



Tutorial: Topology, Waves, and the Refractive Index

S. A. R. Horsley¹

Received: 9 February 2023 / Accepted: 23 April 2023 / Published online: 20 June 2023
© The Author(s) 2023

Abstract

This tutorial is divided into two parts: the first examines the application of topology to problems in wave physics. The origins of the Chern number are reviewed, where it is shown that this counts the number of critical points of a complex tangent vector field on the surface. We then show that this quantity arises naturally when calculating the dispersion of modes in any linear system, and give examples of its application to find one-way propagating interface modes in both continuous and periodic materials. The second part offers a physical interpretation for the Chern number, based on the idea that the critical points which it records can be understood as points where the refractive index vanishes. Using the theory of crystal optics, we show that when the refractive index vanishes in a *complex valued* direction, the wave is forced to circulate in only one sense, and this is the origin of the one-way propagation of topological interface states. We conclude by demonstrating that this idea of ‘zero refractive index in a complex direction’ can be used as a shortcut to find acoustic and electromagnetic materials supporting one-way interface states.

Keywords Wave propagation · Topology · Metamaterials

1 Introduction

Topology is the study of whether objects can be smoothly transformed into each other. Sometimes these ‘objects’ are extremely abstract mathematical ideas, sometimes they’re not: Can I uncoil this garden hose without removing the end from the bucket? Can I untangle these necklaces without undoing the clasps? Can I wrap this map of the world onto a globe? These are all problems of topology. Related topological questions appear across physics, where wave dispersion surfaces [1], knotted fluid flow lines [2], and electromagnetic fields [3] can all be grouped according to whether, or not they can be smoothly transformed into one another.

In physics and engineering there has been a recent burst of activity, applying topology to control waves. Specifically, topology has been applied to design materials, stipulating what happens at their interface without having to know anything about what the interface is like. For instance, take two homogeneous lumps of elastic stuff. We can ensure that vibrational

✉ S. A. R. Horsley
s.horsley@exeter.ac.uk

¹ School of Physics and Astronomy, University of Exeter, Stocker Road, EX4 4QL Devon, UK

waves can be trapped at the interface formed when we stick them together, irrespective of how messy our joinery is!

The basic idea is this: take a material where wave propagation can be specified in terms of a conserved wave vector \mathbf{k} . For topology to be at all powerful, we need to be able to treat the components of \mathbf{k} as the coordinates on a closed surface. In many cases this is possible. We can, for example, wrap the first Brillouin zone onto a torus, whenever the material is periodic.

Imagine that attached to each point on this closed surface is a vector $|\psi\rangle$; the solution to the wave equation for each particular value of \mathbf{k} . Now, taking two different materials we can construct two closed surfaces (e.g. two tori), upon each of which there is a different form of the wave, $|\psi\rangle$ or $|\psi'\rangle$. The question is whether it is possible to smoothly deform one of these wave fields into the other, a question topologists have already developed the necessary tools to answer, at least in the negative!

If the two waves *cannot* be smoothly deformed into one another then something non-smooth and perhaps ‘interesting’ must happen when we try, something that will occur for instance at an interface, in the transition region between the two materials. This ‘interesting’ thing turns out to be the presence of one or more interface states, where the wave is trapped in the transition region. Topology therefore guarantees the presence of interface states between two materials, without the physicist ever having to consider whether the interface is flat, rough, curved, sharp, narrow, or wide.

This is an odd business for most physicists and engineers, who are used to caring about details! A graded index fiber optic cable, for instance, must be made with precision, confining light rays with a particular spatial distribution of refractive index to minimize dispersion. Here we have a completely different kind of theory, where the design process doesn’t even mention the details of the region where the mode is to be confined. This peculiar insensitivity to the form of the interface (often called ‘topological protection’) was perhaps first appreciated by Volkov and co-workers, who were concerned with the physics of electrons around electronic contacts [4, 5]. The significance of topology was appreciated later [6, 7], in connection with earlier work on the quantum hall effect [8, 9].

Since the discovery of negative refraction [10] and transformation optics [11], and the rapid development of metamaterials [12], it has been widely appreciated that one wave is like any other. Although quantum mechanics has its peculiarities, there is nothing fundamentally different between the Dirac equation, Schrödinger equation, the classical Maxwell equations, or the equations of elasticity. The same topological arguments have therefore been applied to design of periodic electromagnetic materials supporting unidirectional interface states, dubbed ‘photonic topological insulators’ [13], which has led to the fields of topological photonics [14] and acoustics [15].

The fact that we can use metamaterials to realise a wide range of material parameters means there is actually more to explore in these classical systems (the distribution of atoms in an ordinary crystal is not easy to specify on the scale of an electron wavelength!). Work on topological photonic and acoustic materials is therefore able to investigate effects that could not be observed in condensed matter systems (see the discussion in e.g. [16] and [17]), and even allows the exploration of active topological non-Hermitian materials, where the material can amplify an absorb the wave in a controlled way [18] (which is also extremely difficult to mimic in electronic systems).

In the author’s opinion, several things are opaque in this subject. Firstly, while the Chern classes are by now a familiar tool for classifying the topology of a wave field $|\psi\rangle$, their origin and connection to other characteristic classes—such as the more familiar Euler class—is never explained in terms palatable to the physicist. Secondly, besides giving a “plausible”

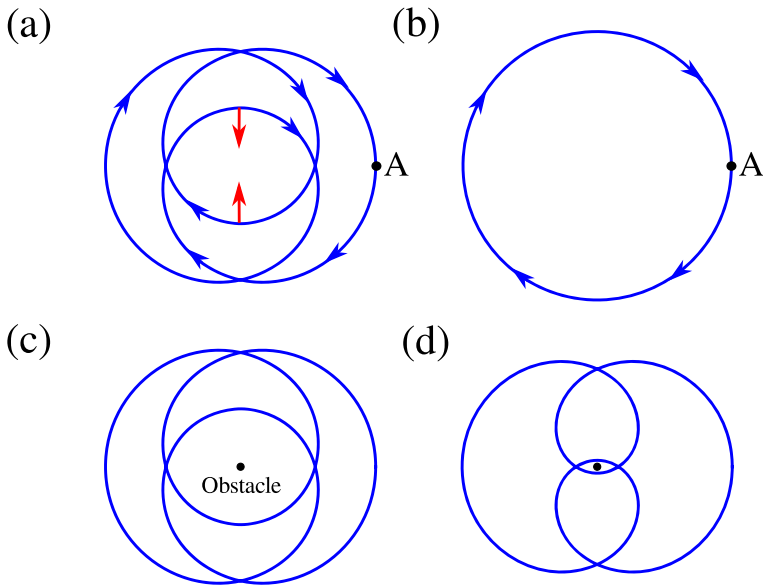


Fig. 1 Without any obstacles, any closed path can be deformed into a simple loop. By pulling along the indicated red arrows, the path in (a) can be unwound into the simple path shown in (b). In panel (c) an obstacle is placed in the centre of the path. As the path winds twice around the obstacle, one reaches an impasse as shown in (d)

intuitive explanation for its truth, the connection between integrals of the Chern class and the number of interface states is never proved in a straightforward way. The first part of the tutorial clarifies both of these points. Finally there is the problem of a physical interpretation for these topological calculations. In the final part we connect a non-zero Chern number to the existence of peculiar points of vanishing refractive index, where the wave is forced to circulate in only one sense (e.g. only clockwise).

2 Winding Numbers of Paths

We'll begin with the topological classification of curves in terms of a 'winding number'. Imagine unwinding a ball of string, one end of which (A) is attached to a wall. With the string we trace out a path, closing it by returning to A and tying the ends. An example is sketched as the blue curve in Fig. 1a.

If there are no obstacles, we can always move the string until it unwinds into a single loop. Pulling in the direction indicated in Fig. 1a, the string can be untwisted into the single loop shown in Fig. 1b. The two configurations are therefore *topologically equivalent*. This situation changes if the space contains an obstacle¹. Suppose there is a tree at the position shown by the black dot in Fig. 1c. Whether we can deform one arrangement of string into another is now determined by the *winding number*, v : the number of times the string encircles the tree. As we know from experience, the winding number cannot be changed by any continuous re-positioning of the string, to change it we must cut the string (or cut down the tree). A difference in the winding number between two configurations of string indicates that they

¹ The obstacle is equivalent to the removal of a point from the plane.

are *topologically in-equivalent*, and ν can be used to classify configurations of string that cannot be continuously changed into one another.

To calculate the winding number, suppose we position the tree at the origin of the coordinate system, integrating the change in polar angle $d\theta$ as we follow the string. After dividing by 2π , the integral counts the number of times the path encircles the obstacle. Each point on the path thus picks out an angle θ on a circle. The winding number counts the number of times this point covers the circle, adding $+1$ for each anticlockwise circuit and -1 for each clockwise one.

It is simplest to write the change in angle $d\theta$ using the complex number $z = x + iy = re^{i\theta}$. The polar angle is then simply $\theta = \text{Im}[\log(z)]$ and the winding number can be written as a line integral of a vector,

$$\nu = \frac{1}{2\pi} \oint_{\text{Path}} d\theta = \frac{1}{2\pi} \oint_{\text{Path}} \nabla \text{Im} [\log(x + iy)] \cdot d\mathbf{x} \tag{1}$$

where $d\mathbf{x} = dx \mathbf{e}_x + dy \mathbf{e}_y$. Having written ν as a line integral, there is a second equivalent way to write (1). Using Stokes’ theorem the one dimensional line integral can be written as a two dimensional surface integral over the enclosed region S_P ,

$$\nu = \frac{1}{2\pi} \int_{S_P} \nabla \times \mathbf{A} \cdot d\mathbf{S} \tag{2}$$

where S_P may be quite a strange origami-like surface, like that enclosed by the blue curve in Fig. 1a. This is an important development that we’ll see again. The winding number (our topological invariant) now appears as a net ‘magnetic flux’ through the surface S_P , something which common to all the topological invariants considered here. The ‘vector potential’ associated with this flux is defined as

$$\mathbf{A} = \nabla \text{Im} [\log(x + iy)]. \tag{3}$$

But having said all this, it now seems as though we made a mistake. The integral (2) is surely always zero, because the curl of any gradient is zero!

We didn’t make a mistake. We have uncovered an extremely important subtlety that appears again and again in topology. The concerning result $\nabla \times \nabla f = 0$ requires f to be a proper function. Every point (x, y) must be associated with a single number $f(x, y)$. This isn’t true for the ‘function’, $\theta = \text{Im}[\log(x + iy)]$, which can take *any* value at the origin.² This defect in the vector potential (3) is known as a critical point. The ‘flux’ in equation (2) records the presence of such critical points, and is confined to the obstacle, where both θ and \mathbf{A} are undefined. The winding number thus only depends on the number of times the surface S_P cuts through the critical point $\mathbf{x} = \mathbf{0}$. The fact that the curl of (3) is zero at all points *except* where there is an obstacle is actually essential for the winding number to be insensitive to deformations of the path, and hence for ν to be a topological invariant.

3 The Euler Characteristic: Winding Numbers of Surfaces

Having given a topological categorization of curves, we move up a dimension to classify the ‘winding’ of closed two dimensional surfaces, like those shown in Fig. 2. Just as we classified

² Mathematicians refer to $d\theta$ as closed (has zero curl away from the origin) yet inexact (isn’t the gradient of a proper function) one-form. De Rham cohomology relates the topology of a space to the existence of such forms.

the path of our string by mapping it to a point on a circle and counting the total number of revolutions, we associate each point on a surface S to an equivalent point on a *sphere*, and count the number of times the sphere is covered as we move over the surface.

We make the connection between a point on an arbitrary surface S and a point on a sphere through the surface normal \mathbf{n} . For each coordinate (x_1, x_2) on our surface we find those coordinates (θ, ϕ) on the sphere where the surface normal vector takes the same value,

$$\mathbf{n}(x_1, x_2) = \sin(\theta) [\cos(\phi)\mathbf{e}_x + \sin(\phi)\mathbf{e}_y] + \cos(\theta)\mathbf{e}_z. \tag{4}$$

as indicated in Fig. 2. In Sec. 2 we mapped the curve onto a circle and calculated the winding number through integrating the angle swept out around the circle. Here we instead integrate up the *solid angle* $d\Omega = \sin(\theta) d\theta d\phi$ swept out on the sphere as we move over the area

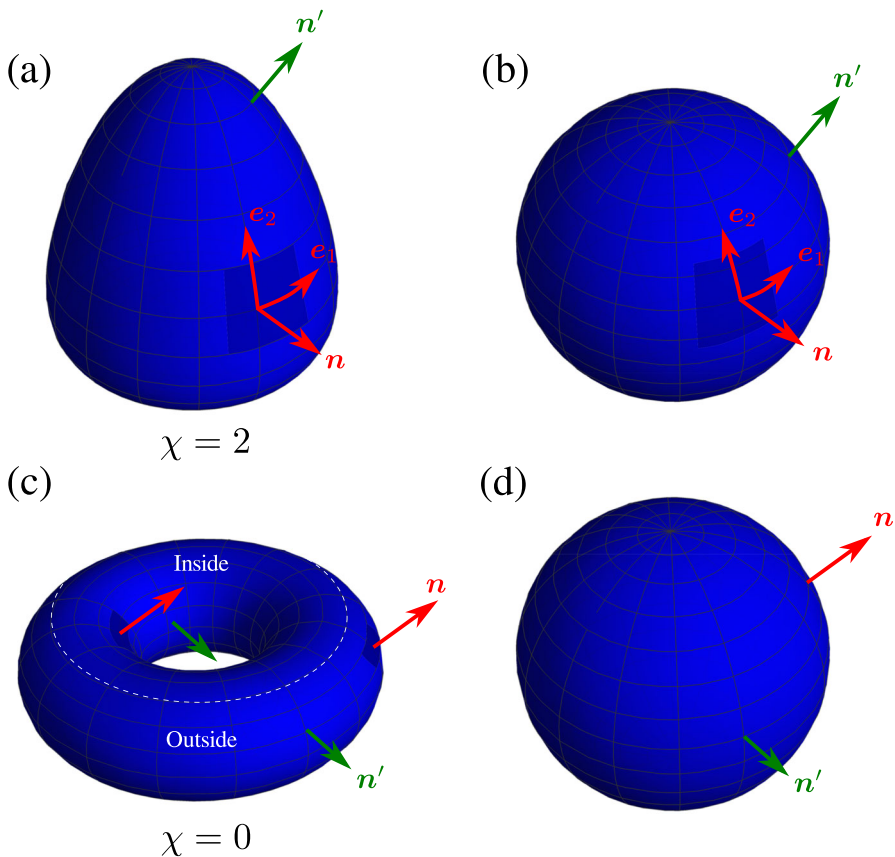


Fig. 2 Two dimensional surfaces can be classified by counting the number of times the surface normal \mathbf{n} covers a sphere, which equals half of the *Euler characteristic*, χ . Panels (b) and (d) show arbitrary surface normal vectors \mathbf{n} and \mathbf{n}' on a sphere, picking out points on the two surfaces in (a) and (c) respectively. Every surface normal on the sphere corresponds to only one point on the egg-like surface in (a), meaning $\chi = 2$. Meanwhile a point on the sphere corresponds to *two* points on the torus: one ‘inside’, and one ‘outside’. Following \mathbf{n} around the torus we see that the outside and inside regions cover the sphere in opposite directions, one undoing the other such that $\chi = 0$

$dx_1 dx_2$ on our arbitrary surface S . A useful expression for $d\Omega$ can be found through transforming the expression for the solid angle from spherical to surface coordinates,

$$\begin{aligned} d\Omega &= \mathbf{n} \cdot \left(\frac{\partial \mathbf{n}}{\partial \theta} \times \frac{\partial \mathbf{n}}{\partial \phi} \right) d\theta d\phi = \mathbf{n} \cdot \left(\frac{\partial \mathbf{n}}{\partial \theta} \times \frac{\partial \mathbf{n}}{\partial \phi} \right) \left(\frac{\partial \theta}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \theta}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx_1 dx_2 \\ &= \mathbf{n} \cdot \left(\frac{\partial \mathbf{n}}{\partial x_1} \times \frac{\partial \mathbf{n}}{\partial x_2} \right) dx_1 dx_2. \end{aligned} \tag{5}$$

By analogy with our calculation of the winding number of a curve (1), we count the number of times ν the surface normal \mathbf{n} wraps around the sphere, simply integrating the solid angle element (5) over the surface S , and dividing the result by 4π

$$\nu = \frac{\chi}{2} = \frac{1}{4\pi} \int_S d\Omega = \frac{1}{4\pi} \int_S \mathbf{n} \cdot \left(\frac{\partial \mathbf{n}}{\partial x_1} \times \frac{\partial \mathbf{n}}{\partial x_2} \right) dx_1 dx_2. \tag{6}$$

The winding number is half the *Euler characteristic*, $\chi = 2\nu$ which is the topological invariant typically used to classify closed surfaces. We should remember that this is simply a way of expressing the winding number. Equation (6) is the essence of the topological classification of surfaces: there is no way to smoothly change one surface into another if their normal vectors cover a sphere a different number of times. Two such incompatible surfaces are shown in Fig. 2a and c.

There is a more interesting way to write the Euler characteristic (6), that replaces the change in the surface normal with the local surface curvature. To make this transition we take some point on the surface, and use it as origin of a Cartesian coordinate system, where the $z = 0$ plane is tangent to the surface, as shown in Fig. 3. Close to this point, the surface shape satisfies

$$z - h(x, y) = 0 \tag{7}$$

where $h(x, y)$ is the height of the surface above the tangent plane. By definition the height and its gradient vanish at the origin. The surface normal is now proportional to the gradient of the above equation for the surface height (7),

$$\mathbf{n} = N (\mathbf{e}_z - \nabla h(x, y)). \tag{8}$$

where the scalar N ensures normalization, $\mathbf{n} \cdot \mathbf{n} = 1$ and equals unity at $x = y = 0$, where $\nabla h = 0$. Using the x and y coordinates in the formula for the element of solid angle (5), and substituting the above expression for the surface normal vector (8), we can re-write the element of solid angle at $x = y = z = 0$ in terms of the Hessian (the matrix of second derivatives with respect to x and y) of the surface height,

$$\begin{aligned} d\Omega &= \mathbf{n} \cdot \frac{\partial \mathbf{n}}{\partial x} \times \frac{\partial \mathbf{n}}{\partial y} dx dy = \mathbf{e}_z \cdot (\nabla \partial_x h \times \nabla \partial_y h) dx dy \\ &= \det \left(\begin{array}{cc} \frac{\partial^2 h}{\partial x^2} & \frac{\partial^2 h}{\partial x \partial y} \\ \frac{\partial^2 h}{\partial x \partial y} & \frac{\partial^2 h}{\partial y^2} \end{array} \right) dx dy. \end{aligned} \tag{9}$$

The determinant of the Hessian is positive for surfaces that are locally elliptic paraboloids, and negative for hyperbolic paraboloids. It equals the inverse product of the two principal radii of curvature, $\det[\partial^2 h / \partial x_i \partial x_j] = (R_1 R_2)^{-1}$, which is known as the surface's *Gaussian curvature* K (see Fig. 3). This argument can be carried out at every point on the surface. Summing the results we find the Euler characteristic (6) can also be re-written as an integral of the curvature of the surface

$$\chi = \frac{1}{2\pi} \int_S K dA \tag{10}$$

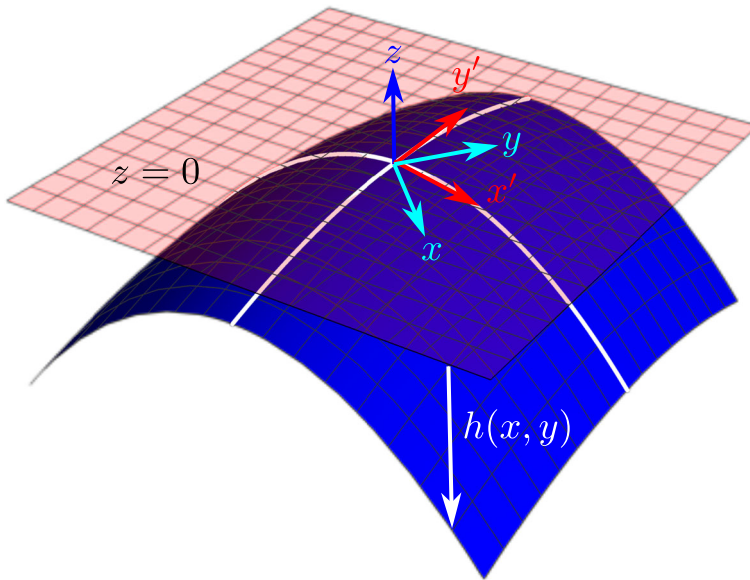


Fig. 3 At any point on a surface the infinitesimal element of solid angle $d\Omega$ appearing in (6) can be re-written in terms of the Hessian of the surface height function $h(x, y)$. Rotating the in-plane coordinates to x', y' to align with the principal axes of curvature (white lines), the Hessian is diagonalized with eigenvalues equal to the inverses of the two radii of curvature (the radii of the two circles that approximate the white curves in the diagram)

where dA is an infinitesimal element of surface area. Equation (10) is the famous *Gauss–Bonnet theorem*, and is quite a remarkable expression. The surface curvature integrated over any closed surface always equals a multiple of 2π . We can continuously deform the surface, changing the distribution of surface curvature, but—so long as we don't tear a new hole in the surface—every region of increased curvature is unavoidably balanced by regions where it is reduced.

In parallel with our earlier discussion of one dimensional curves, we can write the Gauss–Bonnet theorem in a third equivalent form, as integral of an effective magnetic flux passing through the closed surface. This is achieved through introducing a pair of orthonormal tangent vectors on the surface, \mathbf{e}_1 and \mathbf{e}_2 . The surface normal is everywhere given by the cross product between these tangent vectors

$$\mathbf{n} = \mathbf{e}_1 \times \mathbf{e}_2. \tag{11}$$

Substituting expression (11) for the surface normal into our expression for the solid angle element (5), we see that it equals the curl of a vector

$$\begin{aligned} d\Omega &= \mathbf{n} \cdot \frac{\partial \mathbf{n}}{\partial x_1} \times \frac{\partial \mathbf{n}}{\partial x_2} dx_1 dx_2 = \left[\left(\mathbf{n} \cdot \frac{\partial \mathbf{e}_2}{\partial x_2} \right) \left(\mathbf{n} \cdot \frac{\partial \mathbf{e}_1}{\partial x_1} \right) - \left(\mathbf{n} \cdot \frac{\partial \mathbf{e}_1}{\partial x_2} \right) \left(\mathbf{n} \cdot \frac{\partial \mathbf{e}_2}{\partial x_1} \right) \right] dx_1 dx_2 \\ &= \left[\frac{\partial \mathbf{e}_1}{\partial x_1} \cdot \frac{\partial \mathbf{e}_2}{\partial x_2} - \frac{\partial \mathbf{e}_1}{\partial x_2} \cdot \frac{\partial \mathbf{e}_2}{\partial x_1} \right] dx_1 dx_2 \\ &= \left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} \right) dx_1 dx_2. \end{aligned} \tag{12}$$

where the components of this two component surface ‘vector potential’ A_i are defined as

$$A_j = \mathbf{e}_1 \cdot \frac{\partial \mathbf{e}_2}{\partial x_j}. \tag{13}$$

Note that, despite appearances there is no bias between the two tangent vectors, and the vector potential can be written equivalently as $A_i = -\mathbf{e}_2 \cdot \partial_i \mathbf{e}_1$, due to the normalization condition $\mathbf{e}_1 \cdot \mathbf{e}_2 = 0$. Although we have called A_j a ‘vector potential’ due to the analogous quantity in physics, more precisely the quantity A_j is a *connection* on the surface, a quantity from differential geometry that characterises how the basis vectors change from point to point (see Appendix A for details).

Our three different expressions for $d\Omega$ show that the Gaussian curvature both expresses a change in the surface normal \mathbf{n} , and an effective ‘magnetic field’ (the curl of the connection) due to the change in the surface tangent vectors \mathbf{e}_i

$$K = \mathbf{n} \cdot (\nabla \times \mathbf{A}) = \mathbf{n} \cdot \left(\frac{\partial \mathbf{n}}{\partial x_1} \times \frac{\partial \mathbf{n}}{\partial x_2} \right) \tag{14}$$

Using the above expression for the solid angle in terms of the vector potential (12) we can also write the Gauss–Bonnet theorem as the integral of a ‘magnetic flux’ passing through the surface

$$\chi = \frac{1}{2\pi} \int_S \nabla \times \mathbf{A} \cdot d\mathbf{S}. \tag{15}$$

where the vector surface area element is given by $d\mathbf{S} = dx_1 dx_2 \mathbf{n}$. Note that the Euler characteristic (15) now takes an identical form to the winding number of a one dimensional curve (2). Just as we saw there, the integral (15) doesn’t seem right. Stokes’ theorem tells us that the integral of a curl over a surface equals a line integral around the surface boundary. But these closed surfaces have no boundary! So surely the integral is zero.

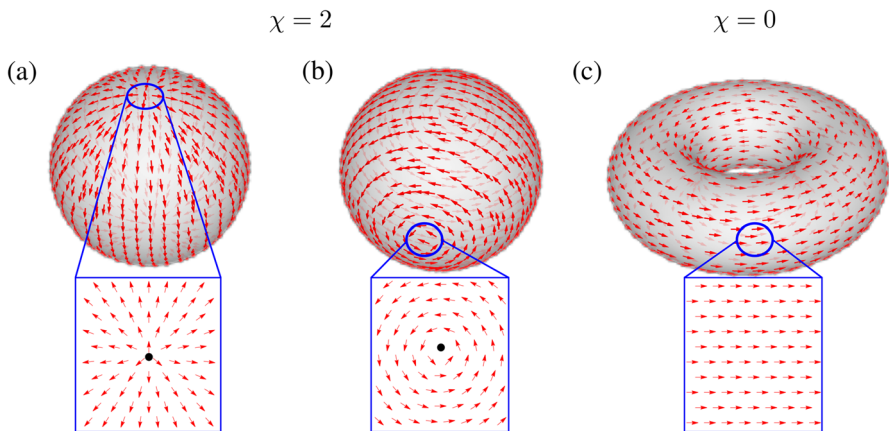


Fig. 4 Instead of the surface normal \mathbf{n} , we can use tangent vector fields $\mathbf{e}_1, \mathbf{e}_2$ to calculate the Euler characteristic. The Euler characteristic χ can then be written as the surface integral of a ‘magnetic flux’ as in (15). The topology of the surface is determined by the number of critical points, where either of the tangent vectors are undefined. When $\chi \neq 0$, as for the sphere in panels (a) and (b), every choice of tangent vector field will exhibit critical points, shown here as the black dots in the zoomed in regions. Only when $\chi = 0$ is it possible to have a tangent vector that is well defined at all points. An example is shown on the surface of the torus in panel (c)

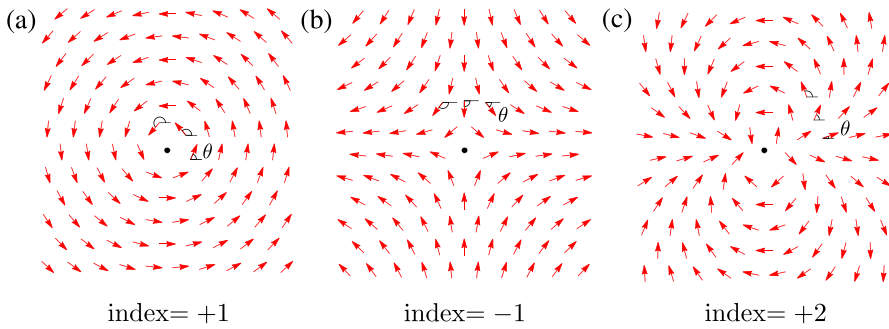


Fig. 5 A critical point of a normalized vector field is a point where its direction is undefined. Critical points are classified in terms of their index, which equals the number of times the vector rotates as we move around the critical point, being positive for an anti-clockwise rotation. In (a–c) we show three critical points (black dots), each of a different index

But, after Sec. 2 we are prepared for this puzzle. Stokes’ theorem can only be applied if the vector potential A is well defined over the whole surface. We have an integral of something that *looks* like a curl over a closed surface, but it isn’t the curl of anything at some discrete points (critical points) on the surface. These critical points are familiar for the polar and azimuthal unit vectors on a sphere, which are both undefined at the poles: it is not always possible to have tangent vectors e_i that are normalized, orthogonal, and everywhere well defined³. In general, tangent vectors on a closed surface exhibit *critical points* where they do not have a well defined direction, as illustrated in Fig. 4. If we exclude a small region surrounding each of these critical points from the surface integral in (15) (A is now well defined everywhere), we can transform (15) to a sum of line integrals encircling the critical points and take the limit as the area of each of these excluded regions tends to zero. The contribution of these infinitesimal excluded regions is not zero. Each of these points will contribute a multiple of 2π ,

$$\oint A_j dx_j = \oint e_1 \cdot \frac{\partial e_2}{\partial x_j} dx_j = \oint e_1 \cdot de_2 = \oint d\theta = 2\pi n. \tag{16}$$

The last three steps follow from the infinitesimal change in the tangent vector, $e_1 \cdot de_2 = d\theta$, where $d\theta$ is angle by which the vector e_2 rotates due to the infinitesimal displacement dx_j . The integer n is known as the *index* of the critical point. Examples of critical points of different index are shown in Fig. 5.

From Eqns. (15) and (16) we conclude that the Euler characteristic of a surface records the critical points of any tangent vector field on the surface. The sum of the indices n_{x_c} of all the critical points x_c of any surface tangent vector field equals the Euler characteristic

$$\chi = \sum_{x_c} n_{x_c}. \tag{17}$$

This result is known as the *Poincaré–Hopf theorem*. As the Euler characteristic is a topological invariant we can thus conclude that however we deform the tangent vector field on a closed

³ The impossibility of having an everywhere well defined tangent vector field on the surface of an even-dimensional sphere is known as the *Hairy ball theorem*. We are familiar with this in everyday life: it is impossible to comb the hair on a sphere to lie flat without introducing a crown, around which the hair swirls or diverges.

surface, the critical points cannot all be eliminated, unless the surface has the topology of a torus, $\chi = 0$.

Example: The Euler Characteristics of the Torus and the Sphere

Points on the surfaces of both a torus and a sphere can be parameterized in terms of two cyclic coordinates, x_1 and x_2 ,

$$\mathbf{r}(x_1, x_2) = a \cos(x_1) \mathbf{e}_z + (R + a \sin(x_1)) (\cos(x_2) \mathbf{e}_x + \sin(x_2) \mathbf{e}_y), \tag{18}$$

where R is the distance from the origin to the centre of the torus ‘tube’ of radius a . The surface makes a topological transition when the distance of the centre of the tube from the origin equals its radius $R = a$, at which point the innermost circle of points on the torus becomes a single point at the origin, and the topology changes to that of a sphere. For $R < a$ the x_1 coordinate has the reduced range $x_1 \in [-\arcsin(R/a), \pi + \arcsin(R/a)]$, which becomes $x_1 \in [0, \pi]$ when $R = 0$, where (18) describes the surface of a sphere.

Tangent vectors on the surface can be found through differentiating (18) with respect to the two coordinates, which after normalization gives the orthogonal pair of vectors

$$\begin{aligned} \mathbf{e}_1 &= -\sin(x_2) \mathbf{e}_x + \cos(x_2) \mathbf{e}_y \\ \mathbf{e}_2 &= -\sin(x_1) \mathbf{e}_z + \cos(x_1) (\cos(x_2) \mathbf{e}_x + \sin(x_2) \mathbf{e}_y). \end{aligned} \tag{19}$$

When the surface has the topology of a torus, the tangent vectors (19) are uniquely defined at all points, and therefore the Euler characteristic equals zero

$$\chi = \frac{1}{2\pi} \sum_n \oint_{C_n} A_j dx_j = 0 \quad (\text{torus}) \tag{20}$$

where the sum runs over the critical points (of which there are non in the case of a torus), each encircled by C_n .

When the surface becomes a sphere ($R = 0$), the two points $x_1 = 0$ and $x_1 = \pi$ are critical points of the tangent vectors (19), being isolated points where the tangent vectors take many possible values, depending on how we approach the point. Using expressions (19) we can calculate the ‘vector potential’ from our earlier formula (13),

$$\begin{aligned} A_1 &= \mathbf{e}_1 \cdot \frac{\partial \mathbf{e}_2}{\partial x_1} = 0 \\ A_2 &= \mathbf{e}_1 \cdot \frac{\partial \mathbf{e}_2}{\partial x_2} = \cos(x_1). \end{aligned} \tag{21}$$

The Euler characteristic then equals the sum of the line integrals of A_j around the critical points

$$\chi = \frac{1}{2\pi} \left[\int_0^{2\pi} A_2(x_1 = 0) dx_2 + \int_{2\pi}^0 A_2(x_1 = \pi) dx_2 \right] = 2 \quad (\text{sphere}) \tag{22}$$

where the line integral around the south pole is taken in the opposite direction, due to the reversal of the surface normal. We have shown the Euler characteristic of a sphere equals 2, as expected from the observation that χ is twice the winding number of the surface normal around a sphere.

4 The Berry Connection and the Euler Characteristic

So far we’ve illustrated something of the basics of topology, but have not made much connection with physics, which is supposed to be why we’re here! We can begin to see the connection by re-writing the formula for the Euler characteristic (15) in terms of a single complex tangent vector field (a so-called complex line bundle)

$$|\psi\rangle = \frac{1}{\sqrt{2}}(\mathbf{e}_1 + i\mathbf{e}_2), \tag{23}$$

where the state is normalized such that $\langle\psi|\psi\rangle = 1$ and we have adopted the bra-ket notation for vectors and inner products. In terms of this complex vector, the ‘vector potential’ (13) appearing in the Gauss–Bonnet theorem takes a simpler form

$$A_j = \mathbf{e}_1 \cdot \frac{\partial \mathbf{e}_2}{\partial x_j} = -i\frac{1}{2}(\mathbf{e}_1 - i\mathbf{e}_2)\frac{\partial}{\partial x_j}(\mathbf{e}_1 + i\mathbf{e}_2) = -i\left\langle\psi\left|\frac{\partial}{\partial x_j}\right|\psi\right\rangle \tag{24}$$

where we used the normalization conditions $\mathbf{e}_i \cdot \mathbf{e}_i = 1$, which implies $\mathbf{e}_i \cdot \partial_j \mathbf{e}_i = 0$. The final expression on the right of (24) can be recognised at once as the *Berry connection* [19]. An analogous quantity appears in quantum theory, and that tells us how to transport a quantum mechanical state vector $|\psi\rangle$ around a space of parameters x_i . For our tangent vector $|\psi\rangle$, the curl of the corresponding Berry connection is—via (14)—simply the Gaussian curvature of the surface.

Using the complex vector $|\psi\rangle$, we can—again, analogous to quantum theory—understand it as an eigenvector of a Hermitian operator $\hat{L} = i\mathbf{n} \times$,

$$\hat{L} = i\mathbf{n} \times \rightarrow \hat{L}|\psi\rangle = |\psi\rangle.. \tag{25}$$

To understand the origins of this operator note that $\mathbf{n} \times (\mathbf{e}_1 + i\mathbf{e}_2) = -i(\mathbf{e}_1 + i\mathbf{e}_2)$. Having introduced the operator \hat{L} we can calculate the Gauss–Bonnet theorem in yet another way! Not only does the Euler characteristic record critical points of the complex vector $|\psi\rangle$ through the Berry connection (24), but these critical points arise from the properties of the operator \hat{L} , of which $|\psi\rangle$ is an eigenfunction.

To see this we calculate the curl of the Berry connection (24)—which can be used in the formula for the Euler characteristic (15)—and use the eigenvalue relation (25) to replace derivatives of the vector with those of the operator \hat{L} ,

$$\begin{aligned} \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} &= -i\left[\left\langle\frac{\partial \psi}{\partial x_1}\left|\chi\right\rangle\left\langle\chi\left|\frac{\partial \psi}{\partial x_2}\right\rangle - \left\langle\frac{\partial \psi}{\partial x_2}\left|\chi\right\rangle\left\langle\chi\left|\frac{\partial \psi}{\partial x_1}\right\rangle\right]\right. \\ &= -i\left[\left\langle\psi\left|\frac{\partial \hat{L}}{\partial x_1}\left|\chi\right\rangle\left\langle\chi\left|\frac{\partial \hat{L}}{\partial x_2}\right|\psi\right\rangle - \left\langle\psi\left|\frac{\partial \hat{L}}{\partial x_2}\left|\chi\right\rangle\left\langle\chi\left|\frac{\partial \hat{L}}{\partial x_1}\right|\psi\right\rangle\right]\right. \\ &= \mathbf{n} \cdot \left(\frac{\partial \mathbf{n}}{\partial x_1} \times \frac{\partial \mathbf{n}}{\partial x_2}\right) = K \end{aligned} \tag{26}$$

where the final line demonstrates the consistency of our different approaches to the Euler characteristic. To derive (26) we used the completeness relation $|\psi\rangle\langle\psi| + |\phi\rangle\langle\phi| + |\chi\rangle\langle\chi| = 1$, where the remaining eigenvectors of \hat{L} are $|\phi\rangle = (\mathbf{e}_1 - i\mathbf{e}_2)/\sqrt{2}$ (eigenvalue -1) and $|\chi\rangle = \mathbf{n}$ (eigenvalue 0). We also applied the result $\partial_i \hat{L}|\chi\rangle + \hat{L}\partial_i|\chi\rangle = 0$.

Equation (26) shows that the curl of the Berry connection is related to the ‘winding’ of the operator \hat{L} . In general this is difficult to picture, but here it is simply another way of telling us the element of solid angle swept out by the surface normal \mathbf{n} on the sphere. We have thus

found yet another method for calculating the Euler characteristic! Not only is it the integral of the curl of the Berry connection (*the Berry curvature*) associated with the complex tangent vector field $|\psi\rangle$, divided by 2π , this can also be written in terms of matrix elements of the operator \hat{L} (25) defined over the surface.

In physics, the integral of the Berry connection $\oint A_i dx_i$ is the phase shift a quantum mechanical wave function obtains after being adiabatically moved through a parameter space with coordinates x_i [19]. Our surface tangent ‘state vector’ $|\psi\rangle$, undergoes the same phase shift as we encircle a critical point. We can see this by examining how the vector changes as we move a small distance on the surface

$$\langle \psi | (|\psi\rangle + d|\psi\rangle) = 1 + \frac{1}{2} (\mathbf{e}_1 - i\mathbf{e}_2) \cdot (-\mathbf{e}_2 + i\mathbf{e}_1) d\theta = (1 + id\theta) = e^{id\theta}. \tag{27}$$

where $d\theta$ is the change in the angle of the basis vectors $\mathbf{e}_{1,2}$ as they rotate around the surface normal \mathbf{n} . Using our earlier equation for the line integral of the vector potential around a critical point (16), we see from the above that the phase accumulated around a critical point is $\oint A_j dx^j = 2\pi n$, where n is the critical point index.

Example: Electromagnetic Polarization

Free space electromagnetic waves are transverse, with the Fourier amplitudes of the fields obeying $\mathbf{k} \cdot \vec{\mathbf{E}} = \mathbf{k} \cdot \vec{\mathbf{H}} = 0$. For a monochromatic field of frequency ω , the length of the wavevector is also fixed by the free space dispersion relation $\mathbf{k}^2 = k_0^2 = \omega^2/c^2$, which defines the surface of a sphere of radius k_0 . Monochromatic radiation is therefore defined by a set of Fourier amplitudes that both live on, and are tangent to a sphere in \mathbf{k} -space.

The electric field in free space can thus be written as an integral over the surface of this sphere,

$$\mathbf{E}(\mathbf{x}) = \int_0^\pi \sin(\theta_k) d\theta_k \int_0^{2\pi} d\phi_k \mathbf{e}(\theta_k, \phi_k) \mathcal{E}(\theta_k, \phi_k) e^{ik_0\mathbf{n}(\theta_k, \phi_k)\cdot\mathbf{x}} \tag{28}$$

where \mathbf{e} represents the direction of the electric field, and \mathcal{E} the scalar amplitude of the wave propagating in direction $\mathbf{n} = \mathbf{k}/k_0$. The polarization vector is chosen to satisfy $\mathbf{e} \cdot \mathbf{e}^* = 1$ and $\mathbf{n} \cdot \mathbf{e} = 0$.

If \mathbf{e} is real valued, the polarization is linear for all directions of propagation. In this case we can use it as the tangent vector \mathbf{e}_1 in (13), with $\mathbf{e}_2 = \mathbf{n} \times \mathbf{e}_1$. The integral of the curl of the ‘vector potential’ defined in (13), $\nabla \times \mathbf{A}$, over the entire sphere in \mathbf{k} space will therefore equal 2, the Euler characteristic of the sphere. *This means that linear polarized fields must have at least one critical point (index 2) in the electric field, as a function of direction, although in general there will be many.* The same argument can also be applied to circularly polarized fields, where at every point the direction of the electric field is of the form $\mathbf{e} = (\mathbf{e}_1 \pm i\mathbf{e}_2)/\sqrt{2}$. As described below (27), there must be a phase vortex around each of these critical points, which can be understood as the Berry phase.

As a concrete example, take radiation from a dipole with dipole moment \mathbf{p} , where the far field electric field takes the form

$$\mathbf{E}(\mathbf{x}) \sim \frac{k_0^2}{4\pi\epsilon_0 r} (\mathbf{p} - \mathbf{e}_r(\mathbf{e}_r \cdot \mathbf{p})) e^{i(k_0 r - \omega t)}. \tag{29}$$

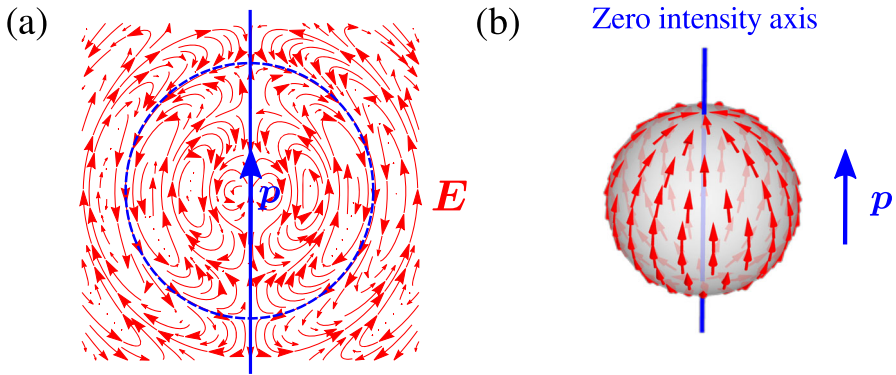


Fig. 6 The nodes in the radiation pattern from an electric dipole are a consequence of topology. (a) Snapshot of the oscillating electric field E from a dipole moment p . Enclosing the dipole in a sphere (the blue curve, for example) there will always be points where the electric field is normal to the sphere, making its direction tangent to the sphere undefined (a critical point). (b) Far away from the dipole, the observation direction picks out a Fourier component of the field (28) and the electric field is everywhere tangent to the sphere, with critical points along the axis of the dipole

In the far field ($r \rightarrow \infty$), the direction of observation selects one direction of propagation in the expansion (28) and we can therefore see that the vector e on the sphere is

$$e = \frac{p - e_r(e_r \cdot p)}{\sqrt{p^2 - (e_r \cdot p)^2}} \tag{30}$$

This vector has two critical points of index $+1$, each when the radial unit vector points along the axis of the dipole: $e_r(e_r \cdot p) = p$. We can thus see that the nodes in the radiation pattern from a simple dipole—something we are familiar with from a first course in electromagnetism and shown in Fig. 6—can be understood to be a necessary consequence of the topology of a sphere. Were the dispersion surface to have a different topology (in a hyperbolic material, for example [20]), these properties would change. This line of argument has far reaching consequences for the polarization of scattered light, governing the polarization of sunlight [21] and multipole scattering in metamaterials [22].

5 Characteristic Classes and Physics

We have now developed several methods, all for calculating the same thing: the Euler characteristic (winding number) of a two dimensional surface. This limitation in part occurs because we always considered *tangent* vectors $e_{1,2}$; the so-called *tangent bundle*. As the critical points of the tangent vectors are a direct reflection of the number of times the surface normal wraps around a sphere, we were stuck with the Euler characteristic. But there is nothing stopping us from adapting the same formulae to calculate *different* topological invariants both for higher dimensional surfaces and for vectors on the surface that are not related to the tangent vectors in any way.

To understand the generalization to higher dimensions we need to introduce some terminology. Although we didn't name them as such, we have so-far been looking at integrals of *characteristic classes*. In (15) we integrated the tangent vectors' *Euler class*, $\nabla \times A/2\pi$ where $A_i = e_1 \cdot \partial_i e_2$, which yields a topological invariant from the *real* vectors $e_{1,2}$ on

the surface. Meanwhile, when we did the same thing for the single *complex* vector field $|\psi\rangle = (\mathbf{e}_1 + i\mathbf{e}_2)/\sqrt{2}$ (23) (a so-called line bundle) we integrated the tangent vectors' *first Chern class*,

$$1^{\text{st}} \text{ Chern class : } c_1 = \frac{1}{2\pi} \left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} \right), \quad \text{where } A_j = -i\langle\psi|\partial_j|\psi\rangle \quad (31)$$

which is the Berry curvature divided by 2π , and yields a topological invariant from *complex* vectors on the surface. These are two examples of *characteristic classes*. For the special case of tangent vectors on a two dimensional surface, integrals of the Euler and first Chern classes give the same result: the Euler characteristic.

The *Chern classes* are just one type of characteristic class. Each Chern class is a 'closed but not exact' expression depending on the Berry connection. This is a generalization of what we've already seen in two dimensions, namely $\nabla \times \mathbf{A}$ 'looks like a curl' (it is closed), but fails to be the curl of any properly defined vector at the critical points on the surface (it is not exact). These critical points are a direct reflection of the topology of both the vector field, and the shape of the surface. The non-zero winding number in both our 1D (1) and 2D (15) examples is equivalent to summing the indices of these critical points. Each of the *Chern classes* beyond the first is a 'closed but not exact' expression that does exactly the same thing in higher dimensions, each being integrated over ever higher dimensional regions: the first Chern class being associated with two dimensional integrals, the second with four dimensions, the third with six dimensions, and so on.

To illustrate the point, let's look at the second Chern class. We consider a four dimensional surface, to which we attach a pair of complex vector fields, $|1\rangle$ and $|2\rangle$ (as opposed to the single vector $|\psi\rangle$ used in two dimensions). As explained in Appendix A, when dealing with a set of N complex vector fields, each component of the Berry connection A_i becomes an $N \times N$ matrix. The Berry curvature is also replaced by the two index 'curvature form' Ω_{ij} , each component of which is—in this particular case—a 2×2 matrix

$$\Omega_{ij} = \frac{1}{2} \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} + i[A_i, A_j] \right). \quad (32)$$

The commutator is defined as $[A_i, A_j] = A_i A_j - A_j A_i$, and represents the difference between Ω_{ij} and the Berry curvature encountered in the previous section.

Expressed in terms of the curvature (32), the *second* Chern class—which is to be integrated over a four dimensional surface—is simply required to be a 'closed but not exact' scalar expression that can be integrated over the surface⁴. As Ω_{ij} has only two of the requisite four spatial indices we must therefore consider the *square* of the Berry curvature,

$$2^{\text{nd}} \text{ Chern class : } c_2 = \frac{1}{8\pi^2} \text{tr}[\Omega^2] = \frac{1}{8\pi^2} \epsilon_{ijkl} \text{tr}[\Omega_{ij}\Omega_{kl}] \quad (33)$$

The 'square' of the curvature given above is not defined as an ordinary matrix product, but as an antisymmetrized product: the 'wedge product' in the terminology of differential forms. This antisymmetrized product is given in the final step of (33), where ϵ_{ijkl} is the completely anti-symmetric unit tensor (repeated indices are summed over), and 'tr' is a trace over the matrix left after the sum over the spatial indices i, j, k, l . The pre-factor of $1/8\pi^2$

⁴ Note that the higher order Chern classes are more concisely expressed in terms of *differential forms* [52]. Here the equivalent of the curvature of the connection (32) is a two-form $\Omega = dA + iA \wedge A$, and the second Chern class (33) is written $c_2 = \Omega \wedge \Omega / 8\pi^2$. The Chern-Simons form (35) is almost always written in this language as $S = dA \wedge A + \frac{2}{3} A \wedge A \wedge A$.

is analogous to the factor of $1/2\pi$ in (31), ensuring the result of integrating (33) over the surface is an integer.

Substituting the curvature (32) into the definition of the second Chern class (33) we see that c_2 can be written as a divergence

$$\begin{aligned}
 c_2 &= \frac{1}{32\pi^2} \epsilon_{ijkl} \operatorname{tr} \left[\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} + i[A_i, A_j] \right] \left[\frac{\partial A_l}{\partial x_k} - \frac{\partial A_k}{\partial x_l} + i[A_k, A_l] \right] \\
 &= \frac{1}{8\pi^2} \epsilon_{ijkl} \frac{\partial}{\partial x_i} \operatorname{tr} \left[A_j \frac{\partial A_l}{\partial x_k} + \frac{2i}{3} A_j A_k A_l \right].
 \end{aligned}
 \tag{34}$$

As we hoped, we have something that is ‘closed but not exact’. As with the expressions we have met for the winding number of a curve (2), and the Euler characteristic (15), the integral of c_2 over any closed four dimensional surface *appears* to be zero. Equation (34) takes the form of a divergence, and the divergence theorem tells us its integral equals an integral over a boundary, which vanishes for any closed surface! Yet again this is not the case: c_2 is non-zero due to the critical points of the so-called *Chern–Simons* form [23]

$$S_i = \epsilon_{ijkl} \left[A_j \partial_k A_l + \frac{2i}{3} A_j A_k A_l \right]
 \tag{35}$$

where we have introduced the compact notation $\partial_i \equiv \partial/\partial x_i$, which we shall use wherever convenient from hereon. The Chern–Simons form exhibits critical points where the basis vectors $|1\rangle$ and $|2\rangle$ become undefined. We should note that the Chern–Simons form appears in several places in physics, including in the next section, and is e.g. an important object in topological field theory [24].

The second Chern class is another example of a quantity that is ‘closed but not exact’. The integral of the second Chern class depends on the integer number of critical points of the Chern Simons form and is thus a number that cannot be continuously changed. The pattern evident in the first two Chern classes, (31) and (33) continues into higher dimensions, with e.g. the 3rd Chern class given in terms of the *cube* of Ω_{ij} , which can be written as the divergence of a Chern–Simons form containing higher powers of the matrices A_i and their derivatives. Note that there are no Chern classes associated with odd dimensional surfaces, simply because they are all zero!

In recent years, Chern classes above the first have been used to design waveguides [25] and acoustic lattices [26], where the dimensions are typically ‘synthetic’, being system parameters such as the resonant frequency.

Example: the First Chern Class Applied to Distinguish Eigenmodes

As an illustration of the application of characteristic classes in physics, we can calculate the first Chern class for an eigenmode $|\psi\rangle$ of a system parameterized by coordinates, x_1 and x_2 that cover a closed surface. Suppose we have two such eigenmodes, $|\psi\rangle$ and $|\phi\rangle$, of a linear operator \hat{L} ,

$$\begin{aligned}
 \hat{L}(x_1, x_2) |\psi\rangle &= \lambda |\psi\rangle \\
 \hat{L}(x_1, x_2) |\phi\rangle &= \lambda' |\phi\rangle.
 \end{aligned}
 \tag{36}$$

By analogy with the discussion of the Sec. 4 we can define separate Berry connections for each of these eigenmodes,

$$A_{\psi,i} = -i\langle\psi|\frac{\partial}{\partial x_i}|\psi\rangle, \quad A_{\phi,i} = -i\langle\phi|\frac{\partial}{\partial x_i}|\phi\rangle. \tag{37}$$

and ask the question of whether the state $|\psi\rangle$ can be continuously deformed into $|\phi\rangle$. If the integral of the first Chern class, known as the first Chern number Ch_1 ,

$$\text{Ch}_{1,\psi} = \frac{1}{2\pi i} \int \nabla \times \langle\psi|\nabla|\psi\rangle dx_1 dx_2 = \frac{1}{2\pi} \int \nabla \times \mathbf{A}_\psi dx_1 dx_2 \tag{38}$$

is a different integer for modes $|\phi\rangle$ and $|\psi\rangle$ then the answer to this question is no.

We can gain some useful insights if, as in Sec. 4, we relate the first Chern class to derivatives of the operator \hat{L} , rather than the state $|\psi\rangle$. To do this we first differentiate the eigenvalue equation (36) to find the overlap between an arbitrary eigenmode $|n\rangle$ and the derivative of $|\psi\rangle$, in terms of the derivative of \hat{L} ,

$$\left\langle n \left| \frac{\partial \psi}{\partial x_i} \right\rangle = \frac{\langle n | \frac{\partial \hat{L}}{\partial x_i} | \psi \rangle}{\lambda - \lambda_n} \tag{39}$$

where we assume that $|n\rangle$ and $|\psi\rangle$ are non-degenerate eigenstates. Taking the curl of the first of the two Berry connections (37) and applying the above identity (39) then leads to the following expression for the Berry curvature,

$$\begin{aligned} \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} &= -i \sum_{|n\rangle \neq |\psi\rangle} \left[\left\langle \frac{\partial \psi}{\partial x_1} \middle| n \right\rangle \left\langle n \middle| \frac{\partial \psi}{\partial x_2} \right\rangle - \left\langle \frac{\partial \psi}{\partial x_2} \middle| n \right\rangle \left\langle n \middle| \frac{\partial \psi}{\partial x_1} \right\rangle \right] \\ &= -i \sum_{|n\rangle \neq |\psi\rangle} \left[\frac{\langle \psi | \frac{\partial \hat{L}}{\partial x_1} | n \rangle \langle n | \frac{\partial \hat{L}}{\partial x_2} | \psi \rangle - \langle \psi | \frac{\partial \hat{L}}{\partial x_2} | n \rangle \langle n | \frac{\partial \hat{L}}{\partial x_1} | \psi \rangle}{(\lambda - \lambda_n)^2} \right] \end{aligned} \tag{40}$$

where in the first step we used the completeness of states $\sum_n |n\rangle\langle n| = 1$, noting that there is zero contribution to the sum from the state $|\psi\rangle$. Not also that for brevity we have dropped the subscript ‘ ψ ’ from the vector potential.

In (26) of Sec. 4 we related the Gaussian curvature to derivatives of a Hermitian operator $\hat{L} = i \mathbf{n} \times$. In the same way, here we find the more general concept of the Berry curvature is given in terms of derivatives of an arbitrary linear operator \hat{L} , of which $|\psi\rangle$ is an eigenstate. We can see from (40) that an eigenvalue degeneracy, $\lambda = \lambda_n$ between any of the levels $|n\rangle$ and the state of interest $|\psi\rangle$, leads to points of singular Berry curvature. Such points are critical points of the Berry connection, arising due to the indeterminacy in the eigenvector $|\psi\rangle$. However, degeneracies are not the only points of non-zero Berry curvature. Equation (40) shows that the curvature is generally non-zero whenever the derivatives of the linear operator, $\partial \hat{L} / \partial x_i$, have complex off diagonal matrix elements with the state of interest $|\psi\rangle$.

An important corollary of (40) is that, if we sum the Berry curvature over all system states $|\psi\rangle$, we obtain zero

$$-i \sum_{|\psi\rangle} \sum_{|n\rangle \neq |\psi\rangle} \left[\frac{\langle \psi | \frac{\partial \hat{L}}{\partial x_1} | n \rangle \langle n | \frac{\partial \hat{L}}{\partial x_2} | \psi \rangle - \langle \psi | \frac{\partial \hat{L}}{\partial x_2} | n \rangle \langle n | \frac{\partial \hat{L}}{\partial x_1} | \psi \rangle}{(\lambda - \lambda_n)^2} \right] = 0 \tag{41}$$

because the two terms in the numerator are now equal, cancelling due to the summation. *This simple observation means that the sum of the first Chern numbers for all the eigenstates of a system is always zero.*

5.1 Example: Polarization Eigenstates in Anisotropic Materials

Take an electromagnetic wave propagating through an anisotropic, non-magnetic material. We shall use the Chern number to characterize a family of these materials, to tell us how many times *all* possible polarizations are explored as we run through the material parameters.

Assuming a permeability $\mu_0 \mathbf{1}$, a Hermitian (lossless) permittivity,

$$\epsilon = \epsilon_0 \begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} & 0 \\ \epsilon_{yx} & \epsilon_{yy} & 0 \\ 0 & 0 & \epsilon_{zz} \end{pmatrix} = \epsilon_0 \begin{pmatrix} \epsilon_{\parallel} & \mathbf{0} \\ \mathbf{0} & \epsilon_{zz} \end{pmatrix}, \tag{42}$$

and propagation along the z -axis, Maxwell’s equations for a wave of fixed frequency ω reduce to

$$\begin{aligned} \frac{k}{\omega} \mathbf{e}_z \times \mathbf{E} &= \mu_0 \mathbf{H} \\ \frac{k}{\omega} \mathbf{e}_z \times \mathbf{H} &= -\epsilon_0 \epsilon_{\parallel} \cdot \mathbf{E} \end{aligned} \tag{43}$$

where k the propagation constant of the wave. The problem is now to find the propagation constant and the electric field vector for a given set of material parameters ϵ_{\parallel} . Eliminating the magnetic field from the two equations (43) we can reduce the problem to an eigenvalue equation

$$\epsilon_{\parallel} \cdot \mathbf{E} = n^2 \mathbf{E} \tag{44}$$

where $n = ck/\omega$ is the refractive index. Therefore the eigenvectors and eigenvalues of the in-plane permittivity tensor determine the polarization and refractive index of the wave, respectively. Interestingly, (44) is equivalent to the Schrödinger equation for a spin 1/2 particle in a magnetic field (see e.g. [19]). With this in mind we re-write the in-plane permittivity tensors in terms of the Pauli spin matrices, $\sigma = \sigma_x \mathbf{e}_x + \sigma_y \mathbf{e}_y + \sigma_z \mathbf{e}_z$, as follows

$$\begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} \\ \epsilon_{yx} & \epsilon_{yy} \end{pmatrix} = \begin{pmatrix} V_0 + V \cos(\theta) & V \sin(\theta) e^{-i\phi} \\ V \sin(\theta) e^{i\phi} & V_0 - V \cos(\theta) \end{pmatrix} = V_0 \mathbf{1} + V \mathbf{n} \cdot \boldsymbol{\sigma}. \tag{45}$$

where \mathbf{n} is a unit vector, here parameterized by the spherical coordinates, θ and ϕ . While the angle θ determines the orientation of the principal axes of the permittivity, ϕ governs the gyrotropy of the medium [27]. We can thus visualise these forms of lossless in-plane permittivity in terms of two real numbers, V_0 and V , and the coordinates on a sphere, θ and ϕ . Interestingly, the *eigenvalues* of (45) (the refractive index squared) are independent of the coordinates θ and ϕ ,

$$n^2 = V_0 \pm V. \tag{46}$$

while the *eigenvectors*—defining the *polarization* of the electric field—depend *only* on the spherical coordinates

$$|\psi_+\rangle = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) e^{i\phi} \end{pmatrix}, \quad |\psi_-\rangle = \begin{pmatrix} \sin(\theta/2) e^{-i\phi} \\ -\cos(\theta/2) \end{pmatrix}. \tag{47}$$

These two states can be pictured as points on the *Bloch sphere* shown in Fig. 7. Keeping the refractive index fixed, the two polarizations can be interchanged by modifying the material parameters (45) according to the substitution $\theta \rightarrow \pi - \theta$ and $\phi \rightarrow \phi + \pi$.

For both states (47), the Berry connection contains only a single component,

$$A_{+,2} = -i \langle \psi_+ | \frac{\partial}{\partial \phi} | \psi_+ \rangle = \sin^2(\theta/2) = -A_{-,2} \tag{48}$$

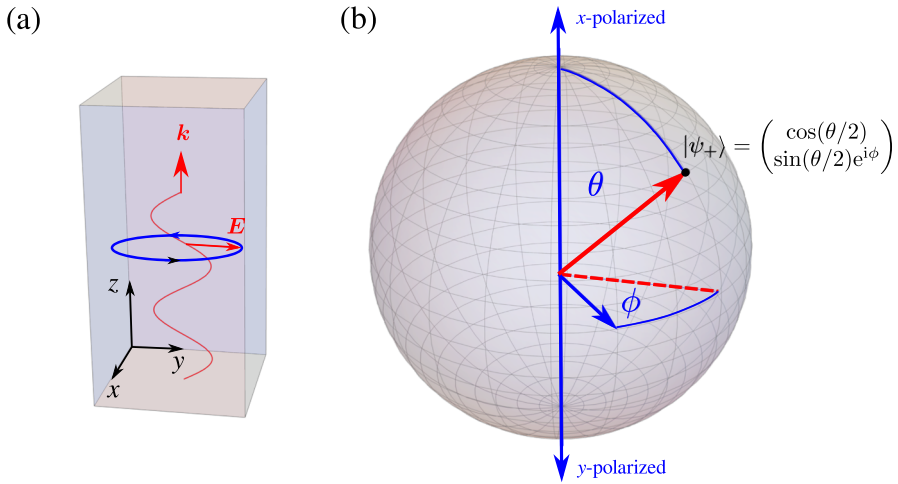


Fig. 7 (a) An electromagnetic wave propagating through a dielectric (here along z), normal to the plane of anisotropy (here the x - y plane) will generally have two different propagation constants $k = n\omega/c$, each corresponding to a different, generally elliptical, polarization, as given by (46) and (47). (b) We can visualise the polarizations on the Bloch sphere, with θ and ϕ labelling the different material parameters in (45). For $\theta = 0, \pi$ the principal axes of the permittivity tensor align with the x and y axes, and the wave is either x or y polarized. Meanwhile, for $\theta = \pi/2$ the principal axes are diagonal and the wave polarization varies between diagonal and circular, depending on the gyrotropy ϕ

which has an $n = 1$ critical point at the south pole, $\theta = \pi$. The result $A_{+,2} = -A_{-,2}$ is also in agreement with Sec. 4, where we found the sum of the Berry curvature over all eigenstates should vanish. From the above equation for the Berry connection, the Berry curvature is found to equal

$$\frac{\partial A_{+,2}}{\partial \theta} - \frac{\partial A_{+,1}}{\partial \phi} = \sin(\theta)/2. \tag{49}$$

The integral of the curvature (49) over all values of θ and ϕ equals the first Chern number, here Ch_+ , which in this system is the number of times the polarization covers the Bloch sphere

$$Ch_+ = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin(\theta) = 1 = -Ch_- \tag{50}$$

A Chern number of unity tells us that, for a fixed value of the refractive index squared, $n^2 = V_0 \pm V$, the family of in-plane permittivity tensors (45) cover all possible electromagnetic polarizations exactly once. The difference in sign between Ch_+ and Ch_- means that the two polarization eigenstates (47) wind in opposite senses around the Bloch sphere as we change the material parameters. This argument fails when we look at isotropic materials where $V = 0$, in which case the ‘gap’ between the two values of refractive index (46) closes, and the permittivity (45) no longer depends on the spherical coordinates. At this point the two polarizations (47) have degenerate values of the refractive index, meaning all polarizations are solutions to (44), whatever the value of V_0 .

Had we parameterized the permittivity differently, as e.g.

$$\begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} \\ \epsilon_{yx} & \epsilon_{yy} \end{pmatrix} = \begin{pmatrix} V_0 \cos(\phi) + V \cos(\theta) & V \sin(\theta) \\ V \sin(\theta) & V_0 \cos(\phi) - V \cos(\theta) \end{pmatrix} \tag{51}$$

the refractive index would have been $n^2 = V_0 \cos(\phi) \pm V$, with eigenstates independent of ϕ , e.g. $|\psi_+\rangle = (\cos(\theta/2), \sin(\theta/2))^T$. For this family of permittivity tensors the Berry connection (48) therefore vanishes, and no area is covered on the Bloch sphere. The Chern numbers thus vanish, $Ch_+ = Ch_- = 0$. This tells us that for anisotropic media, gyrotropy is therefore essential to realise all possible polarization eigenstates.

6 Chern Numbers and Dispersion Relations

We introduced the ‘characteristic classes’ to set up what is probably the most striking application of topology in wave physics: the prediction of interface modes, and in particular the possibility of these interface modes being constrained to propagate in one direction only. In this section we prove the relation between the integral of the first Chern class, and the prediction of interface modes between different materials.

Take some planar material that supports waves. Perhaps electromagnetic waves in a periodic array of pillars, or elastic waves in a plate. Suppose the system is homogeneous, or at least periodic, so the modes $|n_k\rangle$ can be labelled with a wave–vector k . The modes will generally be the solution to some eigenvalue equation

$$\hat{L}(k)|n_k\rangle = \lambda_{n,k}|n_k\rangle \tag{52}$$

where the integer n labels the different branches of the dispersion relation. At this stage we make no assumption about the meaning of $\lambda_{n,k}$; it could be a frequency, a wave vector component, or a material parameter.

In the case of homogeneous media, the wave–vector ranges over all pairs of real values, which can be mapped to a sphere via the stereographic projection (so long as care is taken at infinity [28], see Fig. 8a). For periodic media the Bloch vector has components ranging over the first Brillouin zone, the edges of which can be connected to each other through the addition of a reciprocal lattice vector, thus forming the surface of a torus (see Fig. 8b). In both cases we are dealing with the situation described in the previous sections: a single complex vector field $|n_k\rangle$ defined over a closed two dimensional surface, the points of which are specified by the vector k .

In a nutshell Suppose we have a system where the different branches of the dispersion relation (the $\lambda_{n,k}$ for different n) are separated by gaps. For instance $\lambda_{2,k}$ takes values in some range, separated above and below by a gap before the next lower range of values $\lambda_{1,k}$, and larger range of values $\lambda_{3,k}$. The idea is that we focus on one of these gaps in the spectrum. We classify all the branches of the dispersion relation below this gap in terms of the integral of the first Chern class (the first Chern number) over all wave–vectors

$$Ch_1 = \frac{1}{2\pi i} \int_S \nabla_k \times \langle n_k | \nabla_k | n_k \rangle dk_1 dk_2 = \frac{1}{2\pi} \int_S \nabla \times A_n dk_1 dk_2. \tag{53}$$

As shown in Fig. 8, the surface S is a sphere for homogeneous media, and a torus for periodic materials.

Now take two *different* materials, A and A' that have a shared gap in their eigenvalue spectrum. If we sum the first Chern numbers Ch_1 for the all the modes $|n_k\rangle$ and $|n'_k\rangle$ below the gap of interest and find the answers for the two materials are not the same, *there is no way to smoothly change the material from A to A' , without closing the gap of interest in the eigenvalue spectrum.*

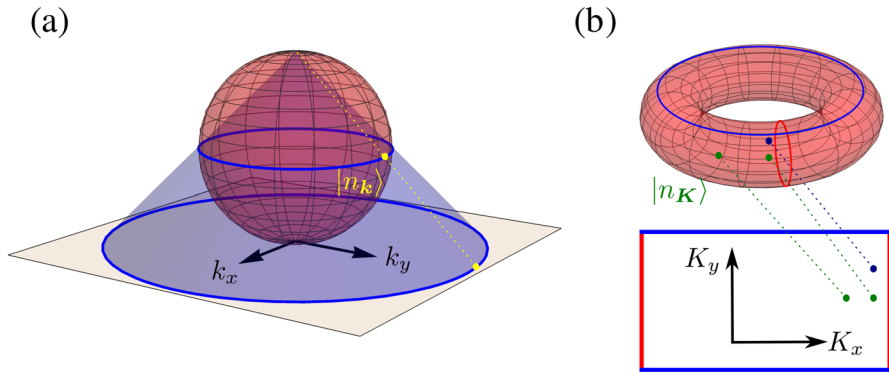


Fig. 8 The modes $|n_k\rangle$ of both homogeneous and periodic two dimensional media can be understood as a set of vectors on a closed surface, as discussed in Sec. 5. For homogeneous media the wave-vector \mathbf{k} ranges over all pairs of real values, which—via the stereographic projection—can be mapped onto the surface of a sphere, as shown in (a). For periodic media, the Bloch vector \mathbf{K} is equivalent after the addition of a reciprocal lattice vector. The edges of the first Brillouin zone (differing as they do by a reciprocal lattice vector) can thus be identified, e.g. the two blue lines in panel (b) represent the same physical states. Connecting the edges of the Brillouin zone yields a torus, as shown in panel (b)

To see this, note that to change the Chern number for any one of the branches of the dispersion relation we must introduce or remove critical points from the Berry connection A_n . From our formula for the Berry connection (40) we know that new critical points arise whenever another branch of the dispersion relation becomes degenerate with branch n . But we know from (41) that the total Berry curvature is zero for all the branches of the dispersion relation. Therefore the sum of the Chern numbers below the gap cannot be changed through the closure of any gap, except the one we are interested in. This argument holds, however we choose to change material A into A' .

Therefore, if we join the two materials together, in the transition region between the two there must be a region of space where the gap in the spectrum closes such that propagation is allowed. This region is where interface modes can be trapped. To add some meat onto the bones of this idea, we now give an argument based around a paper of the author [29], which was in turn adapted to Maxwell’s equations from [30].

6.1 Volovik’s Mode Counting Argument:

Suppose we have a system where the material varies smoothly as a function of x , as shown in Fig. 9. Asymptotically we have material A as x becomes large and positive and material A' when x is large and negative. Along the y -axis, the system remains translationally invariant so that we can replace y with the Fourier variable k . Due to the lack of translational invariance along x , the wave operator cannot be written in terms of the Fourier variable k_x , but depends on both $-i\partial_x$ and x . The eigenvalue equation is thus changed from (52) to

$$\hat{L}(-i\partial_x, x, k)|n_k\rangle = \lambda_{n,k}|n_k\rangle. \tag{54}$$

The Green function for this equation obeys

$$\left[\hat{L}(-i\partial_x, x, k) - \lambda \right] \mathbf{G}(x, x', k, \lambda) = \mathbf{1} \delta(x - x') \tag{55}$$

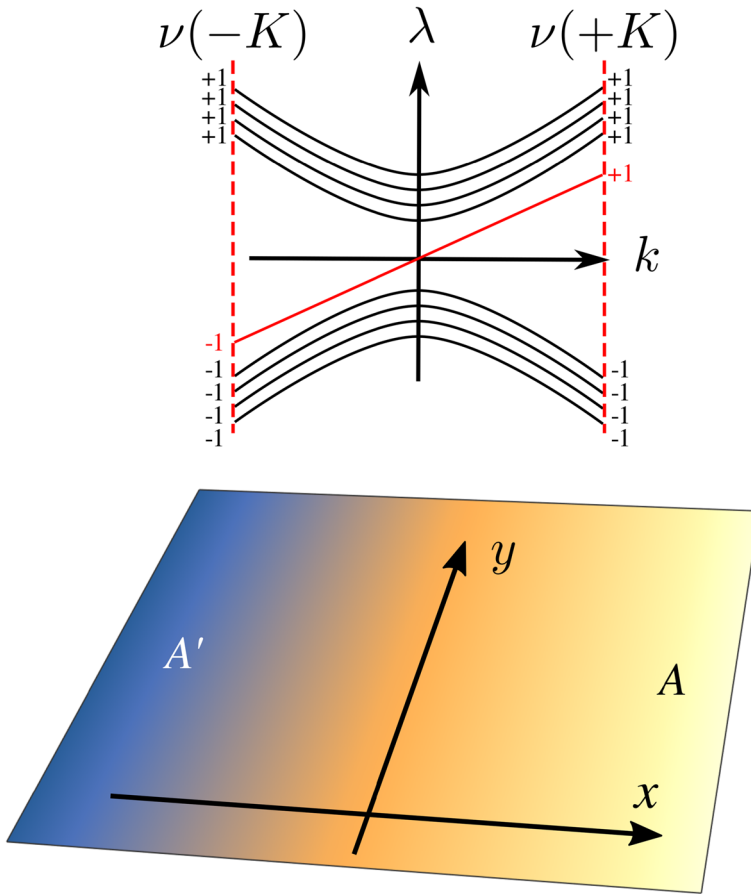


Fig. 9 The spectral asymmetry $\nu(k)$ defined in (57) is half the difference between the number of modes with positive and negative eigenvalue, λ for a fixed value of $k_y = k$. If a single mode crosses $\lambda = 0$ with increasing k (shown here in red) then the difference $\nu(+K) - \nu(-K)$ equals ± 1 . If the asymptotic materials A and A' both have a gap in their eigenvalue spectrum around $\lambda = 0$, the mode that crosses $\lambda = 0$ must be an interface mode, confined to the inhomogeneous region between the blue and yellow regions indicated in the figure

and contains information about all of the eigenmodes. Here we make use of the Green function to characterize how the eigenvalues change with respect to the y component of the wavevector, k . Using the completeness of the eigenmodes of (54), $\sum_n \langle x|n\rangle \langle n|x'\rangle = \mathbf{1} \delta(x - x')$ we can expand the Green function in the eigenmode basis, yielding the ‘bilinear expansion’ [31]

$$G(x, x', k, \lambda) = \sum_n \frac{\langle x|n\rangle \langle n|x'\rangle}{\lambda_{n,k} - \lambda} \dots \tag{56}$$

For us λ is a complex number so that the denominator in (56) is typically non-zero.

As discussed above, we assume that when taken as homogeneous media, the two asymptotic materials, A and A' have a shared gap in their eigenvalue spectrum that includes $\lambda = 0$. Therefore, if we find a solution to (54) with $\lambda = 0$ we know that it must be confined to the region in space where the material is in the process of changing from A to A' , i.e. it is an interface mode. We count these modes through introducing the *spectral asymmetry*, which

records the difference in the number of modes with positive and negative eigenvalues (see Fig. 9),

$$v(k) = \frac{1}{2} \sum_n \text{sign}[\lambda_{n,k}] \tag{57}$$

this number is a kind of tripwire, changing whenever a mode of the system crosses $\lambda = 0$. The spectral asymmetry can be calculated in terms of the Green function as follows,

$$\begin{aligned} v(k) &= \text{Re} \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} \text{Tr}[G(x, x, k, i\lambda)] = \sum_n \text{Tr}[\langle x|n\rangle\langle n|x\rangle] \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} \frac{\lambda_{n,k}}{\lambda_{n,k}^2 + \lambda^2} \\ &= \frac{1}{2} \sum_n \text{sign}[\lambda_{n,k}] \end{aligned} \tag{58}$$

where the capitalized ‘Tr’ has been used for the sake of brevity. It means both a trace over matrix indices, *and* an integration over position x .

We assume the material parameters change on a length scale that is large compared to all other system scales. Equation (58) also shows us we only need the behaviour of the Green function in the neighbourhood of $x = x'$. The differential equation (55) can therefore be expanded to leading order in the distance $x - x'$. Making the change of variables from x and x' to the average position $X = (x + x')/2$, and separation $\xi = x - x'$, we expand the linear operator \hat{L} to leading order in ξ and ∂_X ,

$$\begin{aligned} \hat{L}(-i\partial_x, x, k) &= \hat{L}\left(-i\partial_\xi - i\frac{1}{2}\partial_X, X + \frac{1}{2}\xi, k\right) \\ &\sim \hat{L}(-i\partial_\xi, X, k_y) - \frac{i}{4} \left[\frac{\partial}{\partial X} \frac{\partial \hat{L}}{\partial k_x} + \frac{\partial \hat{L}}{\partial k_x} \frac{\partial}{\partial X} \right] + \frac{1}{4} \left[\xi \frac{\partial \hat{L}}{\partial X} + \frac{\partial \hat{L}}{\partial X} \xi \right]. \end{aligned} \tag{59}$$

where derivatives with respect to k_x indicate a derivative of the operator with respect to the first argument of \hat{L} . Everywhere in (59) the operator \hat{L} depends only on derivatives ∂_ξ and average coordinate X . Performing a Fourier transform of (59) over ξ , the wave operator \hat{L} becomes a simpler object involving only two differential operators

$$\int d\xi e^{-ik_x\xi} \hat{L}(-i\partial_x, x, k) \sim L(k_x, X, k) + \frac{i}{2} \left[\frac{\partial L}{\partial X} \frac{\partial}{\partial k_x} - \frac{\partial L}{\partial k_x} \frac{\partial}{\partial X} \right]. \tag{60}$$

Staying in k_x space, we can now solve for the Green function of (60) to the same order, writing it as the sum of a zeroth and first order term $G \sim G_0 + G_1$. The zeroth order Green function is a solution to (55) with the operator \hat{L} replaced by the first term in (60). This is simply the Green function for an infinite homogeneous medium with the local material properties at the average position X ,

$$G_0(k_x, X, k_y, \lambda) = [L(k_x, X, k_y) - \lambda]^{-1} \tag{61}$$

Meanwhile the first order correction to G is the solution to

$$[L(k_x, X, k) - \lambda] G_1 = -\frac{i}{2} \left[\frac{\partial L}{\partial X} \frac{\partial}{\partial k_x} - \frac{\partial L}{\partial k_x} \frac{\partial}{\partial X} \right] G_0 \tag{62}$$

which can be solved using the zeroth order Green function (61), giving the correction

$$G_1(k_x, X, k, \lambda) = -\frac{i}{2} G_0 \left[\frac{\partial L}{\partial X} \frac{\partial G_0}{\partial k_x} - \frac{\partial L}{\partial k_x} \frac{\partial G_0}{\partial X} \right]. \tag{63}$$

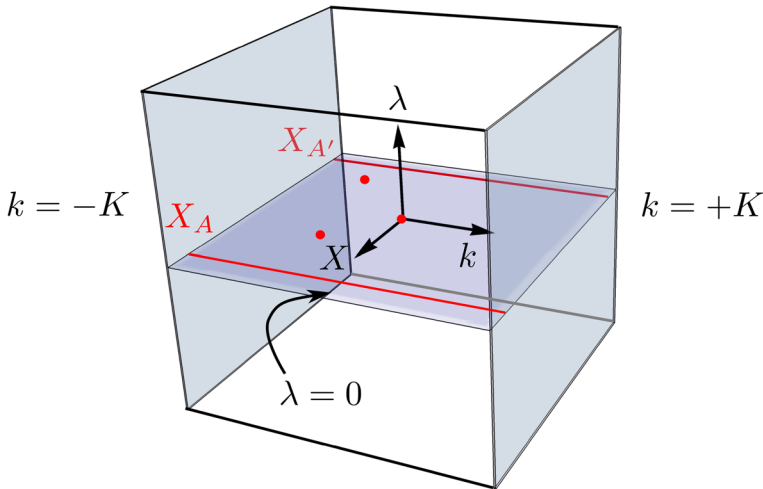


Fig. 10 Illustration of the integrals (64) and (66). The spectral asymmetry is calculated as a three dimensional surface integral (here shown as a shaded 2D plane), at fixed k . The number of modes crossing $\lambda_{n,k} = 0$ equals the difference in the spectral asymmetry $\nu(+K) - \nu(-K)$ (the two shaded planes). Provided the Green function vanishes for large X, k_x , and λ the integral can be replaced with the closed cubic surface integral shown above. As the only contributions to the integral are the critical points of A_i on the $\lambda = 0$ surface (occurring whenever the band gap closes, and shown as red dots here), we can deform the closed surface into two parallel surfaces at fixed X , where the material parameters no longer change

Summing Eqns. (61) and (63) we have an approximate expression for the Green function, valid when there is a slow change of material parameters with respect to position.

With our approximate expression for the Green function we can now compute the spectral asymmetry via (58). To keep the expressions relatively simple, we assume the spectral asymmetry computed from G_0 is zero. This assumption is easily relaxed, and means that we only consider materials on the way from A to A' that—were they a homogeneous medium—would have a symmetric eigenvalue spectrum around $\lambda_{n,k} = 0$ ⁵. The spectral asymmetry is then determined by the first order correction to the Green function. Substituting the first order correction to the Green function (63) into (58) yields

$$\begin{aligned}
 \nu(k) &= \text{Im} \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} \int_{-\infty}^{\infty} \frac{dk_x}{2\pi} \int_{-\infty}^{\infty} \frac{dX}{2} \text{tr} \left[G_0 \frac{\partial G_0^{-1}}{\partial X} \frac{\partial G_0}{\partial k_x} - G_0 \frac{\partial G_0^{-1}}{\partial k_x} \frac{\partial G_0}{\partial X} \right] \\
 &= \text{Im} \int \frac{d^3x}{8\pi^2} \text{tr} \left[G_0 \frac{\partial G_0^{-1}}{\partial x_1} G_0 \frac{\partial G_0^{-1}}{\partial x_2} G_0 \frac{\partial G_0^{-1}}{\partial x_0} - G_0 \frac{\partial G_0^{-1}}{\partial x_2} G_0 \frac{\partial G_0^{-1}}{\partial x_1} G_0 \frac{\partial G_0^{-1}}{\partial x_0} \right]
 \end{aligned}
 \tag{64}$$

where we introduced the four coordinates $(x_0, x_1, x_2, x_3) = (\lambda, X, k_x, k)$, with the integration carried out over the first three coordinates, $d^3x = dx_0 dx_1 dx_2$.

Now the topology appears! Equation (64) is actually an integral of the Chern–Simons form (35) in disguise. To see this, we define a connection A_i with matrix components, as

⁵ In the more general case, where the eigenvalues are not symmetrically distributed around the gap at $\lambda \sim 0$ we have to consider *two* contributions to the spectral asymmetry.

described in Sec. 4 and Appendix A,

$$A_i = iG_0 \frac{\partial G_0^{-1}}{\partial x_i} = iH_0^{-1} \frac{\partial H_0}{\partial x_i} + iU_0^\dagger \frac{\partial U_0}{\partial x_i}. \tag{65}$$

where we have written the matrix G_0^{-1} in polar form, $G_0^{-1} = H_0 U_0$, where H_0 is Hermitian and U_0 unitary. Note that the connection (65) has zero curvature (167), and that the final term in (65) involving the unitary matrix is that found in (175) of Appendix A, corresponding to a rotation of the basis vectors. It is this final term that winds around the critical points of the connection A_i , analogous to the winding of the Berry phase around critical points of a tangent vector on a surface.

The difference between the spectral asymmetry between the fixed values $k = +K$ and $-K$ tells us the number of modes N that cross $\lambda_{n,k} = 0$. After taking this difference, the open surface integral in (64) can be replaced with an integral over a closed three dimensional surface see Fig. 10), and the number of modes N can be written as ⁶

$$N = \nu(+K) - \nu(-K) = \frac{i}{24\pi^2} \text{Im} \oint d^3x \epsilon_{ijk3} \text{tr} [A_i A_j A_k].$$

$$= \frac{1}{8\pi^2} \text{Im} \oint d^3x \epsilon_{ijk3} \text{tr} \left[\frac{\partial A_j}{\partial x_i} A_k + \frac{2i}{3} A_i A_j A_k \right] \tag{66}$$

Note the sign of N indicates the direction in which the modes cross $\lambda_{n,k} = 0$.

After comparison with Sec. 4 we can see that (66) is the integral of the Chern–Simons form over a closed three dimensional surface, divided by $8\pi^2$. As described there, this boundary integral is equal to the four dimensional bulk integral of the second Chern class (33)! However, due to the form of the vector potential (65), the ‘curvature’ Ω_{ij} defined in (32) appears to vanish identically everywhere, and once again we have an integral that looks like it should be exactly zero!

It’s the same old story: the integral is recording the discrete critical points of A_i , which occur when $G_0^{-1} = [L - i\lambda]$ cannot be inverted (i.e. both the coordinate λ is zero, and one or more pairs of eigenvalues $\lambda_{n,k}$ are zero). *This means that the number of interface modes N in our inhomogeneous system that pass through zero eigenvalue is determined by the number of times the ‘gap’ in the eigenvalue spectrum of $L(k_x, X, k)$ closes.* This is a satisfying result: as we deform material A into material A' we must close the gap N times in order to have N interface modes crossing $\lambda_{n,k} = 0$.

So long as our integration volume encloses all of the critical points of A_i we won’t change the predicted number of interface modes, N . Therefore—as illustrated in Fig. 10—we can replace the integral over the constant k surfaces in (66) with an equivalent one over constant X surfaces,

$$N = \frac{1}{24\pi^2} \text{Im} \oint d^3x \epsilon_{i1jk} \text{tr} [A_i A_j A_k] \tag{67}$$

We now only need understand the dependence of the operator L on the wave–vector k in the two asymptotic homogeneous materials A and A' , and no longer need to consider the interface. It is also clear from the above argument that our integral over the coordinate λ on

⁶ To form the closed surface integral the integral of the ‘vector potential’ (65) must vanish over the bounding surfaces of constant X , k_x , and λ . For large fixed λ this is automatic. For the large fixed k_x boundary we require periodic boundary conditions or diverging operator L with increasing k_x . But for the large fixed X boundary we require the operator L to diverge as $|X| \rightarrow \infty$. Thus, e.g. increasing x we assume material A is continuously changed into material A' , then remaining unchanged over a long distance, before L ultimately diverges at infinity.

which the Green function depends is somewhat redundant. The critical points of A_i always occur in the plane of $\lambda = 0$. Substituting the expression (65) for A_i into (67), we can perform the integral over λ exactly using the following result

$$\int_{-\infty}^{\infty} d\lambda \operatorname{tr} \left[(L - i\lambda)^{-2} \frac{\partial L}{\partial k_x} (L - i\lambda)^{-1} \frac{\partial L}{\partial k} \right] = 2\pi \sum_{n < 0, m \neq n} \frac{\langle n | \frac{\partial L}{\partial k_x} | m \rangle \langle m | \frac{\partial L}{\partial k} | n \rangle - \langle n | \frac{\partial L}{\partial k} | m \rangle \langle m | \frac{\partial L}{\partial k_x} | n \rangle}{(\lambda_{n,k} - \lambda_{m,k})^2} \tag{68}$$

which comes from an expansion of the operator L in terms of its eigemodes $L = \sum_n \lambda_{n,k} |n\rangle \langle n|$ and an application of Cauchy’s integral formula. After applying (68) to our integral (67), the number of interface modes equals the difference

$$N = \bar{v}(X_{A'}) - \bar{v}(X_A) \tag{69}$$

where we have defined

$$\bar{v}(x) = \frac{1}{2\pi i} \sum_{n < 0, m \neq n} \int_S d^2k \frac{\langle n | \frac{\partial L}{\partial k_x} | m \rangle \langle m | \frac{\partial L}{\partial k} | n \rangle - \langle n | \frac{\partial L}{\partial k} | m \rangle \langle m | \frac{\partial L}{\partial k_x} | n \rangle}{(\lambda_{n,k} - \lambda_{m,k})^2} \tag{70}$$

and e.g. X_A indicates a position away from the inhomogeneity where the material parameters are those of material A .

Comparison with our expression for the Berry curvature (40), and the first Chern number (38) shows that $\bar{v}(x)$ is equal to the sum of all the Chern numbers for the bands below $\lambda_{n,k} = 0$,

$$\bar{v}(x) = \sum_{n < 0} \operatorname{Ch}_{1,n}(x) \tag{71}$$

which will be an integer so long as the surface S appearing in (70) can be made closed (as shown in Fig. 8). The number of interface modes N arising from changing the material from A to A' is thus equal to the difference in Chern numbers between the two media, summed over all the bands below the gap in the eigenvalue spectrum,

$$N = \sum_{n < 0} [\operatorname{Ch}_{1,n}(X_{A'}) - \operatorname{Ch}_{1,n}(X_A)]. \tag{72}$$

We now have the striking result mentioned at the beginning of this section. The number of interface modes trapped between materials A and A' is determined by the difference in their Chern numbers. These are computed over a closed surface parameterized by the wave–vector, \mathbf{k} , and summed over all branches of the dispersion relation below the gap in the spectrum where there are the interface modes of interest.

So far in the tutorial we have gradually introduced the machinery of topology in order to get to this amazing result, the origin of which is not often discussed in metamaterials textbooks or papers. We now consider a few simple examples.

Example: Counting One Mode

As our first example we give a simple one dimensional calculation of the spectral asymmetry (58) in terms of the Green function, and relate it to a winding number.

Suppose we have a mode that satisfies the linear dispersion relation $k = \alpha k_0$. We can see immediately that exactly one mode crosses from negative k_0 to positive k_0 as k is increased from $-\infty$ to $+\infty$. One choice of linear equation that gives such a dispersion relations is,

$$-i\alpha^{-1} \frac{\partial \phi}{\partial y} = k_0 \phi. \tag{73}$$

Comparing to the previous section we see that the linear operator is $\hat{L} = -i\alpha^{-1} \partial_y$, and the eigenvalue is the wavenumber, $\lambda_k = k_0$.

We can now calculate the spectral asymmetry as a function of the wave-vector, k . For a fixed value of k the Green function is simple

$$(\alpha^{-1}k - \lambda) G(k, \lambda) = 1 \rightarrow G(k, \lambda) = \frac{1}{\alpha^{-1}k - \lambda}. \tag{74}$$

According to (58), the spectral asymmetry equals the real part of the integral of the Green function over purely imaginary values of λ . For our simple Green function (74), this is simply the derivative of the phase of G

$$\text{Re}[G(k, i\lambda)] = \text{Im} \frac{d}{d\lambda} \log \left(\frac{1}{\alpha^{-1}k - i\lambda} \right) = \frac{d}{d\lambda} \arg[G(k, i\lambda)] \tag{75}$$

The integral of (75) over λ is thus simply the change in the phase of the Green function between the end points of the integral. The phase angle is defined relative to the critical point in (75) that is at $k, \lambda = 0$ (see Fig. 11), where the Green function diverges, and is exactly where our mode crosses $\lambda_k = 0$!

Integrating from $\lambda = -\infty$ to $+\infty$, this change of angle is $\pm\pi$ depending on the sign of $\alpha^{-1}k$. Applying (58) we thus find the spectral asymmetry equals

$$\nu(k) = \text{Re} \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} G(i\lambda) = \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} \frac{d}{d\lambda} \arg[G(k, i\lambda)] = \frac{1}{2} \text{sign}[\alpha^{-1}k] \tag{76}$$

Assuming $\alpha > 0$ and taking the difference between the spectral asymmetry at fixed values $k = +K$ and $k = -K$, the number of modes equals the winding number of the argument of

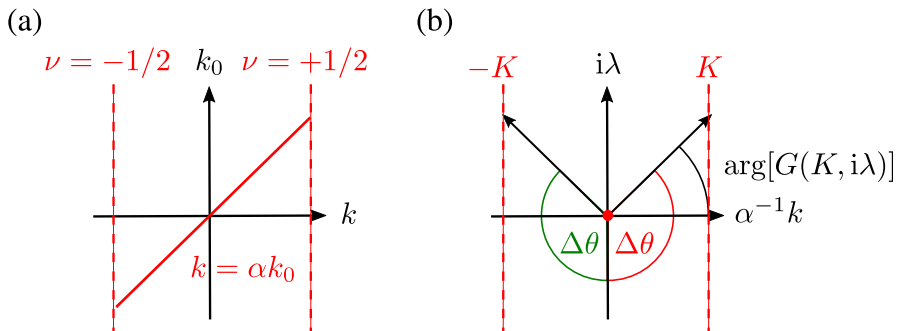


Fig. 11 Counting one mode: the first order equation (73) has wave-like solutions with a dispersion relation $k = \alpha k_0 = \alpha \omega/c$. Panel (a) shows the dispersion relation for $\alpha > 0$, indicating the spectral asymmetry, which is $+1/2$ for positive k and $-1/2$ for negative k . (b) The spectral asymmetry $\nu(K)$ can be written as the change in the argument ($\Delta\theta$) of the Green function (74) as λ varies from $-\infty$ to $+\infty$, divided by 2π . The difference $\nu(K) - \nu(-K)$ thus equals the number of times the argument of the Green function winds around the critical point at $\alpha^{-1}k + i\lambda = 0$

the Green function

$$\begin{aligned}
 N &= \nu(+K) - \nu(-K) \\
 &= \frac{1}{2\pi} \left[\arg[G(K, i\infty)] - \arg[G(K, -i\infty)] + \arg[G(-K, i\infty)] - \arg[G(-K, -i\infty)] \right] \\
 &= 1
 \end{aligned}
 \tag{77}$$

This is a simple case of the winding number given in Sec. 6: it is the topological invariant counting the critical points of the Green function, which here occurs where the mode crosses $\lambda_k = 0$.

Example: A Lattice of Resonators

As a second example, we take an array of coupled resonators, as sketched in Fig. 12a. These ‘resonators’ are a discrete approximation to a continuous system. For instance, the amplitude of a single resonator can represent the pressure in an acoustically resonant hole; the electric polarization of a dielectric particle; the displacement of an elastic rod; or simply the extension of a spring.

As we are dealing with a *discrete* lattice of resonators, the continuum theory given in Sec. 6 does not obviously apply. Yet the reasoning can be simply adapted. For instance we can expand the continuous wave in an orthonormal basis of N functions, $|\psi\rangle = a_0|0\rangle + a_1|1\rangle + \dots + a_{N-1}|N-1\rangle$, with the basis functions representing e.g. the modes of the lattice before a perturbation is applied, or the modes of individual resonators in the tight

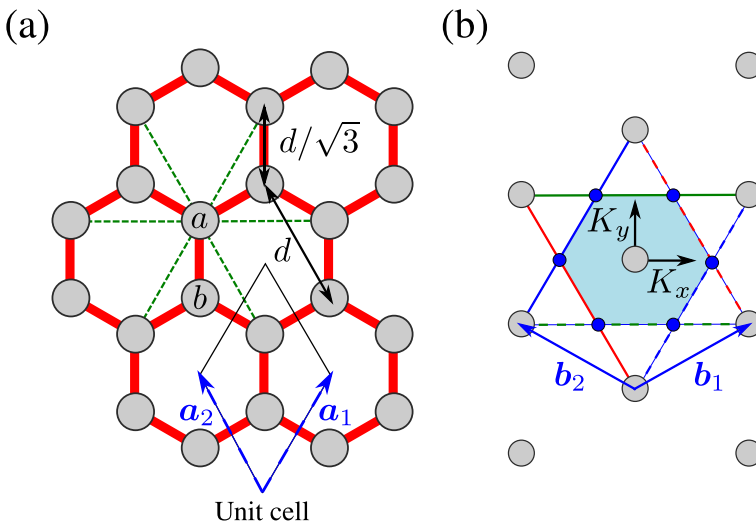


Fig. 12 In some cases wave equations can be approximated as a discrete system of coupled resonators. A periodic honeycomb lattice of resonators (lattice constant d , lattice vectors a_1 and a_2) is drawn in panel (a) where there are two resonators, ‘a’ and ‘b’ per unit cell, separated by a distance $d/\sqrt{3}$. In the Haldane model each resonator is coupled to both its three nearest neighbours (red lines) and its six next nearest neighbours (green dashed lines). The corresponding reciprocal space is sketched in panel (b), where the grey circles indicate multiples of the reciprocal lattice vectors b_1 and b_2 . The first Brillouin zone is shaded in blue, with the dashed and solid boundaries of the same colour connected by the addition of a reciprocal lattice vector. The six corners of the Brillouin zone where the coupling $g(\mathbf{K})$ vanishes are indicated as blue dots

binding approximation [32]. Assuming the basis itself obeys $\nabla_{\mathbf{K}} \times \langle n | \nabla_{\mathbf{K}} | m \rangle = 0$, the Berry curvature of a given mode $|\psi\rangle$ is given entirely in terms of the expansion coefficients, $-i\langle \psi | \nabla_{\mathbf{K}} | \psi \rangle = -i\nabla_{\mathbf{K}} \times \sum_n a_n^* \nabla_{\mathbf{K}} a_n$. This is the expression we would obtain for the Berry curvature of an N component complex vector $|\psi\rangle = (a_0, a_1, \dots, a_{N-1})^T$ in a discrete system.

In our simplified model we assume a tight binding approximation, where each resonator only has one possible mode (i.e. frequencies are such that higher order modes do not contribute). Due to the periodicity of the lattice, the unit cell contains all the degrees of freedom. Therefore a unit cell containing a single resonator has only one effective degree of freedom and will therefore exhibit a dispersion relation $\omega(\mathbf{K})$ with only one branch (band). As the theory given in Sec. 6 depends on the closure of a gap between two or more bands, we consider a lattice with two resonators (labelled a and b) per unit cell.

Labelling the amplitudes of the two resonators in each point in the lattice as $a_{n,m}$ and $b_{n,m}$, the equations of motion can be written in the general form

$$\begin{aligned} \ddot{a}_{n,m} + \omega_a^2 a_{n,m} &= \sum_{n',m'} [\alpha_{n'-n,m'-m} b_{n',m'} + \beta_{n'-n,m'-m} a_{n',m'}] \\ \ddot{b}_{n,m} + \omega_b^2 b_{n,m} &= \sum_{n',m'} [\alpha_{n-n',m-m'} a_{n',m'} + \gamma_{n'-n,m'-m} b_{n',m'}] \end{aligned} \tag{78}$$

where the ‘ a ’ and ‘ b ’ resonant frequencies are ω_a and ω_b respectively. The amplitude of the cross-coupling between resonators is given by $\alpha_{n,m}$, and the coupling between like resonators is given by $\beta_{n,m}$ and $\gamma_{n,m}$. It is assumed that the self coupling between like resonators vanishes $\gamma_{0,0} = \beta_{0,0} = 0$, as these terms are equivalent to a modification of the resonant frequencies $\omega_{a,b}$. From hereon we will work at a fixed frequency ω , where the coupling constants can take complex values.

We now compute the first Chern number for an infinite periodic system of resonators, writing the resonator amplitudes in accordance with Bloch’s theorem, and taking a fixed frequency of oscillation ω

$$\begin{aligned} a_{n,m} &= e^{i[\mathbf{K} \cdot (na_1 + ma_2) - \omega t]} a \\ b_{n,m} &= e^{i[\mathbf{K} \cdot (na_1 + ma_2) - \omega t]} b \end{aligned} \tag{79}$$

where \mathbf{K} is the Bloch vector, and \mathbf{a}_1 and \mathbf{a}_2 are the real space lattice vectors (see Fig. 12). Substituting (79) into (78), the infinite set equations of motion reduce to a set of two coupled linear equations

$$\begin{aligned} [\omega_a^2 - f(\mathbf{K})] a - g(\mathbf{K}) b &= \omega^2 a \\ [\omega_b^2 - h(\mathbf{K})] b - g^*(\mathbf{K}) a &= \omega^2 b \end{aligned} \tag{80}$$

where we have defined the three Bloch-vector dependent coupling functions

$$\begin{aligned} f(\mathbf{K}) &= \sum_{n',m' \neq 0,0} \beta_{n',m'} e^{i\mathbf{K} \cdot (n'\mathbf{a}_1 + m'\mathbf{a}_2)} \\ g(\mathbf{K}) &= \alpha_{0,0} + \sum_{n',m' \neq 0,0} \alpha_{n',m'} e^{i\mathbf{K} \cdot (n'\mathbf{a}_1 + m'\mathbf{a}_2)} \\ h(\mathbf{K}) &= \sum_{n',m' \neq 0,0} \gamma_{n',m'} e^{i\mathbf{K} \cdot (n'\mathbf{a}_1 + m'\mathbf{a}_2)}. \end{aligned} \tag{81}$$

Equation (80) is almost in the form required by the theory described in Sec. 6. One wrinkle is that we do not yet satisfy the assumption that the eigenvalues of the operator \hat{L} are sym-

metrically distributed around zero (the trace of \hat{L} does not vanish). To avoid complicating the discussion, we assume the two resonators in the unit cell are identical: $\omega_a = \omega_b$ and $h(\mathbf{K}) = f(\mathbf{K})$. Subtracting ω_a^2 from both sides of (80) and writing $\lambda = \omega^2 - \omega_a^2$ then gives us

$$\begin{pmatrix} -f(\mathbf{K}) & -g(\mathbf{K}) \\ -g^*(\mathbf{K}) & f(\mathbf{K}) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix} \tag{82}$$

The operator \hat{L} is now traceless and can be written in terms of Pauli matrices, as in the example application of the Chern number given in Sec. 5,

$$\hat{L} = -f(\mathbf{K}) \sigma_z - g_1(\mathbf{K}) \sigma_x + g_2(\mathbf{K}) \sigma_y \tag{83}$$

where f is a real function, and $g = g_1 + ig_2$. The Chern number for the operator (83) records the same thing as in the example of Sec. 5, namely the number of times the vector

$$\mathbf{n} = \frac{-g_1 \mathbf{e}_x + g_2 \mathbf{e}_y - f \mathbf{e}_z}{\sqrt{f^2 + g_1^2 + g_2^2}} \tag{84}$$

covers the unit sphere as \mathbf{K} is varied over the first Brillouin zone, i.e. the winding of the Hamiltonian around the point $f = g = 0$. Comparison with the example of Sec. 5 shows that the two eigenstates of \hat{L} are given by (47) with eigenvalues $\lambda = \omega^2 - \omega_a^2 = \pm[f^2 + g_1^2 + g_2^2]$. As in that example, the eigenstates depend only on the spherical angles θ and ϕ of the vector \mathbf{n} , which are here identified as $\cos(\theta) = -f/[f^2 + g_1^2 + g_2^2]^{1/2}$ and $\exp(i\phi) = (-g_1 + ig_2)/[g_1^2 + g_2^2]^{1/2}$. The Berry connection is also given in Sec. 5, by (48).

From the example of Sec. 5.1, we know that the points of vanishing cross coupling between the two resonators in the unit cell, $g(\mathbf{K}) = 0$ are the critical points of the Berry connection. These correspond to the North or South pole of the unit sphere defined by (84), depending on the sign of f . We therefore see that the Chern number depends in an important way on the coupling functions between the same resonators in each unit cell, f and g . If, for example f is always positive then \mathbf{n} will only ever explore the lower half of the unit sphere and the Chern number will therefore always be zero. Between such lattices of resonators there can never be the non-trivial interface states predicted above.

To ensure that the Chern number does not vanish, we take an approach that is close to the famous example known as the ‘Haldane model’ developed by F. Duncan Haldane in the late 1980s [8]. We consider the special case of a lattice with hexagonal symmetry and lattice constant d^7 . Taking two identical resonators per unit cell, we construct a honeycomb lattice where each ‘ a ’ resonator is separated by $d/\sqrt{3}$ from three nearest neighbour ‘ b ’ resonators, as shown in Fig. 12a. Assuming only nearest neighbour coupling between the a and b resonators, $\alpha_{0,0} = \alpha_{1,0} = \alpha_{0,1} = \alpha \exp(-iK_y d/\sqrt{3})$ of equal strength $\alpha_{1,0} = \alpha_{0,1} = \alpha$, the coupling function $g(\mathbf{K})$ reduces to

$$\begin{aligned} g(\mathbf{K}) &= \alpha e^{-iK_y d/\sqrt{3}} \left[1 + e^{i\mathbf{K} \cdot \mathbf{a}_1} + e^{i\mathbf{K} \cdot \mathbf{a}_2} \right] \\ &= \alpha e^{-iK_y d/\sqrt{3}} \left[1 + 2 \cos(K_x d/2) e^{i\frac{\sqrt{3}d}{2} K_y} \right] \end{aligned} \tag{85}$$

This quantity vanishes at the 6 corner points on the boundary of the first Brillouin zone (see Fig. 13a): when the K_x component of the Bloch vector equals $\pm 2\pi/3d$, or $\pm 4\pi/3d$, with the K_y component equal to $\pm 2\pi/\sqrt{3}d$, or 0 respectively. These 6 points can be grouped into two lots of three, where one group can be obtained from the other via the substitution $\mathbf{K} \rightarrow -\mathbf{K}$,

⁷ Lattice vectors $\mathbf{a}_1 = \frac{d}{2}(\mathbf{e}_x + \sqrt{3}\mathbf{e}_y)$, and $\mathbf{a}_2 = \frac{d}{2}(-\mathbf{e}_x + \sqrt{3}\mathbf{e}_y)$

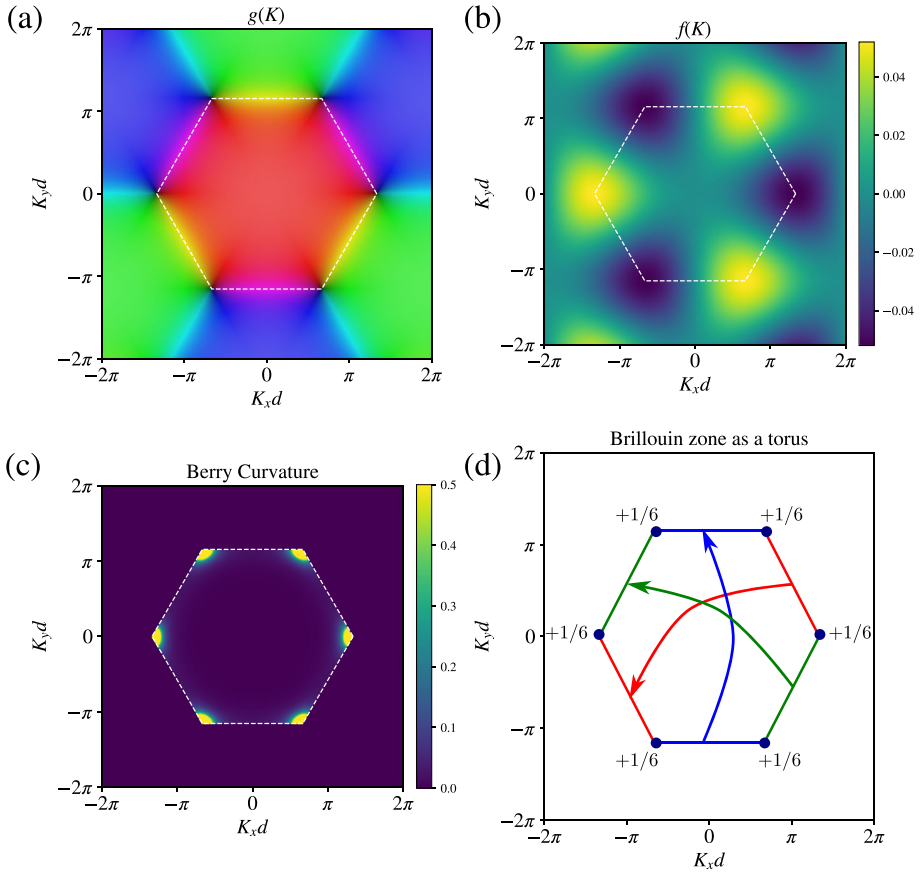


Fig. 13 Coupling functions and Berry curvature for the honeycomb lattice shown in Fig. 12, for parameters $\beta = 0.01$ and $\alpha = 1$. The hexagon indicated in all four plots shows the boundary of the first Brillouin zone. Panel (a) shows a phase plot of the complex coupling $g(\mathbf{K})$ between ‘a’ and ‘b’ resonators (color indicates phase and saturation, magnitude). Panel (b) shows the real valued coupling $f(\mathbf{K})$ between next–nearest–neighbour resonators. Panel (c) shows the Berry curvature, evaluated within the first Brillouin zone. Panel (d) indicates the ‘folding’ rules for mapping the first Brillouin zone onto a torus, with the total Berry curvature around each corner of the zone indicated

equivalent to a complex conjugation of (85). Being equivalent to a complex conjugation, the phase of $g(\mathbf{K})$ winds in an opposite sense around these two sets of points. Therefore, if $f(\mathbf{K}) = f(-\mathbf{K})$ (as it does if e.g. there is no coupling beyond nearest neighbour, $f = 0$), the winding number around these critical points will cancel, yielding a Chern number of zero!

To ensure that the Chern number does not vanish, we introduce a complex next–nearest neighbour coupling that breaks time reversal symmetry. We suppose that each resonator couples to its six equivalents at a distance a away with strength $\beta_{1,0} = -\beta_{-1,0} = i\beta$, $\beta_{0,1} = -\beta_{0,-1} = -i\beta$, and $\beta_{1,-1} = \beta_{-1,1} = -i\beta$

$$\begin{aligned}
 f(\mathbf{K}) &= -i\beta \left[-e^{i\mathbf{K}\cdot\mathbf{a}_1} + e^{-i\mathbf{K}\cdot\mathbf{a}_1} + e^{i\mathbf{K}\cdot\mathbf{a}_2} - e^{-i\mathbf{K}\cdot\mathbf{a}_2} + e^{i\mathbf{K}\cdot(\mathbf{a}_1-\mathbf{a}_2)} - e^{-i\mathbf{K}\cdot(\mathbf{a}_1-\mathbf{a}_2)} \right] \\
 &= 2\beta \left[\sin(K_x d) - \sin((K_x + \sqrt{3}K_y)d/2) - \sin((K_x - \sqrt{3}K_y)d/2) \right] \quad (86)
 \end{aligned}$$

which is an odd function of \mathbf{K} , meaning—via the argument above—that the total Berry curvature does not vanish.

Figure 13 shows the Berry curvature in the first Brillouin zone for the coupling functions (85) and (86) (positive α and β), which now corresponds to a Chern number of +1. From the relationship between interface states and Chern numbers derived above we can thus see that there will always be an interface state trapped between a lattice with non-zero next-nearest neighbour coupling (86), and one with $f(\mathbf{K}) = 0$, where the Chern number vanishes.

Example: Electromagnetic Waves in a Gyrotropic Medium

As a final example we consider a continuous system where the eigenvalue λ , appearing in Sec. 6, is a material parameter rather than frequency. This example is based on the results in [29].

Take an electromagnetic material where the relative magnetic permeability μ is a real scalar, and the permittivity ϵ is a tensor of the form

$$\epsilon = \begin{pmatrix} \epsilon_{\parallel} & \mathbf{0} \\ \mathbf{0} & \epsilon_{\perp} \end{pmatrix} \tag{87}$$

where ϵ_{\parallel} is a 2×2 matrix representing the anisotropic permittivity in the x - y plane of propagation. Taking propagation in the x - y plane, and assuming TM polarization, where $\mathbf{H} = (h/\eta_0)\mathbf{e}_z$, where $\eta_0 = \sqrt{\mu_0/\epsilon_0}$, Maxwell's equations take the form

$$\begin{aligned} -i\nabla \times \mathbf{E} &= k_0\mu h\mathbf{e}_z \\ i\nabla h \times \mathbf{e}_z &= k_0\epsilon_{\parallel} \cdot \mathbf{E}. \end{aligned} \tag{88}$$

To simplify the discussion, consider media where the two diagonal elements of ϵ_{\parallel} are equal to each other and also equal to the scalar permeability μ . This allows us to write $\epsilon_{\parallel} = \epsilon'_{\parallel} + \mathbf{1}\lambda$ and $\mu = \lambda$, where ϵ'_{\parallel} has zeros on the diagonal. When $\epsilon'_{\parallel} = 0$, the material is *impedance matched* ($\mu = \epsilon = \lambda$), and the field behaves as if the distance has been rescaled by a factor of λ . This is a simple example of ‘transformation optics’ [11, 33]. When the impedance matching condition is satisfied, there is a propagating wave with some wave number $|\mathbf{k}|$ for every value of λ , with positive λ corresponding to positive index media, and negative λ negative index media. There is thus no ‘gap’ in the λ spectrum. Through introducing a *gyrotropy* [27] parameterized by α

$$\epsilon'_{\parallel} = \begin{pmatrix} 0 & -i\alpha(x) \\ i\alpha(x) & 0 \end{pmatrix} \tag{89}$$

we break time reversal symmetry and open up a ‘gap’, where a range of λ values correspond to materials where no wave can propagate. Here we assume that α does not depend on the wave-vector \mathbf{k} of the electromagnetic field as it does in chiral media, thus the material breaks time reversal symmetry. A similar analysis to that given below can be carried out for a chiral medium where α depends on the out-of-plane propagation constant, which is described in the next section. We now use the mode counting argument given above to count the number of interface modes that cross this spectral gap in an inhomogeneous material.

It is assumed that the gyrotropy, α varies with position, as discussed in the theory given above. Maxwell's equations can then be written as an eigenvalue problem equivalent to that

given in (54) at the beginning of the theory of interface modes given in Sec. 6

$$\begin{pmatrix} 0 & i\alpha(x_1) & i\partial_2 \\ -i\alpha(x_1) & 0 & -i\partial_1 \\ i\partial_2 & -i\partial_1 & 0 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ h \end{pmatrix} = \lambda \begin{pmatrix} E_x \\ E_y \\ h \end{pmatrix} \tag{90}$$

where we used the dimensionless coordinates $(x_1, x_2) = k_0(x, y)$. At large $|x|$, where the material becomes homogeneous, the operator \hat{L} can be written in Fourier space as

$$L = \begin{pmatrix} 0 & i\alpha & -k_2 \\ -i\alpha & 0 & k_1 \\ -k_2 & k_1 & 0 \end{pmatrix} \tag{91}$$

which has vanishing trace, as assumed in Sec. 6.

Finding the eigenvalue of (91) is to ask the question “what is the material parameter λ for a medium with gyrotropy α and a wave with wave-vector $\mathbf{k} = (k_1, k_2)$?”. From the eigenvectors and eigenvalues of this operator we can now compute the Chern number of the lower propagation branch shown in Fig. 14. The eigenvectors and eigenvalues are

$$\lambda = \pm\sqrt{k^2 + \alpha^2}, \quad |\psi_{\pm}\rangle = \frac{1}{\sqrt{2k}\sqrt{k^2 + \alpha^2}} \begin{pmatrix} -\lambda k_2 - i\alpha k_1 \\ \lambda k_1 - i\alpha k_2 \\ k^2 \end{pmatrix} \tag{92}$$

and

$$\lambda = 0, \quad |\psi_0\rangle = \frac{1}{\sqrt{k^2 + \alpha^2}} \begin{pmatrix} k_1 \\ k_2 \\ i\alpha \end{pmatrix} \tag{93}$$

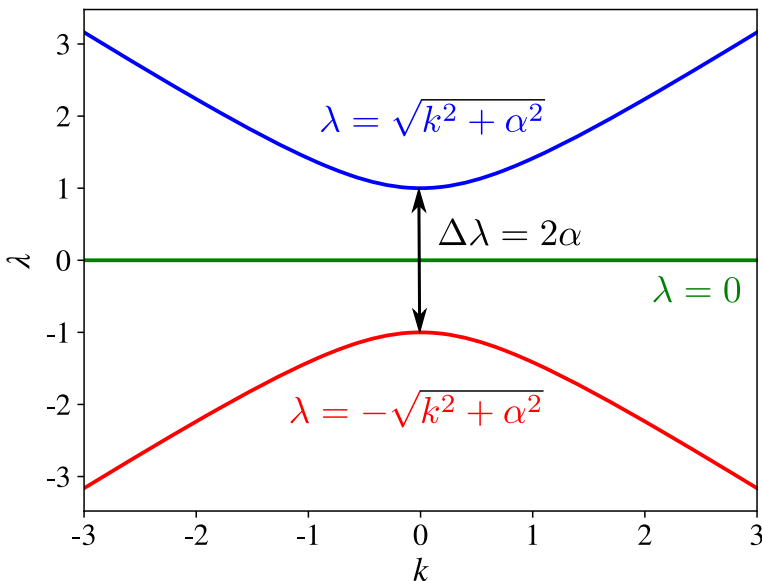


Fig. 14 The material parameters $\lambda = \epsilon = \mu$, calculated as a function of the wavenumber $k = (k_1^2 + k_2^2)^{1/2}$ and fixed gyrotropy $\alpha = 1$. There are three modes with $\lambda = \pm(k^2 + \alpha^2)^{1/2}$ and $\lambda = 0$. A non-zero value of the gyrotropy opens up a ‘gap’ in the spectrum: a range of λ between positive and negative index materials, where there are no propagating solutions

where $k^2 = k_1^2 + k_2^2$ ⁸.

We now count the interface modes that arise from the change in the gyrotropy from $-ve$ to $+ve$. The final result of Sec. 6 says that the number of these modes is given by the difference in the Chern numbers for the homogeneous media either side of the interface. To calculate this we use the Berry connection for the state $|\psi_{-}\rangle$, wrapping the infinite k -space onto the sphere using the stereographic projection (see the schematic in Fig. 8a)

$$\mathcal{K} = k_x + ik_y = k e^{i\phi} = \frac{\sin(\theta) e^{i\phi}}{1 - \cos(\theta)} = \cot(\theta/2) e^{i\phi}. \tag{94}$$

so that on the sphere the state $|\psi_{-}\rangle$ takes the form

$$|\psi_{-}\rangle = \frac{1}{\sqrt{2}\sqrt{\cot^2(\theta/2) + \alpha^2}} \begin{pmatrix} \sqrt{\cot^2(\theta/2) + \alpha^2} \sin(\phi) - i\alpha \cos(\phi) \\ -\sqrt{\cot^2(\theta/2) + \alpha^2} \cos(\phi) - i\alpha \sin(\phi) \end{pmatrix} \tag{95}$$

This state is undefined at both North ($\theta = 0$) and South ($\theta = \pi$) poles. However, at the North pole the vector is purely real, a defect that is associated with zero Berry curvature (31). The Berry connection A_j computed from (95) has a single relevant component

$$A_\phi = -i\langle\psi_{-}|\frac{\partial}{\partial\phi}|\psi_{-}\rangle = \frac{\alpha}{\sqrt{\cot^2(\theta/2) + \alpha^2}} \tag{96}$$

which vanishes at the North pole, and has a critical point at the South pole. Integrating (96) around the critical point at the South pole $\theta = \pi$ yields a Chern number of ± 1 ,

$$\text{Ch}_1 = -\frac{1}{2\pi} \int_0^{2\pi} A_\phi d\phi = -\text{sign}[\alpha]. \tag{97}$$

Thus from (72) the number of interface modes N supported at an interface between media where the gyrotropy changes from positive to negative sign is

$$N = (+1) - (-1) = 2. \tag{98}$$

This pair of interface modes can be understood as those with vanishing tangential electric and magnetic field at the point where the gyrotropy changes sign on the interface, i.e. the interface modes associated with a perfect electric or magnetic conductor placed at this point in the graded material. An examination of this pair of modes has revealed some of the subtleties in the application of the theory of Sec. 6 to continuous media. The reader is encouraged to consult [34] and [35] for more details.

7 One-way Propagation and the Refractive Index

The application of topology to predict interface modes reveals two remarkable things. Firstly that the abstract mathematics of topology has a rather direct and powerful application to the design of materials. Secondly that there exist interface modes that can propagate in only *one* direction (see, for example Fig. 15). This is unusual behaviour for a wave, to say the least. It means that whatever you put in the way of such an interface mode (a mirror, chocolate,

⁸ Unlike the theory presented in Sec. 6, one of the eigenmodes (92–93) has zero eigenvalue. But if the eigenvalues are all shifted *up* by a small positive number η then the zero eigenvalue is removed from the spectrum. This makes the zeroth order spectral asymmetry discussed previously (below (63)) non-zero, but leaves the *difference* in the spectral asymmetry (69) unaffected, and thus all the results derived above hold.

or an elephant), there is simply no possibility for it to reflect. This is not entirely true for the example of Fig. 15, which has a fixed polarization and can thus be reflected by any polarization converting object, but let’s not let that discourage us.

One problem with these topological arguments is that they do not give us an explanation for *why* there are such interface modes. All we have to go on is an integer that took a long time to calculate. What is it about these particular materials that force the wave to propagate in only one direction? In this section we more fully explore the final example of Sec. 6, using the concept of the refractive index rather than topology. This is based on the findings of [36], and we’ll find that the refractive index concept gives us a different, but complementary way to understand one-way propagation.

The starting point is the Berry connection (96) for a gyrotropic medium. As we established in the previous section, the Chern number (97), $Ch_1 = -\text{sign}[\alpha]$ records the single critical point in the Berry connection, which is at the South pole ($\theta = \pi$) of the sphere onto which k -space has been stereographically projected. As shown in Fig. 8), the South pole of the sphere is the origin of k -space. *So what is happening at this critical point in the Berry connection?*

To answer this, let’s return to Maxwell’s Eqns. (88), setting $\mu = \lambda$. We use the same set of material parameters as we did our earlier discussion, $\epsilon_{\parallel} = \lambda \mathbf{1} + i\alpha \mathbf{e}_z \times$. Substituting this in our earlier form of Maxwell’s equations (88), the gradient of the out of plane magnetic field is governed by

$$\nabla h = -ik_0(\lambda \mathbf{e}_z \times \mathbf{E} - i\alpha \mathbf{E}) \tag{99}$$

As we saw earlier in Sec. 4, the complex vectors $\mathbf{e}_{\pm} = (\mathbf{e}_x \pm i\mathbf{e}_y)/\sqrt{2}$ are eigenvectors of the cross product $\mathbf{e}_z \times \mathbf{e}_{\pm} = \mp i\mathbf{e}_{\pm}$. Therefore, taking the inner product of (99) with \mathbf{e}_+ simplifies the equation to

$$\mathbf{e}_+ \cdot \nabla h = \frac{1}{\sqrt{2}} \left(\frac{\partial h}{\partial x} + i \frac{\partial h}{\partial y} \right) = k_0 (\lambda - \alpha) \mathbf{e}_+ \cdot \mathbf{E} \tag{100}$$

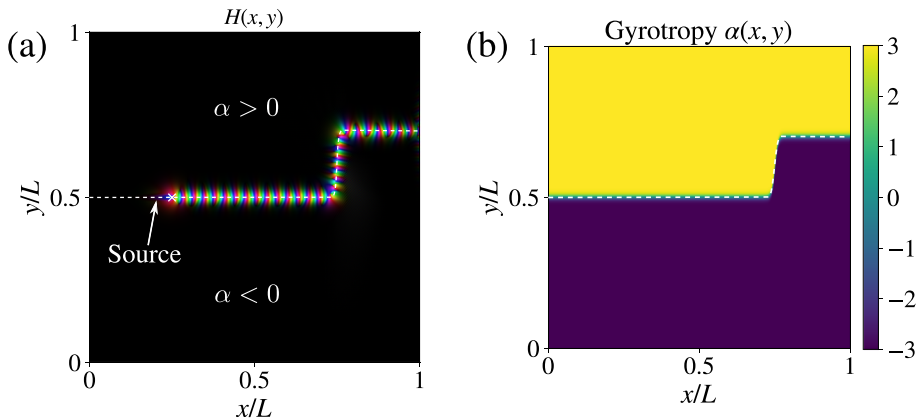


Fig. 15 One-way propagating electromagnetic interface waves, from a source in a medium with graded gyrotropy, as described in the final example of Sec. 6 (simulated using program given in Appendix B). Panel (a) shows a phase plot of the out of plane magnetic field for the profile of gyrotropy given in (b), with constant diagonal part of the permittivity $\epsilon = \lambda = 2.5 + 0.001i$. Away from the interface the value of the gyrotropy exceeds the permittivity and we are in the gap indicated in Fig. 14. A close examination of the interface mode shows that it exhibits a beating as it propagates. This is the interference between the two interface modes predicted in (98)

and with e_- it simplifies to

$$e_- \cdot \nabla h = \frac{1}{\sqrt{2}} \left(\frac{\partial h}{\partial x} - i \frac{\partial h}{\partial y} \right) = -k_0 (\lambda + \alpha) e_- \cdot E \tag{101}$$

Equations (100) and (101) are important. The critical point of the Berry connection is at the point $k = 0$, which from the dispersion plot given in Fig. 14, corresponds to the material parameters $\lambda = \pm\alpha$ (depending on whether we are computing the Chern number of the upper, or lower band of propagation, respectively). From Eqns. (100) and (101)) we can see that at these critical points the magnetic field obeys the equation

$$\frac{\partial h}{\partial x} \pm i \frac{\partial h}{\partial y} = 0 \quad (\lambda = \pm\alpha) \tag{102}$$

These are the *Cauchy–Riemann* equations from complex analysis (see e.g. [37])! These are fulfilled by analytic functions of either $Z^* = x - iy$, in the case $\lambda = \alpha$, or $Z = x + iy$, when $\lambda = -\alpha$. Therefore the critical points of the Berry connection in a gyrotropic medium—which are, of course the reason the Chern number is non-zero—correspond to *those points where the wave is an analytic function of position*. At these points the wave depends solely on either the complex number Z , or on Z^* , depending on the band of interest and the sign of the gyrotropy.

The physical significance of the wave becoming an analytic function is clear if we consider a Taylor expansion of the out of plane magnetic field, h around some point $Z_0 = x_0 + iy_0$ in the plane. Assuming h is a function of Z , and expressing the complex number in terms of polar coordinates (r, θ) centred at the point of expansion

$$h(x + iy) = \sum_{n=0}^{\infty} h_n (Z - Z_0)^n = \sum_{n=0}^{\infty} h_n r^n e^{in\theta}. \tag{103}$$

This expansion in powers of $\exp(i\theta)$ is equivalent to expanding the wave in terms of its component angular momenta. Noting that the terms in the series each evolve in time as $\exp(i(n\theta - \omega t))$, we see that each term rotates with a fixed angular velocity $\dot{\theta} = \omega/n$. As the field must be everywhere finite (assuming the material is simply connected), n is always positive in the series (103). Therefore *a wave that is given as an analytic function of position rotates in only one sense; anti-clockwise in the case of (103), and as shown in Fig. 16b.*

To emphasize the point, compare this to the expansion of a generic function of x and y ,

$$h(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} h_{n,m} x^n y^m = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} h_{n,m} r^{n+m} \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)^n \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^m \tag{104}$$

which—as well as containing two summation indices rather than one—contains both positive and negative powers of $\exp(i\theta)$, meaning that there are component waves that can rotate in both senses around the origin. This is illustrated in Fig. 16a. *The critical point of the Berry connection calculated in (96) therefore corresponds to a set of material parameters where the wave can only circulate one way.*

7.1 Critical Points and the Refractive Index

We have just established that in a gyrotropic medium, the points where $\lambda = \pm\alpha$ are where the wave behaves as an analytic function of position. As the length of the wave-vector also vanishes at this point, $k = 0$, it is reminiscent of a point of vanishing refractive index. Indeed,

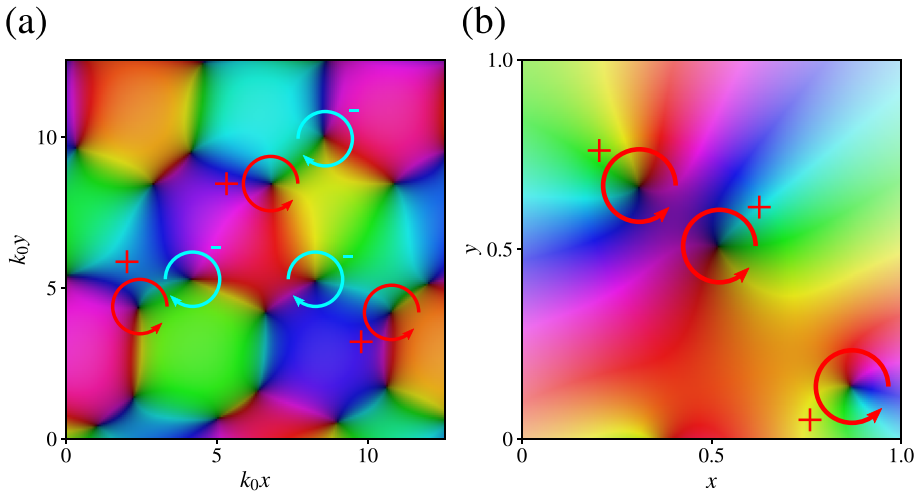


Fig. 16 Difference between (a) a random superposition of plane waves of wavenumber k_0 , and (b) an analytic function of $Z = x + iy$. Arrows indicate the direction in which the phase increases around the zeros. In panel (a) we plot the complex function obtained through adding together 8 plane waves of random complex amplitude, propagating at angles $\{0, \pi/4, \pi/2, 3\pi/4, \dots\}$. In panel (b) we plot the function $\Pi_n(z - z_n)$ for 8 randomly generated complex numbers z_n . While the sum of waves generates zeros around which the phase circulates in either sense, the analytic function always exhibits circulation in an anti-clockwise sense

the behaviour of the field can be connected to the study of wave propagation in anisotropic materials, and the critical points of the Berry connection can be understood as an unusual kind of point of vanishing refractive index. We shall show that this finding can be used as a shortcut to materials where there are one-way interface states.

To simplify the discussion we consider the propagation of a transverse magnetic field $\mathbf{H} = H\mathbf{e}_z$ in the x - y plane of a homogeneous non-magnetic ($\mu = 1$) material. Combining the two Maxwell equations given by (88), the electric field can be eliminated, leaving a second order equation for the out of plane magnetic field amplitude h

$$\nabla \times (\boldsymbol{\epsilon}^{-1} \cdot \nabla h \times \mathbf{e}_z) = k_0^2 h \mathbf{e}_z. \tag{105}$$

As the material is homogeneous, we can write the magnetic field in the form $h = \mathbf{e}_z h_0 \exp(i\mathbf{k} \cdot \mathbf{x})$, where $\mathbf{n}(\theta) = \cos(\theta)\mathbf{e}_x + \sin(\theta)\mathbf{e}_y$ and $k = (k_x^2 + k_y^2)^{1/2}$, thus eliminating the derivatives from (105). Assuming a material of the same form as in (87), where the only off-diagonal elements are ϵ_{xy} and ϵ_{yx} , (105) becomes an equation for the θ dependent refractive index $n = k/k_0$

$$(\mathbf{n} \times \mathbf{e}_z) \cdot \boldsymbol{\epsilon}_{\parallel}^{-1} \cdot (\mathbf{n} \times \mathbf{e}_z) = \left(\frac{k_0}{k}\right)^2 = \frac{1}{n(\theta)^2} \tag{106}$$

This equation defines the refractive index as a function of the propagation angle θ in the x - y plane. Equation (106) is a special case of the defining equation for the *refractive index ellipsoid*, used in the optics of three dimensional crystals [38].

It is illustrative to re-write the dispersion relation (106) in terms of the eigenvalues, ϵ_i and eigenvectors, \mathbf{e}_i of the in-plane permittivity, obeying $\boldsymbol{\epsilon}_{\parallel} \cdot \mathbf{e}_i = \lambda_i \mathbf{e}_i$. In terms of these

quantities, the inverse of the in-plane permittivity is given by,

$$\epsilon_{\parallel}^{-1} = \frac{1}{\epsilon_1} \mathbf{e}_1 \otimes \mathbf{e}_1^* + \frac{1}{\epsilon_2} \mathbf{e}_2 \otimes \mathbf{e}_2^* \tag{107}$$

where we've assumed a Hermitian (and hence lossless) permittivity tensor. Substituting (107) into the dispersion relation (106), the angle dependent refractive index can then be written as

$$n(\theta) = \sqrt{\frac{\epsilon_1 \epsilon_2}{\epsilon_2 |(\mathbf{e}_1 \times \mathbf{e}_z) \cdot \mathbf{n}|^2 + \epsilon_1 |(\mathbf{e}_2 \times \mathbf{e}_z) \cdot \mathbf{n}|^2}}, \tag{108}$$

where we have taken the positive root (although in some important cases the negative root should be taken [10]).

Take a moment to dwell on the dependence of the refractive index (108) on the permittivity tensor. As the permittivity is Hermitian, the eigenvectors are orthonormal $\mathbf{e}_i^* \cdot \mathbf{e}_j = \delta_{ij}$, and form a complete set $\mathbf{1} = \mathbf{e}_1^* \otimes \mathbf{e}_1 + \mathbf{e}_2^* \otimes \mathbf{e}_2$. Therefore, when the two eigenvalues are equal $\epsilon_1 = \epsilon_2 = \epsilon$, the denominator on the right of (108) simply equals ϵ . Such a medium is an isotropic dielectric in the plane of propagation, and the refractive index reduces to the textbook expression, $n(\theta) = \sqrt{\epsilon}$, shown as the dashed circle in Fig. 17a. As shown in the figure, when $\epsilon \rightarrow 0$, this dispersion circle closes to a point and the refractive index vanishes.

Meanwhile, when the two eigenvalues differ $\epsilon_2 > \epsilon_1$ and the eigenvectors are real, the refractive index varies between its largest value $\sqrt{\epsilon_2}$ (propagation along \mathbf{e}_1) and its smallest value $\sqrt{\epsilon_1}$ (propagation along \mathbf{e}_2). The angle dependence of the refractive index $n(\theta)$ now either traces out an ellipse (Fig. 17b), when ϵ_1 and ϵ_2 are both positive, or a hyperbola (Fig. 17c), when ϵ_1 and ϵ_2 have different signs, constituting a hyperbolic material.⁹ At the transition between elliptical and hyperbolic dispersion, the smallest eigenvalue passes through zero $\epsilon_1 \rightarrow 0$. In this case the eccentricity of the dispersion ellipse tends to unity, and the ellipse is compressed into a line, as shown in Fig. 17b. Again *the refractive index vanishes, but now this only occurs for one direction of propagation*. For instance taking $\mathbf{e}_1 = \mathbf{e}_x$ and $\mathbf{e}_2 = \mathbf{e}_y$ we have,

$$n(\theta) = \sqrt{\frac{\epsilon_1 \epsilon_2}{\epsilon_2 \sin^2(\theta) + \epsilon_1 \cos^2(\theta)}} = \begin{cases} \sqrt{\epsilon_2} \rightarrow 0 & \text{(Propagation along } x \text{)} \\ \sqrt{\epsilon_1} \neq 0 & \text{(Propagation along } y \text{)}. \end{cases} \tag{109}$$

We can therefore see that setting one of the eigenvalues of the in-plane permittivity to zero makes the refractive index in the direction perpendicular to the corresponding eigenvector vanish. In this way it is possible to have zero refractive index for only one direction of propagation.

The situation becomes more interesting when the eigenvectors \mathbf{e}_i are complex and the eigenvalues are positive. Now the propagation vector \mathbf{n} , appearing in the denominator of (108), can never be completely parallel (or orthogonal) to either of the complex vectors, $\mathbf{e}_1 \times \mathbf{e}_z$ or $\mathbf{e}_2 \times \mathbf{e}_z$. As a consequence the denominator can *never* vanish, even if one of the eigenvalues are zero! Therefore, if we let the smallest eigenvalue ϵ_1 alone tend to zero, the numerator of (108) is zero, making the refractive index $n(\theta)$ vanish for all directions of propagation θ ! As shown in Fig. 17d, in this limit the dispersion ellipse closes to a point, like that of an isotropic zero index medium, where the permittivity tensor as a whole vanishes. However, the behaviour is more subtle now.

⁹ See [20] for details about these fascinating materials.

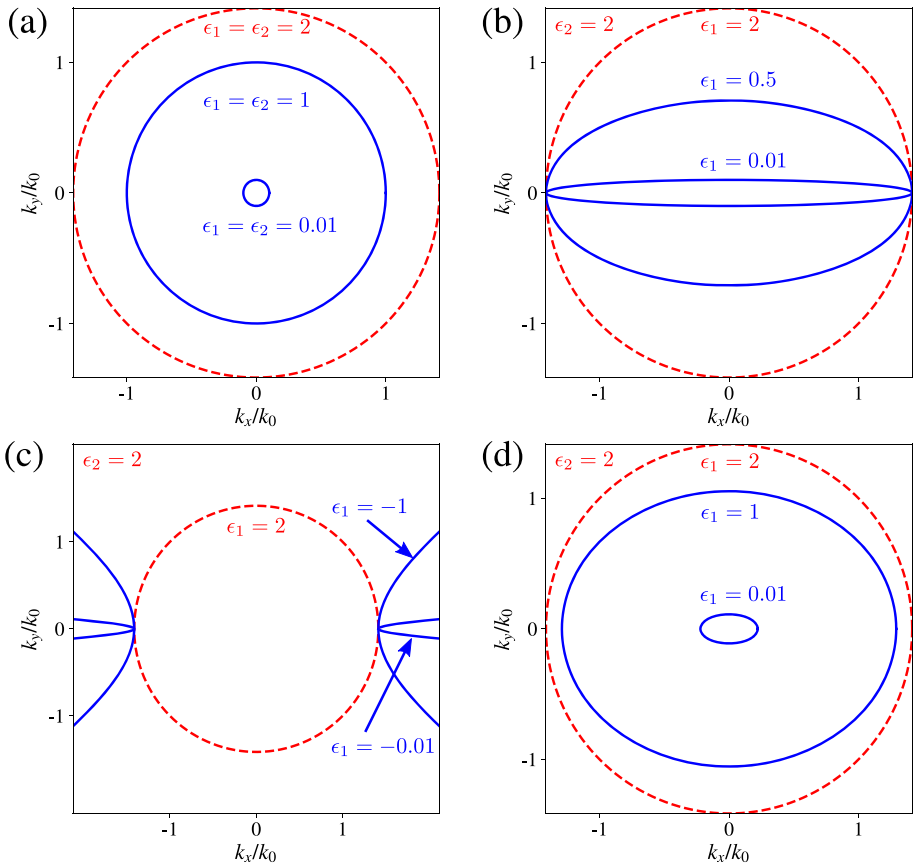


Fig. 17 Refractive index surfaces in planar media: (a) Isotropic medium with equal eigenvalues. The dispersion surface closes to a point as the eigenvalues are reduced to zero. In panel (b) we have an anisotropic medium with eigenvectors $\mathbf{e}_1 = \mathbf{e}_x$ and $\mathbf{e}_2 = \mathbf{e}_y$, and positive eigenvalues. The dispersion surface now forms an ellipse, with eccentricity approaching unity as one of the eigenvalues approaches zero. In panel (c) the medium is hyperbolic with the same eigenvectors as (b). The dispersion surface is now open, with a range of angles where propagation is not allowed. Finally, panel (d) shows the dispersion surface for complex eigenvectors $\mathbf{e}_1 = (2\mathbf{e}_x + i\mathbf{e}_y)/\sqrt{5}$ and $\mathbf{e}_2 = (\mathbf{e}_x + 2i\mathbf{e}_y)/\sqrt{5}$. Unlike the case of real eigenvectors, the dispersion surface closes to a point rather than a line as only *one* of the eigenvalues approaches zero

If we return to the defining equation for the refractive index (106) and multiply through by k , we have

$$(\mathbf{k} \times \mathbf{e}_z) \cdot \left(\frac{1}{\epsilon_1} \mathbf{e}_1 \otimes \mathbf{e}_1^* + \frac{1}{\epsilon_2} \mathbf{e}_2 \otimes \mathbf{e}_2^* \right) \cdot (\mathbf{k} \times \mathbf{e}_z) = k_0^2. \tag{110}$$

Given that the magnitude of the wave number k_0 is fixed, as $\epsilon_1 \rightarrow 0$ the first term in the brackets of (110) dominates and we are left with the condition

$$\mathbf{e}_1 \cdot (\mathbf{k} \times \mathbf{e}_z) \rightarrow 0 \tag{111}$$

an equation that could be equivalently written as $(\mathbf{e}_z \times \mathbf{e}_1) \cdot \nabla h = 0$, i.e. the refractive index is zero in the $\mathbf{e}_z \times \mathbf{e}_1^*$ direction. For real \mathbf{e}_1 , this indicates the squashing of the dispersion ellipse into a line, as shown in Fig. 17. When the first eigenvector is a complex vector e.g.

$\mathbf{e}_1 = \mathbf{e}_+ = (\mathbf{e}_x + i\mathbf{e}_y)/\sqrt{2}$, the refractive index is zero in a complex direction, and our condition reduces to $\partial_x h + i\partial_y h = 0$, which are the Cauchy–Riemann conditions (102) found earlier.¹⁰

Therefore, even though the dispersion surface in the limit $\epsilon_1 \rightarrow 0$ shown in Fig. 17d looks like that of an isotropic medium where the refractive index vanishes, the behaviour of the wave is quite different. Rather than uniformly stretch the wavelength to infinity as would happen in an isotropic medium, instead the wave is forced to propagate with only one sense of circulation. Although analytic functions diverge at infinity and are therefore inadmissible in a bulk material, this behaviour is revealed at a boundary with another material (see Sec. 8), where e.g. an interface state $\exp(-ik(x - iy))$ ($y > 0$) would be an allowed solution, whereas the counter propagating wave $\exp(ik(x + iy))$ would not. This unusual kind of zero index material exhibits one-way propagation where the wave obeys the Cauchy–Riemann conditions. This is what the defect in the Berry connection (96), and the non-zero Chern number (97) is recording.

8 Applications

We now give applications in three different wave physics regimes where we can enforce one-way propagation through simply demanding that the refractive index is zero in a complex direction.

8.1 General Electromagnetic Media

We can use this idea of a ‘vanishing index in a complex direction’ to extend the discussion of Sec. 7 from gyrotropic media, to general electromagnetic materials. In an arbitrary material, in the absence of any sources, and at a fixed frequency ω , Maxwell’s equations are given by

$$\begin{aligned} \nabla \times \mathbf{E} &= i\omega\mathbf{B} \\ \nabla \times \mathbf{H} &= -i\omega\mathbf{D}. \end{aligned} \tag{112}$$

We take a general lossless linear material, where the constitutive relations are given by

$$\begin{aligned} \mathbf{D} &= \epsilon_0[\boldsymbol{\epsilon} \cdot \mathbf{E} + \boldsymbol{\xi} \cdot \eta_0\mathbf{H}] \\ \mathbf{B} &= \mu_0[\boldsymbol{\mu} \cdot \mathbf{H} + \boldsymbol{\xi}^\dagger \cdot \eta_0^{-1}\mathbf{E}] \end{aligned} \tag{113}$$

where the three 3×3 tensors $\boldsymbol{\epsilon}$ and $\boldsymbol{\mu}$ are Hermitian, and the bi-anisotropy tensor $\boldsymbol{\xi}$ is arbitrary. The Hermitian property of the permittivity and permeability, and the appearance of $\boldsymbol{\xi}$ and $\boldsymbol{\xi}^\dagger$ ensure that the material does not absorb wave energy [39].

A compact and useful way to write Maxwell’s equations (112) is in the form of a six-vector $(\mathbf{E}, \mathbf{h})^T$, as follows

$$\begin{pmatrix} \mathbf{0} & i\nabla \times \\ -i\nabla \times & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{h} \end{pmatrix} = k_0 \begin{pmatrix} \boldsymbol{\epsilon} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^\dagger & \boldsymbol{\mu} \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{h} \end{pmatrix} \tag{114}$$

where, as in the previous sections $\mathbf{h} = \eta_0\mathbf{H}$. As discussed in [29, 40–42], equation (114) has a great deal in common with the Dirac equation [43], where the operator on the left hand side

¹⁰ Other choices of complex vectors also yield the Cauchy–Riemann conditions, but with the coordinates x and y rescaled.

is analogous to the operator $\alpha \cdot \hat{p}$, and the right hand side matrix contains terms analogous to the mass, energy, and an external gauge field.

We now restrict propagation to the x - y plane. With this assumption, the curl of the fields can be written in terms of derivatives of the in-plane field components e.g. \mathbf{E}_{\parallel} , and the out of plane ones e.g. $E_z = E$. For example, $\nabla \times \mathbf{E} = \nabla \times \mathbf{E}_{\parallel} + \nabla E \times \mathbf{e}_z$. With this assumption, the in-plane part of the left hand side of (114) depends only on the gradient of the out of plane field components E and h . To isolate these parts of the field, we take an inner product of (114) with the six-vector

$$V = \begin{pmatrix} v_E \\ 0 \\ v_H \\ 0 \end{pmatrix} \tag{115}$$

where $v_{E,H}$ are two arbitrary vectors lying in the x - y plane¹¹. This inner product yields the single scalar equation for the derivatives of the out-of-plane field

$$i [v_E \cdot \nabla h \times \mathbf{e}_z - v_H \cdot \nabla E \times \mathbf{e}_z] = k_0 [(v_E \cdot \boldsymbol{\epsilon} + v_H \cdot \boldsymbol{\xi}^\dagger) \cdot \mathbf{E} + (v_H \cdot \boldsymbol{\mu} + v_E \cdot \boldsymbol{\xi}) \cdot \mathbf{h}]. \tag{116}$$

This equation can be reduced to a simple gradient of a combination of out of plane field components if the two vectors $v_{E,H}$ are chosen as parallel. We write $v_{E,H} = \alpha_{E,H} \mathbf{e} \times \mathbf{e}_z$ where $\alpha_{E,H}$ are scalar quantities, \mathbf{e} is a unit vector ($\mathbf{e} \cdot \mathbf{e}^* = 1$). We then have,

$$i \mathbf{e} \cdot \nabla [\alpha_E h - \alpha_H E] = k_0 (\mathbf{e} \times \mathbf{e}_z) \cdot [(\alpha_E \boldsymbol{\epsilon} + \alpha_H \boldsymbol{\xi}^\dagger) \cdot \mathbf{E} + (\alpha_H \boldsymbol{\mu} + \alpha_E \boldsymbol{\xi}) \cdot \mathbf{h}]. \tag{117}$$

The left hand side of this equation is a generalization of Eqns. (100) and (101) discussed in Sec. 7. Setting the right hand side of (117) to zero picks out a set of material parameters such that the refractive index is zero in the direction \mathbf{e} . In order for this to hold, we must impose two conditions on the material tensors

$$\begin{aligned} (\mathbf{e} \times \mathbf{e}_z) \cdot (\alpha_E \boldsymbol{\epsilon} + \alpha_H \boldsymbol{\xi}^\dagger) &= 0 \\ (\mathbf{e} \times \mathbf{e}_z) \cdot (\alpha_H \boldsymbol{\mu} + \alpha_E \boldsymbol{\xi}) &= 0. \end{aligned} \tag{118}$$

For the particular case of $\mathbf{e} = \mathbf{e}_+ = (\mathbf{e}_x + i\mathbf{e}_y)/\sqrt{2}$, (118) provides a large family of material parameters where the out of plane field component $\alpha_E h - \alpha_H E$ behaves as an analytic function of position; thus circulating in only one sense and exhibiting unidirectional interface states. Note that in the particular case where we take $\alpha_H = 0$ and $\boldsymbol{\xi} = 0$, (118) reduces to the zero-index condition for gyrotropic media $\mathbf{e}_+ \times \mathbf{e}_z \cdot \boldsymbol{\epsilon} = 0$ identified above in Eqns. (99–101).

Example: The Cauchy–Riemann conditions in free space There is an interesting special case of conditions (118), where it can be applied to electromagnetic near fields propagating in *free space*. At first this seems counter intuitive: surely we need a material if the wave is to exhibit something as strange as complex analyticity! But suppose we consider a free space electromagnetic wave propagating out of the plane, at fixed wave-vector k_z . The dispersion relation of such a mode is given by

$$k_x^2 + k_y^2 + k_z^2 = k_0^2 \rightarrow k_0 = \pm \sqrt{k^2 + k_z^2} \tag{119}$$

¹¹ Recall that the topological arguments described in Secs. 5 and 6 are only applicable to planar systems. This is also true for the arguments based on the refractive index given here.

which is of the same form as for a gyrotropic medium (92), and for a relativistic particle: we have two bands of propagation, separated by a ‘gap’ $\Delta k_0 = 2k_z$ equivalent to a mass, and due to the propagation constant k_z , as shown in Fig. 18. The system also has exactly the same behaviour under time reversal. Just as for the gyrotropy constant α , if we reverse the direction of time the out of plane wave-vector k_z changes sign, although such a sign change does not affect the dispersion relation (119). As we shall see, in a particular polarization basis, out of plane propagation is completely equivalent to gyrotropy.

As above, consider a system that is translationally invariant along the z axis. Instead of taking the field as uniform in z , we assume propagation with wave vector component k_z . Separating out the in-plane derivatives as $\nabla = \nabla_{\parallel} + e_z \partial_z$, this modifies Maxwell’s equations (112) to

$$\begin{aligned} \nabla_{\parallel} \times \mathbf{E} + ik_z e_z \times \mathbf{E} &= i\omega \mathbf{B} \\ \nabla_{\parallel} \times \mathbf{H} + ik_z e_z \times \mathbf{H} &= -i\omega \mathbf{D} \end{aligned} \tag{120}$$

Comparison with the constitutive relations (113) we can see that out of plane propagation is equivalent to adding an anti-symmetric contribution to the bi-anisotropy tensor $\Delta \xi = (k_z/k_0) e_z \times$. This effective contribution to the bi-anisotropy is what allows us to fulfil the zero index condition (118), even in free space.

Taking free space $\epsilon = \mu = \mathbf{1}_3$, and the effective bi-anisotropy $\xi = (k_z/k_0) e_z \times$, the zero index conditions become

$$(e \times e_z) \cdot (\alpha_E \epsilon + \alpha_H \xi^\dagger) = (e_y -e_x \ 0) \begin{pmatrix} \alpha_E & \alpha_H \left(\frac{k_z}{k_0}\right) & 0 \\ -\alpha_H \left(\frac{k_z}{k_0}\right) & \alpha_E & 0 \\ 0 & 0 & \alpha_E \end{pmatrix} = 0 \tag{121}$$

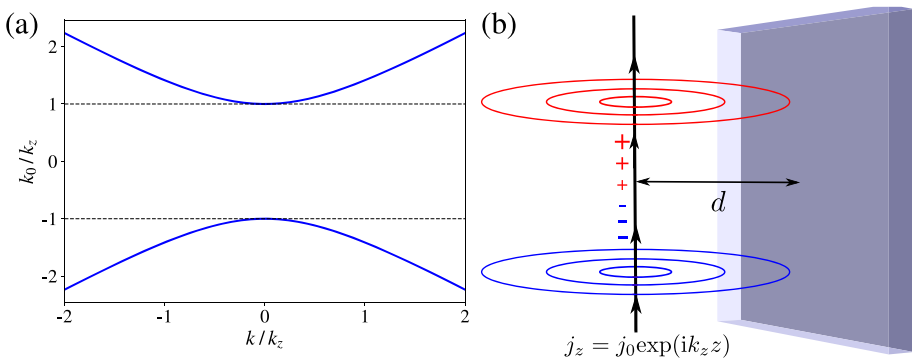


Fig. 18 Free space exhibits a similar dispersion relation to that of a gyrotropic medium, see e.g. Fig. 14, and in some cases shows ‘zero index in a complex direction’. (a) For a fixed out of plane propagation constant k_z the frequency is $k_0 = \pm(k^2 + k_z^2)^{1/2}$, with a ‘gap’ in the dispersion relation where $k_0 < |k_z|$. The wave becomes an analytic function of position when $k_0 = |k_z|$ (dashed lines). (b) An imagined experiment where an oscillating line current $j_z = j_0 \exp(ik_z z)$ excites electromagnetic waves close to a planar dielectric, a distance d away, where one-way propagation is evident when $k_0 < |k_z|$



Fig. 19 Phase plot of the electric E , magnetic h , and the two ‘circular polarizations’ $E \pm ih$ for an oscillating line source of wave number k_z in front of a metal with $\epsilon = -2$ (see Fig. 18b). As the wavenumber approaches the free space wave number $k_0/k_z \rightarrow 1$, the system approaches the ‘zero index’ point shown in Fig. 18a, and the field components $E \pm ih$ behave as analytic functions of Z and Z^*

and

$$(\mathbf{e} \times \mathbf{e}_z) \cdot (\alpha_H \boldsymbol{\mu} + \alpha_E \boldsymbol{\xi}) = (e_y \ -e_x \ 0) \begin{pmatrix} \alpha_H & -\alpha_E \left(\frac{k_z}{k_0}\right) & 0 \\ \alpha_E \left(\frac{k_z}{k_0}\right) & \alpha_H & 0 \\ 0 & 0 & \alpha_H \end{pmatrix} = 0. \quad (122)$$

In both equations (121) and (122) the matrix in the middle equation represents an effective permittivity for the combination of fields $\alpha_E h - \alpha_H E$, analogous to the gyrotropic permittivity defined below (88). Choosing the combination of fields where $\alpha_E = 1$ and $\alpha_H = i$, these conditions become *exactly* the same as for a gyrotropic medium. For the complex direction

$e = e_+$, the zero index conditions (121–122) are fulfilled when

$$k_z = -k_0. \tag{123}$$

For a wave propagating out of the plane with a wave–vector obeying (123), the linear combination of fields ¹² $h - iE = -i(E + ih)$ behaves as an analytic function of position in the x – y plane, and is thus forced to circulate in only one sense (anti–clockwise). Meanwhile the polarization $h + iE$ circulates in the opposite sense (clockwise). The sense of rotation is reversed for both polarizations when we take the opposite direction of out of plane propagation, $k_z = +k_0$.

Figure 18 shows an example where the electromagnetic field from an oscillating line source (see schematic in Fig. 18b) has been calculated analytically in terms of the Fresnel coefficients of an isotropic half space of permittivity $\epsilon = -2$ and permeability $\mu = 1$ (see e.g. [44] for details). As the wave–number k_z of the line current approaches k_0 we can see that a confined mode (a surface plasmon) emerges, which—in terms of the field components $E + iH$ and $E - iH$ —can only propagate in one direction on the interface (Fig. 19).

8.2 Continuous Elastic Media

Having shown the applicability of our zero index condition to general electromagnetic materials, we give an example for another kind of wave. In the theory of elasticity, the equation of motion is a continuous version of the Newtonian equation of motion $F = ma$ [45],

$$\rho \frac{\partial^2 U}{\partial t^2} = -\rho \omega^2 U = \nabla \cdot \sigma. \tag{124}$$

where the local force density is the divergence of the stress tensor $\nabla \cdot \sigma \equiv \partial_i \sigma_{ij}$, the material mass density is ρ , and the local displacement of the material from its equilibrium position is U . We assume an elastic wave of fixed frequency, which gives the middle equation in (124) where we applied $\partial_t^2 \rightarrow -\omega^2$.

We cannot use (124) without a constitutive relation between the stress and the displacement. More precisely it is the relative displacement of different parts of a body—the strain, $u_{ij} = (\partial_i U_j + \partial_j U_i)/2$ —rather than an overall displacement that gives rise to stress, and for linear elastic materials the stress and strain are related by the rank four *stiffness tensor* C_{ijkl} ,

$$\sigma_{ij} = C_{ijkl} u_{kl}. \tag{125}$$

The derivatives of the local displacement field are thus governed by the inverse stiffness tensor, known as the *compliance tensor* C_{ijkl}^{-1}

$$u_{kl} = \frac{1}{2} \left(\frac{\partial U_k}{\partial x_l} + \frac{\partial U_l}{\partial x_k} \right) = C_{kl ij}^{-1} \sigma_{ij} \tag{126}$$

which obeys

$$C_{ijkl}^{-1} C_{kl pq} = \delta_{ip} \delta_{jq}. \tag{127}$$

As in the theory of Sec. 7, we assume the field is independent of the z coordinate (e.g. confinement in an elastic plate, or waveguide), and propagates solely in the x – y plane. This

¹² For propagation in the x – y plane, this combination of fields represents the amplitude of the circular polarization that rotates anti–clockwise over time.

means that the z components of the strain tensor simplify to

$$u_{13} = \frac{1}{2} \frac{\partial U_3}{\partial x} \quad u_{23} = \frac{1}{2} \frac{\partial U_3}{\partial y} \quad u_{33} = \frac{\partial U_3}{\partial z} = 0. \tag{128}$$

Taking $k = 3$ in the constitutive relation (126), and using the simplified form of the strain tensor (128), the spatial derivatives of the out of plane displacement are given in terms of the stress tensor by,

$$\frac{\partial U_3}{\partial x_l} = 2C_{3lij}^{-1} \sigma_{ij}. \tag{129}$$

Contracting both sides of (129) with the unit vector \mathbf{e} (components e_l), we obtain an equation analogous to (117), telling us the derivative of the out of plane displacement in the \mathbf{e} direction

$$\mathbf{e} \cdot \nabla U_3 = 2C_{3lij}^{-1} e_l \sigma_{ij}, \tag{130}$$

This is what we were after! For elasticity, this allows us to set the derivative of the wave amplitude in a given direction to zero. In order that the elastic refractive index vanish in direction \mathbf{e} (i.e. U_3 is stretched to uniformity in this direction), the \mathbf{e} vector must be a zero eigenvector of the compliance tensor

$$C_{3lij}^{-1} e_l = 0. \tag{131}$$

To enforce the Cauchy–Riemann conditions, \mathbf{e} must be one of the complex vectors \mathbf{e}_\pm , which implies that a stiffness tensor C_{ijkl} supporting this kind of propagation must also be complex. Yet in lossless systems the stiffness tensor is usually taken as a real symmetric object obeying $C_{ijkl} = C_{klij}$. However, just as in electromagnetism when dealing with monochromatic waves, the stiffness tensor can take complex values. A general lossless linear elastic medium has a *Hermitian*, rather than real symmetric stiffness tensor (see e.g. [46]), obeying $C_{ijkl} = C_{klij}^*$. Just as in electromagnetism, such complex valued stiffness tensors arise in systems where time reversal symmetry has been explicitly broken due to e.g. an externally applied magnetic field (for example, magnetostriction effects), or motion of the medium.

Example: An elastic material exhibiting the Cauchy–Riemann conditions Condition (131) can be used to design an elastic medium where a one-way propagating elastic wave is trapped at its interface. To show this we start from an isotropic elastic material, which has the following form of stiffness tensor in terms of the bulk K and shear G moduli [45]

$$C_{ijkl} = \left(K - \frac{2G}{3} \right) \delta_{ij} \delta_{kl} + 2G \delta_{ik} \delta_{jl}. \tag{132}$$

Take propagation in the x – y plane and shear displacement solely in the z direction. The components of the strain are $u_{3i} = (1/2) \partial_i U_3$, as identified above in (128). For the isotropic medium (132) the first bulk modulus dependent term does not contribute to the stress, which is related to the strain by a diagonal 2×2 matrix containing the stiffness tensor elements $C_{3131} = 2G$ and $C_{3232} = 2G$,

$$\begin{pmatrix} \sigma_{31} \\ \sigma_{32} \end{pmatrix} = \begin{pmatrix} C_{3131} & C_{3132} \\ C_{3231} & C_{3232} \end{pmatrix} \begin{pmatrix} u_{31} \\ u_{32} \end{pmatrix} = \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix} \begin{pmatrix} \partial_1 U_3 \\ \partial_2 U_3 \end{pmatrix}. \tag{133}$$

and similarly for inverse relation containing the compliance tensor

$$\begin{pmatrix} \partial_1 U_3 \\ \partial_2 U_3 \end{pmatrix} = 2 \begin{pmatrix} C_{3131}^{-1} & C_{3132}^{-1} \\ C_{3231}^{-1} & C_{3232}^{-1} \end{pmatrix} \begin{pmatrix} \sigma_{31} \\ \sigma_{32} \end{pmatrix} = \begin{pmatrix} \frac{1}{G} & 0 \\ 0 & \frac{1}{G} \end{pmatrix} \begin{pmatrix} \sigma_{31} \\ \sigma_{32} \end{pmatrix}. \tag{134}$$

With this form of material parameters it is not possible to fulfil our zero index condition (131), without sending the shear modulus to infinity, which in optics is equivalent to an *isotropic* zero index medium.

By analogy with the discussion of gyrotropic electromagnetic materials in Sec. 7, we make the components $C_{3132}^{-1} = -i\alpha/2$ of the compliance tensor non-zero and purely imaginary.

$$\begin{pmatrix} \partial_1 U_3 \\ \partial_2 U_3 \end{pmatrix} = \begin{pmatrix} \kappa & -i\alpha \\ i\alpha & \kappa \end{pmatrix} \begin{pmatrix} \sigma_{31} \\ \sigma_{32} \end{pmatrix} \tag{135}$$

where we set the ‘diagonal’ of the compliance tensor as $C_{3131}^{-1} = C_{3232}^{-1} = \kappa/2$. Figure 20 gives a sketch of what the stress–strain relationship is like in such a material, with the direction of the in–plane stress circulating over a single cycle of the wave. Performing an inner product of both sides of (135), we find the derivative of the out of plane displacement in the $e_{\pm}^* = e_{-}$ direction

$$\frac{\partial U_3}{\partial x} + i \frac{\partial U_3}{\partial y} = (\kappa - \alpha) (\sigma_{31} + i \sigma_{32}). \tag{136}$$

which equals zero when $\alpha = \kappa$ (analogous to the $\alpha = \lambda$ point in the dispersion relation of the gyrotropic medium shown in Fig. 14), at which point the wave becomes an analytic function of position. The stiffness tensor C_{ijkl} corresponding to the choice (135) has components

$$\begin{pmatrix} C_{3131} & C_{3132} \\ C_{3231} & C_{3232} \end{pmatrix} = \frac{2}{\kappa^2 - \alpha^2} \begin{pmatrix} \kappa & i\alpha \\ -i\alpha & \kappa \end{pmatrix} \tag{137}$$

showing that the stiffness tensor must have very large components close to the zero index points $\kappa = \pm\alpha$ (just as it must for an isotropic zero index elastic medium).

Now we consider a planar elastic medium with a stiffness tensor of the form (137), and solve the equation of motion (124), assuming the material is terminated by vacuum. Given

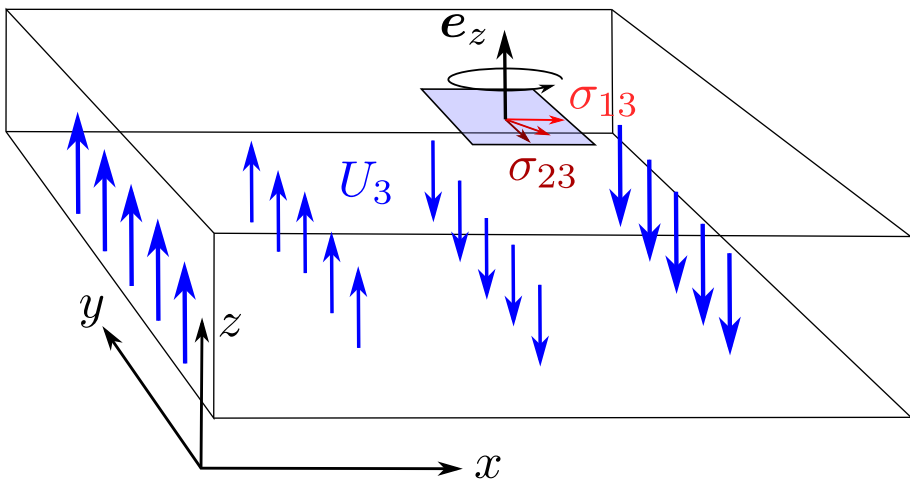


Fig. 20 Elastic materials with ‘zero index in a complex direction’. In an ordinary infinite isotropic elastic material, a shear wave polarized along e_z and propagating in the x direction gives rise to an off–diagonal stress σ_{13} , which means that a small area element pointing along e_z is subject to a force along the x axis. The peculiar zero index materials studied here, and defined by (135), have a stress–strain relation that circulates over time. During a single cycle the shear force will rotate from the x –axis to the y –axis, to the $-ve$ x –axis, and so on

that the displacement field has only a single non-zero component $\mathbf{U} = U_3 \mathbf{e}_z$, and only the components σ_{12} and σ_{13} of the stress are non-zero, the equation of motion (124) reduces to

$$\nabla \cdot \boldsymbol{\sigma} + \rho \omega^2 \mathbf{U} = \frac{\partial \sigma_{13}}{\partial x} + \frac{\partial \sigma_{23}}{\partial y} + \rho \omega^2 U_3 = 0 \tag{138}$$

For a homogeneous medium, where the parameters κ and α in (137) are independent of position, the equation of motion becomes the Helmholtz equation

$$\nabla^2 U_3 + \frac{\rho \omega^2}{\kappa} (\kappa^2 - \alpha^2) U_3 = 0 \tag{139}$$

which is identical to that for a scalar wave in a material with wave number $|\mathbf{k}| = \omega \sqrt{\rho (\kappa^2 - \alpha^2)} / \kappa$. Superficially the wave appears to behave as if in an isotropic zero index medium, and as $\alpha \rightarrow \pm \kappa$, the dispersion circle closes to a point as shown in Fig. 17a and d. As discussed previously, the one-way propagation is only evident at inhomogeneities, e.g. interfaces.

If the material has an interface with vacuum, with surface normal \mathbf{e}_y , the stress components $\mathbf{e}_y \cdot \boldsymbol{\sigma}$ will equal zero,

$$\sigma_{23} = \frac{1}{\kappa^2 - \alpha^2} \left(-i\alpha \frac{\partial U_3}{\partial x} + \kappa \frac{\partial U_3}{\partial y} \right) = 0. \tag{140}$$

Assuming that the elastic medium occupies $y < 0$, and vacuum occupies $y > 0$, we can write an interface state as $U_3 = \exp(ikx + \beta y)$, where β is the real and positive decay constant of the wave into the elastic medium. Demanding that the normal stress vanishes as in (140) relates the decay constant and the propagation constant

$$k = -\frac{\kappa}{\alpha} \beta \tag{141}$$

implying $k < 0$ when $\alpha > 0$, and $k > 0$ when $\alpha < 0$. As the decay constant β must be positive, the interface state can only satisfy the boundary condition (140) for one direction of propagation. When our zero index condition is satisfied $\alpha = \kappa$, the displacement U_3 becomes a function of $x + iy$, i.e. an analytic function, obeying the Cauchy–Riemann conditions and circulating in only an anti-clockwise sense.

To show that this mode is a solution to the equations of elasticity (124), we finally need to verify that the boundary condition (141) is consistent with the dispersion relation derived from (139),

$$k^2 = k^2 - \beta^2 = \frac{\rho \omega^2}{\kappa} (\kappa^2 - \alpha^2) \rightarrow k^2 = \rho \kappa \omega^2 \tag{142}$$

so that when $\alpha > 0$ the mode satisfies

$$k = -\sqrt{\rho \kappa} \omega, \tag{143}$$

which is independent of α ! Through demanding that the refractive index of the elastic mode vanish in a complex direction we have thus found a one-way interface state with a linear dispersion, independent of the parameter α , exactly as found for gyrotropic electromagnetic materials using a topological argument.

8.3 Periodic Media

Finally, let’s apply this idea to a periodic planar material. As the material is not homogeneous, it is not obvious whether the concept of the refractive index can be applied at all. The closest

we can get is to consider the Bloch vector \mathbf{K} , which is analogous to the wave-vector \mathbf{k} in a homogeneous medium. In this case, zero index in a given direction occurs when \mathbf{K} vanishes along one or more directions. Alternatively we can expand the dispersion relation around points \mathbf{K}_b on the Brillouin zone boundary, $\mathbf{K} = \mathbf{K}_b + \delta\mathbf{K}$. The change in the mode's frequency $\delta\omega$ as a function of the deviation $\delta\mathbf{K}$ from the zone boundary can then also be considered analogous to the dispersion relation in a homogeneous medium, and when one or more components of $\delta\mathbf{K}$ vanish, this is analogous to a point of zero index.

We consider the two dimensional Helmholtz equation, governing the behaviour of a TE polarized electromagnetic wave in a periodic permittivity profile $\epsilon(\mathbf{x})$

$$[\nabla^2 + k_0^2\epsilon(\mathbf{x})]\phi(\mathbf{x}) = 0 \tag{144}$$

although the same equation can also describe elastic and acoustic pressure waves. Suppose the profile $\epsilon(\mathbf{x})$ is such that two modes, $\phi_1(\mathbf{x})$ and $\phi_2(\mathbf{x})$, have degenerate frequencies ω at point \mathbf{K} in the Brillouin zone. To (144) we add a perturbation $\delta\epsilon$ to the permittivity. Then to examine small deviations away from this point in the Brillouin zone we expand the field as a sum of the two modes

$$\phi(\mathbf{x}) = a_1(\mathbf{x})\phi_1(\mathbf{x}) + a_2(\mathbf{x})\phi_2(\mathbf{x}) \tag{145}$$

where the expansion coefficients a_1 and a_2 vary slowly in position compared to the two solutions ϕ_1 and ϕ_2 . Substituting (145) into (144), and dropping derivatives of $a_{1,2}$ beyond the first we have,

$$2\nabla\phi_1 \cdot \nabla a_1 + 2\nabla\phi_2 \cdot \nabla a_2 + k_0^2\delta\epsilon[a_1\phi_1 + a_2\phi_2] = -2k_0\delta k_0\epsilon[a_1\phi_1 + a_2\phi_2] \tag{146}$$

For a fixed value of \mathbf{K} , the non-degenerate modes of (144) obey $\int \epsilon\phi_i\phi_j^*d^2\mathbf{x} = \delta_{ij}$, where the integral is taken over a unit cell of the medium. We are free to choose our two degenerate modes $\phi_{1,2}$ to obey the same condition. Taking the inner product of (146) with ϕ_1^* and ϕ_2^* , and neglecting the variation of the expansion coefficients $a_{1,2}$ over the unit cell we obtain two equations that can be written as a single vector differential equation

$$-\mathbf{i}\boldsymbol{\alpha} \cdot \nabla |\psi\rangle + m|\psi\rangle = \frac{\delta k_0}{k_0}|\psi\rangle \tag{147}$$

where the 'wavefunction' is defined as $|\psi\rangle = (a_1, a_2)^T$, and we have introduced three matrices α_j that form the vector of matrices $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$,

$$\alpha_j = -\frac{i}{k_0^2} \begin{pmatrix} \int \phi_1^* \partial_j \phi_1 d^2\mathbf{x} & \int \phi_1^* \partial_j \phi_2 d^2\mathbf{x} \\ \int \phi_2^* \partial_j \phi_1 d^2\mathbf{x} & \int \phi_2^* \partial_j \phi_2 d^2\mathbf{x} \end{pmatrix} \tag{148}$$

and the 'mass' matrix

$$m = -\frac{1}{2} \begin{pmatrix} \int \phi_1^* \delta\epsilon \phi_1 d^2\mathbf{x} & \int \phi_1^* \delta\epsilon \phi_2 d^2\mathbf{x} \\ \int \phi_2^* \delta\epsilon \phi_1 d^2\mathbf{x} & \int \phi_2^* \delta\epsilon \phi_2 d^2\mathbf{x} \end{pmatrix}. \tag{149}$$

The latter is named as such due to the similarity between (147) and the Dirac equation (see [47] and [48] for a complementary discussion).

The two-fold degeneracy at the point \mathbf{K} is assumed to arise from a symmetry of the lattice¹³. Assuming a $2\pi/N$ rotational symmetry, the two degenerate eigenfunctions will either be invariant under the rotation, or will become mixed up by it, in the same way as the components of a two dimensional vector after the application of the rotation matrix \mathbf{R} . We take the modes ϕ_1 and ϕ_2 to be eigenfunctions of this rotation matrix, which has eigenvalues

¹³ For a discussion of degeneracies, symmetries, and group theory see e.g. [53] and [54]

$\exp(\pm 2\pi i/N)$. With this choice the modes transform under rotation as $\phi_1 \rightarrow \exp(2\pi i/N) \phi_1$ and $\phi_2 \rightarrow \exp(-2\pi i/N) \phi_2$ ¹⁴.

We can use this behaviour of the modes under rotation to deduce the form of the matrix elements appearing in Eqns. (148) and (149). For instance, the off-diagonal elements of the ‘mass’ matrix must take the same value if we use the rotated coordinate system $\mathbf{x}' = \mathbf{R}^T \cdot \mathbf{x}$,

$$\begin{aligned} \int \phi_1^*(\mathbf{x}) \delta\epsilon(\mathbf{x}) \phi_2(\mathbf{x}) d^2\mathbf{x} &= \int \phi_1^*(\mathbf{R} \cdot \mathbf{x}') \delta\epsilon(\mathbf{R} \cdot \mathbf{x}') \phi_2(\mathbf{R} \cdot \mathbf{x}') d^2\mathbf{x} \\ &= e^{-4\pi i/N} \int \phi_1^*(\mathbf{x}') \delta\epsilon(\mathbf{x}') \phi_2(\mathbf{x}') d^2\mathbf{x}' \end{aligned} \tag{150}$$

where \mathbf{R} is the two dimensional rotation matrix, for rotation by $2\pi/N$, and we assume that $\delta\epsilon$ takes the same form after this rotation. Equation (150) implies that the integral is zero, unless we have an $N = 2$ fold symmetry, leaving only the possibility of $N = 3, 4,$ and 6 fold lattice symmetry¹⁵. We assume $N > 2$. Similarly the diagonal elements of the α matrix must obey

$$\begin{aligned} -i \int \phi_1^*(\mathbf{x}) \nabla \phi_1(\mathbf{x}) d^2\mathbf{x} &= -i \int \phi_1^*(\mathbf{R} \cdot \mathbf{x}') \mathbf{R}^T \cdot \nabla' \phi_1(\mathbf{R} \cdot \mathbf{x}') d^2\mathbf{x} \\ &= -i \mathbf{R}^T \cdot \int \phi_1^*(\mathbf{x}') \nabla' \phi_1(\mathbf{x}') d^2\mathbf{x}' \end{aligned} \tag{151}$$

implying that the matrix element written on the left of (151) is an eigenfunction of the rotation matrix \mathbf{R}^T , with unit eigenvalue. As the eigenvalues of the rotation matrix are $\exp(\pm 2\pi i/N)$, the matrix element itself must be zero! The two conditions (150) and (151) imply that the Dirac-like equation (147) takes the form

$$\begin{pmatrix} 0 & -\mathbf{e} \cdot i\nabla \\ -\mathbf{e}^* \cdot i\nabla & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} + \begin{pmatrix} m & 0 \\ 0 & -m \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \lambda \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \tag{152}$$

where the complex vector \mathbf{e} is defined as the integral

$$\mathbf{e} = -\frac{i}{k_0^2} \int \phi_1^*(\mathbf{x}) \nabla \phi_2(\mathbf{x}) d^2\mathbf{x}. \tag{153}$$

The ‘mass’, m equals the difference in the ‘averaged’ values of the permittivity perturbation

$$m = \frac{1}{4} \left[\int \phi_2^* \delta\epsilon \phi_2 d^2\mathbf{x} - \int \phi_1^* \delta\epsilon \phi_1 d^2\mathbf{x} \right] \tag{154}$$

and the ‘energy’ eigenvalue λ equals the relative frequency shift of the mode plus the sum of the ‘averaged’ permittivity values,

$$\lambda = \frac{\delta k_0}{k_0} + \frac{1}{4} \left[\int \phi_1^* \delta\epsilon \phi_1 d^2\mathbf{x} + \int \phi_2^* \delta\epsilon \phi_2 d^2\mathbf{x} \right]. \tag{155}$$

We have now reduced our Dirac-like equation (147) to (152): the exact form of the two dimensional Dirac equation [43]. In the special case where $m = \lambda$ (analogous to the point where $E = mc^2$ for a relativistic particle), (152) implies that

$$\mathbf{e} \cdot \nabla a_1 = 0 \tag{156}$$

¹⁴ Note this automatically makes the modes obey the orthogonality relation given below (146), via the same argument as given in (150).

¹⁵ See the crystallographic restriction theorem [55], which shows that only the rotation groups $C_1, C_2, C_3, C_4,$ and C_6 are consistent with translational symmetry.

i.e. the ‘refractive index’ is again zero in the complex direction \mathbf{e} ! In the case of periodic medium this means that the dispersion surface in the vicinity of the point \mathbf{K} in the Brillouin zone behaves similarly to the zero index limit of a homogeneous medium, as shown in Fig. 17.

But what is the value of \mathbf{e} ? This too can be deduced using the same symmetry arguments as above. The complex vector \mathbf{e} must obey

$$\begin{aligned} \mathbf{e} &= -\frac{i}{k_0^2} \int \phi_1^*(\mathbf{x}) \nabla \phi_2(\mathbf{x}) \, d^2\mathbf{x} = -\frac{i}{k_0^2} \mathbf{R}^T \cdot \int \phi_1^*(\mathbf{R} \cdot \mathbf{x}') \nabla' \phi_2(\mathbf{R} \cdot \mathbf{x}') \, d^2\mathbf{x}' \\ &= -\frac{i}{k_0^2} e^{-\frac{4\pi i}{N}} \mathbf{R}^T \cdot \int \phi_1^*(\mathbf{R} \cdot \mathbf{x}') \nabla' \phi_2(\mathbf{R} \cdot \mathbf{x}') \, d^2\mathbf{x}', \end{aligned}$$

and \mathbf{e} must therefore be an eigenvector of the inverse rotation matrix with eigenvalue $\exp(4\pi i/N)$. This is possible for an $N = 3$ fold symmetry, where the vector $\mathbf{e} = v(\mathbf{e}_x + i\mathbf{e}_y)$ is such an eigenvector (v is a positive real constant). Therefore, for the case of a doubly degenerate point in the Brillouin zone, with three fold symmetry, (156) reduces to the Cauchy–Riemann equations, $\partial a_1 / \partial \mathcal{Z}^* = 0$. This means that the spatially varying envelope of the wave in the lattice becomes an analytic function of position. On top of the standing wave at e.g. a point \mathbf{K} on the Brillouin zone boundary, we thus have a one–way circulation of the wave leading again to one–way interface states. Again, this reproduces the same result that would be obtained from a topological analysis of the wave in the vicinity of the degeneracy in the Brillouin zone.

Note that unlike the case of homogeneous media, here we only considered a region close to some point of the dispersion relation. Therefore our theory says nothing about the total number of interface states. From the perspective of a topological calculation the analogue is the ‘valley’ Chern number [49] (computed as an integral over a region of the Brillouin zone, rather than the full zone), which we can thus see records the presence of such points of zero index.

Example: The Jackiw–Rebbi state and Cauchy–Riemann conditions Through introducing a slow spatial variation of the perturbation $\delta\epsilon$, one–way modes (in the limited sense discussed immediately above) can be confined to propagate within the region of the inhomogeneity. We now show this for the special case of an interface between two materials where the wave behaves as an analytic function of \mathcal{Z} and \mathcal{Z}^* respectively.

For a uniform perturbation to the permittivity $\delta\epsilon$ and fixed propagation constant \mathbf{k} , the Dirac equation (147) reduces to

$$\begin{aligned} \begin{pmatrix} m & -iv \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \\ -iv \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) & -m \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} &= \begin{pmatrix} m & v(k_x + ik_y) \\ v(k_x - ik_y) & -m \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \\ &= \lambda \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \end{aligned} \tag{157}$$

The eigenvalues of (157) therefore fix the propagation constant \mathbf{k} to obey $\lambda = \pm(v^2\mathbf{k}^2 + m^2)^{1/2}$. We have seen this form of dispersion relation many times now! Not only is this the counterpart of the relativistic dispersion relation, $E = \pm(m^2c^4 + p^2)^{1/2}$, the same form governs the elastic (142), and electromagnetic (92) waves discussed above. As in all those cases the dispersion relation is that illustrated in Fig. 14: we have two regions of allowed propagation, where the magnitude of the ‘energy’ is greater than the rest energy, separated by a band gap. At the edge of the band gap the wave becomes an analytic function of position and accordingly exhibits one–way propagation.

Suppose that $\delta\epsilon$ is such that it changes sign under an inversion of the (x, y, z) coordinate system $\delta\epsilon \rightarrow -\delta\epsilon$. Such an inversion reverses the sense of rotation, and must interchange the two modes ϕ_1 and ϕ_2 (i.e. it is equivalent to swapping the eigenvectors of the rotation matrix). Thus, for this particular form of perturbation

$$\int \phi_1^* \delta\epsilon \phi_1 d^2\mathbf{x} = - \int \phi_2^* \delta\epsilon \phi_2 d^2\mathbf{x} \tag{158}$$

implying that Eqns. (154) and (155) simply reduce to, $\lambda = \delta k_0/k_0$, and $m = (1/2) \int \phi_2^* \delta\epsilon \phi_2 d^2\mathbf{x}$. If m changes as a function of x alone, homogeneous at infinity and smoothly changing from $-\lambda$ to $+\lambda$, then there is a general solution to (157)

$$|\psi\rangle = e^{-\frac{1}{v} \int_0^x m(x') dx' + ik_y y} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \tag{159}$$

which holds only for $\lambda = -vk_y$, and hence only a negative phase velocity. Assuming $|m| = \lambda$ as $|x| \rightarrow \infty$, the mode (159) takes the form

$$|\psi\rangle \propto e^{-|k_y|(|x|-iy)} \tag{160}$$

i.e. an analytic function of \mathcal{Z} or \mathcal{Z}^* , depending on which side of the interface we are considering. This mode is a special case of the one-way propagation Jackiw–Rebbi mode of the two dimensional Dirac equation [50]. Again, the idea of enforcing one-way propagation of a wave through demanding analyticity—in this case finding unidirectional interface state through connecting materials where the wave behaves as a function of \mathcal{Z} and \mathcal{Z}^* —is a simple shortcut to findings that are ordinarily connected with topological arguments [51].

9 Concluding Remarks

Topology is a deep subject that a physicist can easily get lost in. At the beginning of the tutorial we spent some time building up some small foundations of topology, for the special case of vectors living on closed surfaces. As many physicists do not receive any education in this area of mathematics, we hope that this will be a useful introduction, clarifying the origins of the infamous “Chern number”, in addition to the connection between this invariant and the number of trapped wave at an interface.

Although powerful, in the author’s view it is also problematic that the prediction of topological interface states lacks any local information about either the wave behaviour, or the properties of the interface. This makes it very difficult to understand the origins of these one-way interface states. What is it about the material that leads to this one way propagation, and could we have predicted these states without using topology?

The Chern number records the number and type of critical points in the Berry connection over e.g. the torus corresponding to the first Brillouin zone. In the second half of the tutorial we showed in some examples that these critical points correspond to points where the refractive index is zero. In terms of crystal optics the refractive index at these points vanishes in a complex direction, e.g. $e_x + ie_y$. This is equivalent to the wave satisfying the Cauchy–Riemann conditions and thus circulating in only one direction, which is the origin of the one-way propagation of the interface states found from a topological calculation. Finding these zero index points can thus be used as a shortcut to find one-way propagating interface modes, as shown in our examples in electromagnetic materials, elastic continua, and periodic materials. The reader may find this a useful alternative to standard topological calculations.

Acknowledgements SARH acknowledges funding from the Royal Society and TATA (RPG-2016-186) as well as helpful discussions with James Capers and Dean Patient.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Notes on Differential Geometry

In this appendix we give a sketch of the relationship between differential geometry and the Berry connection, explaining how the Berry connection and the associated Berry curvature should be generalized when we are dealing with an arbitrary number of basis vectors on a surface of arbitrary dimension. Note that, as in the main text, the summation convention between repeated vector indices is usually assumed, but is sometimes not used, for clarity.

To start, let’s consider the case of real vectors. In Secs. 2 and 3 we referred to the quantity A in (13) as a ‘vector potential’, given its similar role in electromagnetism. More precisely A_j is the *connection* on the surface, telling us how the vector basis changes as we move from point to point. To understand this, take a tangent vector V on one of the two dimensional surfaces discussed in Sec. 3. Expanding the vector in terms of its two components V_1, V_2 in the orthonormal basis e_1, e_2 , its differential with respect to the coordinates x_j is

$$\frac{\partial V}{\partial x_j} = \sum_{n=1,2} \left[\frac{\partial V_n}{\partial x_j} e_n + V_n \frac{\partial e_n}{\partial x_j} \right] = \sum_{n=1,2} \left[\frac{\partial V_n}{\partial x_j} + \sum_{m=1,2} e_n \cdot \frac{\partial e_m}{\partial x_j} V_m \right] e_n. \tag{161}$$

This expression contains two terms: the first records the change in the *vector components*, and the second records the change in the *basis directions*. This second term is the connection (a three index object $\Gamma_{nmj} = e_n \cdot \partial_j e_m$), telling us how the direction of the vector changes even when its components V_n are constant. Equation (161) is a particular case of the *covariant derivative* used in differential geometry [23]. Taking the inner product of (161) with respect to the two basis vectors, we have two equations expressing the rate of change of the vector in the two directions,

$$\begin{pmatrix} e_1 \cdot \frac{\partial V}{\partial x_j} \\ e_2 \cdot \frac{\partial V}{\partial x_j} \end{pmatrix} = \frac{\partial}{\partial x_j} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} + A_j \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}. \tag{162}$$

where the expression for the ‘vector potential’, $A_j = e_1 \cdot \partial_j e_2$ (introduced in (13)) has been applied. It is thus clear that, besides the constant matrix $i\sigma_y$, the connection is ‘the vector potential’. Rewriting the two real equations as a single complex one, and using the equivalent expression for the Berry connection (24), we can see even more clearly that A_j represents the connection on the surface

$$(e_1 + ie_2) \cdot \frac{\partial V}{\partial x_j} = \frac{\partial}{\partial x_j} (V_1 + iV_2) - iA_j (V_1 + iV_2). \tag{163}$$

Here the connection on the complex line bundle $e_1 - ie_2$ simply equals $-iA_j$. Suppose we want to move the vector and keep it pointing in the same direction. This is known as *parallel*

transport. For this we need the left hand sides of (162) and (163) to be zero, which means that the change in the vector components V_n must compensated by the change in the basis vectors. The equation for parallel transport is thus

$$\frac{\partial}{\partial x_j}(V_1 + iV_2) - iA_j(V_1 + iV_2) = 0. \tag{164}$$

The *Berry curvature* measures the change in the complex function $V = V_1 + iV_2$ as we attempt to keep it parallel and move it around an infinitesimal loop in the (x_1, x_2) coordinates. We can see this directly if we calculate the change in the logarithm of the vector as we parallel transport it around a closed trajectory,

$$\Delta \log(V) = \oint \frac{1}{V} \frac{\partial \log(V)}{\partial x_j} dx_j = i \oint A_j dx_j = i \int \left(\frac{\partial A_1}{\partial x_2} - \frac{\partial A_2}{\partial x_1} \right) d^2x. \tag{165}$$

Writing $V = |V|\exp(i\phi)$, and substituting it in (165) we have for a small loop, the change in the phase of the vector equals

$$\Delta\phi = \left(\frac{\partial A_1}{\partial x_2} - \frac{\partial A_2}{\partial x_1} \right) \times \text{Area} = \Omega_{ij}\epsilon_{ij} \times \text{Area}. \tag{166}$$

where we have introduced the ‘*curvature form*’

$$\Omega_{ij} = \frac{1}{2} \left(\frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i} \right) \tag{167}$$

When translated back into the real vector components $V_{1,2}$, (166) equates to a rotation of the vector by an angle $\Delta\phi$ after translation around the loop.

Let’s now generalise our concept of the connection and the curvature to N complex vector fields $|n\rangle$ attached to a surface of any dimension. Writing the complex vector on the surface as $V = \sum_n V_n |n\rangle$ and differentiating with respect to the coordinate x_j ,

$$\begin{aligned} \frac{\partial V}{\partial x_j} &= \sum_{n=1}^N \frac{\partial V_n}{\partial x_j} |n\rangle + \sum_{n=1}^N V_n \frac{\partial}{\partial x_j} |n\rangle \\ &= \sum_{n=1}^N \left[\frac{\partial V_n}{\partial x_j} + \sum_{m=1}^N \langle n | \frac{\partial}{\partial x_j} |m\rangle V_m \right] |n\rangle. \end{aligned} \tag{168}$$

Comparison of Eqs. (168) with (163) shows the term $\langle n | \partial_j |m\rangle$ is the generalization of the term iA_j . However the components of A are no longer numbers at each point on the surface, but $N \times N$ matrices with elements

$$A_j^{(nm)} = i\langle n | \partial_j |m\rangle \tag{169}$$

an expression that reduces to the ordinary Berry connection (24) when there is only one complex vector on the surface. We can again parallel transport the vector V by setting the left hand side of (168) to zero, giving

$$\frac{\partial V_n}{\partial x_j} = i \sum_{m=1}^N A_j^{(nm)} V_m. \tag{170}$$

The analogue of the Berry curvature can again be found through integrating the change in the vector as we move around a small closed loop. As we can’t use the Stokes theorem for ordinary vectors for the connection $A_j^{(nm)}$, we take the loop to be spanned by the two

infinitesimal vectors dx_j and dy_j , and take the difference between parallel transport along these two vectors in opposite order. First translating along dx_j , the vector changes to

$$\begin{aligned} V_n(x_j + dx_j) &= V_n(x_j) + dx_l \frac{\partial V_n(x_j)}{\partial x_l} \\ &= V_n(x_j) + i dx_l \sum_{m=1}^N A_l^{(nm)} V_m(x_j). \end{aligned} \tag{171}$$

and then along dy_j we have

$$\begin{aligned} V_n(x_j + dx_j + dy_j) &= V_n(x_j + dx_j) + dy_l \frac{\partial V_n(x_j + dx_j)}{\partial x_l} \\ &= V_n(x_j) + i dx_l \sum_{m=1}^N A_l^{(nm)} V_n(x_j) + i dy_l \sum_{m=1}^N A_l^{(nm)} V_m(x_j) \\ &\quad + i dx_q dy_l \sum_{m=1}^N \left[\frac{\partial A_q^{(nm)}}{\partial x_l} + i \sum_{p=1}^N A_q^{(np)} A_l^{(pm)} \right] V_m(x_j) \end{aligned} \tag{172}$$

Equation (172) tells us the vector after translation around half the loop. Swapping the order of the vectors dx_l and dy_m , then gives the change in the vector for traversing half the loop in the opposite direction. Taking the difference thus gives the change in the vector after completing the full loop. This gives,

$$\Delta V_n = i dA_{ql} \sum_m \left[\frac{\partial A_q^{(nm)}}{\partial x_l} - \frac{\partial A_l^{nm}}{\partial x_q} + i \sum_p [A_q^{(np)} A_l^{(pm)} - A_l^{(np)} A_p^{(pm)}] \right] V_m \tag{173}$$

where the area element is $dA_{ql} = \frac{1}{2}[dx_q dy_l - dx_l dy_q]$. Comparing (173) to the similar expression (165), we can identify the generalization of the curvature form (167) as

$$\Omega_{ij} = \frac{1}{2} \left(\frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i} + i[A_i, A_j] \right) \tag{174}$$

where we have suppressed the matrix indices of the connection A_i , and $[,]$ is a matrix commutator.

Finally, suppose that the basis vectors can be written as a unitary transformation from a basis $|\vec{m}\rangle$ that does not vary over the surface, $|n\rangle = \sum_m U_{nm} |\vec{m}\rangle$. Substituting this basis into the expression for the connection (169), yields

$$A_j^{(nm)} = i \sum_{m,p} U_{mn}^* \partial_j U_{mp} \langle \vec{m} | \vec{p} \rangle \rightarrow A_j = iU^\dagger \partial_j U \tag{175}$$

where in the final step we used the orthogonality condition $\langle \vec{m} | \vec{p} \rangle = \delta_{mp}$ and suppressed the matrix indices. The reader can verify that substituting the special form of connection (175) into the curvature form (174) gives zero.

Appendix B: Finite Difference Calculation of Interface States

The reader may find it useful to have a simple program where these ‘topologically robust’ interface states can be simulated. Below we give a Python program that was implemented

as a Jupyter notebook to produce Fig. 15. This solves Maxwell's equations in the spatially varying gyrotropic medium discussed in the final example of Sec. 6. Explicitly it solves the equation

$$\nabla \cdot (\epsilon^{-1} \cdot \nabla h) + k_0^2 h = \nabla \times (\epsilon^{-1} \eta_0 \mathbf{j}) = S(\mathbf{x}) \quad (176)$$

where $S(\mathbf{x})$ is the spatial distribution of the source of waves. Writing

$$\epsilon = \begin{pmatrix} \lambda & -i\alpha \\ i\alpha & \lambda \end{pmatrix} \rightarrow \epsilon^{-1} = \frac{1}{\lambda^2 - \alpha^2} \begin{pmatrix} \lambda & i\alpha \\ -i\alpha & \lambda \end{pmatrix} \quad (177)$$

the equation our program has to solve becomes

$$\nabla \cdot \left(\frac{\lambda}{\lambda^2 - \alpha^2} \cdot \nabla h \right) + i\mathbf{e}_z \cdot \nabla \times \left(\frac{\alpha}{\lambda^2 - \alpha^2} \nabla h \right) + k_0^2 h = S(\mathbf{x}) \quad (178)$$

This is the equation solved by the program below.

```
import numpy as np
import scipy.sparse as sp
import scipy.sparse.linalg as spla
%matplotlib widget

# Parameters
k0=1.0 # Wavenumber
lm=2.0*np.pi/k0 # Wavelength
L=20*lm # LxL simulation box
N=800 # NxN points

# Simulation grid
xv=np.linspace(0,L,N,endpoint=True)
yv=np.linspace(0,L,N,endpoint=True)
X,Y=np.meshgrid(xv,yv)

# Coordinate difference
dx=xv[1]-xv[0]
dy=yv[1]-yv[0]

# Identity matrices
IN=sp.identity(N)
INN=sp.identity(N*N)

# Material parameters
Apml=1.0 # Maximum amplitude of Im[eps] of absorbing layers
dpml=lm # Decay length of absorbing layers
# Absorbing layers to terminate simulation
def pml(x,y):
    pmlx=1j*Apml*(np.exp(-x/dpml)+np.exp((x-L)/dpml))
    pmlly=1j*Apml*(np.exp(-y/dpml)+np.exp((y-L)/dpml))
    return pmlx+pmlly

# Diagonal permittivity components
def eps(x,y):
```

```

    return 2.5+0.001*1j+pml(x,y)

dalphi=0.2*lm # length scale of interface transition

# Position of interface y(x)
def bdry(x):
    return 0.5*L+0.1*L*(np.tanh((x-0.75*L)/dalphi)+1.0)

# Gyrotropy, alpha
def alpha(x,y):
    return 3.0*np.tanh((y-bdry(x))/dalphi)

# Source
x0=0.25*L # x position
y0=0.5*L # y position
d=0.1*lm # size
def J(x,y):
    return (0.1/(dx*dy))*np.exp(-((x-x0)/d)**2 -
        ((y-y0)/d)**2)

# Construct differential operators
# Material matrices
epv=eps(X,Y)
alv=alpha(X,Y)
t1=epv/(epv**2 - alv**2)
t2=alv/(epv**2 - alv**2)
aop=sp.diags(t1.flatten())
bop=sp.diags(t2.flatten())

# 1 Dimensional derivatives
dx=(np.eye(N,k=1)-np.eye(N,k=0))/dx
dx=(np.eye(N,k=0)-np.eye(N,k=-1))/dx
dy=(np.eye(N,k=1)-np.eye(N,k=0))/dy
dy=(np.eye(N,k=0)-np.eye(N,k=-1))/dy

# 2 dimensional derivatives
Dx=sp.kron(IN,dx)
Dx=sp.kron(IN,dx)
Dy=sp.kron(dy,IN)
Dy=sp.kron(dy,IN)

# Helmholtz operator
H1=Dx@Dx+Dy@Dy
H2=1j*(Dx@bop@Dy - (Dy@bop@Dx))
H3=(k0**2)*INN
Hop=H1+H2+H3

# Source array
Jv=J(X,Y).flatten()

```

```
# Find fields
```

```
Hv=spla.spsolve(Hop,Jv) # Find H-field
```

```
Hv=Hv.reshape(N,N) # Reshape vector to NxN grid
```

The final quantity 'Hv' is an $N \times N$ complex array of field values at the different x and y coordinates, generated from a source at position 'x0','y0'.

References

1. Lu, L., Joannopoulos, J., Soljacić, M.: Topological photonics. *Nat. Phot.* **8**, 821 (2014)
2. Arnold, V.I.: *Topological Methods in Hydrodynamics*, Springer-Verlag (1998)
3. Irvine, W., Bouwmeester, D.: Linked and knotted beams of light. *Nat. Phys.* **4**, 716 (2008)
4. Volkov, B.A., Pankratov, O.A.: Two-dimensional massless electrons in an inverted contact. *JETP Lett.* **42**, 145 (1985)
5. Pankratov, O., Pakhomov, S., Volkov, B.: Supersymmetry in heterojunctions: Bandinverting contact on the basis of $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ and $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$. *Solid State Commun.* **61**, 93 (1987)
6. Kane, C.L., Mele, E.J.: Z_2 Topological order and the quantum spin hall effect. *Phys. Rev. Lett.* **95**, 146802 (2005)
7. Bernevig, B.A., Zhang, S.-C.: Quantum spin hall effect. *Phys. Rev. Lett.* **96**, 106802 (2006)
8. Haldane, F.D.M.: Model for a quantum hall effect without landau levels: condensed-matter realization of the "Parity Anomaly". *Phys. Rev. Lett.* **61**, 2015 (1988)
9. Thouless, D.J., Kohmoto, M., Nightingale, M.P., den Nijs, M.: Quantized hall conductance in a two-dimensional periodic potential. *Phys. Rev. Lett.* **49**, 405 (1982)
10. Pendry, J.B.: Negative refraction makes a perfect lens. *Phys. Rev. Lett.* **85**, 3966 (2000)
11. Pendry, J.B., Schurig, D., Smith, D.R.: Controlling electromagnetic fields. *Science* **312**, 1780 (2006)
12. Zheludev, N., Kivshar, Y.: From metamaterials to metadevices. *Nature Mater.* **11**, 917 (2012)
13. Haldane, F., Raghu, S.: Possible realization of directional optical waveguides in photonic crystals with broken time-reversal symmetry. *Phys. Rev. Lett.* **100**, 013904 (2008)
14. Ozawa, T., Price, H.M., Amo, A., Goldman, N., Hafezi, M., Lu, L., Rechtsman, M.C., Schuster, D., Simon, J., Zilberberg, O Carusotto, I.: Topological photonics. *Rev. Mod. Phys.* **91**, 015006 (2019)
15. Ma, G., Xiao, M., Chan, C.T.: Topological phases in acoustic and mechanical systems. *Nat. Rev. Phys.* **1**, 281 (2019)
16. Kim, M., Jacob, Z., Rho, J.: Recent advances in 2d, 3d and higher-order topological photonics. *Light Sci. Appl.* **9**, 130 (2020)
17. Mordechai, S., Bandres, M.A.: Topological photonics: Where do we go from here? *Nanophotonics* **10**, 425 (2021)
18. Ghatak, A., Brandenbourger, M., van Wezel, J., Coullais, C.: Observation of non-hermitian topology and its bulk-edge correspondence in an active mechanical metamaterial. *Proc. Natl. Academy Sci.* **117**, 29561 (2020)
19. Berry, M.V.: Quantal phase factors accompanying adiabatic changes. *Proc. Roy Soc. A* **392**, 45 (1984)
20. Poddubny, A., Iorsh, I., Belov, P., Kivshar, Y.: Hyperbolic metamaterials. *Nat. Photon.* **7**, 948 (2013)
21. Berry, M.V., Dennis, M.R., Lee, R.L.: Polarization singularities in the clear sky. *New J. Phys.* **6**, 162 (2004)
22. Liu, W., Liu, W., Shi, L., Kivshar, Y.: Topological polarization singularities in metaphotonics. *Nanophotonics* **10**, 1469 (2021)
23. Nakahara, M.: *Geometry, Topology, and Physics*, Taylor and Francis (2003)
24. Hu, S.: *Lecture Notes in Chern-Simons-Witten Theory*, World Scientific (2001)
25. Lu, L., Gao, H., Wang, Z.: Topological one-way fiber of second chern number. *Nat. Commun.* **9**, 5385 (2018)
26. Chen, Z.-C., Zhu, W., Tan, Y., Wang, L., Ma, G.: Acoustic realization of a four-dimensional higher-order chern insulator and boundary-modes engineering. *Phys. Rev. X* **11**, 011016 (2021)
27. Landau, L.D., Lifshitz, E.M., Pitaevskii, L.P.: *Electrodynamics of Continuous Media*, Butterworth-Heinemann (2004)
28. Silveirinha, M.G.: Chern invariants for continuous media. *Phys. Rev. B* **92**, 125153 (2015)
29. Horsley, S.A.R.: Topology and the optical Dirac equation. *Phys. Rev. A* **98**, 043837 (2018)
30. Volovik, G.: *The Universe in a Helium Droplet*, Oxford University Press (2009)
31. Lanczos, C.: *Linear Differential Operators*, Society for Industrial and Applied Mathematics (1996)

32. Harrison, W.A.: *Electronic Structure and Properties of Solids*, Dover (2012)
33. Leonhardt, U.: Optical conformal mapping. *Science* **312**, 1777 (2006)
34. Shastri, K., Abdelrahman, M., Monticone, F.: Nonreciprocal and topological plasmonics. *Photonics* **8**, 133 (2021)
35. Ali Hassani Gangaraj, S., Monticone, F.: Physical violations of the bulk-edge correspondence in topological electromagnetics. *Phys. Rev. Lett.* **124**, 153901 (2020)
36. Horsley, S.A.R., Woolley, M.: Zero-refractive-index materials and topological photonics. *Nat. Phys.* **17**, 348 (2021)
37. Needham, T.: *Visual Complex Analysis*, Clarendon Press (1998)
38. Mansuripur, M.: *Classical Optics and its Applications*, Cambridge University Press (2002)
39. Mackay, T.G., Lakhtakia, A.L.: *Electromagnetic Anisotropy and Bianisotropy: A Field Guide*, World Scientific Publishing (2010)
40. Barnett, S.M.: Optical Dirac equation. *New J. Phys.* **16**, 093008 (2014)
41. Horsley, S.A.R.: Indifferent electromagnetic modes: bound states and topology. *Phys. Rev. A* **100**, 053819 (2019)
42. Mechelen, T.V., Jacob, Z.: Photonic Dirac monopoles and skyrmions: spin-1 quantization. *Opt. Mater. Express* **9**, 95 (2019)
43. Thaller, B.: *The Dirac Equation*, Springer (2013)
44. Barnes, B., Horsley, S.A.R., Vos, W.L.: Classical antennas, quantum emitters, and densities of optical states. *J. Opt.* **22**, 073501 (2020)
45. Landau, L.D., Lifshitz, E.M.: *Theory of Elasticity*, Butterworth-Heinemann (2007)
46. Norris, A.N., Shuvalov, A.L., Kutsenko, A.A.: Analytical formulation of three-dimensional dynamic homogenization for periodic elastic systems. *Proc. Roy. Soc. A* **468**, 1629 (2012)
47. Makwana, M.P., Craster, R.V.: Geometrically navigating topological plate modes around gentle and sharp bends. *Phys. Rev. B* **98**, 184105 (2018)
48. Makwana, M.P., Craster, R.V.: Designing multidirectional energy splitters and topological valley super-networks. *Phys. Rev. B* **98**, 235125 (2018)
49. Zhang, F., MacDonald, A.H., Mele, E.J.: Valley chern numbers and boundary modes in gapped bilayer graphene. *Proc. Natl. Academy Sci.* **110**, 10546 (2013)
50. Jackiw, R., Rebbi, C.: Solitons with fermion number 1/2. *Phys. Rev. D* **13**, 3398 (1976)
51. Hasan, M.Z., Kane, C.L.: Colloquium: Topological Insulators. *Rev. Mod. Phys.* **82**, 3045 (2010)
52. Lovelock, D., Rund, H.: *Tensors, Differential Forms, and Variational Principles* Dover (1990)
53. Landau, L.D., Lifshitz, E.M.: *Quantum Mechanics*, Butterworth-Heinemann (2003)
54. Hamermesh, M.: *Group Theory and its Applications to Physical Problems*, Dover (2012)
55. Landau, L.D., Lifshitz, E.M.: *Statistical Physics: Part 1*, Butterworth-Heinemann (2005)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.