Edge Modes, Degeneracies, and Topological Numbers in Non-Hermitian Systems

Daniel Leykam,1 Konstantin Y. Bliokh,2,3 Chunli Huang,1,4 Y. D. Chong,1,5 and Franco Nori2,6
1Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore
2CEMS, RIKEN, Wako-shi, Saitama 351-0198, Japan
3Nonlinear Physics Centre, RSPE, The Australian National University, Canberra, Australian Capital Territory 0200, Australia
4Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan
5Centre for Disruptive Photonic Technologies, Nanyang Technological University, Singapore 637371, Singapore
6Physics Department, University of Michigan, Ann Arbor, Michigan 48109-1040, USA

(Received 17 October 2016; published 23 January 2017)

We analyze chiral topological edge modes in a non-Hermitian variant of the 2D Dirac equation. Such modes appear at interfaces between media with different “masses” and/or signs of the “non-Hermitian charge.” The existence of these edge modes is intimately related to exceptional points of the bulk Hamiltonians, i.e., degeneracies in the bulk spectra of the media. We find that the topological edge modes can be divided into three families (“Hermitian-like,” “non-Hermitian,” and “mixed”); these are characterized by two winding numbers, describing two distinct kinds of half-integer charges carried by the exceptional points. We show that all the above types of topological edge modes can be realized in honeycomb lattices of ring resonators with asymmetric or gain-loss couplings.

DOI: 10.1103/PhysRevLett.118.040401

Introduction.—There is presently enormous interest in two groups of fundamental physical phenomena: (i) topological edge modes in quantum Hall fluids and topological insulators [1–3], which are Hermitian, and (ii) novel effects in non-Hermitian wave systems (including \(PT\)-symmetric systems) [4–6]. Both types of phenomena have been studied in the context of quantum as well as classical waves, and both are deeply tied to the geometrical features of spectral degeneracies. In the Hermitian case, the common degeneracies are Dirac points: linear band crossings (generically, in a 3D parameter space), which separate distinct topological phases and mark the birth or destruction of topological edge modes [7,8]. Non-Hermitian systems, however, exhibit a distinct class of spectral degeneracies known as exceptional points (EPs), which are branch points in a 2D parameter space where the Hamiltonian becomes nondiagonalizable [4,9,10].

In Hermitian systems, the bulk-edge correspondence relations that give rise to topological edge modes are typically based upon the Berry connection and rely on the eigenvector orthogonality granted by Hermiticity [1–3]. Is there a generalization of the bulk-edge correspondence to non-Hermitian systems [11]? When sufficiently weak non-Hermiticity (e.g., loss) is introduced to topological insulator models, the edge modes can retain some of their original characteristics [12,13]. On the other hand, certain non-Hermitian models with chiral symmetry can support anomalous edge modes that have no clear relationship to Hermitian topological edge modes [14–16]. These modes are embedded within a complex gapless band structure and appear in the vicinity of EPs; however, it is not known whether they can be related to model-independent bulk topological invariants similar to those in Hermitian systems.

This Letter aims to shed light on the nature of topological edge modes in non-Hermitian quantum systems. In contrast to previous studies based on lattice models [11–22], we focus on a non-Hermitian continuum model. This is motivated by the fact that, in the Hermitian case, many model-independent features of topological edge modes can be understood in terms of the generic properties of continuum models, such as the Dirac equation in various dimensions [7,23–25]. For example, zero-energy Jackiw-Rebbi end modes of the 1D Dirac equation [23,24] underpin end modes of the Su-Schrieffer-Heeger lattice model [26,27].

Our continuum model consists of a 2D non-Hermitian Hamiltonian that is linear in both \(k_x\) and \(k_y\) and possesses a tunable mass parameter \(m\), similar to the 2D Dirac equation. The bulk band structure is complex and possesses a pair of EPs (branch points). Along interfaces between media with different “masses” and/or signs of the non-Hermiticity, we find that there exist zero-energy chiral edge modes. Remarkably, the appearance of these edge modes and their regions of existence are fully determined by the EPs in the bulk spectra of the media. We show that these modes can be classified as “Hermitian-like,” “non-Hermitian,” and “mixed,” using two topological numbers. The first number is related to the chirality of the eigenstates (i.e., the sign of the Berry curvature), while the second one characterizes the chirality of the EP [9,10,28–30]. The non-Hermitian and mixed edge modes resemble the “anomalous” edge modes found in Ref. [16]. Moreover, we are able to enumerate the zero-energy edge modes by using an index.
theorem, a variant of the Aharonov-Casher theorem for the 2D Dirac equation in a vector potential [31]. Finally, we show that a lattice counterpart of this continuum model, including the anomalous edge modes, can be realized using honeycomblike arrays of ring resonators with non-Hermitian couplings [32,33,37].

Non-Hermitian Hamiltonian.—Our model is based on the following non-Hermitian Hamiltonian, defined on a 2D momentum space $\mathbf{k} = (k_x, k_y)$:

$$\hat{H} = \begin{pmatrix} k_x - i k_y & m \\ m & -k_x + i k_y \end{pmatrix} \equiv (k_x - i k_y) \hat{\sigma}_z + m \hat{\sigma}_y \equiv \mathbf{B} \cdot \hat{\sigma}. \quad (1)$$

Here, $\hat{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ denotes the vector of Pauli matrices (permuted cyclically for later convenience), and $\mathbf{B} = (B_x, B_y, 0)$ is an effective complex “magnetic field,” which will be used in the subsequent Berry-phase analysis. The Hamiltonian $\hat{H}$ contains three continuously tunable real parameters: the momenta $k_x$ and $k_y$, and $m$ (assumed real), which mixes the two spinor components, and which we call “mass” for convenience. The parameter $s = \pm 1$, which we will regard as a “non-Hermitian charge,” determines the sign of the imaginary part, such that $[\hat{H}(s)]^\dagger = \hat{H}(-s)$.

The Hamiltonian (1) involves only two Pauli matrices and has the chiral symmetry $\{\hat{H}, \hat{\sigma}_y\} = 0$. It is also $PT$ symmetric (where $T$ involves complex conjugation and momenta reversal, while $P$ is the reflection $x \to -x$) and can hence have real eigenvalues [5,6]. The eigenvalues of $\hat{H}$ are

$$\lambda^\pm = \pm \sqrt{\mathbf{B} \cdot \mathbf{B}} = \pm \sqrt{m^2 + (k_x - i k_y)^2}, \quad (2)$$

and its (non-normalized) eigenvectors are

$$\psi^\pm = \begin{pmatrix} 1 \\ B_y/(B_x + \lambda^\pm) \end{pmatrix}. \quad (3)$$

The complex spectrum (2) is shown in Fig. 1. Along the $k_y$ axis, the real part of the spectrum $\text{Re}(\lambda)$ is gapped for $-|m| < k_x < |m|$ and ungapped for $|k_x| > |m|$. There are two EPs at $k_{\text{EP}} = (0, \pm |m|)$, separating the “gapped” and “ungapped” $k_x$ domains.

Unlike Hermitian degeneracies, EPs involve the coalescence of eigenvectors, not just the eigenvalues $\lambda^\pm(k_{\text{EP}}) = 0$. $\hat{H}(k_{\text{EP}})$ is defective and has a single chiral eigenmode (an eigenvector of $\hat{\sigma}_z$):

$$\psi(k_{\text{EP}}) = \begin{pmatrix} 1 \\ i\chi_{\text{EP}} \end{pmatrix}, \quad (4)$$

where $\chi_{\text{EP}} = \pm \text{sgn}(sm)$ is the chirality of the EP [9,10,28–30].

Chiral edge modes.—We translate Eq. (1) to a Schrödinger wave equation by taking $\mathbf{k} = -i\nabla$ and allowing the mass $m$ and/or non-Hermitian charge $s$ to vary with position (though we still assume that $s$ only takes the values $\pm 1$):

$$\hat{H} = [-i\partial_x - s(x, y)\partial_y] \hat{\sigma}_z + m(x, y) \hat{\sigma}_y. \quad (5)$$

Consider an interface between two uniform media with different $m$ and/or $s$. For now, let the interface be along the line $x = 0$, such that $m = m_1, s = s_1$ for $x < 0$ (medium 1) and $m = m_2, s = s_2$ for $x > 0$ (medium 2). We seek edge modes that propagate along $y$ and are normalizable along $x$:

$$\psi_{\text{edge}} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \begin{pmatrix} e^{ik_y + \kappa_1}, & \text{Re}(\kappa_1) < 0, x > 0 \\ e^{ik_y + \kappa_2}, & \text{Re}(\kappa_2) > 0, x < 0 \end{pmatrix}. \quad (6)$$

Substituting Eq. (6) into Eq. (5), we find the zero-energy edge modes $\lambda_{\text{edge}} = 0$ which exist when the following real equations are satisfied:

$$-\kappa_1 = s_1 k \pm m_1, \quad -\kappa_2 = s_2 k \pm m_2. \quad (7)$$

For $\kappa_1 < 0$ and $\kappa_2 > 0$, there can be zero, one, or two solutions to Eq. (7) for each $k$. The number of solutions also depends on $m_{1,2}$ and $s_{1,2}$. Like the eigenmodes at the EPs of the bulk system, these edge modes are chiral, satisfying $\beta/\alpha = \pm i$. Similar to Eq. (4), we define the mode chirality as $\chi_{\text{edge}} = \text{Im}(\beta/\alpha)$.

We first examine the two simplest cases.

(A) The media have equal charges $s_1 = s_2 = s$ and opposite masses $m_1 = -m_2 = m$. In this case, there is one zero-energy edge mode for each $k \in (-|m|, |m|)$ and no edge modes for all other $k$. This $k$ range corresponds to the $k_x$ domain with the gapped bulk spectra $\text{Re}(\lambda^\pm)$ between the two EPs (Fig. 1). This domain includes $k = 0$, which is the Hermitian limit where Eq. (5) reduces to the Jackiw-Rebbi model for 1D Dirac modes [23,24]. Thus, this is a family of non-Hermitian edge modes that are continuant from the Hermitian Jackiw-Rebbi edge modes. The mode chirality is $\chi_{\text{edge}} = \text{sgn}(m)$, independent of $s$.

(B) The media have equal “masses” $m_1 = m_2 = m$ but opposite “charges” $s_1 = -s_2 \equiv s$. In this case, there are two edge modes in the domain $k \in \text{sgn}(s)(|m|, \infty)$. This
corresponds to one of the $k_s$ domains with the ungapped \( \text{Re}(\lambda^\pm) \) bulk spectra. The two edge modes have opposite chiralities $\chi_{\text{edge}} = \pm 1$. Unlike case (A), these modes are essentially non-Hermitian. First, they are asymmetric in $k$ and do not exist in the Hermitian limit $k = 0$. Second, the modes are defective: the corresponding left eigenvectors [right eigenvectors of $\tilde{H}(-k) = \tilde{H}^\dagger(k)$] do not exist.

When $|m_1| \neq |m_2|$, the situation is more complicated. Figure 2 shows the edge modes for varying $m_2$, with $s_1 = s_2 = 1$ and $m_1 > 0$. For $m_2 > m_1$, there is one edge mode for each $k \in (m_1, m_2)$, as shown in Fig. 2(a). For $m_2 < m_1$, there is one edge mode for each $k \in (-m_2, -m_1)$, as shown in Figs. 2(b)–2(d); this includes the special case (A) discussed above. For certain values of $k$, $\text{Re}(\lambda^\pm)$ is gapped in one medium and ungapped in the other medium. We call such zones and the corresponding edge modes mixed. When $m_2 > 0$, there are only positive or only negative values of $k$, as shown in Figs. 2(a) and 2(b). In Fig. 2, we also indicate the chiralities of the EPs in the two media $\chi^{\text{EP}}$ and the chiralities of the edge modes $\chi_{\text{edge}}$. Notably, the edge modes always connect a pair of EPs with the same chirality, while the modes themselves have the opposite chirality.

We summarize the conditions under which zero-energy edge modes exist using the phase diagrams in Fig. 3. Here, we fix $m_1 > 0$ and $s_1 = 1$, and use $k/m_1$ and $m_2/m_1$ as plot axes. The red (blue) regions show where there is a single edge mode with $\chi_{\text{edge}} = + 1$ ($\chi_{\text{edge}} = -1$). Figure 3(a) shows the case where the two media have equal non-Hermitian charge $s$, with the special case (A) lying on the $m_2/m_1 = -1$ line and the Jackiw-Rebbi model [23,24] lying on the $k = 0$ line. Figure 3(b) shows the opposite-charge case; it also contains a purple region indicating two edge modes with $\chi_{\text{edge}} = \pm 1$, which includes the case (B) on the line $m_2/m_1 = 1$.

We will now show that these phase diagram features—i.e., the number of zero-energy edge modes and under what conditions they appear—can be understood from the topological properties of Eq. (1).

**Winding numbers.**—Since one family of edge modes can be continued to Jackiw-Rebbi modes [23,24], and the termination points of the edge modes are EPs of the bulk spectrum, we can guess that the edge modes can be characterized by bulk topological invariants [1–3]. Along the interface, the conserved $k_s$ plays the role of a tunable parameter for calculating a 1D winding number. However, it turns out that two winding numbers are needed to fully describe the edge modes in the non-Hermitian case.

Previous researchers [11,12,16,38] have focused on the winding numbers of the eigenvectors $\psi^\pm$. However, we...
emphasize that encircling an EP (branch point) swaps the bands, so that two loops in parameter space are required to return to the original state (with a π geometric phase gained) [4,9,10,16,38–40]. Hence, there is no globally smooth way to define two distinct bands for Eq. (3).

One way to resolve this band-labeling ambiguity is to consider winding numbers associated with the complex magnetic field B defined in Eq. (1), which has no discontinuities. We take a spherical-like representation $B = B(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$, where both the “magnitude” $B = \lambda^+$ and the “angles” $(\theta, \phi)$ are complex [41]. The chiral symmetry of $H$ constrains $B$ to the plane $\theta = \pi/2$, so only $B$ and $\phi$ vary with $k$.

We now introduce the winding number

$$w_1 = \frac{1}{2\pi} \int_{k_y=-\infty}^{k_y=+\infty} \nabla_k \phi \cdot dk,$$

where the integral is taken along a $k_y$ line with fixed $k_x$. This winding number originates from a non-Hermitian generalization of the Berry phase, describing the effects of varying direction of $B$ [41]. It is equivalent to the winding numbers used in Refs. [11,12,38]. Applying Eq. (8) to Eq. (1), we find [33]

$$w_1 = \begin{cases} -\frac{1}{2} \text{sgn}(m) & |k_y| < |m| \\ 0 & |k_y| > |m|. \end{cases}$$

This explains the edge modes in case (A), corresponding to the $m_2/m_1 = -1$ line in Fig. 3(a). The difference in the topological numbers of the two media is $\Delta w_1 = w_1(m_2) - w_1(m_1) = \text{sgn}(m_1)$; accordingly, we observe a single edge mode of chirality $\chi_{\text{edge}} = \text{sgn}(m_1)$ for $|k_y| < |m_1|$.

For other parameter choices, $\Delta w_1$ can be fractional. For edge modes shown in Fig. 2(a), we find $\Delta w_1 = -1/2$, and for Fig. 2(b), $\Delta w_1 = 1/2$. Edge modes in these cases resemble the anomalous edge modes found in Ref. [16]. Clearly, $w_1$ alone is insufficient to characterize these modes, which are asymmetric in $k$.

To classify the anomalous edge modes, we introduce a second winding number using the complex magnitude of B: $B = \lambda^+$. Near each EP, the eigenvalues form “half-vortices”: $\lambda^+ \propto \pm \sqrt{|k - k_{\text{EP}}|} \exp[\pm i\text{Arg}(k - k_{\text{EP}})]/2$, where $s/2$ is the vortex topological charge. We define

$$w_2 = \frac{1}{2\pi} \int_{k_y=-\infty}^{k_y=+\infty} \nabla_k \text{Arg}(\lambda^+) \cdot dk,$$

where the integral is again taken similarly to Eq. (8). For the spectrum (2), we find [33]

$$w_2 = \begin{cases} 0 & |sk_y| < |m| \\ \frac{1}{2} \text{sgn}(sk_y) & |sk_y| > |m|. \end{cases}$$

This winding number has the required asymmetry in $k_y$. Whenever $w_2 \neq 0$, there are branch cuts in $\lambda^+$, and $\dot{H}$ cannot be continuously deformed into a gapped Hermitian system. Unlike $w_1$, which is a generalization of the Berry phase, the $w_2$ winding number is specific to non-Hermitian systems and has no direct Hermitian counterpart.

Using $\Delta w_1$ and $\Delta w_2$, we can completely characterize the edge modes shown in Fig. 3. First, for $w_2 = 0$, the existence of Hermitian-like edge modes (and their chirality) is determined by $\Delta w_1$. Second, for $\Delta w_2 \neq 0$, the number of anomalous (non-Hermitian and mixed) edge modes is $2|\Delta w_2|$, while $\text{sgn}(\Delta w_2)$ determines whether they are localized to the left or right edge of medium 1. In Fig. 3, the anomalous non-Hermitian edge modes only exist on the right edge when $\Delta w_3 < 0$. In particular, the purple region in Fig. 3(b) corresponds to $\Delta w_2 = -1$, and accordingly there are two anomalous edge modes with opposite chiralities ($\Delta w_1 = 0$), and both are defective. Thus, the winding numbers $w_{1,2}$ provide the bulk-edge correspondence for the non-Hermitian Hamiltonian (5) and describe topological properties of the edge modes Fig. 3.

Since $w_1$ and $w_2$ only change when $k_y$ crosses an EP, we can identify the “topological charges” of the individual EPs as $(q_1, q_2) = \frac{1}{2}(\pm |m|, \pm x)$. There are four inequivalent non-Hermitian degeneracies, in contrast to the two inequivalent Hermitian degeneracies. This is a consequence of the richer morphologies of complex vector fields that parametrize non-Hermitian Hamiltonians [42].

Index theorem.—Another way to analyze the zero-energy modes (zero modes) of the non-Hermitian Hamiltonian (5) is to consider the Hermitian Hamiltonian

$$\hat{H} = \hat{H}^\dagger \hat{H}.$$

Zero modes of $\hat{H}$ are also zero modes of $\hat{H}$, and vice versa. When $s = \pm 1$ is a constant, we find that

$$\hat{H} = -i \nabla - \partial_x A(x,y) - i \partial_y B(x,y),$$

where $B(x,y) = \partial_x A_y - \partial_y A_x$ and $A = (0, m)$. This is a Pauli-type Hamiltonian for a nonrelativistic particle in a matrix-valued vector potential [43].

The normalizable zero modes of $\hat{H}$ can now be counted by an “index-theorem” argument [31]. The result is that there are $N = \lfloor |\Phi|/2\pi \rfloor$ such modes, where $\Phi$ is the total flux of $B$. This holds for arbitrary complex analytic mass fields $m(x,y)$. For the previously considered special case of two media with a straight interface, there is a flux of $(m_2 - m_1)$ per unit length along the domain wall, implying that the zero modes occupy a $k$ range of $\Delta k = m_2 - m_1$, in precise agreement with Figs. 2 and 3(a) (see details in Ref. [33]).

Discussion.—We have analyzed a 2D non-Hermitian continuum model that exhibits different types of zero-energy
edge modes, which can be classified using two half-integer-valued winding numbers calculated from the complex bulk band structure. These are inherently associated with topological properties of bulk eigenmodes and non-Hermitian degeneracies (EPs) in the band structure. One family of edge modes includes the well known (Hermitian) Jackiw-Rebbi zero modes [23,24]. However, the classification also contains essentially non-Hermitian edge modes that cannot be continued into Jackiw-Rebbi-type edge modes; these seem to be continuum counterparts of the anomalous edge modes recently encountered in certain 1D non-Hermitian lattice models [14–16].

The three families of non-Hermitian topological edge modes can be realized in a non-Hermitian 2D photonic resonator lattice [6,14,15,44–47], with non-Hermiticity introduced through either asymmetric scattering between clockwise and anticlockwise modes [4–6,14,15] or amplifying or lossy inter-resonator coupling [45–47]. In the Supplemental Material [33], we show that lattice and interface orientations can be chosen to yield different values of \( w_1 \), \( w_2 \) and, correspondingly, different families of zero-energy edge modes [37].

We have focused on the case of two uniform media separated by the line \( x = 0 \). For other orientations of a straight interface, we obtain similar phase diagrams, taking \( k = k \hat{x} \) where \( k \) is the wave vector parallel to the interface. The index-theorem derivation of the number of normalizable zero modes is even more general and applies to arbitrary analytic mass fields. The above features, and comparisons with previously known examples, suggest that the variety of chiral edge modes and topological numbers found in this work may be generic to a wide class of non-Hermitian wave systems.

The idea of this research was conceived by K. Y. B., C. Y. D., and D. L. at the workshop “Topological States of Light and Beyond” in the Center for Theoretical Physics of Complex Systems, Institute for Basic Science, Daejeon, South Korea. We are grateful to A. Miroshnichenko, A. B. Khanikaev, and S. Flach for their hospitality and inspiring discussions. This research was supported by the Singapore National Research Foundation (Grant No. NRFF2012-02), the Singapore Ministry of Education (MOE) Academic Research Fund Tier 2 (Grant No. MOE2015-T2-2-008), the RIKEN Interdisciplinary Theoretical Science Research Group (iTHES) Project, the Multi-University Research Initiative (MURI) Center for Dynamic Magneto-Optics via the Air Force Office of Scientific Research (AFOSR) (Grant No. FA9550-14-1-0040), Grant-in-Aid for Scientific Research (A), Core Research for Evolutionary Science and Technology (CREST), the John Templeton Foundation, and the Australian Research Council.


[37] Similar to the doubling of Hermitian Dirac points required by the Nielsen-Ninomiya theorem, a lattice regularization doubles the degeneracies and introduces a large-wave-number cutoff to the edge modes; see Ref. [33].