On the effective models of spin-orbit coupling in a two-dimensional electron gas

K. V. Samokhin*

Department of Physics, Brock University, St. Catharines, Ontario L2S 3A1, Canada

We use the method of invariants to derive one- and two-band effective Hamiltonians of a noncentrosymmetric two-dimensional electron gas, in the presence of magnetic field. A complete classification of the antisymmetric spin-orbit and magnetic coupling terms near the Γ point is developed for all two-dimensional crystal symmetries. The effective Hamiltonian depends on the symmetry of the Bloch bands at the Γ point, which is described by one of the double-valued corepresentations of the two-dimensional magnetic point group. In some bands, the spin-orbit coupling is cubic in the electron momentum and the effective Zeeman interaction is strongly anisotropic. As an example of a two-band effective Hamiltonian, we introduce a simple model of a topological insulator with the intraband and interband spin-orbit coupling and investigate its bulk and boundary properties.

Keywords: spin-orbit coupling; multiband effective Hamiltonian; method of invariants; topological insulator

I. INTRODUCTION

Two-dimensional (2D) electron systems with the spin-orbit (SO) coupling have been one of most popular fields of research in recent years, driven primarily by their applications to spintronics\textsuperscript{1,2} and topological superconductivity.\textsuperscript{3} The long list of materials of interest includes graphene,\textsuperscript{4} monolayer transition-metal dichalcogenides,\textsuperscript{5} conducting interfaces or surfaces of insulating oxides,\textsuperscript{6} ultra-thin metal films on various substrates,\textsuperscript{7} and others. In most of these systems the inversion symmetry is broken due to the presence of a substrate, or due to different nature of the materials sandwiching the conducting layer.

The SO coupling in a noncentrosymmetric crystal lifts the spin degeneracy of the electron states almost everywhere in the Brillouin zone, producing the Bloch bands characterized by a complex spin texture and a nontrivial momentum-space topology. This is responsible for a number of remarkable effects, both in normal and superconducting states, such as the quantum spin Hall effect in topological insulators,\textsuperscript{8} the spin field-effect transistor,\textsuperscript{9} the magnetoelectric effect,\textsuperscript{10} and the unusual nonuniform superconducting states in the presence of magnetic field or even without any field.\textsuperscript{11}

Theoretical understanding of the SO coupling-controlled physics in noncentrosymmetric materials is based on the Rashba model, see Ref. 12 and also Ref. 13 for a review. In the Rashba model and its generalizations, the SO coupling is described by the terms in the effective Bloch Hamiltonian which are odd in the electron momentum (in the original model,\textsuperscript{12} the SO coupling is linear in $\mathbf{k}$). Such terms, known as the “asymmetric” or “antisymmetric” SO coupling, are allowed by symmetry and can be justified microscopically using the $\mathbf{k} \cdot \mathbf{p}$ perturbation theory.\textsuperscript{14} While there exists an extensive literature focusing on the crystal structures and geometries relevant to semiconducting systems, see, e.g., Ref. 15, a complete classification of the phenomenological effective Hamiltonians in noncentrosymmetric 2D crystals, which includes both the SO coupling and the magnetic field interactions, seems to be missing. The goal of the present paper is to fill this gap.

The structure of the paper is as follows. In Sec. II, we discuss the standard Rashba model and its different interpretations. In Secs. III-V, we use the method of invariants to derive the one-band and two-band effective Hamiltonians in the vicinity of the Γ point. This is done for all crystal symmetries in 2D, taking into account an external magnetic field. We show that the structure of the effective Hamiltonian crucially depends on the Γ-point corepresentation of the 2D magnetic point group and in many cases differs significantly from the standard Rashba form. Our investigation of two-band effective Hamiltonians is motivated by the fact that the popular models of topological insulators, such as the Kane-Mele model\textsuperscript{16} and the Bernevig-Hughes-Zhang model,\textsuperscript{17} rely on presence of the SO coupling in a two-band, or a two-sublattice, system. In Sec. VI, we introduce a toy model of a 2D TR-invariant topological insulator and study its bulk and boundary properties, emphasizing the qualitatively different effects of the intraband and interband SO coupling. In the Appendices, we present the details of the calculations, in particular, the microscopic justification of the various terms in the phenomenological effective Hamiltonians and also a summary of the properties of the modified Dirac Hamiltonian. Throughout the paper we use the units in which $\hbar = 1$, neglecting, in particular, the difference between the quasiparticle momentum and the wave vector.

* E-mail: kirill.samokhin@brocku.ca
II. RASHBA MODEL

The electron-lattice SO coupling combined with inversion symmetry breaking lifts the twofold spin degeneracy of the Bloch states at the wave vector $k = (k_x, k_y)$. In the case of one spin-degenerate band, these effects are described by the following effective Hamiltonian, known as the Rashba model:

$$
\hat{H}_{\text{Rashba}} = \epsilon(k)\hat{\sigma}_0 + \gamma(k)\hat{\sigma} + \mu_B H\hat{\sigma},
$$

where $\hat{\sigma}_0$ is the unit matrix and $\hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3)$ are the Pauli matrices. The second term is the antisymmetric SO coupling, characterized by the pseudovector $\gamma$, which is invariant under the operations of the 2D point group, satisfies $\gamma(-k) = -\gamma(k)$, and vanishes in a centrosymmetric crystal or in the absence of the SO coupling. In the original Rashba model, $\gamma$ one has $\gamma(k) = \gamma_0(k_ye_1 - k_xe_2)$. The last term in Eq. (1) describes the Zeeman interaction of the electron spin with a uniform external magnetic field (the spin magnetic moment is given by $-\mu_B\hat{\sigma}$, where $\mu_B$ is the Bohr magneton, the electron charge is $\epsilon$, and the Landé factor is set to 2). The orbital effects of the magnetic field will not be considered in this paper; if needed, they can be included via the Peierls substitution, whereby $k$ is replaced by the operator $k + (\epsilon/hc)\mathbf{A}(r)$, where $\mathbf{k} = -i\nabla$ and $\mathbf{A}$ is the vector potential. Given the popularity of the the Rashba model and its wide range of applications, it makes sense to revisit its microscopic justification and clarify the meaning of the various terms in Eq. (1).

We consider a quasi-2D electron gas in the $xy$ plane. The potential $U(x, y, z)$ is assumed to be periodic in the $x$ and $y$ directions, but confining in the $z$ direction. Focusing on the applications to electrons on a substrate or a surface, or in a conducting interface between two different materials, we further assume that $U(x, y, -z) \neq U(x, y, z)$, so that the system lacks inversion symmetry. Neglecting the electron-electron interactions, the lattice vibrations, disorder, and setting $\mathbf{H} = 0$, the microscopic Hamiltonian has the following form:

$$
\hat{H} = \frac{\hat{p}^2}{2m} + U(r) + \frac{\hbar}{4m^2c^2}\hat{\sigma}[\nabla U(r) \times \hat{p}],
$$

where $\hat{p} = -i\hbar\nabla$ is the momentum operator and the last term describes the electron-lattice SO coupling. Regarding the symmetry of $U(r)$, the symmetry operations (rotations and reflections) leaving the 2D crystal lattice invariant form the point group of the crystal, which we denote by $G$.

There are ten crystallographic point groups in 2D, also known as the rosette groups: five cyclic groups $C_n$ and five dihedral groups $D_n$, with $n = 1, 2, 3, 4, 6$. The group $C_n$ has $n$ elements and is generated by the rotation $C_{nz}$ about the $z$ axis by an angle $2\pi/n$, while the group $D_n$ has $2n$ elements and is generated by $C_{nz}$ and also by the reflection in a vertical plane (we choose $\sigma_y$ – the reflection in the $y = 0$ plane – as the second generating element). In addition to the rosette group operations, in the absence of an external magnetic field the Hamiltonian (2) is also invariant under the time reversal (TR) operation $K$.

There are several conceptually different ways to construct the effective models of the SO coupling in an inversion-asymmetric system, depending on how Eq. (2) is partitioned to represent various physical interactions. One possible approach is to represent the microscopic Hamiltonian in the form $\hat{H} = \hat{H}_0 + \hat{H}_{SOC}$, where $\hat{H}_0$ comprises the first two terms in Eq. (2), while $\hat{H}_{SOC}$ is the third term. Assuming that the SO coupling is switched on in the presence of an inversion-asymmetric crystal potential, one first finds the eigenstates of $\hat{H}_0$, which are twofold degenerate due to spin, and then calculates the matrix elements of $\hat{H}_{SOC}$ in this basis. In this approach, $\hat{H}_{Rashba}$ is a $2 \times 2$ matrix in the spin space, $\epsilon(k)$ is the band dispersion in a noncentrosymmetric crystal without the SO coupling, and $\gamma(k)$ describes the effects of the SO coupling. Such perturbative treatment of the SO coupling might be problematic, especially in compounds containing heavy elements.

Alternatively, one can split the Hamiltonian (2) into the inversion-symmetric and inversion-antisymmetric parts: $\hat{H} = \hat{H}_s + \hat{H}_a$, both including the SO coupling. The eigenstates of $\hat{H}_s$ are twofold degenerate due to the combined symmetry operation $K_1$, called conjugation ($I$ is the space inversion operation). Then, one calculates the matrix elements of $\hat{H}_a$ in the eigenbasis of $\hat{H}_s$. In this approach, $\hat{H}_{Rashba}$ is a $2 \times 2$ matrix in the conjugacy space, $\epsilon(k)$ is the band dispersion in a centrosymmetric lattice potential with the SO coupling, and $\gamma(k)$ describes the effects of inversion symmetry breaking. While this approach may be legitimate in three dimensional (3D) crystals with a strong SO coupling, the possibility to treat $\hat{H}_a$ as a perturbation in a 2D electron gas on a substrate is questionable.

In this paper, we use another approach, which is based on the method of invariants in the $k \cdot p$ perturbation theory. In this approach, one starts with the exact basis of the Bloch states at the $\Gamma$ point at zero magnetic field, which includes both the SO coupling and the inversion symmetry breaking, and then constructs the effective Hamiltonian using an expansion in powers of $k$ and $H$ in the vicinity of the $\Gamma$ point. Assuming that the Bloch states at the $\Gamma$ point are $d_T$-fold degenerate, the effective Hamiltonian is given by a $d_T \times d_T$ matrix. We shall see that in our 2D case, $d_T$ is always equal to two, however, the effective Hamiltonian does not necessarily have the simple form (1).
III. BLOCH STATES AT THE Γ POINT

Let us first look in detail at the symmetry of the the Bloch states at \( k = 0 \) in the absence of magnetic field. The reduced Hamiltonian at the \( \Gamma \) point has the same form as Eq. (2) and its full symmetry group is given by

\[
G_{k=0} = \mathbb{C} + K \mathbb{C},
\]

where the TR operation \( K \) commutes with all elements of the rosette group \( G \). Since \( K \) is antiunitary, \( G_{k=0} \) is a Type II magnetic, or Shubnikov, point group, see, e.g., Ref. 22. At other TR invariant momenta, which satisfy \( -k = k + G \), where \( G \) is a reciprocal lattice vector, the full symmetry group also contains \( K \) and has a form similar to Eq. (3), with the rosette group replaced by one of its subgroups.

The Bloch states at \( k = 0 \) transform according to the irreducible double-valued corepresentations (coreps) of the magnetic group (3). The coreps can be obtained from the double-valued irreducible representations (irreps) of the unitary component \( \mathbb{C} \) using a standard procedure\(^{22,23} \) and are classified into three cases, A, B, or C, which determine whether or not the antiunitary symmetry brings about an additional degeneracy and also the type of this degeneracy. While there is no additional degeneracy due to TR in Case A (i.e. Case A coreps are the same as the usual irreps), in the other two cases the coreps have twice the dimension of the irreps they are constructed from, leading to an additional degeneracy of the “doubling” type in Case B or the “pairing” type in Case C. In the latter case, the coreps are obtained by combining two complex-conjugate irreps.

Since each rosette group is isomorphic to a certain 3D point group, its double-valued coreps can be easily found, with the results listed in Table I (our notations for the irreps are the same as in Ref. 24). Due to the absence of inversion symmetry, the coreps do not have a definite parity. It is important to note that all double-valued coreps of the rosette groups are 2D, therefore the electron bands are twofold degenerate at the \( \Gamma \) point. The Bloch states \(| n, s \rangle \equiv |n, s \rangle \) are labelled by the band index \( n \) and the additional index \( s = 1, 2 \), which distinguishes two orthonormal states within the same band:

\[
|n, 1\rangle, \quad |n, 2\rangle = K |n, 1\rangle.
\]

Due to the presence of the SO coupling, the Kramers index \( s \) is not the same as the electron spin projection. If the states \(|n, 1\rangle\) and \(|n, 2\rangle\) form the basis of a corep described by \( 2 \times 2 \) matrices \( \hat{D}_n(g) \), where \( g \in \mathbb{C} \), then we have

\[
g |n,s\rangle = \sum_{s'} |n,s'\rangle \hat{D}_{n,s',s}(g).
\]

Using also \( \hat{D}_n(K) = -i \hat{\sigma}_2 \), one can obtain the corep matrices for all elements of the magnetic group (3).

If the \( \Gamma \)-point corep is equivalent to the spin-1/2 corep, then the band is called “pseudospin band” and one can put

\[
\hat{D}_n(g) = \hat{D}^{(1/2)}(R)\]

for \( g = R \) or \( IR \), where \( \hat{D}^{(1/2)}(R) = e^{-i\theta (n \hat{\sigma})/2} \) is the spin-1/2 representation of a counterclockwise rotation \( R \) through an angle \( \theta \) about an axis \( n \). In general, however, the Bloch states at the \( \Gamma \) point do not transform under the rosette group operations in the same way as the basis spinors \(|\uparrow\rangle\) and \(|\downarrow\rangle\) (the eigenstates of the spin operator \( \hat{s}_z \)), which means that the \( \Gamma \)-point corep is not equivalent to the spin-1/2 corep and

\[
\hat{D}_n(g) \neq \hat{D}^{(1/2)}(R)
\]

for some \( g \). We call such bands “nonpseudospin bands”. The representation matrices for the nonpseudospin double-valued coreps are shown in Table II, see Appendix A for details.

At a generic momentum in the 2D Brillouin zone, such that \( k \) is invariant neither under TR nor any rosette group operations, the full symmetry group of the reduced Hamiltonian is given by \( G_{\hat{k}} = C_1 \). This group does not contain any antiunitary elements, therefore the Bloch states are classified according to the usual double-valued irreps of \( C_1 \). Since there is only one such irrep, namely \( \Gamma_2 \), which is one-dimensional (1D), the bands are nondegenerate at a generic \( k \).

IV. ONE-BAND EFFECTIVE HAMILTONIAN

The band structure in the vicinity of the \( \Gamma \) point can be described by an effective \( k \)-dependent Hamiltonian, whose eigenvalues correspond to the band energies. Including the Zeeman interaction, the effective Hamiltonian becomes
TABLE I: The double-valued coreps of the rosette groups at the Γ point. Each rosette group is isomorphic to some noncentrosymmetric 3D point group $G_{3D}$. All coreps are 2D. The Case C coreps are obtained by pairing two complex-conjugate 1D irreps. The last column shows whether the Bloch states at the Γ point transform as the basis spinors.

| $G$ | $G_{3D}$ | corep | corep case | pseudospin |
|-----|----------|-------|-----------|------------|
| $C_1$ | $C_1$ | $\Gamma_2$ | B | Y |
| $C_2$ | $C_2$ | $(\Gamma_3, \Gamma_4)$ | C | Y |
| $C_3$ | $C_3$ | $(\Gamma_4, \Gamma_5)$ | B | N |
| $C_4$ | $C_4$ | $(\Gamma_5, \Gamma_6)$ | C | Y |
| $C_5$ | $C_5$ | $(\Gamma_7, \Gamma_8)$ | C | Y |
| $C_6$ | $C_6$ | $(\Gamma_9, \Gamma_{10})$ | C | N |
| $D_1$ | $C_3$ | $\Gamma_3, \Gamma_4$ | C | Y |
| $D_2$ | $C_{2v}$ | $\Gamma_5$ | A | Y |
| $D_3$ | $C_{3v}$ | $\Gamma_4$ | A | Y |
| $D_4$ | $C_{4v}$ | $\Gamma_6$ | A | Y |
| $D_5$ | $C_{6v}$ | $\Gamma_7$ | A | Y |
| $D_6$ | $C_{6v}$ | $\Gamma_9$ | A | N |

TABLE II: The nonpseudospin double-valued corep matrices at the Γ point ($g$ denotes the generators of the rosette group $G$).

| $G$ | corep | $\hat{D}_1(g)$ |
|-----|-------|---------------|
| $C_3$ | $\Gamma_6$ | $\hat{D}(C_{3k}) = -\sigma_0$ |
| $C_4$ | $(\Gamma_7, \Gamma_8)$ | $\hat{D}(C_{4k}) = -\hat{D}^{(1/2)}(C_{4k})$ |
| $C_6$ | $(\Gamma_9, \Gamma_{10})$ | $\hat{D}(C_{6k}) = -\hat{D}^{(1/2)}(C_{6k})$ |
| $D_3$ | $(\Gamma_5, \Gamma_6)$ | $\hat{D}(C_{3k}) = -\sigma_0, \hat{D}(\sigma_y) = \hat{D}^{(1/2)}(C_{2k})$ |
| $D_4$ | $\Gamma_7$ | $\hat{D}(C_{4k}) = -\hat{D}^{(1/2)}(C_{4k}), \hat{D}(\sigma_y) = \hat{D}^{(1/2)}(C_{2y})$ |
| $D_5$ | $\Gamma_8$ | $\hat{D}(C_{6k}) = -\hat{D}^{(1/2)}(C_{6k}), \hat{D}(\sigma_y) = \hat{D}^{(1/2)}(C_{2y})$ |
| $D_6$ | $\Gamma_9$ | $\hat{D}(C_{6k}) = -\hat{D}^{(1/2)}(C_{6k}), \hat{D}(\sigma_y) = \hat{D}^{(1/2)}(C_{2y})$ |

also a function of the magnetic field $H$. In the one-band case, i.e., if just one Γ-point corep is taken into account, the effective Hamiltonian is given in the basis (4) by a $2 \times 2$ Hermitian matrix $\hat{H}(k, H)$, which depends on $k$ and $H$ as parameters and has the following general form (in this section we drop the band index $n$):

$$\hat{H}(k, H) = \varepsilon(k, H)\sigma_0 + \Lambda(k, H)\sigma,$$

(7)

where the functions $\varepsilon$ and $\Lambda = (\Lambda_1, \Lambda_2, \Lambda_3)$ contain the effects of the SO coupling and inversion symmetry breaking. Diagonalizing this Hamiltonian, one obtains two bands

$$\xi_{\lambda}(k, H) = \varepsilon(k, H) + \lambda|\Lambda(k, H)|,$$

(8)

labelled by the “helicity” index $\lambda = \pm$. The helicity bands are nondegenerate, except the points where all three components of $\Lambda$ simultaneously vanish. In this section, we use a phenomenological, symmetry-based, approach known as the method of invariants$^{21}$ to find the form of $\varepsilon$ and $\Lambda$. The parameters of the effective Hamiltonian can be calculated microscopically, using the $k \cdot p$ perturbation theory, see Appendix B.

Under an operation $g$ from the rosette group $G$, the basis functions are transformed into $g|s⟩ = \sum_s' |s'⟩D_{s's}(g)$, see Eq. (5). On the other hand, the matrix elements of $\hat{H}(k, H)$ in the “old” basis must be equal to the matrix elements
of the effective Hamiltonian with the transformed parameters $\hat{H}(gk, gH)$ in the “new” basis. In this way, we obtain the following constraint on the matrix $\hat{H}$:

$$\hat{H}(k, H) = \hat{D}^\dagger(g) \hat{H}(gk, H) \hat{D}(g).$$

(9)

Regarding the invariance under TR operation, the symmetry of the matrix $\hat{H}(k, H)$ is the same as that of the matrix representation of an observable $O$ which depends on $k$ and $H$ and satisfies the following condition: $K\hat{O}(k, H)K^{-1} = \hat{O}(-k, -H)$. Using $K|i\rangle = \sum_{i_1} |i_1\rangle D_{i_1i}(K)$, where $\hat{D}(K)$ is the representation of TR in an arbitrary basis $|i\rangle$, and the antunitarity property $\langle i|K^{-1}|j\rangle = \langle i|K|j\rangle^*$, we obtain for the matrix elements of the operator $\hat{O}$:

$$\langle i|\hat{O}(k, H)|j\rangle = \langle i|K^{-1}\hat{O}(-k, -H)K|j\rangle = \sum_{j_1} D_{j_1j}^\ast(K) \langle i|K^{-1}\hat{O}(-k, -H)|j_1\rangle = \sum_{i,j_1} D_{i_1i}(K) D_{j_1j}^\ast(K) \langle i_1|\hat{O}(-k, -H)|j_1\rangle^*.$$ 

Therefore,

$$\hat{H}(k, H) = \hat{D}^\dagger(K) \hat{H}^\ast(-k, -H) \hat{D}^\ast(K),$$

(10)

where $\hat{D}(K) = -i\hat{\sigma}_2$ in the basis (4). Note that the constraints (9) and (10) are completely general and can also be applied to multiband effective Hamiltonians. If $N$ bands (coreps) are included in $\hat{H}$, then $\hat{D}(g)$ and $\hat{D}(K)$ become $2N \times 2N$ block-diagonal matrices, see Sec. V for $N = 2$.

Substituting Eq. (7) in the TR invariance constraint (10), we obtain

$$\varepsilon(k, H) = \varepsilon(-k, -H), \quad \Lambda(k, H) = -\Lambda(-k, -H),$$

(11)

while the point-group invariance condition (9) yields

$$\varepsilon(k, H) = \varepsilon(g^{-1}k, g^{-1}H), \quad \Lambda(k, H) = \hat{R}(g)\Lambda(g^{-1}k, g^{-1}H).$$

(12)

The $3 \times 3$ orthogonal matrix $\hat{R}$ depends on the $\Gamma$-point corep and is defined by the following expression:

$$\hat{D}^\dagger(g)\hat{\sigma}_\mu \hat{D}(g) = \sum_\nu \hat{R}_{\mu\nu}(g) \hat{\sigma}_\nu.$$ 

(13)

Here and below we use the Greek indices $(\mu, \nu = 1, 2, 3)$ to label the Pauli matrices in the corep space and the corresponding components of $\Lambda$, and the Latin indices $(i, j = x, y, z)$ to label the components of vectors and pseudovectors, such as $k$ and $H$, in physical space. It follows from Eqs. (11) and (12) that the helicity bands (8) satisfy $\xi_\delta(k, H) = \xi_\delta(-k, -H)$ and $\xi_\delta(k, H) = \xi_\delta(g^{-1}k, g^{-1}H)$. When applying the rosette group invariance conditions one should keep in mind that the proper and improper symmetry elements act differently on the vector $k$ and the pseudovector $H$: for a proper rotation $g = R$, we have $g^{-1}k = R^{-1}k$ and $g^{-1}H = R^{-1}H$, but for an improper rotation $g = IR$, we have $g^{-1}k = -R^{-1}k$ and $g^{-1}H = R^{-1}H$.

In all pseudospin bands, the corep matrix has the form (6) and, using the identity

$$\hat{D}^{(1/2)}(R)\hat{\sigma}_\mu \hat{D}^{(1/2)}(R) = \sum_\nu R_{\mu\nu} \hat{\sigma}_\nu,$$

where $\hat{R}$ is the $3 \times 3$ rotation matrix, we find that $\hat{R}(g) = \hat{R}(g)$ for all elements of the rosette group. This is also valid in most nonpseudospin bands, as one can check using the corep matrices from Table II. However, there are four exceptional coreps in trigonal and hexagonal crystals, in which $\hat{R}(g) \neq \hat{R}(g)$ for some elements of the rosette group. For $\Gamma_6$ of $C_3$, we have $\hat{R}(C_{3e}) = \hat{1}$ (the $3 \times 3$ unit matrix); for $(\Gamma_{11}, \Gamma_{12})$ of $C_6$, we have $\hat{R}(C_{6z}) = \hat{R}(C_{2z})$; for $(\Gamma_5, \Gamma_6)$ of $D_3$, we have $\hat{R}(C_{3e}) = \hat{1}$ and $\hat{R}(\sigma_y) = \hat{R}(C_{2y})$; and for $\Gamma_4$ of $D_6$, we have $\hat{R}(C_{6z}) = \hat{R}(C_{2z})$ and $\hat{R}(\sigma_y) = \hat{R}(C_{2y})$. We shall see below that the effective Hamiltonian in these bands has a form which is considerably different from the non-exceptional bands.

We focus on the response of the system to a weak magnetic field and expand the effective Hamiltonian in powers of $H$:

$$\varepsilon(k, H) = \varepsilon(k) + H\ell(k) + O(H^2),$$

$$\Lambda_\nu(k, H) = \gamma_\nu(k) + \sum_i H_i \mu_{i\nu}(k) + O(H^2),$$

(14)
where $\epsilon$, $\ell$, $\gamma$, and $\bar{\mu}$ are real functions of the wave vector. While the first term in $\mathbf{A}$ reproduces the antisymmetric SO coupling in Eq. (1), the interaction with an external magnetic field does not, in general, have the simple Zeeman form.

### A. Antisymmetric SO coupling

From the requirements of TR and point-group invariance, see Eqs. (11) and (12), we obtain:

$$\epsilon(-k) = \epsilon(k), \quad \epsilon(k) = \epsilon(g^{-1}k),$$

i.e., $\epsilon$ is an invariant scalar, for which one can put $\epsilon(k) = k^2/2m^*$, where $m^*$ is the effective mass. For the antisymmetric SO coupling, we have

$$\gamma(-k) = -\gamma(k), \quad \gamma(k) = \hat{R}(g)\gamma(g^{-1}k).$$

In the non-exceptional bands we have $\hat{R}(g) = \hat{R}(g)$ for all $g$, therefore $\gamma$ is an invariant pseudovector. In contrast, in the four exceptional cases we have $\hat{R}(g) \neq \hat{R}(g)$ for some $g$, so that $\gamma$ does not transform as an invariant pseudovector. Expressions for the antisymmetric SO coupling in the vicinity of the $\Gamma$ point for all rosette groups are listed in Table III (for non-exceptional bands, these expressions are the same as in Ref. 25), with an example of the calculation given in Appendix C. For a classification of the antisymmetric SO coupling in 3D crystals, see Refs. 26 and 19.

Note that for $G = C_2, C_4, C_6, D_2, D_4,$ and $D_6$ the SO coupling is “planar”, in the sense that $\gamma_3$ vanishes at all $k$. This symmetry-imposed constraint is due to the presence of the rotation $C_{2z}$ in these rosette groups. Indeed, putting $g = C_{2z}$ in Eq. (16) and observing that $\hat{R}(C_{2z}) = \hat{R}(C_{2z})$ for all cores, we obtain: $\gamma_3(k) = C_{2z}\gamma_3(C_{2z}^{-1}k) = \gamma_3(-k) = -\gamma_3(k)$, therefore $\gamma_3(k) = 0$.

Also note that in the exceptional bands, the linear in $k$ terms are not allowed by symmetry and the antisymmetric SO coupling is cubic in $k$. This should be contrasted with the “cubic Rashba” model discussed in the context of quasi-2D semiconductor quantum wells, or surface states in rare-earth compounds, in which the linear in momentum terms are numerically small due to material-specific reasons.

### B. Magnetic field coupling

The magnetic field enters two different terms in the effective Hamiltonian (14): the term containing $\hat{\mu}(k)$, which formally resembles the usual Zeeman coupling, albeit with an anisotropic magneton, and also the term containing $\ell(k)$, which is absent in the standard Rashba model (1). Both terms originate from the microscopic Zeeman interaction modified by the SO coupling and can be obtained using the $k \cdot p$ perturbation theory, see Appendix B. From Eqs. (11) and (12), we obtain the following symmetry constraints:

$$\ell(-k) = -\ell(k), \quad \ell(k) = \hat{R}(g)\ell(g^{-1}k)$$

and

$$\hat{\mu}(-k) = \hat{\mu}(k), \quad \hat{\mu}(k) = \hat{R}(g)\hat{\mu}(g^{-1}k)\hat{R}^{-1}(g)$$

for all $g \in G$.

We see that $\ell(k)$ transforms as an invariant pseudovector in all bands, pseudospin and nonpseudospin. It has the same structure as $\gamma(k)$ in the non-exceptional bands, with the results listed in Table IV. In particular, $\ell_z(k) = 0$ if the rotation $C_{2z}$ is a symmetry element. Note that the absence of inversion symmetry is important: if the point group contained $g = I$, then both $\gamma$ and $\ell$ would vanish identically, according to Eqs. (16) and (17).

It follows from Eq. (18) that the effective Zeeman coupling $\hat{\mu}(k)$ is an invariant second-rank tensor only in non-exceptional bands, when $\hat{R}(g) = \hat{R}(g)$. In these bands, we have $\hat{\mu}(k) = \hat{R}(g)\hat{\mu}(g^{-1}k)\hat{R}^{-1}(g)$, therefore the simplest expression for the effective Zeeman coupling has the usual, i.e., isotropic and momentum-independent, form: $\mu_{\nu}(k) = \mu_{\nu}(\delta_{\nu\nu})$. The effective “magneton” $\mu_{\nu}$ includes the effects of the crystal field and the SO coupling and does not have to be equal to the Bohr magneton $\mu_B$. If needed, one can use a more general expression

$$\hat{\mu}(k) = \begin{pmatrix} \mu_\perp & 0 & 0 \\ 0 & \mu_\perp & 0 \\ 0 & 0 & \mu_z \end{pmatrix},$$

where $\mu_\perp$ and $\mu_z$ are real functions of the wave vector.
TABLE III: Lowest-order polynomial expressions for the antisymmetric SO coupling \( (a_i \text{ are real constants, } b_i \text{ are complex constants, and } k_\pm = k_x \pm ik_y) \). Asterisks label the exceptional nonpseudospin bands, in which \( \mathcal{R}(g) \neq \mathcal{R}(g) \) for some elements of the rosette group.

| \( G \) | \( \Gamma \) | \( \gamma(k) \) |
|---|---|---|
| \( C_1 \) | \( \Gamma_2 \) | \((a_1 k_x + a_2 k_y)e_1 + (a_3 k_x + a_4 k_y)e_2 + (a_5 k_x + a_6 k_y)e_3\) |
| \( C_2 \) | \( \Gamma_3, \Gamma_4 \) | \((a_1 k_x + a_2 k_y)e_1 + (a_3 k_x + a_4 k_y)e_2\) |
| \( C_3 \) | \( \Gamma_5, \Gamma_6 \) | \((a_1 k_x + a_2 k_y)e_1 + (-a_2 k_x + a_1 k_y)e_2\) |
| \( C_6 \) | \( \Gamma_7, \Gamma_8 \) | \((a_1 k_x + a_2 k_y)e_1 + (-a_2 k_x + a_1 k_y)e_2\) |
| \( D_1 \) | \( \Gamma_9, \Gamma_{10} \) | \((a_1 k_x + a_2 k_y)e_1 + (-a_2 k_x + a_1 k_y)e_2\) |
| \( D_2 \) | \( \Gamma_11, \Gamma_{12} \) | \((b_1 k_x^3 + b_2 k_x^3)e_1 + (b_2 k_x^3 + b_2 k_x^3)e_2\) |
| \( D_3 \) | \( \Gamma_2 \) | \(a_1 k_y e_1 + a_2 k_y e_2 + a_3 k_y e_3\) |
| \( D_4 \) | \( \Gamma_5 \) | \(a_1 k_y e_1 + a_2 k_y e_2\) |
| \( D_5 \) | \( \Gamma_6 \) | \(a_1 k_y e_1 - k_x e_2\) |
| \( D_6 \) | \( \Gamma_7 \) | \(a_1 k_y e_1 - k_x e_2\) |
| \( D_7 \) | \( \Gamma_8 \) | \(a_1 k_y e_1 - k_x e_2\) |
| \( D_8 \) | \( \Gamma_9 \) | \(ia_1 (k_3 e_1 - k_x e_2)\) |

TABLE IV: Lowest-order polynomial expressions for the magnetic coupling \( \ell(k) \) \( (a_i \text{ are real constants, } b_i \text{ are complex constants, and } k_\pm = k_x \pm ik_y) \).

| \( G \) | \( \ell(k) \) |
|---|---|
| \( C_1 \) | \((a_1 k_x + a_2 k_y)\hat{x} + (a_3 k_x + a_4 k_y)\hat{y} + (a_5 k_x + a_6 k_y)\hat{z}\) |
| \( C_2 \) | \((a_1 k_x + a_2 k_y)\hat{x} + (a_3 k_x + a_4 k_y)\hat{y}\) |
| \( C_3 \) | \((a_1 k_x + a_2 k_y)\hat{x} + (-a_2 k_x + a_1 k_y)\hat{y} + (b_1 k_x^3 + b_2 k_x^3)\hat{z}\) |
| \( C_4 \) | \((a_1 k_x + a_2 k_y)\hat{x} + (-a_2 k_x + a_1 k_y)\hat{y}\) |
| \( C_5 \) | \((a_1 k_x + a_2 k_y)\hat{x} + (-a_2 k_x + a_1 k_y)\hat{y}\) |
| \( D_1 \) | \(a_1 k_y \hat{x} + a_2 k_y \hat{y} + a_3 k_y \hat{z}\) |
| \( D_2 \) | \(a_1 k_y \hat{x} + a_2 k_y \hat{y}\) |
| \( D_3 \) | \(a_1 (k_3 \hat{x} - k_y \hat{y}) + ia_2 (k_3^3 - k_3^3)\hat{z}\) |
| \( D_4 \) | \(a_1 (k_y \hat{x} - k_x \hat{y})\) |
| \( D_5 \) | \(a_1 (k_y \hat{x} - k_x \hat{y})\) |

which distinguishes between the basal-plane and z-axis Zeeman couplings, but is still \( \mathbf{k} \)-independent.

In contrast, in the exceptional bands in trigonal and hexagonal crystals the effective Zeeman coupling is essentially anisotropic. In the \( \Gamma_6 \) bands for \( G = C_3 \) and in the \( (\Gamma_{11}, \Gamma_{12}) \) bands for \( G = C_6 \) we have

\[
\hat{\mu}(k) = 
\begin{pmatrix}
\beta_1 k_x^2 + \beta_1^* k_x^2 & \beta_2 k_x^2 + \beta_2^* k_x^2 & 0 \\
 i\beta_1 k_x^2 - i\beta_1^* k_x^2 & i\beta_2 k_x^2 - i\beta_2^* k_x^2 & 0 \\
 0 & 0 & \mu_z
\end{pmatrix},
\]

where \( \beta_{1,2} \) are complex constants and \( \mu_z \) is a real constant, while in the \( (\Gamma_5, \Gamma_6) \) bands for \( G = D_3 \) and in the \( \Gamma_9 \)
bands for $\mathcal{G} = D_6$ we have

$$
\hat{\mu}(k) = \begin{pmatrix}
\alpha_1(k_x^2 - k_y^2) & 2\alpha_2 k_x k_y & 0 \\
-2\alpha_1 k_x k_y & \alpha_2(k_x^2 - k_y^2) & 0 \\
0 & 0 & \mu_z
\end{pmatrix},
$$

(21)

where $\alpha_{1,2}$ and $\mu_z$ are real constants, see Appendix C for an example of the calculation. Note that the basal-plane components of $\hat{\mu}$ vanish at $k_x = k_y = 0$ in all four exceptional cases for the symmetry reasons. Similar phenomenon has also been found in 3D trigonal and hexagonal centrosymmetric materials, where the transverse components of the effective Zeeman coupling vanish along the main symmetry axis, which leads to qualitative changes in the temperature dependence of the spin susceptibility in the superconducting state.\textsuperscript{28} The vanishing of the Zeeman coupling perpendicular to the trigonal axis has also been reported in bismuth.\textsuperscript{29}

### C. Example: $\mathcal{G} = D_3$

This rosette group describes, for instance, the symmetry of a quasi-2D semiconducting quantum well grown along the [111] direction from a bulk semiconductor with diamond or zinc-blende structure,\textsuperscript{15} or the surface states of the topological insulator Bi$_2$Te$_3$ (Ref. 30). According to Table I, there are two types of bands in these materials: the pseudospin bands which correspond to the $\Gamma_4$ corep, and the nonpseudospin bands which correspond to the $(\Gamma_5, \Gamma_6)$ corep.

Assuming that the electron gas is placed in a magnetic field $\mathbf{H} \perp \hat{z}$, we obtain from Tables III and IV, and also Eqs. (19) and (21), the following expressions for the effective one-band Hamiltonians:

$$
\hat{\mathcal{H}} = \frac{k^2}{2m^*} + b(H_x k_y - H_y k_x) + (a_1 k_y + \mu_\perp H_x)\hat{\sigma}_1 + (-a_1 k_x + \mu_\perp H_y)\hat{\sigma}_2 + ia_2(k_+^3 - k_-^3)\hat{\sigma}_3
$$

(22)
in the $\Gamma_4$ bands, and

$$
\hat{\mathcal{H}} = \frac{k^2}{2m^*} + b(H_x k_y - H_y k_x) + ia_3(k_+^3 - k_-^3)\hat{\sigma}_3 + [ia_1(k_+^3 - k_-^3) + a_1 H_x(k_+^3 - k_-^3) - 2\alpha_1 H_y k_z k_y]\hat{\sigma}_1 + [a_2(k_+^3 + k_-^3) + 2\alpha_2 H_x k_x k_y + a_2 H_y(k_x^2 - k_y^2)]\hat{\sigma}_2
$$

(23)
in the $(\Gamma_5, \Gamma_6)$ bands.

The matrices (22) and (23) can be easily diagonalized producing two helicity bands $\xi_{\pm}$, see Eq. (8). Assuming $\mathbf{H} = H\hat{x}$, we have

$$
\xi_{\pm}(k, \mathbf{H}) = \frac{k^2}{2m^*} + b H k_y \pm \left[ a_1^2 k_x^2 + (a_1 k_y + \mu_\perp H)^2 + 4a_2^2(k_+^3 - 3k_2^2 k_y)^2 \right]^{1/2}
$$

(24)
in the $\Gamma_4$ case, and

$$
\xi_{\pm}(k, \mathbf{H}) = \frac{k^2}{2m^*} + b H k_y \pm \left\{ [2a_1^2(k_+^3 - 3k_2^2 k_y) + a_1 H(k_+^3 - k_-^3)^2] + [2a_2^2(k_+^3 - 3k_2^2 k_y) + 2\alpha_2 H k_x k_y]^2 + 4a_2^2(k_+^3 - 3k_2^2 k_y)^2 \right\}^{1/2}
$$

(25)
in the $(\Gamma_5, \Gamma_6)$ case. In the absence of magnetic field, the helicity bands have a sixfold rotation symmetry, but at $\mathbf{H} \neq 0$, this symmetry is lost as the bands get shifted and distorted, as shown in Fig. 1. One can see from Eqs. (24) and (25) that there is a qualitative difference between the magnetic response in the $\Gamma_4$ and $(\Gamma_5, \Gamma_6)$ bands. In the former case the band degeneracy is lifted by the field, whereas in the latter case the bands remain degenerate at the $\Gamma$ point even in a nonzero field, because of the vanishing of the effective Zeeman interaction at $k = 0$. The degeneracy may be lifted by the terms nonlinear in $\mathbf{H}$, which we neglected.

### V. TWO-BAND EFFECTIVE HAMILTONIAN

The symmetry analysis of the previous section can be straightforwardly extended to the multiband case. Let us consider two twofold degenerate bands corresponding to two $\Gamma$-point coreps, which can be the same or different. In
FIG. 1: The Fermi surfaces \( \xi_{\pm}(k, H) = \mu \) (\( \mu \) is the chemical potential) of the helicity bands \((24)\), for \( H = 0 \) (top panel) and for \( H \neq 0 \) (bottom panel). The Fermi surfaces in the \((\Gamma_5, \Gamma_6)\) case, see Eq. \((25)\), exhibit a similar behaviour, but the bands remain degenerate at \( k = 0 \) in the presence of \( H \parallel \hat{x} \).

the \( \Gamma\)-point basis \( \{|1, 1\}, |1, 2\}, |2, 1\}, |2, 2\} \), see Eq. \((4)\), the effective Hamiltonian is given by a \( 4 \times 4 \) matrix:

\[
\hat{\mathcal{H}}(k, H) = \begin{pmatrix}
\hat{\mathcal{H}}_{11}(k, H) & \hat{\mathcal{H}}_{12}(k, H) \\
\hat{\mathcal{H}}_{21}(k, H) & \hat{\mathcal{H}}_{22}(k, H)
\end{pmatrix},
\]

(26)

where \( \hat{\mathcal{H}}_{nn'} \) \( (n, n' = 1, 2) \) are \( 2 \times 2 \) matrices in the Kramers space satisfying the Hermiticity condition \( \hat{\mathcal{H}}_{nn'}(k, H) = \hat{\mathcal{H}}_{nn'}^\dagger(k, H) \). The invariance requirements under the rosette group and TR operations, Eqs. \((9)\) and \((10)\), hold in the two-band case as well, with \( \hat{\mathcal{D}} \) having the following block-diagonal form:

\[
\hat{\mathcal{D}}(g) = \begin{pmatrix}
\mathcal{D}_1(g) & 0 \\
0 & \mathcal{D}_2(g)
\end{pmatrix}, \quad \hat{\mathcal{D}}(K) = \begin{pmatrix}
-i\hat{\sigma}_2 & 0 \\
0 & -i\hat{\sigma}_2
\end{pmatrix},
\]

(27)

which corresponds to a four-dimensional reducible corep.

The Hamiltonian \((26)\) corresponds to two coupled generalized Rashba models. The intraband \( (n = n') \) and interband \( (n \neq n') \) components can be represented in the form similar to Eq. \((7)\):

\[
\hat{\mathcal{H}}_{nn}(k, H) = \mathcal{H}_{nn}(k, H)\hat{\sigma}_0 + \mathcal{\Lambda}_n(k, H)\hat{\sigma} \quad \hat{\mathcal{H}}_{12}(k, H) = \mathcal{\bar{\epsilon}}(k, H)\hat{\sigma}_0 + \mathcal{\bar{\Lambda}}(k, H)\hat{\sigma}, \quad \hat{\mathcal{H}}_{21} = \hat{\mathcal{H}}_{12}^\dagger.
\]

(28)

Here \( \mathcal{\varepsilon}_{1,2} \) and \( \mathcal{\Lambda}_{1,2} \) are real, but \( \mathcal{\bar{\varepsilon}} \) and \( \mathcal{\bar{\Lambda}} \) can be complex. The symmetry constraints for the intraband components have the same form as Eqs. \((11)\) and \((12)\), namely,

\[
\mathcal{\varepsilon}_n(k, H) = \mathcal{\varepsilon}_n(g^{-1}k, g^{-1}H), \quad \mathcal{\Lambda}_n(k, H) = \mathcal{\Lambda}_n(g^{-1}k, g^{-1}H),
\]

(29)

where the \( 3 \times 3 \) orthogonal matrix \( \mathcal{\hat{R}}_n \) is given by Eq. \((13)\) in the \( n \)th band, and

\[
\mathcal{\varepsilon}_n(k, H) = \mathcal{\varepsilon}_n(-k, -H), \quad \mathcal{\Lambda}_n(k, H) = -\mathcal{\Lambda}_n(-k, -H).
\]

(30)

For the interband components, the constraints become more complicated. From TR invariance, we have

\[
\mathcal{\bar{\varepsilon}}(k, H) = \mathcal{\bar{\varepsilon}}^*(-k, -H), \quad \mathcal{\bar{\Lambda}}_n(k, H) = -\mathcal{\bar{\Lambda}}_n^*(-k, -H),
\]

(31)
whereas the rosette group invariance requires that
\[ \tilde{\varepsilon}(\mathbf{k}, \mathbf{H}) \sigma_0 + \tilde{\Lambda}(\mathbf{k}, \mathbf{H}) \sigma = \varepsilon(g \mathbf{k}, g \mathbf{H}) [\hat{D}_{\tau^1} \tilde{\varepsilon}(g \mathbf{k}, g \mathbf{H})] + \tilde{\Lambda}(g \mathbf{k}, g \mathbf{H}) [\hat{D}_{\tau^1} \tilde{\varepsilon}(g \mathbf{k}, g \mathbf{H})]. \tag{32} \]

In particular, if both bands have the same symmetry, i.e., correspond to the same corep, then \( \hat{D}_1 = \hat{D}_2 \) and the last expression takes the same form as Eq. (29).

\[ \varepsilon_{\text{int}}(\mathbf{k}, \mathbf{H}) = \varepsilon_{\text{int}}(\mathbf{k}) + b_n(H_0 k_y - H_y k_x) + (a_n k_y + \mu_{\perp n} H_x) \sigma_1 + (-a_n k_x + \mu_{\perp n} H_y) \sigma_2, \tag{33} \]

for \( \mathbf{H} \perp \hat{z} \). We assume that
\[ \varepsilon_1(\mathbf{k}) = \frac{k^2}{2m_1}, \quad \varepsilon_2(\mathbf{k}) = \varepsilon_b + \frac{k^2}{2m_2}, \tag{34} \]

where \( \varepsilon_b > 0 \) is the band gap and the effective masses \( m_1 \) and \( m_2 \) can be positive or negative, corresponding to electron- or hole-like bands, respectively.

For the interband components, we have
\[ \tilde{\varepsilon}(\mathbf{k}, \mathbf{H}) = \tilde{\varepsilon}(\mathbf{k}) + \mathbf{H} \tilde{\mathbf{\ell}}(\mathbf{k}) + O(H^2), \]
\[ \tilde{\Lambda}_\nu(\mathbf{k}, \mathbf{H}) = \tilde{\gamma}_\nu(\mathbf{k}) + \sum_i H_i \tilde{\mu}_i(\mathbf{k}) + O(H^2), \tag{35} \]

where the \( \tilde{\gamma} \) term will be called the “interband SO coupling” (which is not quite accurate, because the SO coupling affects all terms in the effective Hamiltonian). It follows from Eq. (32) that \( \tilde{\varepsilon} \) is an invariant scalar and \( \tilde{\Lambda} \) is an invariant pseudovector only if both bands correspond to the same corep. In general, we have
\[ \tilde{\varepsilon}(\mathbf{k}) = \tilde{\varepsilon}^*(\mathbf{k}), \quad \tilde{\varepsilon}(\mathbf{k}) = \pm \tilde{\varepsilon}(C_{4z}^{-1} \mathbf{k}), \quad \tilde{\varepsilon}(\mathbf{k}) = \tilde{\varepsilon}(\sigma_y^{-1} \mathbf{k}), \tag{36} \]
\[ \tilde{\gamma}(\mathbf{k}) = -\tilde{\gamma}^*(\mathbf{k}), \quad \tilde{\gamma}(\mathbf{k}) = \pm C_{4z} \tilde{\gamma}(C_{4z}^{-1} \mathbf{k}), \quad \tilde{\gamma}(\mathbf{k}) = C_{2y} \tilde{\gamma}(\sigma_y^{-1} \mathbf{k}), \tag{37} \]
\[ \tilde{\mathbf{\ell}}(\mathbf{k}) = -\tilde{\mathbf{\ell}}^*(\mathbf{k}), \quad \tilde{\mathbf{\ell}}(\mathbf{k}) = \pm C_{4z} \tilde{\mathbf{\ell}}(C_{4z}^{-1} \mathbf{k}), \quad \tilde{\mathbf{\ell}}(\mathbf{k}) = C_{2y} \tilde{\mathbf{\ell}}(\sigma_y^{-1} \mathbf{k}), \tag{38} \]

and
\[ \tilde{\mu}(\mathbf{k}) = \tilde{\mu}^*(\mathbf{k}), \quad \tilde{\mu}(\mathbf{k}) = \pm \tilde{R}(C_{4z}) \tilde{\mu}(C_{4z}^{-1} \mathbf{k}) \tilde{R}^{-1}(C_{4z}), \quad \tilde{\mu}(\mathbf{k}) = \tilde{R}(C_{2y}) \tilde{\mu}(\sigma_y^{-1} \mathbf{k}) \tilde{R}^{-1}(C_{2y}), \tag{39} \]

instead of Eqs. (15), (16), (17), and (18), respectively. The upper (lower) sign refers to the bands having the same (different) symmetry. This sign difference, combined with the fact that the interband couplings are not necessarily real functions of \( \mathbf{k} \), produces a richer variety of solutions than in the intraband case.

For instance, one can see from Eq. (39) that the interband Zeeman coupling has a simple form
\[ \tilde{\mu}(\mathbf{k}) = \begin{pmatrix} \tilde{\mu}_{\perp} & 0 & 0 \\ 0 & \tilde{\mu}_{\perp} & 0 \\ 0 & 0 & \tilde{\mu}_z \end{pmatrix}, \tag{40} \]

with real \( \tilde{\mu}_{\perp} \) and \( \tilde{\mu}_z \), only if both bands have the same symmetry. In contrast, if the bands correspond to different coreps, then the interband Zeeman coupling is necessarily momentum-dependent and vanishes at \( |k_x| = |k_y| \):
\[ \tilde{\mu}(\mathbf{k}) = \begin{pmatrix} \tilde{\alpha}_{\perp} (k_x^2 - k_y^2) & 0 & 0 \\ 0 & \tilde{\alpha}_{\perp} (k_x^2 - k_y^2) & 0 \\ 0 & 0 & \tilde{\alpha}_z (k_x^2 - k_y^2) \end{pmatrix}, \tag{41} \]

Due to a large number of possibilities, here we consider only the case of the rosette group \( \mathcal{G} = D_4 \), which describes, for instance, the symmetry of a conducting interface between two insulating oxides LaAlO_3 and SrTiO_3 (Ref. 6). There are two types of bands, corresponding to \( \Gamma_6 \) or \( \Gamma_7 \) coreps. Using the corep matrices from Table II, we obtain that \( \mathcal{R}_n(C_{4z}) = \tilde{R}(C_{4z}) \) and \( \mathcal{R}_n(\sigma_y) = \tilde{R}(C_{2y}) \), therefore, according to Eq. (29), \( \varepsilon_n \) is an invariant scalar and \( \Lambda_n \) is an invariant pseudovector, for all two-band combinations. Expanding \( \varepsilon_n \) and \( \Lambda_n \) in powers of a weak magnetic field, similarly to Eq. (14), and using Tables III and IV, and also Eq. (19), we arrive at the following expressions for the intraband components of the effective Hamiltonian:
switch to a different basis, the topological properties of the system, in particular, the spectrum of the boundary modes, it is more convenient to use the coefficients $a_0(k^2_\parallel + k^2_\perp)$ instead of $a_1(k^0_\parallel k^1_\perp)$, so that the first index in $\tilde{E}$ is $\tilde{a}_1$ and one can label the bands in such a way that $\tilde{E}$ has the same eigenvalues and does not have a constant term, even though the constraints (36) permit a real $k$-independent solution in the case of same-symmetry bands. The reason is that such terms cannot appear in the framework of the $k \cdot p$ perturbation theory, see Appendix B.

### VI. MODEL OF A TOPOLOGICAL INSULATOR

Let us show how a two-band model introduced in Sec. VA can describe a TR invariant topological insulator. At $H = 0$, the effective Hamiltonian (26) becomes

$$\tilde{H}(k) = \begin{pmatrix} \epsilon_1(k)\sigma_0 + \gamma_1(k)\sigma & \tilde{c}(k)\sigma_0 + \tilde{\gamma}(k)\sigma \\ \tilde{c}(k)\sigma_0 + \tilde{\gamma}(k)\sigma & \epsilon_2(k)\sigma_0 + \gamma_2(k)\sigma \end{pmatrix}. \tag{42}$$

where $\epsilon_{1,2}$ are given by Eq. (34), the intraband SO coupling has the form $\gamma_{1,2}(k) = a_{1,2}(k_y e_1 - k_x e_2)$, see Table III, and the interband couplings depend on the band symmetries and can be found in Table V. In Appendix E, we discuss a different (“Dirac”) representation of the two-band Hamiltonian using the $4 \times 4$ matrices.

The bulk spectrum of the model (42) consists of four bands, which are nondegenerate at a general $k$, but remain twofold degenerate at $k = 0$ (and other TR invariant momenta) for symmetry reasons, see Sec. III. Due to the TR symmetry, the band energies are even in $k$. Indeed, if $\Psi$ is an eigenvector of the matrix (42) corresponding to an eigenvalue $E(k)$, then its TR partner $\Psi = \mathcal{D}(k)\Psi^*$, with $\mathcal{D}(k)$ given by Eq. (27), is an eigenvector of $\tilde{H}(k)$, because $\tilde{H}(-k)\Psi = \mathcal{D}(k)\tilde{H}^*(k)\Psi^* = E(k)\Psi$ [here we used Eq. (10)]. Thus, $\tilde{H}(k)$ and $\tilde{H}(-k)$ have the same eigenvalues and one can label the bands in such a way that $E_i(k) = E_i(-k)$, where $i = 1, 2, 3, 4$, with $E_1(0) = E_2(0)$ and $E_3(0) = E_4(0)$. Explicit analytical solutions for the band energies and the wave functions can only be obtained after some simplifying assumptions, see Sec. VI.A below.

The Hamiltonian has the form (42) in the $\Gamma$-point basis

$$\{ |1, 1 \rangle, |1, 2 \rangle, |2, 1 \rangle, |2, 2 \rangle \}, \tag{43}$$

where the first index in $|n, s \rangle$ labels the bands, whereas the second one is the Kramers index. To investigate the topological properties of the system, in particular, the spectrum of the boundary modes, it is more convenient to switch to a different basis,

$$\{ |1, 1 \rangle, |2, 2 \rangle, |1, 2 \rangle, |2, 1 \rangle \}, \tag{44}$$

in which the Hamiltonian becomes

$$\hat{H} = \frac{\xi}{2} + \frac{k^2}{2M} + \hat{H}_1 + \hat{H}_2, \tag{45}$$

where

$$\hat{H}_1(k) = \begin{pmatrix} \hat{h}(k) & 0 \\ 0 & \hat{h}(k) \end{pmatrix}, \quad \hat{H}_2(k) = \begin{pmatrix} 0 & \hat{m}(k) \\ \hat{m}^\dagger(k) & 0 \end{pmatrix}. \tag{46}$$

The notations are as follows:

$$\hat{h}(k) = \begin{pmatrix} \xi & \tilde{\gamma}_1 - i \tilde{\gamma}_2 \\ \tilde{\gamma}_1 + i \tilde{\gamma}_2 & -\xi \end{pmatrix}, \quad \hat{m}(k) = \begin{pmatrix} \xi & \tilde{\gamma}_1 + i \tilde{\gamma}_2 \\ \tilde{\gamma}_1 - i \tilde{\gamma}_2 & -\xi \end{pmatrix}. \tag{47}$$

### TABLE V: The interband couplings of the two-band effective Hamiltonian near the $\Gamma$ point, for $G = D_4$: $\tilde{a}_i$ are real constants.

| bands | $\tilde{\epsilon}(k)$ | $\tilde{\gamma}(k)$ | $\tilde{\epsilon}(k)$ |
|-------|----------------------|----------------------|----------------------|
| $\Gamma_6, \Gamma_6$ | $\tilde{a}_0(k^2_\parallel + k^2_\perp)$ | $\tilde{a}_1(k_\parallel e_1 - k_\perp e_2) + i \tilde{a}_2(k^2_\parallel - k^2_\perp)k_\parallel k_\perp e_3$ | $\tilde{a}_3(k_\parallel k_\perp e_1 - k_\perp k_\parallel e_2)$ |
| $\Gamma_7, \Gamma_7$ | $\tilde{a}_0(k^2_\perp - k^2_\parallel)$ | $\tilde{a}_1(k_\parallel e_1 + k_\perp e_2) + i \tilde{a}_2k_\parallel k_\perp e_3$ | $\tilde{a}_3(k_\parallel k_\perp e_1 + k_\perp k_\parallel e_2) + i \tilde{a}_4k_\parallel k_\perp e_3$|

where $\tilde{a}_\parallel$ and $\tilde{a}_\perp$ are real constants. Representative expressions for the other interband couplings satisfying the constraints (36), (37), and (38) are given in Table V, see Appendix D for an example of the calculation. Note that, unlike the intraband SO coupling, the interband SO coupling is a complex function of $k$ and does not have a definite parity. Also note that $\tilde{\epsilon}(k)$ cannot contain a constant term, even though the constraints (36) permit a real $k$-independent solution in the case of same-symmetry bands. The reason is that such terms cannot appear in the framework of the $k \cdot p$ perturbation theory, see Appendix B.
\[ \hat{m}(\mathbf{k}) = \begin{pmatrix} \gamma_{1,1} - i\gamma_{1,2} & \xi + \tilde{\gamma}_3 \\ \xi + \tilde{\gamma}_3 & \gamma_{2,1} + i\gamma_{2,2} \end{pmatrix}, \] (48)

and

\[ \xi = \frac{k^2}{2m^*} - \frac{E_b}{2}, \quad \frac{1}{M} = \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \quad \frac{1}{m^*} = \frac{1}{2} \left( \frac{1}{m_1} - \frac{1}{m_2} \right). \]

We used the fact that \( \tilde{\gamma} \) and \( \gamma_{1,2} \) are real, whereas \( \tilde{\gamma}_3 \) is imaginary, see Table V. In the basis (44), the matrix representation of TR is given by \( \hat{D}(K) = -i\hat{s}_2 \otimes \hat{s}_3 \), where \( \hat{s} \) are the Pauli matrices, and we obtain from Eq. (10) that

\[ \hat{h}(\mathbf{k}) = \hat{s}_3 \hat{h}^\dagger(-\mathbf{k}) \hat{s}_3, \] (49)

while \( \hat{m} \) satisfies

\[ \hat{m}(\mathbf{k}) = -\hat{s}_3 \hat{m}^\dagger(-\mathbf{k}) \hat{s}_3. \] (50)

Thus, the \( 2 \times 2 \) Hamiltonians \( \hat{h} \) and \( \hat{m} \) are TR partners.

The advantage of the representation (45) is that \( \hat{H}_1 \) has a block-diagonal structure which can be connected with some well-known models of topological superconductors and insulators. The matrices \( \hat{h} \) and \( \hat{m} \) have the same form as the Bogoliubov-de Gennes Hamiltonian of a chiral p-wave superconductor (for one spin projection), with \( \mathcal{E}_b/2 \) playing the role of the chemical potential. This system is known to have topologically nontrivial bulk states with gapless boundary modes. \( \mathcal{E}_b \) is in the limit \( m^* \to \infty \), \( \hat{H}_1 \) takes the form of a massive Dirac Hamiltonian with the mass given by \( \mathcal{E}_b/2 \), and if the band gap changes sign as a function of position, then one arrives at a 2D version of the Volkov-Pankratov model of a topological insulator. If \( m^* \) is finite, then \( \hat{H}_1 \) becomes a “modified” Dirac Hamiltonian, see Appendix F, and the first three terms in Eq. (45) are similar to the Bernevig-Hughes-Zhang model. \( ^{17} \)

**A. Reduced model**

Based on the analogy with the models of topological superconductors and insulators, one can expect that our Hamiltonian (45) may exhibit nontrivial topology in the bulk, with gapless modes at the boundary. To develop an analytically treatable model, we make some simplifying assumptions. According to Appendix F, \( \hat{H}_1 \) describes a topologically nontrivial system only when \( m^* > 0 \). This condition is satisfied if \( m_1 > 0 \) and \( m_2 < 0 \), which corresponds to a “band inversion” situation. Furthermore, one can put \( m_1 = -m_2 = m^* \) and \( \gamma_1(\mathbf{k}) = \gamma_2(\mathbf{k}) = a(k_y \mathbf{e}_1 - k_x \mathbf{e}_2) \), without affecting the topological properties of the system.

We keep the terms of the lowest (linear) order in \( \mathbf{k} \) in \( \hat{H}_2 \) and also assume that both bands have the same symmetry, therefore \( \gamma(\mathbf{k}) = a(k_y \mathbf{e}_1 - k_x \mathbf{e}_2) \), see Table V. In this way, dropping the constant \( \mathcal{E}_b/2 \), we obtain the following model Hamiltonian in the basis (44):

\[ \hat{\mathcal{H}} = \hat{H}_1 + \hat{H}_2, \quad \hat{H}_1 = \begin{pmatrix} \hat{h} & 0 \\ 0 & \hat{\bar{h}} \end{pmatrix}, \quad \hat{H}_2 = \begin{pmatrix} 0 & \hat{m} \\ \hat{m}^\dagger & 0 \end{pmatrix}, \] (51)

where

\[ \hat{h} = \xi \hat{s}_3 + a(k_y \hat{s}_1 - k_x \hat{s}_2), \quad \hat{\bar{h}} = \xi \hat{s}_3 + \bar{a}(k_y \hat{s}_1 + k_x \hat{s}_2), \quad \hat{m} = a(k_y \hat{s}_0 + ik_x \hat{s}_3), \]

and \( \xi = k^2/2m^* - \mathcal{E}_b/2 \). Without loss of generality, we assume that \( a, \bar{a} > 0 \). The structure of Eq. (51) suggests that, while the interband SO coupling enters the chiral Hamiltonians \( \hat{h} \) and \( \hat{\bar{h}} \) and gives a “topological character” to the system, the effects of the intraband SO coupling, which are contained in \( \hat{H}_2 \), can be treated as a perturbation.

**B. Bulk properties; \( \mathbb{Z}_2 \) invariant**

Let us start with the bulk properties of our model. To calculate the bulk spectrum, it is more convenient to return to the basis (43), in which the Hamiltonian takes the form

\[ \hat{\mathcal{H}} = \begin{pmatrix} \xi & 0 \\ 0 & -\xi \end{pmatrix} \otimes \hat{s}_0 + \begin{pmatrix} a & \bar{a} \\ \bar{a} & a \end{pmatrix} \otimes \hat{q}, \] (52)
where \( \hat{q} = k_y \hat{\sigma}_1 - k_x \hat{\sigma}_2 \). The 2×2 Kramers blocks in Eq. (52) commute and the Hamiltonian can be easily diagonalized. Introducing the eigenstates of \( \hat{q} \):

\[
\chi_+(k) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -ie^{-i\varphi_k} \end{pmatrix},
\chi_-(k) = \frac{1}{\sqrt{2}} \begin{pmatrix} -ie^{i\varphi_k} \\ 1 \end{pmatrix},
\]

with \( \varphi_k = \arg(k_x + ik_y) \), which correspond to the eigenvalues \( \pm k = \pm \sqrt{k_x^2 + k_y^2} \), we obtain the following eigenstates of the Hamiltonian (52):

\[
\Psi_1 = \begin{pmatrix} w_- \\ w_+ \end{pmatrix} \otimes \chi_-, \quad \Psi_2 = \begin{pmatrix} w_- \\ -w_+ \end{pmatrix} \otimes \chi_+, \quad \Psi_3 = \begin{pmatrix} w_+ \\ -w_- \end{pmatrix} \otimes \chi_-, \quad \Psi_4 = \begin{pmatrix} w_+ \\ w_- \end{pmatrix} \otimes \chi_+,
\]

where

\[
w_\pm = \frac{1}{\sqrt{2}} \sqrt{1 \pm \frac{\xi}{\sqrt{\xi^2 + \tilde{a}^2 k^2}}}.
\]

The corresponding eigenvalues form four isotropic bands:

\[
E_1(k) = -E_+, \quad E_2(k) = -E_-, \quad E_3(k) = E_-, \quad E_4(k) = E_+,
\]

where

\[
E_\pm(k) = \sqrt{\xi^2(k) + \tilde{a}^2 k^2} \pm ak.
\]

The spectrum qualitatively depends on the relative strength of the intraband and interband SO coupling. If \( a < \tilde{a} \), then the bulk spectrum is gapped, but if \( a > \tilde{a} \), then the bulk gap closes, see Fig. 2. We assume that the interband SO coupling is stronger than the intraband one, so that the model describes an insulator with the bands 1 and 2 occupied. The twofold degeneracy of the bands is lifted by the intraband SO coupling everywhere, except the TR invariant point \( k = 0 \).

The eigenstates (54) do not depend on \( a \), therefore one can expect that the bulk topology is independent of the strength of the intraband SO coupling (as long as the bulk spectrum remains gapped). The Hamiltonian given by Eqs. (52) or (51) corresponds to the symplectic symmetry class AII and is characterized in 2D by a \( \mathbb{Z}_2 \) topological invariant.\(^{34}\) Among the several equivalent representations of this invariant,\(^{35}\) it is the Pfaffian form proposed in Ref. 36 that is the easiest to calculate for our model. We introduce the “overlap” matrix

\[
M_{nm}(k) = \langle \Psi_n(k) | K | \Psi_m(k) \rangle,
\]
where \( n \) and \( m \) label the occupied bands. The overlap matrix is not necessarily unitary, but is antisymmetric at all \( \mathbf{k} \) due to the condition \( K^2 = -1 \). In our case, there are two occupied bands, the TR acts in the basis (43) as follows: 
\[
K \Psi = 1 \otimes (-i \hat{s}_2) \Psi^*,
\]
and, using the the fact that \( K \chi_+(\mathbf{k}) = \chi_- (\mathbf{k}) \) and \( K \chi_- (\mathbf{k}) = -\chi_+ (\mathbf{k}) \), we obtain from Eq. (54) the following expression for the overlap matrix:
\[
\hat{M}(\mathbf{k}) = \frac{\xi}{\sqrt{\xi^2 + \hat{a}^2 k^2}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.
\]

The Pfaffian of this matrix is given by
\[
\text{Pf} \hat{M}(\mathbf{k}) = -\frac{\xi(\mathbf{k})}{\sqrt{\xi^2(\mathbf{k}) + \hat{a}^2 k^2}}.
\]

According to Ref. 36, the \( Z_2 \) invariant \( \nu \) is equal to one half of the number of sign changes of Pf \( \hat{M} \) along the boundary of a half of the Brillouin zone (HBZ), the latter being defined in such a way that it does not contain time-reversed momenta \( \mathbf{k} \) and \(-\mathbf{k} \) at the same time.

In our model, the Brillouin zone extends to the whole momentum plane and one can choose, for instance, the \( k_x = 0 \) line as the HBZ boundary. The zeros of the Pfaffian (56) are the same as those of \( \xi(\mathbf{k}) \), namely it vanishes on a circle of radius \( m^* \mathcal{E}_b \), but only if \( m^* \) and \( \mathcal{E}_b \) have the same sign. Thus, we arrive at the following result:
\[
\nu = \begin{cases} 
1, & \text{if } m^* \mathcal{E}_b > 0, \\
0, & \text{if } m^* \mathcal{E}_b < 0.
\end{cases}
\]

In the atomic limit, we have \( \xi(\mathbf{k}) = \text{const} \) and the Pfaffian does not change sign, therefore, \( \nu = 0 \) corresponds to a topologically trivial insulator. In the topologically nontrivial case \( \nu \neq 0 \), the invariant (57) counts the number of Kramers-degenerate pairs of the boundary zero modes, see the next subsection. Note that \( a \) and \( \hat{a} \) do not explicitly enter Eq. (57), therefore one can neglect the intraband SO coupling \( a \) without affecting the topological properties.

C. Boundary modes

Let us now put the model (51) in a half-space \( x \geq 0 \) and assume that the wave functions vanish at \( x = 0 \). The momentum along the boundary is a good quantum number and, making the replacement \( k_x \rightarrow \hat{k}_x = -i \nabla_x \) in Eq. (51), we see that \( \hat{h} \) is the same as \( \hat{H}_+ \) from Appendix F 2, while \( \hat{h} \) is the same as \( \hat{H}_- \), with the parameters given by \( \alpha = 1/2 m^* \), \( \beta = -\mathcal{E}_b/2 \), and \( \gamma = \hat{a} \). Without the intraband SO coupling, \( \hat{h} \) and \( \hat{h} \) decouple and the Hamiltonian (51) becomes a modified Dirac Hamiltonian with \( m^* \mathcal{E}_b > 0 \), which has two counterpropagating linear boundary modes given by
\[
E(k_y) = -\hat{a} k_y, \quad \hat{E}(k_y) = \hat{a} k_y,
\]
see Eqs. (F13) and (F14). At \( k_y = 0 \), the two branches cross and form a Kramers pair of the boundary zero modes, whose wave functions have the following form:
\[
\Psi_0(\mathbf{r}) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} F(x), \quad \Psi_0(\mathbf{r}) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} F(x). \quad (59)
\]

Here \( F(x) \) is localized near the boundary, vanishes at \( x = 0 \), and can be chosen to be real, see Eqs. (F15) and (F16). If one takes into account the intraband SO coupling, then it is easy to check that all matrix elements of \( \hat{H}_2(k_y = 0) = i \hat{s}_2 \otimes \hat{s}_3 (a \nabla_x) \) in the subspace spanned by the states (59) vanish. Therefore, the TR-degenerate boundary zero modes survive at \( a \neq 0 \), which is consistent with the observation that the variation of \( a \) does not affect the bulk topological invariant (57).

The boundary zero modes are only destroyed by a TR symmetry-breaking perturbation, for instance, by an external magnetic field. Assuming that the field is parallel to the \( xy \) plane, so that one does not have to worry about the
orbital effects, and keeping the terms of the zeroth order in \( k \) in the intraband [Eq. (14)] and interband [Eq. (35)] magnetic couplings, we obtain an additional term in the Hamiltonian (51):

\[
\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{H}_3,
\]

where

\[
\hat{H}_3 = \begin{pmatrix}
0 & \tilde{\mu}_\perp H_- & \mu_\perp H_- & 0 \\
\tilde{\mu}_\perp H_+ & 0 & 0 & \mu_\perp H_+ \\
\mu_\perp H_- & 0 & 0 & \tilde{\mu}_\perp H_+ \\
0 & \mu_\perp H_+ & \tilde{\mu}_\perp H_- & 0
\end{pmatrix}
\]

(60)

describes the effective Zeeman coupling in the basis (44). Here \( H_\pm = H_x \pm iH_y \) and we used Eqs. (19) and (40). The matrix elements of Eq. (60) in the zero-mode subspace have the following form:

\[
\langle \Psi_0 | \hat{H}_3 | \Psi_0 \rangle = -\mu_\perp H_x, \quad \langle \Psi_0 | \hat{H}_3 | \bar{\Psi}_0 \rangle = i\mu_\perp H_y, \quad \langle \bar{\Psi}_0 | \hat{H}_3 | \Psi_0 \rangle = i\mu_\perp H_y, \quad \langle \bar{\Psi}_0 | \hat{H}_3 | \bar{\Psi}_0 \rangle = \tilde{\mu}_\perp H_x.
\]

Therefore, in the presence of a magnetic field the boundary modes are pushed apart and away from the zero energy, with the energy splitting given by

\[
\delta E(k_y = 0) = 2\sqrt{\mu_\perp^2 H_x^2 + \mu_\perp^2 H_y^2}.
\]

Note that the field normal to the boundary affects the zero modes only via the interband Zeeman coupling.

VII. CONCLUSIONS

We derived the phenomenological effective Hamiltonians for 2D electrons in the presence of the SO coupling and a magnetic field, by using the method of invariants. The SO coupling of an arbitrary strength is taken into account exactly, by considering the double-valued corepresentations of the magnetic rosette groups. Focusing on the vicinity of the \( \Gamma \) point, the effective Hamiltonian is represented by an expansion in powers of the electron momentum \( k \) and the magnetic field \( H \). In the one-band case, we developed a complete classification of the effective Hamiltonians for all 2D crystal symmetries. In the two-band case, due to a large number of possible band combinations, we focused on the example of the rosette group \( D_4 \).

We obtained various terms in the effective Hamiltonians representing the intraband and interband antisymmetric SO coupling and the magnetic field interactions. While in some cases these terms reproduce the structure of the standard Rashba model, in others, most notably in trigonal and hexagonal crystals, they are significantly different. In the exceptional nonpseudospin bands, namely, \( \Gamma_6 \) for \( G = C_3, (\Gamma_{11}, \Gamma_{12}) \) for \( G = C_6, (\Gamma_5, \Gamma_6) \) for \( G = D_3 \), and \( \Gamma_9 \) for \( D_6 \), all components of the intraband SO coupling are cubic in \( k \), whereas the linear terms are absent for the symmetry reasons, and the in-plane components of the effective Zeeman coupling vanish at the \( \Gamma \) point.

As an application of the two-band effective Hamiltonian for \( G = D_4 \), we introduced a simple model of a TR-invariant topological insulator. In this model, the intraband SO coupling mixes two \( 2 \times 2 \) TR-partnered chiral Hamiltonians, whose topological character originates from the interband SO coupling. We calculated the spectrum of the boundary modes, which form Kramers pairs protected by a \( Z_2 \) invariant in the bulk.

Acknowledgments

This work was supported by a Discovery Grant 2021-03705 from the Natural Sciences and Engineering Research Council of Canada.

Appendix A: The \( \Gamma \)-point basis

Since all double-valued coreps of the rosette groups are 2D, the Bloch states at the \( \Gamma \)-point are twofold degenerate. Dropping the band index, these states have the following form:

\[
|1\rangle = \begin{pmatrix} \psi_\uparrow(r) \\ \psi_\downarrow(r) \end{pmatrix}, \quad |2\rangle = K|1\rangle = \begin{pmatrix} -\psi_\uparrow^*(r) \\ \psi_\downarrow^*(r) \end{pmatrix},
\]

(A1)
where \( \psi_\uparrow(\mathbf{r}) \) and \( \psi_\downarrow(\mathbf{r}) \) are lattice-periodic functions, which include the effects of the noncentrosymmetric crystal field and the SO coupling. Under a symmetry operation \( g \), the Bloch states transform as \( g|s\rangle = \sum_{s'}|s'\rangle\mathcal{D}_{ss'}(g) \), where \( \mathcal{D} \) is the corep matrix. Recall that the action of \( g \) on a spinor wave function is given by \( g \left( \begin{array}{c} \psi_\uparrow(\mathbf{r}) \\ \psi_\downarrow(\mathbf{r}) \end{array} \right) = \mathcal{D}^{(1/2)}(g) \left( \begin{array}{c} \psi_\uparrow(g^{-1}\mathbf{r}) \\ \psi_\downarrow(g^{-1}\mathbf{r}) \end{array} \right) \),

where \( \mathcal{D}^{(1/2)}(g) \) is the spin-1/2 representation of \( g \).

Explicit form of the corep matrices can be obtained using the procedure of Ref. 19, with the only modification that the unitary element of the symmetry group in our case is the TR operation \( K_I \), instead of the conjugation operation \( K_I \). The pseudospin bands, the corep matrices are given by \( \mathcal{D}^{(1/2)}(g) \). The results for the nonpseudospin bands are shown in Table II. Note that the corep matrices depend on the choice of the basis: a unitary rotation of the basis, \( |s\rangle \rightarrow |s'\rangle = \sum_{s} |s\rangle U_{ss'} \), produces an equivalent corep with \( \mathcal{D}'(g) = U^{-1}\mathcal{D}(g)U \) and \( \mathcal{D}'(K) = U^{-1}\mathcal{D}(K)U^* \). The “orientation” and the phases of the Bloch basis at the \( \Gamma \)-point can always be chosen to reproduce both the matrices in Table II and the matrix representation of the TR operation, \( \mathcal{D}(K) = -i\hat{\sigma}_2 \).

Let us consider, for example, a trigonal crystal with \( \hat{G} = D_3 \). This group is generated by the rotation \( C_{3z} \) and the reflection \( \sigma_y \) and has three double-valued irreps: \( \Gamma_4 \), which is 2D and is equivalent to the spin-1/2 irrep, and also \( \Gamma_5 \) and \( \Gamma_6 \), which are 1D. The irreps \( \Gamma_5 \) and \( \Gamma_6 \) are complex conjugate to each other and pair up to form a single “physically irreducible” 2D representation of \( D_3 \), or a 2D Case C corep of the magnetic point group \( D_3 + K D_3 \), which is given by the following matrices:

\[
\mathcal{D}(C_{3z}) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathcal{D}(\sigma_y) = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \quad \mathcal{D}(K) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},
\]

where we used the group character tables for \( D_3 \), see, e.g., Refs. 22 and 24. We see that the \( (\Gamma_5, \Gamma_6) \) corep is not equivalent to the spin-1/2 representation. It can be brought by a unitary transformation \( \hat{U} = \exp(-i\pi\hat{\sigma}_1/4) \) to an equivalent form

\[
\mathcal{D}(C_{3z}) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathcal{D}(\sigma_y) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{D}(K) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},
\]

which is the one listed in Table II.

**Appendix B: Effective Hamiltonians from the \( k \cdot p \) perturbation theory**

In this appendix, we outline how the phenomenological effective Hamiltonian discussed in Sec. IV can be obtained, at least in principle, from a microscopic theory. The eigenfunction of Eq. (2) which corresponds to the wave vector \( \mathbf{k} \) has the form \( \psi_\mathbf{k}(\mathbf{r}) = \mathcal{V}^{-1/2}\varphi_\mathbf{k}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} \), where \( \mathcal{V} \) is the system volume. The Bloch factors \( \varphi_\mathbf{k} \) are spin-1/2 spinors with the same periodicity as the crystal lattice, which are eigenstates of the reduced Hamiltonian

\[
\hat{H}_\mathbf{k} = \frac{(\hat{p} + \hbar \mathbf{k})^2}{2m} + U(\mathbf{r}) + \frac{\hbar}{4mc^2}\hat{\sigma}[\nabla U(\mathbf{r}) \times (\hat{p} + \hbar \mathbf{k})] + \mu_B \mathbf{H} \hat{\sigma}
\]

\[
= \hat{H}_0 + \frac{\hbar^2 k^2}{2m} + \delta \hat{H}, \tag{B1}
\]

where \( \hat{H}_0 \) corresponds to \( \mathbf{k} = \mathbf{0} \) and has the same form as Eq. (2),

\[
\delta \hat{H} = \hbar k \hat{\mathbf{v}} + \mu_B \mathbf{H} \hat{\sigma}, \tag{B2}
\]

and \( \hat{\mathbf{v}} = \hat{p}/m + (\hbar/4mc^2)(\hat{\sigma} \times \nabla U) \) is the velocity operator in the presence of the SO coupling. Note that our unperturbed Hamiltonian \( \hat{H}_0 \) includes the electron-lattice SO coupling, which is assumed to be stronger than the Zeeman interaction \( \mu_B \mathbf{H} \hat{\sigma} \). The latter is included in the perturbation \( \delta \hat{H} \) along with the usual \( \mathbf{k} \cdot \mathbf{p} \) term. The orbital effects of the magnetic field are neglected.

In the absence of magnetic field, the energy levels of \( \hat{H}_0 \) are twofold degenerate due to the TR symmetry. The spinors \( \varphi_{\mathbf{k}=0}(\mathbf{r}) \equiv \langle \mathbf{r}|n, s \rangle \) are labelled by the band index \( n \) and the Kramers index \( s = 1, 2 \), and have the form (A1). They transform according to the irreducible double-valued coreps of the magnetic rosette group (3). The corresponding eigenvalues of \( \hat{H}_0 \) are denoted by \( \epsilon_n \).
If the complete set of \(|n, s\) is known, then an exact description of the band structure at \(k \neq 0\) is provided by diagonalizing the infinite-dimensional matrix \((n, s|\hat{H}_k|n', s')\). Its elements satisfy certain constraints imposed by the TR symmetry. It follows from \(K\hat{v}K^{-1} = -\hat{v}\) and \(K\hat{\sigma}K^{-1} = -\hat{\sigma}\) that \(K\hat{\delta}\hat{H}K^{-1} = -\hat{\delta}\hat{H}\). Therefore, using the antiunitarity of the TR operator, \(\langle K|i\rangle = \langle i|K^\dagger\rangle\ast\), and the Hermiticity of \(\delta\hat{H}\), we obtain

\[
\langle n, 2|\delta\hat{H}|n', 2\rangle = \langle n, 1|K^{-1}\delta\hat{H}K|n', 1\rangle\ast = -\langle n, 1|\delta\hat{H}|n', 1\rangle = -\langle n', 1|\delta\hat{H}|n, 1\rangle. \tag{B3}
\]

Similarly, using also the fact that \(K^2 = -1\), we have

\[
\langle n, 1|\delta\hat{H}|n', 2\rangle = -\langle n, 1|(K^{-1}\delta\hat{H}K)\hat{K}|n', 1\rangle = \langle n, 1|K^\dagger\delta\hat{H}|n', 1\rangle = \langle n', 1|\delta\hat{H}K|n, 1\rangle = \langle n', 1|\delta\hat{H}|n, 1\rangle. \tag{B4}
\]

and

\[
\langle n, 2|\delta\hat{H}|n', 1\rangle = \langle n', 2|\delta\hat{H}|n, 1\rangle. \tag{B5}
\]

The above relations are valid for any \(n\) and \(n'\). While the diagonal matrix elements \(\langle n, s|\delta\hat{H}|n, s\rangle\) are real, the off-diagonal ones are complex, in general.

The constraints (B3), (B4), and (B5) impose the following matrix structure on the \(k \cdot p\) Hamiltonian:

\[
\langle n, s|\delta\hat{H}|n', s'\rangle = iQ_{nn'}\delta_{ss'} + R_{nn'}\sigma_{ss'}, \tag{B6}
\]

where \(Q_{nn'}\) and \(R_{nn'}\) are real quantities satisfying \(Q_{nn'} = -Q_{n'n}\) and \(R_{nn'} = R_{n'n}\). Explicitly,

\[
Q_{nn'} = \text{Im} \langle n, 1|\delta\hat{H}|n', 1\rangle, \\
R_{nn'} = \text{Re} \langle n, 1|\delta\hat{H}|n', 2\rangle - \text{Im} \langle n, 1|\delta\hat{H}|n', 2\rangle, \text{Re} \langle n, 1|\delta\hat{H}|n', 1\rangle.
\]

Note that \(Q_{nn'}\) and \(R_{nn'}\) are linear in both \(k\) and \(H\) and, since the states \(|n, s\) do not have definite parity, \(\langle n, s|\psi|n, s'\rangle \neq 0\), in contrast to the centrosymmetric case. For illustration, focusing on just two bands, \(n, n' = 1, 2\), we obtain from Eq. (B6):

\[
\hat{H}_k = \begin{pmatrix}
\epsilon_1 + \frac{\hbar^2k^2}{2m} + r_1\hat{\sigma} & i\hat{q} + \hat{r}\hat{\sigma} & \cdots \\
-i\hat{q} + \hat{r}\hat{\sigma} & \epsilon_2 + \frac{\hbar^2k^2}{2m} + r_2\hat{\sigma} & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix}, \tag{B7}
\]

where \(r_1 = R_{11}, r_2 = R_{22}, \hat{q} = Q_{12},\) and \(\hat{r} = R_{12}\). We use the same notation, \(\hat{H}_k\), for the reduced Hamiltonian (B1) and its matrix representation in the \(\Gamma\)-point basis \(\{|1, 1\rangle, |1, 2\rangle, |2, 1\rangle, |2, 2\rangle\}\).

The effective one-band Hamiltonian can be obtained by eliminating the interband matrix elements \(\langle n, s|\hat{H}_k|n', s'\rangle\) \((n \neq n')\) in any desired order by a unitary transformation. This procedure, which is more generally applicable to the matrix elements connecting different groups of degenerate, or quasi-degenerate, states in an arbitrary Hamiltonian,\textsuperscript{15,37} is variously known in the literature as the Luttinger-Kohn\textsuperscript{38} or Schrieffer-Wolff\textsuperscript{39} transformation and can be traced back to the Foldy-Wouthuysen transformation in quantum electrodynamics.\textsuperscript{40} Let us drop the wave vector subscript and consider a matrix

\[
\hat{H} = \hat{H}_0 + \frac{\hbar^2k^2}{2m} + \zeta\hat{\delta}\hat{H}, \quad (0 \leq \zeta \leq 1), \tag{B8}
\]

which generalizes Eq. (B1). Here \(\hat{H}_0\) is a diagonal matrix with \(\langle n, s|\hat{H}_0|n', s'\rangle = \epsilon_{n}\delta_{nn'}\delta_{ss'}\) and \(\hat{\delta}\hat{H} = \hat{\delta}\hat{H}' + \hat{\delta}\hat{H}''\), with the matrices \(\hat{\delta}\hat{H}'\) and \(\hat{\delta}\hat{H}''\) having only intraband and interband matrix elements, respectively. For instance, keeping just two bands in Eq. (B7), we have

\[
\hat{\delta}\hat{H}' = \begin{pmatrix}
\hat{h}_1 & 0 \\
0 & \hat{h}_2
\end{pmatrix}, \quad \hat{\delta}\hat{H}'' = \begin{pmatrix} 0 & \hat{h}' \\
\hat{h} & 0 \end{pmatrix}, \tag{B9}
\]

where \(\hat{h}_n = r_n\hat{\sigma}\) and \(\hat{\delta} = i\hat{q} + \hat{r}\hat{\sigma}\). To develop a perturbative expansion in \(\delta\hat{H}\), we introduced a bookkeeping parameter \(\zeta\), which will be set to 1 at the end of the calculation. We assume that the eigenvalues of \(\hat{H}_0\) are well separated from each other, so that the energy splittings between the twofold degenerate bands are much larger than the matrix elements of \(\delta\hat{H}'\) and \(\delta\hat{H}''\).
We shall now try to remove the interband matrix elements from the Hamiltonian (B8) by a unitary transformation \( \hat{H} \rightarrow \hat{H}_U = \hat{U} \hat{H} \hat{U}^{-1} \), where \( \hat{U} = \exp(i\hat{S}) \) and \( \hat{S} \) is a Hermitian matrix, which has the same block structure as \( \delta \hat{H}'' \), i.e., no intraband elements. Since \( \hat{S} \) vanishes at \( \zeta = 0 \), it can be sought in the form 
\[ \hat{S} = \zeta \hat{S}_1 + \zeta^2 \hat{S}_2 + \ldots, \]
and we obtain:
\[
\hat{H}_U = \hat{H}_0 + \frac{\hbar^2 k^2}{2m} + \zeta \hat{M}_1 + \zeta^2 \hat{M}_2 + O(\zeta^3),
\]
where
\[
\hat{M}_1 = -i[\hat{H}_0, \hat{S}_1] + \delta \hat{H}' + \delta \hat{H}'',
\]
\[
\hat{M}_2 = -i[\hat{H}_0, \hat{S}_2] - \frac{1}{2}[[\hat{H}_0, \hat{S}_1], \hat{S}_1] - i[\delta \hat{H}', \hat{S}_1] - i[\delta \hat{H}'', \hat{S}_1].
\]

The interband blocks will be removed from \( \hat{M}_1 \) if
\[
-i[\hat{H}_0, \hat{S}_1] + \delta \hat{H}' = 0,
\]
from which we obtain \( \hat{S}_1 \). Similarly, the interband blocks will be removed from \( \hat{M}_2 \) if
\[
-i[\hat{H}_0, \hat{S}_2] - i[\delta \hat{H}'', \hat{S}_1] = 0,
\]
from which we obtain \( \hat{S}_2 \). Repeating this procedure, one can find \( \hat{S}_k \) and \( \hat{M}_k \) by iteration, to any desired order. In particular, \( \hat{M}_1 = \delta \hat{H}' \), \( \hat{M}_2 = -i[\delta \hat{H}'', \hat{S}_1]/2 \), etc. Setting \( \zeta = 1 \), we finally obtain the Hamiltonian which contains only intraband terms:
\[
\hat{H}_U = \hat{H}_0 + \frac{\hbar^2 k^2}{2m} + \delta \hat{H}' - \frac{i}{2}[[\hat{H}_0, \hat{S}_1], \hat{S}_1] + \ldots,
\]
where \( \hat{S}_1 \) satisfies Eq. (B11).

For example, neglecting all bands except the two explicitly shown in Eq. (B7), we seek \( \hat{S}_1 \) in the following form:
\[
\hat{S}_1 = \begin{pmatrix} 0 & \hat{s} \\ \hat{s}^\dagger & 0 \end{pmatrix}.
\]

Solving the equation (B11) with the interband matrix given by Eq. (B9), we find \( \hat{s} = (i/\mathcal{E}_b)\hat{q} \), where \( \mathcal{E}_b = \epsilon_2 - \epsilon_1 \) is the band splitting (we assume \( \mathcal{E}_b > 0 \)). Substituting this into Eq. (B12), we obtain:
\[
\hat{H}_1 = \epsilon_1 + \frac{\hbar^2 k^2}{2m} + r_1 \hat{\sigma} - \frac{1}{\mathcal{E}_b}(\hat{q}^2 + \hat{r}^2)
\]
and
\[
\hat{H}_2 = \epsilon_2 + \frac{\hbar^2 k^2}{2m} + r_2 \hat{\sigma} + \frac{1}{\mathcal{E}_b}(\hat{q}^2 + \hat{r}^2).
\]

These \( 2 \times 2 \) matrices can be interpreted as the effective one-band Hamiltonians incorporating the corrections of the second order in the interband couplings. Since \( r_{1,2}, \hat{q}, \) and \( \hat{r} \) are linear functions of \( k \) and \( \hat{H} \), one can rewrite the expressions (B13) and (B14) in the form
\[
\hat{H}_n(k, \hat{H}) = \epsilon_n(k) + \sum_{ij} A_{n,ij}k_i\hat{\sigma}_j + \sum_{ij} B_{n,ij}\hat{H}_i\hat{\sigma}_j + \sum_{ij} C_{n,ij}\hat{H}_i k_j + O(H^2),
\]
where we included the quadratic in \( k \) terms in \( \epsilon_n(k) \). The coefficients \( A, B, \) and \( C \) contain the matrix elements of the velocity \( \hat{\sigma} \) and the field \( \hat{H} \) in the \( \Gamma \)-point basis.

Thus, we see that the \( k \cdot p \) perturbation theory combined with the unitary transformation eliminating the interband matrix elements reproduces the general structure of the phenomenological one-band effective Hamiltonian, see Eq. (14). In particular, the antisymmetric SO coupling corresponds to the second term in Eq. (B15), whereas the last term reproduces the \( \ell \)-coupling to the magnetic field.

We showed in Sec. IV that in some bands the antisymmetric SO coupling is nonlinear (cubic) in \( k \). To derive such terms using the \( k \cdot p \) perturbation theory, one has include higher orders of the expansion in the interband matrix elements. In order to obtain a two-band effective Hamiltonian of Sec. V, one would have to keep the \( 4 \times 4 \) matrix block shown in Eq. (B7) and use a unitary transformation to eliminate the interband elements connecting the bands 1 and 2 with all other bands.
Appendix C: $\mathcal{G} = D_3$, $(\Gamma_5, \Gamma_6)$ band

As an example of calculation of the antisymmetric SO coupling $\gamma(k)$ and the generalized Zeeman coupling $\hat{\mu}(k)$ in the exceptional bands, let us consider the $(\Gamma_5, \Gamma_6)$ corep in a trigonal crystal with $\mathcal{G} = D_3$. There are two group generators, $g_1 = C_{3z}$ and $g_2 = \sigma_y$, and we obtain from Table II that $R(C_{3z}) = \hat{1}$ and $R(\sigma_y) = \hat{R}(C_{2y})$. Therefore, Eq. (16) yields the following symmetry constraints:

$$\gamma(k) = \gamma(C_{3z}^{-1}k), \quad \gamma(k) = C_{2y}\gamma(\sigma_y^{-1}k). \quad (C1)$$

Introducing $k_{3z} = k_x \pm ik_y$, the first of these constraints becomes $\gamma(k_+, k_-) = \gamma(e^{-2i\pi/3}k_+, e^{2i\pi/3}k_-)$, whose lowest-order odd-degree polynomial solution is given by

$$\gamma(k) = (b_1k_+^3 + b_2^*k_-^3)e_1 + (b_2k_+^3 + b_3^*k_-^3)e_2 + (b_3k_+^3 + b_4^*k_-^3)e_3,$$

where $b_{1,2,3}$ are complex constants. Imposing the second of the constraints $(C1)$, $\gamma_1(k_+, k_-) = -\gamma_3(k_-, k_+)$ and $\gamma_2(k_+, k_-) = \gamma_2(k_-, k_+)$, we arrive at the expression in Table III.

This example makes it clear that the form of $\gamma(k)$ depends on the choice of the second generator of a dihedral group, which is in turn determined by the orientation of the coordinate axes relative to the crystallographic axes. If one chose $g_2 = \sigma_x$ (the reflection in the $x = 0$ plane) then the second constraint would become $\gamma_1(k_+, k_-) = \gamma_2(-k_-, -k_+)$ and $\gamma_2(k_+, k_-) = -\gamma_2(-k_-, -k_+)$, therefore

$$\gamma(k) = ia_1(k_+^3 - k_-^3)e_1 + a_2(k_+^3 + k_-^3)e_2 + a_3(k_+^3 + k_-^3)e_3,$$

where $a_{1,2,3}$ are real constants.

From Eq. (18), the invariance conditions for $\hat{\mu}$ have the following form:

$$\hat{\mu}(k) = \hat{R}(C_{3z})\hat{\mu}(C_{3z}^{-1}k), \quad \hat{\mu}(k) = \hat{R}(C_{2y})\hat{\mu}(\sigma_y^{-1}k)\hat{R}^{-1}(C_{2y}), \quad (C2)$$

where

$$\hat{R}(C_{3z}) = \begin{pmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \hat{R}(C_{2y}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

are the rotation matrices. The constraints (C2) have a block-diagonal solution:

$$\hat{\mu} = \begin{pmatrix} \mu_{x1} & \mu_{x2} & 0 \\ \mu_{y1} & \mu_{y2} & 0 \\ 0 & 0 & \mu_{z3} \end{pmatrix}. \quad (C3)$$

While the equation for $\mu_{z3}$ is trivially satisfied by a constant $\mu_{z3}(k) = \mu_z$, the solutions for the other components are more involved. Introducing $u_1^\pm = \mu_{x1} \pm i\mu_{y1}$ and $u_2^\pm = \mu_{x2} \pm i\mu_{y2}$, we obtain: $u_1^\pm(k) = e^{\pm2i\pi/3}u_1^\pm(C_{3z}^{-1}k)$ and $u_2^\pm(k) = e^{\pm2i\pi/3}u_2^\pm(C_{3z}^{-1}k)$. One can easily see that the basal-plane components of the effective Zeeman coupling necessarily depend on $k$ and vanish at $k = 0$. Using the lowest-order even-degree polynomial solutions for $u_1^\pm(k)$, we have

$$\hat{\mu}(k) = \begin{pmatrix} \beta_1k_+^2 + \beta_2^*k_-^2 & \beta_2k_+^2 + \beta_3^*k_-^2 & 0 \\ i\beta_1k_+^2 - i\beta_2^*k_-^2 & i\beta_2k_+^2 - i\beta_3^*k_-^2 & 0 \\ 0 & 0 & \mu_z \end{pmatrix},$$

where $\beta_{1,2}$ are complex constants and $\mu_z$ is a real constant. Imposing now the second of the constraints (C2), we obtain Eq. (21).

Appendix D: Interband couplings for $\mathcal{G} = D_4$

To illustrate the calculations of the interband couplings in a two-band Rashba model (26) for $\mathcal{G} = D_4$, we consider $\tilde{\epsilon}(k)$ and $\gamma(k)$ in the case when the bands have different symmetry, i.e., when one band corresponds to the $\Gamma_6$ corep
and the other to the $\Gamma_7$ corep. We obtain from Eq. (36) that $\text{Re} \hat{c}$ is even in $k$, whereas $\text{Im} \hat{c}$ is odd. On the other hand, the invariance under $C_{2z}$ means that $\hat{c}(k) = \hat{c}(-k)$, therefore $\hat{c}(k)$ is real. It follows from Eq. (36) that $\hat{c}(k, -, k) = -\hat{c}(-k, +, k)$ and $\hat{c}(k, +, k) = \hat{c}(k, -, k)$. The lowest-order even-degree polynomial solution of these equations has the form $\hat{c} \propto k_x^2 + k_y^2$, therefore $\hat{c}(k) \propto k_x^2 - k_y^2$.

From Eq. (37), we obtain that $\text{Re} \hat{\gamma}$ is odd in $k$, whereas $\text{Im} \hat{\gamma}$ is even. The invariance under $C_{2z}$ means that $\hat{\gamma}_{1,2}(k) = -\hat{\gamma}_{1,2}(-k)$ and $\hat{\gamma}_{3}(k) = \hat{\gamma}_{3}(-k)$, therefore $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are real, but $\hat{\gamma}_3$ is purely imaginary. Introducing $\hat{\gamma}_\pm = \hat{\gamma}_1 \pm i\hat{\gamma}_2$, the invariance conditions under $C_{2z}$ become $\hat{\gamma}_\pm(k, +, k) = \mp i\hat{\gamma}_\pm(-k, +, k)$ and $\hat{\gamma}_3(k, +, k) = -\hat{\gamma}_3(-k, +, k)$. The lowest-order polynomial solution of these equations has the following form: $\hat{\gamma}_\pm = c_\pm k_x^2$ and $\hat{\gamma}_3 = c_1 k_x^2 + c_2 k_y^2$. The invariance under $\sigma_y$ imposes additional constraints: $\hat{\gamma}_\pm(k, +, k) = -\hat{\gamma}_\pm(k, -, k)$ and $\hat{\gamma}_3(k, +, k) = -\hat{\gamma}_3(k, -, k)$. Therefore, $c_+ = -c_-$, $c_2 = -c_1$, and we arrive at the expressions in Table V.

### Appendix E: Two-band Hamiltonian in the $\Upsilon$-matrix representation

Any effective two-band Hamiltonian, see Eq. (26), can be represented as a linear combination of 16 Hermitian $4 \times 4$ matrices as follows:

$$\hat{\mathcal{H}}(k, \mathbf{H}) = \sum_{p, q=0}^{3} \omega_{pq}(k, \mathbf{H}) \hat{\tau}_p \otimes \hat{\sigma}_q,$$  

(E1)

where $\hat{\tau}_0$ and $\hat{\tau}_{1,2,3}$ ($\hat{\sigma}_0$ and $\hat{\sigma}_{1,2,3}$) are the identity matrix and the Pauli matrices acting on the band (Kramers) indices and the coefficients $\omega_{pq}$ are real. The TR invariance constraint (10), with the matrices $\hat{D}(K)$ given by Eq. (27), takes the form

$$\sum_{p, q=0}^{3} \omega_{pq}(k, \mathbf{H}) \hat{\tau}_p \otimes \hat{\sigma}_q = \sum_{p, q=0}^{3} \omega_{pq}(-k, -\mathbf{H}) \hat{U}_K^\dagger \hat{\tau}_p \otimes \hat{\sigma}_q \hat{U}_K,$$

(E2)

where $\hat{U}_K = \hat{\tau}_0 \otimes i\hat{\sigma}_2$. It is straightforward to check that six of the matrices $\hat{\tau}_p \otimes \hat{\sigma}_q$ are invariant under TR, in the sense that $\hat{U}_K \hat{\tau}_p \otimes \hat{\sigma}_q \hat{U}_K^\dagger = \hat{\tau}_p \otimes \hat{\sigma}_q$, whereas the remaining ten change sign. The six TR-even matrices are the unit matrix $\hat{\tau}_0 \otimes \hat{\sigma}_0 \equiv \hat{\Upsilon}_0$ and the five following traceless matrices:

$$\hat{\tau}_1 \otimes \hat{\sigma}_0 \equiv \hat{\Upsilon}_1, \quad \hat{\tau}_3 \otimes \hat{\sigma}_0 \equiv \hat{\Upsilon}_2, \quad \hat{\tau}_2 \otimes \hat{\sigma}_1 \equiv \hat{\Upsilon}_3, \quad \hat{\tau}_2 \otimes \hat{\sigma}_2 \equiv \hat{\Upsilon}_4, \quad \hat{\tau}_2 \otimes \hat{\sigma}_3 \equiv \hat{\Upsilon}_5.$$

(E3)

The matrices $\hat{\Upsilon}_a$ ($a = 1, ..., 5$) satisfy the relations

$$\{ \hat{\Upsilon}_a, \hat{\Upsilon}_b \} = 2\delta_{ab},$$

(E4)

i.e., generate a Clifford algebra. The ten TR-odd matrices can be chosen as

$$\hat{\Upsilon}_{ab} = \frac{1}{2\iota} [\hat{\Upsilon}_a, \hat{\Upsilon}_b] = -i\hat{\Upsilon}_a \hat{\Upsilon}_b, \quad a < b.$$

It is easy to see that $\hat{\Upsilon}_a^2 = \hat{\Upsilon}_b^2 = \hat{\Upsilon}_0$. Note that we use the $\Gamma$-point basis $\{|1, 1\}, |1, 2\rangle, |2, 1\rangle, |2, 2\rangle$ and our nomenclature for the $\Upsilon$ matrices follows Ref. 36.

In terms of the $\Upsilon$ matrices, Eq. (E1) takes the form

$$\hat{\mathcal{H}}(k, \mathbf{H}) = d_0(k, \mathbf{H}) \hat{\Upsilon}_0 + \sum_a d_a(k, \mathbf{H}) \hat{\Upsilon}_a + \sum_{a < b} d_{ab}(k, \mathbf{H}) \hat{\Upsilon}_{ab}.$$  

(E5)

The expansion coefficients here are real and satisfy the following constraints:

$$d_0(-k, -\mathbf{H}) = d_0(k, \mathbf{H}), \quad d_a(-k, -\mathbf{H}) = d_a(k, \mathbf{H}), \quad d_{ab}(-k, -\mathbf{H}) = -d_{ab}(k, \mathbf{H}).$$

(E6)

which follow from the TR invariance requirement (E2). Further constraints are imposed by the point-group symmetries. For example, in the case of $G = D_4$ in zero field we obtain from Eq. (42) and Table V that only the following coefficients are nonzero:

$$d_0 = \frac{\epsilon_1 + \epsilon_2}{2}, \quad d_1 = \iota, \quad d_2 = \frac{\epsilon_1 - \epsilon_2}{2}, \quad d_5 = \iota \gamma_3, \quad d_{13} = \frac{a_1 - a_2}{2} k_y, \quad d_{14} = -\frac{a_1 - a_2}{2} k_x, \quad d_{23} = -\gamma_1, \quad d_{24} = -\gamma_2, \quad d_{35} = \frac{a_1 + a_2}{2} k_x, \quad d_{45} = \frac{a_1 + a_2}{2} k_y.$$
The reduced two-band model of a topological insulator introduced in Sec. VI A, corresponds to
\[
d_0 = \frac{E_b}{2}, \quad d_2 = \xi = \frac{k^2}{2m^*}, \quad d_{23} = -\hat{a}k_y, \quad d_{24} = \hat{a}k_x, \quad d_{35} = ak_x, \quad d_{45} = ak_y,
\]
with all other coefficients equal to zero.

The spectrum of the Hamiltonian (E5) has a simple analytical form in the “Clifford limit”, when Eq. (E5) contains only the matrices forming a Clifford algebra, in addition to the unit matrix. To illustrate this, let us consider a centrosymmetric TR-invariant crystal, in which case the cores and the bands have a definite parity. Setting $H = 0$ and $g = I$ ($I$ is the spatial inversion), the constraint (9) takes the form

\[
\hat{H}(k) = \hat{U}_I^\dagger \hat{H}(-k) \hat{U}_I,
\]
(E7)

where

\[
\hat{U}_I = \hat{D}(I) = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix} \otimes \sigma_0,
\]

and $p_{1,2}$ are the band parities. A straightforward calculation shows that $\hat{\Upsilon}_0$ and $\hat{\Upsilon}_2$ are always inversion-even, in the sense that $\hat{U}^\dagger_I \hat{\Upsilon}_0 \hat{U}_I = \hat{\Upsilon}_0$ and $\hat{U}^\dagger_I \hat{\Upsilon}_2 \hat{U}_I = \hat{\Upsilon}_2$. The other $\Upsilon$ matrices can be even or odd, depending on the relative parity of the bands, i.e. $\hat{U}^\dagger_I \hat{\Upsilon}_a \hat{U}_I = p_1p_2 \hat{\Upsilon}_a$, for $a = 1, 3, 4, 5$. Similarly, $\hat{\Upsilon}_{ab}$ is inversion-even if $a, b \neq 2$, otherwise $\hat{U}^\dagger_I \hat{\Upsilon}_{ab} \hat{U}_I = p_1p_2 \hat{\Upsilon}_{ab}$.

From Eq. (E7) and the TR constraints (E6), we obtain that if the bands have the same parity, then all $d_{ab} = 0$ and the Hamiltonian becomes

\[
\hat{H}(k) = d_0(k) \hat{\Upsilon}_0 + \sum_{a=1}^5 d_a(k) \hat{\Upsilon}_a,
\]
(E8)
i.e., it contains only the Clifford algebra matrices and the unit matrix. This last expression can be easily diagonalized by squaring the second term and using the Clifford algebra relations (E4). In this way, we find that the bands are twofold degenerate and given by

\[
E_{\pm}(k) = d_0(k) \pm \sqrt{\sum_a d_a^2(k)}.
\]
(E9)

If the bands have opposite parity, then $d_a = 0 (a \neq 2)$ and $d_{ab} = 0 (a, b \neq 2)$, therefore

\[
\hat{H}(k) = d_0(k) \hat{\Upsilon}_0 + d_2(k) \hat{\Upsilon}_2 + d_{12}(k) \hat{\Upsilon}_{12} + d_{23}(k) \hat{\Upsilon}_{23} + d_{24}(k) \hat{\Upsilon}_{24} + d_{25}(k) \hat{\Upsilon}_{25}.
\]
(E10)

Since the five matrices $\hat{\Upsilon}_2$, $\hat{\Upsilon}_{12}$, $\hat{\Upsilon}_{23}$, $\hat{\Upsilon}_{24}$, and $\hat{\Upsilon}_{25}$ form a Clifford algebra, the spectrum consists of two twofold degenerate bands:

\[
E_{\pm}(k) = d_0(k) \pm \sqrt{d_2^2(k) + d_{12}^2(k) + d_{23}^2(k) + d_{24}^2(k) + d_{25}^2(k)}.
\]
(E11)
The expressions (E9) and (E11) reproduce, in a rather circuitous way, the well-known statement that the Bloch bands in a centrosymmetric TR-invariant crystal are twofold degenerate at each $k$. Note that, although $\hat{H}(k) \neq \hat{H}(-k)$ if the bands have opposite parity, see Eq. (E10), the band dispersions are always even in $k$.

**Appendix F: Modified Dirac Hamiltonian**

In this Appendix, we review the properties of the modified Dirac Hamiltonian of the form $\hat{H} = \text{diag}(\hat{H}_+, \hat{H}_-)$, with

\[
\hat{H}_{\pm}(k) = \begin{pmatrix} \alpha(k_x^2 + k_y^2) + \beta & \gamma(k_y \pm ik_x) \\ \gamma(k_y \mp ik_x) & -\alpha(k_x^2 + k_y^2) - \beta \end{pmatrix} = g_{\pm}(k) \hat{s},
\]
(F1)

where $\hat{s}$ are the Pauli matrices and $g_{\pm} = (\gamma k_y, \mp \gamma k_x, ak^2 + \beta)$. Note that $\hat{H}_{-}(k) = \hat{s}_3 \hat{H}_{+}^\dagger(-k) \hat{s}_3$, which can be interpreted as a consequence of TR invariance, according to Eq. (49). While the $2 \times 2$ blocks $\hat{H}_{\pm}$ have the same bulk spectrum consisting of two symmetric gapped branches, $\sqrt{\gamma^2 k^2 + (ak^2 + \beta)^2}$ and $-\sqrt{\gamma^2 k^2 + (ak^2 + \beta)^2}$, their opposite chiralities manifest themselves in different spectra of the topologically protected gapless boundary modes.
1. Topology in the bulk

In order to construct a bulk topological invariant for our 2D system, we introduce, following Ref. 32, an auxiliary real variable \( \omega \) and define the Green’s function for \( \hat{H} = \hat{H}_+ \) or \( \hat{H}_- \) as follows: \( \hat{G}(\mathbf{k}, \omega) = [i\omega - \hat{H}(\mathbf{k})]^{-1} \). Then, the topological invariant has the form

\[
N = -\frac{1}{24\pi^2} \int \text{Tr}(\hat{G}d\hat{G}^{-1})^3,
\]

where the integration is performed over \( \omega \) and \( \mathbf{k} = (k_x, k_y) \). For \( \hat{H}(\mathbf{k}) = g(\mathbf{k})\hat{s} \), we obtain:

\[
N = \frac{1}{4\pi} \int d^2k \hat{g} \left( \frac{\partial \hat{g}}{\partial k_x} \times \frac{\partial \hat{g}}{\partial k_y} \right),
\]

where \( \hat{g} = g/|g| \).

Since at \( |\mathbf{k}| \to \infty \), \( \hat{g}_\pm \to (0, 0, \text{sign } \alpha) = \text{const} \), the 2D momentum space is isomorphic to the \( S^2 \) sphere and \( N_\pm \) is an integer (the winding number) equal to the degree of a mapping \( \mathbf{k} \to \hat{g}_\pm(\mathbf{k}) \), i.e. \( S^2 \to S^2 \). After a straightforward calculation, Eq. (F3) yields the following result:

\[
N_\pm = \pm \frac{1}{2}(\text{sign } \beta - \text{sign } \alpha),
\]

therefore \( \hat{H}_+ \) and \( \hat{H}_- \) describe a topologically nontrivial bulk system if \( \alpha \beta < 0 \), then \( |N_\pm| = 1 \). In contrast, if \( \alpha \beta > 0 \), then \( N_\pm = 0 \) and the system is topologically trivial. This conclusion will be confirmed by an explicit calculation of the spectrum of the boundary modes in Sec. F2 below. Also, one can show that the first Chern number, which is given by the Berry flux through the Brillouin zone, is equal to the winding number (F3). For this reason, the Hamiltonians \( \hat{H}_+ \) and \( \hat{H}_- \) describe what is sometimes called the Chern insulators.

The topological invariant (F2) can also be defined for the full modified Dirac Hamiltonian \( \hat{H} = \text{diag}(\hat{H}_+, \hat{H}_-) \) as follows: \( \hat{N} = -(1/24\pi^2) \int \text{Tr}(\hat{G}d\hat{G}^{-1})^3 \), where \( \hat{G}(\mathbf{k}, \omega) = [i\omega - \hat{H}(\mathbf{k})]^{-1} \). It is easy to see that the total Chern number \( \hat{N} = N_+ + N_- = 0 \), according to Eq. (F4). This does not mean, however, that the system is topologically trivial, since one can define another invariant:

\[
\nu = \frac{1}{2} |N_+ - N_-| = \begin{cases} 
 1, & \text{if } \alpha \beta < 0, \\
 0, & \text{if } \alpha \beta > 0.
\end{cases}
\]

This is the \( Z_2 \) invariant counting the number of time-reversed pairs of counterpropagating boundary modes, see the next subsection.

2. Boundary modes

Let us consider a half-plane \( x \geq 0 \) and replace \( \mathbf{k} \to \hat{k} = -i\mathbf{\nabla} \) in Eq. (F1). The momentum \( k_y \) is a good quantum number and the eigenstates of \( \hat{H}_\pm \) have the form \( e^{ik_yv}\psi(x) \). We assume that the wave functions vanish at \( x = 0 \) (a “hard wall” boundary condition). At given \( k_y \), the Hamiltonians become

\[
\hat{H}_\pm(k_y) = \begin{pmatrix}
\alpha(k_y^2 - \nabla_x^2) + \beta & \gamma(k_y \pm \nabla_x) \\
\gamma(k_y \pm \nabla_x) & -\alpha(k_y^2 - \nabla_x^2) - \beta
\end{pmatrix},
\]

whose solutions corresponding to the energy \( E \) and localized near the surface have the form \( \psi(x) = (u, v)^\top e^{-\kappa x} \), where \( \text{Re } \kappa > 0 \). For \( \hat{H}_+ \), we find

\[
\frac{u}{v} = \frac{\gamma(k_y - \kappa)}{E + \alpha \kappa - \beta} = \frac{E - \alpha z + \beta}{\gamma(k_y + \kappa)},
\]

where \( z = \kappa^2 - k_y^2 \). The last expression produces an equation for \( \kappa \), which has two solutions, \( \kappa_{1,2}^2 = k_y^2 + z_{1,2} \), with

\[
z_{1,2} = d \pm \sqrt{d^2 - \frac{\beta^2 - E^2}{\alpha^2}}, \quad d = \frac{\beta}{\alpha} + \frac{\gamma^2}{2\alpha^2}.
\]
Therefore, at given \( k_y \), the general localized eigenstate has the following form:

\[
\psi(x) = C_1 \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} e^{-\kappa_1 x} + C_2 \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} e^{-\kappa_2 x}.
\]  

(F9)

Note that at \( \alpha = 0 \), i.e., in the “Dirac limit”, there is only one localized solution and the hard wall boundary condition is not applicable, because it would lead to \( \psi(x) = 0 \) at all \( x \geq 0 \).

Substituting the solution (F9) into the boundary condition \( \psi(0) = 0 \) and using Eq. (F7), we arrive at the following two equivalent equations for \( E \):

\[
E = -\beta - \alpha(k_y - \kappa_1)(k_y + \kappa_2), \quad E = \beta + \alpha(k_y - \kappa_1)(k_y - \kappa_2).
\]

Adding and subtracting these equations, we obtain:

\[
E = -\alpha(\kappa_1 + \kappa_2)k_y,
\]

which means that \( E \) is an odd function of \( k_y \), and

\[
\kappa_1 \kappa_2 = -\frac{\beta}{\alpha} - k_y^2.
\]

(F11)

Squaring Eq. (F10) and taking into account Eq. (F8), we have \( E^2 = \gamma^2 k_y^2 \), therefore the boundary mode energy has the form \( E(k_y) = \zeta |\gamma| k_y \), where the coefficient \( \zeta \) satisfies \( \zeta^2 = 1 \), does not depend on \( k_y \), and can be calculated at \( k_y = 0 \).

Setting \( k_y = 0 \) in Eq. (F11), we see that the localized solution with \( \text{Re} \kappa_{1,2} > 0 \) exists only if the signs of \( \alpha \) and \( \beta \) are opposite, which agrees with the bulk topological argument given above. Focusing on the topological regime with \( \alpha \beta < 0 \) and introducing a dimensionless parameter

\[
\varrho = \frac{\gamma^2}{2|\alpha \beta|} - 1,
\]

we obtain from Eq. (F8):

\[
\kappa_1(k_y = 0) = \sqrt{\frac{\beta}{|\alpha|}} f^{1/2}(\varrho), \quad \kappa_2(k_y = 0) = \sqrt{\frac{|\beta|}{\alpha}} f^{-1/2}(\varrho),
\]

(F12)

where \( f(\varrho) = \varrho + \sqrt{\varrho^2 - 1} \). If \( \varrho > 1 \), then \( f, \kappa_1, \) and \( \kappa_2 \) are all real positive, and it follows from Eq. (F10) that \( \zeta = -\text{sign}(\alpha) \). If \( |\varrho| < 1 \), then \( f = e^{i\varphi} (0 < \varphi < \pi) \), \( \kappa_2 = \kappa_1^* \) with \( \text{Re} \kappa_{1,2} > 0 \), and we again obtain \( \zeta = -\text{sign}(\alpha) \).

Thus, we finally arrive at the following exact expression for the energy of a single nondegenerate boundary mode of the Hamiltonian \( \hat{H}_+ \):

\[
E_+(k_y) = -\text{sign}(\alpha)|\gamma| k_y = \text{sign}(\beta)|\gamma| k_y.
\]

(F13)

Similarly, for \( \hat{H}_- \) we obtain:

\[
E_-(k_y) = \text{sign}(\alpha)|\gamma| k_y = -\text{sign}(\beta)|\gamma| k_y.
\]

(F14)

Thus, at \( \alpha \beta < 0 \) the modified Dirac Hamiltonian \( \hat{H} \) has two counterpropagating chiral boundary modes, given by Eqs. (F13) and (F14). It is easy to see that these modes can be obtained from one another by TR. Indeed, the Hamiltonians (F6) satisfy \( \hat{H}_-(k_y) = \hat{s}_3 \hat{H}_+^*(-k_y) \hat{s}_3 \). If \( \psi_{+,-k_y}(x) \) is an eigenstate of \( \hat{H}_+ \) corresponding to the eigenvalue \( E_+(k_y) \), then its time-reversed partner \( \psi_{+-k_y}(x) = \hat{s}_3 \psi_{+,-k_y}(x) \) is an eigenstate of \( \hat{H}_- \) corresponding to the eigenvalue \( E_-(k_y) = E_+( -k_y) \).

Regarding the explicit expressions for the boundary mode wave functions, we focus on \( k_y = 0 \) and obtain from Eqs. (F7) and (F12) that the zero mode for \( \hat{H}_+ \) has the form

\[
\psi_+(x) = C \begin{pmatrix} 1 \\ -\text{sign}(\alpha \gamma) \end{pmatrix} (e^{-\kappa_1 x} - e^{-\kappa_2 x}).
\]

(F15)
For $\hat{H}_\perp$, we have

$$\psi_-(x) = C \begin{pmatrix} 1 \\ \text{sign}(\alpha \gamma) \end{pmatrix} \left( e^{-\kappa_1 x} - e^{-\kappa_2 x} \right). \quad \text{(F16)}$$

If we choose the zero-mode wave functions to be real, then the normalization coefficient at $|\varrho| > 1$ is given by $C = \sqrt{\kappa_1 \kappa_2 (\kappa_1 + \kappa_2)/(\kappa_1 - \kappa_2)}$. At $|\varrho| < 1$, one can write $\kappa_{1,2} = \kappa \pm i q$ ($\kappa$, $q$ are real positive), and $C = i \sqrt{\kappa (\kappa^2 + q^2)}/2q^2$. The zero modes (F15) and (F16) are TR partners: $\psi_-(x) = \hat{s}_3 \psi_+^*(x)$. It is also worth noting that the coordinate dependence of the boundary mode qualitatively changes as $|\varrho|$ varies, from a superposition of two exponentials at $|\varrho| > 1$ to a damped sinusoid at $|\varrho| < 1$.