003).
[149] J. Luo, H. Munekata, F. F. Fang, and P. J. Stiles, Phys. Rev. B **41**, 7685 (1990).
[150] M. J. Anderson, F. Perez, and C. A. Ullrich, Phys. Rev. B **104**, 245422 (2021).